N° d'ordre: 03732

THÈSE

Présentée devant

devant l'Université de Rennes 1

pour obtenir

le grade de : DOCTEUR DE L'UNIVERSITÉ DE RENNES 1
Mention MATHÉMATIQUES ET APPLICATIONS

par

Guillaume DUJARDIN

Équipe d'accueil : Equipe IPSO - INRIA Bretagne Atlantique
École Doctorale : Matisse
Composante universitaire : UFR DE MATHÉMATIQUES

Titre de la thèse :

Étude de schémas de discrétisation en temps de l'équation de Schrödinger

soutenue le 13 juin 2008 devant la commission d'examen

M. :	Michel	CROUZEIX	Président
MM. :	Nicolas	LERNER	Rapporteurs
	Alexander	OSTERMANN	
MM. :	Éric	CANCÈS	Examinateurs
	Benoît	GRÉBERT	
	François	CASTELLA	Directeurs
	Erwan	FAOU	

Remerciements

En premier lieu, je tiens à remercier Nicolas Lerner et Alexander Ostermann qui ont accepté de rapporter sur ma thèse. Je remercie également Éric Cancès, Michel Crouzeix et Benoît Grébert qui m'ont fait l'honneur d'accepter d'être membres de mon jury de soutenance. Enfin, mes remerciements vont à mes deux directeurs de thèse, François Castella et Erwan Faou, dont la présence et la disponibilité tout au long de mes années de thèse n'ont jamais été mises en défaut.

Ces trois dernières années ont été pour moi l'occasion de nombreuses rencontres de membres de la communauté mathématique d'Analyse Numérique. Certaines d'entre elles m'ont beaucoup apporté et c'est avec plaisir que je remercie David Cohen, Stéphane Descombes, Ludwig Gauckler, Arieh Iserles, Christian Lubich et Brynjulf Owren pour leur amabilité et leur enthousiasme à chaque occasion.

Choisir un sujet et un encadrement de thèse n'a pas été une chose aisée pour moi. C'est pourquoi je remercie tout particulièrement Michel Crouzeix pour avoir eu l'excellente idée de me mettre en relation il y a un peu plus de trois ans avec l'équipe de Philippe Chartier à l'IRISA/INRIA. En effet, a posteriori, mon IPSO-compatibilité me laisse à penser qu'il eût été difficile pour moi d'intégrer une équipe où mon travail de thèse et mes premiers pas dans le monde de la recherche eussent été plus naturels.

Intégrer l'équipe IPSO, c'est aussi intégrer l'IRISA/INRIA, et bénéficier d'un confort de travail exceptionnel à plusieurs titres. Tout d'abord, la proximité de nombreux chercheurs aux thématiques variées favorise les échanges et les collaborations. Ainsi, je remercie chaleureusement Bruno Sericola pour les nombreuses discussions informatiques que nous avons pu avoir. Le confort de travail à l'IRISA/INRIA doit beaucoup à l'efficacité de son réseau d'assistants. À ce titre, je remercie les assistants avec qui j'ai eu le plaisir de travailler, tant pour la qualité de leur travail que pour leur amabilité. Merci donc à Angélina, Myriam, Fabienne, Sara, Loïc et Laurence. En outre, j'ai eu le plaisir de travailler avec les membres du service des Relations Extérieures, et je tiens à remercier Elisabeth Lebret et Violaine Tygréat pour leur capacité à travailler efficacement dans la bonne humeur. J'ai également bénéficié de la qualité des moyens informatiques mis à ma disposition par l'IRISA/INRIA et en particulier des services proposés par l'Atelier, dont je remercie chacun des membres pour sa patience, son professionnalisme et sa disponibilité. Enfin, une journée à l'IRISA/INRIA ne serait rien sans le sourire de Laurence, de Pierrette ou de Marie-Charlotte à la "CAFET' ". Merci à elles pour leur bonne humeur constante et communicatrice.

Mes années de thèse ont été l'occasion de contacts privilégiés avec les membres de l'IRMAR, de l'UFR de Mathématiques et de l'ÉNS de Cachan. Bien évidemment, mes remerciements vont plus particulièrement aux membres des équipes d'Analyse Numérique et d'Équations aux Dérivées Partielles, que j'ai naturellement plus souvent fréquentés, ainsi qu'aux membres de l'UFR avec qui j'ai eu le plaisir de découvrir l'enseignement supérieur à l'université. Je tiens à remercier en particulier Christophe Cheverry qui a été, trois années durant, mon tuteur de monitorat. En outre, je remercie aussi chaleureusement mes camarades doctorants de l'ADOC, de Binet et d'ailleurs, et plus spécifiquement Fanny, Jimmy et Ludovic qui m'ont accueilli si aimablement au plateau de mathématiques. Enfin, je remercie sincèrement Gilles pour tout ce qu'il a accepté de partager avec moi pendant ces années (oui, je laisse volontairement planer le mystère).

Il me faut remercier le Ministère de la Recherche de de l'Enseignement Supérieur, qui a accepté de financer mes trois années de thèse par un contrat d'allocataire de recherche couplé à un monitorat. L'IRISA/INRIA a également apporté une contribution financière majeure en me donnant accès à ses moyens informatiques et humains ainsi qu'en prenant en charge certaines de mes missions. En outre, je remercie le Conseil Général du Maine-et-Loire qui continue de financer les transports en commun entre Rennes et Angers.

Enfin, je remercie tout particulièrement mes parents, mon frère et ma grand-mère, dont le soutien moral a été constant au cours de ces trois années. Mon tout dernier remerciement va tout naturellement à Gladys, pour sa patience et sa capacité d'adaptation constantes, pour nous avoir fait le fils dont nous rêvions, et pour tout le reste également.

Table des matières

Introduction

Cette thèse consiste en l'analyse de quelques méthodes numériques de résolution d'équations aux dérivées partielles de type Schrödinger. On peut classer ces équations en trois groupes : Schrödinger linéaire avec potentiel multiplicatif, Schrödinger linéaire inhomogène et Schrödinger non linéaire ; dans ce travail, on considère que l'espace physique est le tore de dimension d. Les méthodes numériques analysées sont des méthodes de discrétisation en temps : aucune discrétisation n'est faite en espace. D'une part, nous considérons des méthodes de splitting symplectiques pour l'équation de Schrödinger linéaire avec potentiel. On s'intéresse notamment à l'estimation en temps long de la régularité de la solution numérique. D'autre part, on considère des méthodes de Runge-Kutta exponentielles de collocation pour la résolution d'équations de Schrödinger linéaires inhomogènes et non linéaires. On donne notamment des conditions suffisantes sur les points de collocation pour que les méthodes à s points soient d'ordres s, $s+1$ et $s+2$ dans ce contexte.

Cette thèse se compose de trois chapitres.

Chapitre 1 Le premier chapitre décrit la problématique et les résultats de l'ensemble de la thèse. Dans un premier temps, on présente quelques aspects de l'intégration numérique des systèmes hamiltoniens en dimension finie et en dimension infinie. Ces éléments motivent les travaux présentés dans cette thèse. Ensuite, on introduit les différentes équations de Schrödinger considérées : équation de Schrödinger linéaire avec potentiel, équation de Schrödinger linéaire inhomogène et équation de Schrödinger non linéaire. Celles-ci sont posées sur le tore de dimension d. Enfin, on présente les principaux résultats obtenus. Les résultats sont relatifs aux méthodes de splitting symplectiques pour l'équation de Schrödinger linéaire avec potentiel et concernent notamment l'estimation de la régularité de la solution numérique en temps long via un théorème de forme normale pour le propagateur numérique pour des pas de temps non résonnants. Les résultats concernent les méthodes de Runge-Kutta exponentielles de collocation appliquées aux problèmes de Schrödinger linéaires inhomogènes et aux problèmes de Schrödinger non linéaires. On donne en particulier des conditions suffisantes sur les points de collocation pour que les méthodes à s points soient d'ordre s, $s+1$ et $s+2$ pour des normes de Sobolev bien choisies. On illustre et explique également l'effet des résonances numériques lors de la résolution des problèmes de Schrödinger linéaires inhomogènes par de telles méthodes.

Chapitre 2 Le second chapitre détaille les résultats obtenus relatifs aux méthodes de splitting symplectiques pour l'équation de Schrödinger linéaire avec potentiel. On considère donc l'équation de Schrödinger linéaire avec potentiel multiplicatif et l'on intègre numériquement le problème de Cauchy par une méthode de splitting en temps : on ne fait aucune discrétisation en espace. Dans un premier temps, on montre un théorème de forme normale pour le propagateur de splitting pour des pas de temps non résonnants dans l'asymptotique des petits potentiels réguliers. Pour cela, on considère un potentiel régulier réel dont on mesure la taille par un petit paramètre et l'on adopte une approche perturbative. Ce résultat permet ensuite de démontrer un corollaire portant sur la régularité de la solution numérique obtenue par la méthode de splitting symplectique : dans l'asymptotique des petits potentiels, on montre que certaines semi-normes de la solution numérique demeurent uniformément bornées en temps exponentiellement longs pour des pas de temps non résonnants. Ces résultats sont illustrés par plusieurs expériences numériques.

Chapitre 3 Le dernier chapitre expose les résultats obtenus concernant les méthodes de Runge-Kutta exponentielles de collocation appliquées à l'équation de Schrödinger linéaire inhomogène puis à l'équation de Schrödinger non linéaire. Suivant la définition de ces méthodes et leur analyse dans le cas des problèmes paraboliques, on introduit ces méthodes pour l'équation de Schrödinger linéaire inhomogène et pour l'équation de Schrödinger non linéaire. On donne ensuite dans chacun de ces deux cas des conditions suffisantes sur les points de collocation pour qu'une méthode à s points soit d'ordre s, $s+1$ ou $s+2$ pour des normes de Sobolev bien choisies. En outre, dans le cas de la résolution des problèmes de Schrödinger linéaires inhomogènes par des méthodes de Runge-Kutta exponentielles, on illustre, explique et quantifie l'effet des résonances numériques observées, en particulier sur les courbes d'ordre de ces méthodes. On illustre également ces résultats par plusieurs expériences numériques.

Chapitre 1

Problématique et résultats

Cette partie introduit la problématique et les résultats obtenus dans cette thèse. Dans un premier temps, on présente quelques aspects de l'intégration numérique des systèmes hamiltoniens, en dimension finie ainsi qu'en dimension infinie, dans le but de motiver les travaux exposés dans la thèse. Ensuite, on décrit brièvement les différentes équations de Schrödinger que nous avons considérées. Enfin, on présente l'ensemble des résultats obtenus dans cette thèse : tous ressortent de l'analyse numérique de certaines méthodes de résolution en temps des équations de Schrödinger que l'on vient de présenter (méthodes de splitting symplectiques et méthodes de Runge-Kutta exponentielles).

1.1 Intégration numérique des systèmes hamiltoniens

On présente ici le formalisme et quelques résultats connus relatifs à l'intégration numérique des systèmes hamiltoniens. Dans une première partie, on présente ces résultats dans le cas où l'espace des phases est de dimension finie. Ensuite, on s'intéresse au cas de la dimension infinie.

1.1.1 Le cas de la dimension finie

Définition 1.1.1 *On appelle système hamiltonien (autonome, sous forme canonique) de dimension* $2d$ *(où* $d \in \mathbb{N}^*$*) tout système d'équations différentielles ordinaires s'écri-*

vant sous la forme

$$\left\{\begin{array}{lcl} \dot{p}_1(t) & = & -\dfrac{\partial H}{\partial q_1}(p(t),q(t)) \\ \vdots & \vdots & \vdots \\ \dot{p}_d(t) & = & -\dfrac{\partial H}{\partial q_d}(p(t),q(t)) \\ \dot{q}_1(t) & = & \dfrac{\partial H}{\partial p_1}(p(t),q(t)) \\ \vdots & \vdots & \vdots \\ \dot{q}_d(t) & = & \dfrac{\partial H}{\partial p_d}(p(t),q(t)) \end{array}\right. \tag{1.1}$$

où H est une fonction réelle dérivable définie sur un ouvert de $\mathbb{R}^{2d}$ et $p = (p_1, \dots, p_d)$ et $q = (q_1, \dots, q_d)$ sont deux d-uplets de fonctions réelles dérivables inconnues. La fonction H est alors appelée hamiltonien du système (1.1).

Remarque 1.1.1 *Sous forme matricielle, le système* (1.1) *s'écrit*

$$\begin{pmatrix} \dot{p} \\ \dot{q} \end{pmatrix} = J\nabla H(p,q),$$

où J est la matrice par blocs $\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$, I est la matrice identité de taille d et ∇H est le vecteur gradient de H écrit en colonne.

Dans ce contexte,

Définition 1.1.2 *On appelle transformation symplectique toute application différentiable g d'un ouvert de $\mathbb{R}^{2d}$ dans $\mathbb{R}^{2d}$ telle que, sur cet ouvert,*

$$[D(g)(p,q)]^t J[D(g)(p,q)] = J,$$

où $D(g)(p,q) = \frac{\mathrm{D}g}{\mathrm{D}(p,q)}$ représente la matrice jacobienne de g au point (p,q) et l'exposant t désigne la transposition.

Pour une interprétation géométrique de cette propriété, on pourra se référer à [HLW06], paragraphe VI.2.

Le flot φ_t du système (1.1) jouit de deux propriétés remarquables :

Proposition 1.1.1 *Supposons que H est une fonction de classe C^2 sur un ouvert de $\mathbb{R}^{2d}$, alors*

- *L'hamiltonien H est préservé par le flot φ_t : le long de chaque solution $t \mapsto (p(t), q(t))$ de* (1.1) *telle que $(p(0), q(0)) = (p_0, q_0)$, on a $H(p(t), q(t)) = H(p_0, q_0)$.*
- *Le flot φ_t est une transformation symplectique.*

Le caractère symplectique du flot caractérise de plus le caractère hamiltonien du système. On renvoie à [HLW06] pour plus de précisions.

Considérons maintenant une méthode numérique à un pas pour intégrer le problème (1.1). Pour $h > 0$ (éventuellement assez petit), notons Φ_h le flot numérique définissant la méthode.

Définition 1.1.3 *La méthode définie par Φ_h est dite symplectique lorsque pour tout $h > 0$ (éventuellement assez petit) l'application Φ_h est symplectique au sens de la Définition 1.1.2.*

Lorsque la méthode numérique à un pas choisie pour intégrer le problème (1.1) est symplectique et lorsque l'hamiltonien H de ce problème est suffisamment régulier, l'*analyse rétrograde* permet de montrer que la solution numérique obtenue possède elle aussi une propriété de bonne conservation de H :

Théorème 1.1.1 *Soit $\Omega \subset \mathbb{R}^{2d}$ un ouvert non vide sur lequel la fonction réelle H est analytique. Supposons que la méthode définie par Φ_h est d'ordre $p \in \mathbb{N}^*$. Considérons $(p_0, q_0) \in \Omega$ et posons pour tout $n \in \mathbb{N}$, $(p_{n+1}, q_{n+1}) = \Phi_h(p_n, q_n)$. Supposons que Φ_h est symplectique et qu'il existe un compact $K \subset \Omega$ tel que pour tout $h > 0$ et tout $n \in \mathbb{N}$, $(p_n, q_n) \in K$, alors il existe $h_0 > 0$ tel que*

$$\forall h \in]0, h_0[, \qquad \forall n \in \mathbb{N} \textit{ tel que } nh \leq e^{\frac{h_0}{h}}, \qquad H(p_n, q_n) = H(p_0, q_0) + \mathcal{O}(h^p).$$

Dans la pratique, en dimension finie, on vérifie au cas par cas l'hypothèse d'existence du compact K indépendant de $h > 0$ contenant toutes les solutions numériques issues de (p_0, q_0). Si l'on souhaite prouver un résultat analogue au Théorème 1.1.1 en dimension infinie en utilisant l'analyse rétrograde, il est souhaitable de pouvoir vérifier un équivalent de cette hypothèse de non sortie d'un compact. Pour cela, on peut introduire une norme analytique contrôlant toutes les dérivées de la solution numérique, et la question se pose de l'estimation de la croissance de cette norme en temps long. Le résultat présenté en 1.3.1 répond à cette question pour les méthodes de splitting (symplectiques) pour l'équation de Schrödinger linéaire avec potentiel.

Enfin, en dimension finie comme en dimension infinie, un moyen commode d'obtenir des méthodes numériques symplectiques est de considérer des méthodes de splitting. En effet, si chacun des propagateurs intervenant dans une méthode de splitting est symplectique, alors leur composition est symplectique. Cette remarque est l'une des motivations de l'étude des méthodes de splitting en temps pour l'équation de Schrödinger linéaire avec potentiel que nous présentons en 1.3.1.

1.1.2 Le cas de la dimension infinie

L'intégration numérique des équations aux dérivées partielles hamiltoniennes est un sujet de recherche actuellement très actif. L'analyse numérique des méthodes utilisées est en général menée au cas par cas, en fonction du type de problème hamiltonien auquel elles sont appliquées.

Par exemple, pour l'équation des ondes non linéaire en dimension 1 d'espace, l'étude de semi-discrétisations spectrales est menée dans [HL07] et l'étude du problème complètement discrétisé est menée dans [CHL07]. En particulier, les questions de conservation de l'hamiltonien (mais aussi des actions et des moments) sont étudiées dans ces deux cas et l'intérêt du caractère symplectique des méthodes considérées est illustré et expliqué. L'outil fondamental de l'analyse numérique des méthodes proposée dans [HL07] et [CHL07] est l'utilisation de séries de Fourier modulées en temps. Ce même outil permet également de faire l'analyse des solutions exactes des mêmes équations d'ondes non linéaires (voir [CHL06]). Notons que l'analyse fournie repose dans chaque cas sur une condition de non résonance géométrique liant la distribution des fréquences du système à la partie linéaire de l'équation et à la taille des données initiales considérées.

Dans un esprit similaire, notre travail consiste en l'étude de certaines méthodes numériques appliquées à certaines équations de Schrödinger (on renvoie à la section 1.2 pour une présentation précise des problèmes considérés, détaillant en particulier la nature hamiltonienne de certains d'entre eux). Dans un premier temps, nous portons notre attention sur les méthodes de splitting en temps symplectiques appliquées à l'équation de Schrödinger linéaire avec potentiel et nous analysons la conservation de la régularité des solutions numériques obtenues dans une perspective d'analyse qualitative en temps long. Nous introduisons également une condition de non résonance liant le pas de temps à la distribution des fréquences du système et nous illustrons ses effets sur le comportement en temps long de la régularité de la solution numérique. Les résultats que nous obtenons sont présentés en 1.3.1 et détaillés au Chapitre 2. Dans une seconde partie, nous considérons des méthodes de Runge-Kutta exponentielles que nous appliquons à l'équation de Schrödinger linéaire inhomogène ainsi qu'à l'équation de Schrödinger non linéaire. Nous donnons des conditions suffisantes pour que les méthodes de Runge-Kutta exponentielles de collocation à s points soient d'ordres s, $s+1$ et $s+2$. En particulier, nous expliquons, quantifions et illustrons l'effet des résonances numériques lors de la résolution de problèmes linéaires inhomogènes par de telles méthodes. Les résultats que nous obtenons sont présentés en 1.3.2 et détaillés au Chapitre 3.

Ainsi, à la différence des résultats du Chapitre 2 dont la problématique est le comportement qualitatif en temps long des solutions numériques, les résultats du Chapitre 3 relèvent d'une problématique en temps fini de recherche d'ordre élevé.

1.2 L'équation de Schrödinger

Nous présentons dans cette section les différentes formes de l'équation de Schrödinger auxquelles nous avons appliqué les méthodes numériques dont l'analyse est menée au Chapitre 2 et au Chapitre 3.

1.2.1 Le tore de dimension d

Les équations de Schrödinger que nous considérons (linéaires avec potentiel, linéaires inhomogènes et non linéaires) sont toutes posées sur le tore de dimension $d \in \mathbb{N}^*$. Ainsi, les éléments propres de l'opérateur de Laplace sont parfaitement identifiés, on dispose d'une caractérisation des espaces de Sobolev à l'aide des coefficients de Fourier ainsi que d'algorithmes numériques de transformation de Fourier rapide.

Nous précisons que l'étude des solutions exactes de l'équation de Schrödinger non linéaire dans ce cas (par exemple le caractère localement ou globalement bien posé des problèmes de Cauchy considérés) n'est pas l'objet de notre travail. L'étude des solutions exactes est sensiblement plus complexe que dans le cas où l'espace physique considéré est $\mathbb{R}^d$ tout entier, dans la mesure où les inégalités de Strichartz classiques ne s'appliquent plus. On pourra trouver une étude de ces questions dans [Bou93b, Bou93a] et [Bou94]. Une revue de ces résultats est disponible dans [Gin95] et l'on pourra consulter les références de cet article.

Nous utilisons les notations suivantes :

- Pour tout couple de d-uplets de nombres complexes $u = (u_1, \dots, u_d) \in \mathbb{C}^d$ et $v = (v_1, \dots, v_d) \in \mathbb{C}^d$, on note $u.v = \bar{u}_1 v_1 + \dots + \bar{u}_d v_d$ le produit scalaire hermitien et l'on note $\|(u_1, \dots, u_d)\|_2 = \sqrt{|u_1|^2 + \dots + |u_d|^2}$ la norme hermitienne correspondante.
- Le tore de dimension d est noté $\mathbb{T}^d$.
- On note Δ le Laplacien en dimension d : $\Delta = \sum_{i=1}^d \partial_{x_i}^2$.
- On note $L^2(\mathbb{T}^d)$ (ou parfois simplement L^2) l'espace des classes de fonctions mesurables à valeurs complexes dont le module au carré est sommable sur $\mathbb{T}^d$.
- Pour tout couple (ψ_1, ψ_2) de fonctions de $L^2(\mathbb{T}^d)$, on note
$$\langle \psi_1, \psi_2 \rangle = (2\pi)^{-d} \int_{\mathbb{T}^d} \overline{\psi_1(x)} \psi_2(x) \mathrm{d}x$$
leur produit scalaire hilbertien. On note $\|\cdot\|_{L^2(\mathbb{T}^d)}$ (ou plus simplement $\|\cdot\|_{L^2}$) la norme hilbertienne correspondante.

1.2.2 L'équation de Schrödinger linéaire avec potentiel

Le premier type d'équation de Schrödinger que nous considérons est l'équation de Schrödinger linéaire avec potentiel multiplicatif. Pour une fonction réelle V définie sur $\mathbb{T}^d$, et une fonction complexe φ_0 définie sur $\mathbb{T}^d$ également, on considère le problème de Cauchy suivant :

$$\left\{ \begin{array}{rcll} i\partial_t \varphi(t,x) &=& -\Delta\varphi(t,x) + V(x)\varphi(t,x) & (t,x) \in \mathbb{R} \times \mathbb{T}^d \\ \varphi(0,x) &=& \varphi_0(x) & x \in \mathbb{T}^d \end{array} \right. . \tag{1.2}$$

Cette équation aux dérivées partielles est hamiltonienne au sens où, si l'on décompose les fonctions de $L^2(\mathbb{T}^d)$ en sommes de leurs parties réelles et imaginaires en posant

$$\varphi(x) = u(x) + iv(x) \tag{1.3}$$

et si l'on définit

$$H(u,v) = \frac{1}{2} \int_{\mathbb{T}^d} \big(\|\nabla u(x)\|_2^2 + \|\nabla v(x)\|_2^2 \big) \mathrm{d}x + \frac{1}{2} \int_{\mathbb{T}^d} V(x) \big(u^2(x) + v^2(x) \big) \mathrm{d}x, \tag{1.4}$$

alors sa première ligne de (1.2) se réécrit

$$\left\{\begin{array}{rcl}\partial_t v(t,x) & = & -\dfrac{\partial H}{\partial u}(u,v)(t,x)\\ \partial_t u(t,x) & = & \dfrac{\partial H}{\partial v}(u,v)(t,x)\end{array}\right. \qquad (t,x)\in\mathbb{R}\times\mathbb{T}^d. \tag{1.5}$$

Par analogie avec la situation en dimension finie (c.f. (1.1)), on appelle hamiltonien du système la fonctionnelle H. Comme le veut l'usage en mécanique quantique, on appelle également hamiltonien du système l'opérateur linéaire non borné $\mathcal{H}=-\Delta+V$. Avec la décomposition en parties réelle et imaginaire (1.3), on a la relation suivante entre la fonctionnelle H et l'opérateur $\mathcal{H}$:

$$\langle\varphi,\mathcal{H}\varphi\rangle = 2(2\pi)^{-d}H(u,v). \tag{1.6}$$

1.2.3 L'équation de Schrödinger linéaire inhomogène

La seconde équation de Schrödinger pour laquelle nous analysons des méthodes numériques est une équation de Schrödinger linéaire avec un terme source. Pour une fonction complexe f du temps $t\in\mathbb{R}$ et de l'espace $x\in\mathbb{T}^d$ et une donnée initiale φ_0 complexe définie sur $\mathbb{T}^d$ données, nous considérons le problème de Cauchy suivant :

$$\left\{\begin{array}{rclr} i\partial_t\varphi(t,x) & = & -\Delta\varphi(t,x)+f(t,x) & (t,x)\in\mathbb{R}\times\mathbb{T}^d\\ \varphi(0,x) & = & \varphi_0(x) & x\in\mathbb{T}^d\end{array}\right. . \tag{1.7}$$

Cette équation n'est pas hamiltonienne en général, en raison du terme source.[1]

1.2.4 L'équation de Schrödinger non linéaire

La dernière famille d'équations de Schrödinger que nous considérons est constituée d'équations non linéaires. Pour une fonction complexe g définie sur $\mathbb{R}\times\mathbb{C}$ et une donnée initiale φ_0 complexe définie sur $\mathbb{T}^d$, nous considérons le problème de Cauchy suivant :

$$\left\{\begin{array}{rclr} i\partial_t\varphi(t,x) & = & -\Delta\varphi(t,x)+g(t,u) & (t,x)\in\mathbb{R}\times\mathbb{T}^d\\ \varphi(0,x) & = & \varphi_0(x) & x\in\mathbb{T}^d\end{array}\right. . \tag{1.8}$$

L'étude des solutions exactes de ce problème est menée par exemple dans [Gin95], d'après [Bou93b, Bou93a] et [Bou94].

Une forme typique de la non linéarité g envisagée est

$$g(z)=w(|z|^2)z$$

où w est une fonction régulière sur $\mathbb{R}^+$. Dans ce cas, l'équation (1.8) est hamiltonienne. En effet, si W désigne une primitive de w sur $\mathbb{R}^+$ et si l'on définit la fonctionnelle H par

[1]Bien entendu, l'équation (1.7) est formellement hamiltonienne pour l'hamiltonien $H(t,u,v) = \frac{1}{2}\int_{\mathbb{T}^d}\left(\|\nabla u(x)\|_2^2+\|\nabla v(x)\|_2^2\right)\mathrm{d}x+\int_{\mathbb{T}^d}\Re\Big(f(t,x)\big(u(x)-iv(x)\big)\Big)\mathrm{d}x$, mais l'hamiltonien n'est alors ni indépendant du temps ni a priori constant le long des solutions.

$$H(u,v) = \frac{1}{2}\int_{\mathbb{T}^d} \left(\|\nabla u(x)\|_2^2 + \|\nabla v(x)\|_2^2\right)\mathrm{d}x + \frac{1}{2}\int_{\mathbb{T}^d} W\left(u^2(x) + v^2(x)\right)\mathrm{d}x, \qquad (1.9)$$

alors, à l'aide du changement de variables (1.3), la partie évolutive de l'équation (1.8) s'écrit (1.5).

1.3 Résultats obtenus

Nous présentons ici les principaux résultats obtenus dans cette thèse. Ceux-ci consistent en l'analyse numérique de certaines méthodes de résolution en temps des équations de Schrödinger présentées en 1.2. Dans un premier temps, on présente les résultats obtenus relatifs aux méthodes de splitting appliquées à l'équation de Schrödinger linéaire avec potentiel (1.2). Ceux-ci ont donné lieu à la publication de l'article [DF07] qui constitue le Chapitre 2. Ensuite, on présente les résultats obtenus lors de l'analyse des méthodes de Runge-Kutta exponentielles de collocation appliquées à l'équation de Schrödinger linéaire inhomogène (1.7) ainsi qu'à l'équation de Schrödinger non linéaire (1.8). Ces résultats ont donné lieu à la prépublication de l'article [Duj08] qui constitue le Chapitre 3.

1.3.1 Méthodes de splitting pour l'équation de Schrödinger (1.2)

Nous présentons les résultats obtenus relatifs aux méthodes de splitting en temps appliquées à l'équation de Schrödinger linéaire avec potentiel multiplicatif (1.2), dont le détail, les preuves et les expériences numériques constituent le Chapitre 2 de cette thèse et ont fait l'objet de l'article [DF07]. On introduit dans un premier temps le contexte et les notations propres à ce travail. Ensuite, on présente les hypothèses et résultats obtenus. D'une part, on énonce un théorème de forme normale pour le propagateur de splitting dans l'asymptotique des petits potentiels pour des pas de temps non résonnants (Théorème 1.3.1). D'autre part, on présente un corollaire important de ce théorème qui permet d'estimer la propagation en temps long de la régularité de la solution numérique obtenue par la méthode de splitting, également pour des pas de temps non résonnants (Corollaire 1.3.1). Enfin, on détaille les éléments essentiels de la preuve du Théorème 1.3.1 et du Corollaire 1.3.1.

1.3.1.1 Contexte et notations

On se place ici en dimension $d = 1$ d'espace[2]. Pour tout $k \in \mathbb{Z}$, on note e_k la fonction définie pour $x \in \mathbb{T}$ par $e_k(x) = e^{ikx}$. On identifie chaque fonction φ de $L^2(\mathbb{T}^d)$ avec la suite $(\hat{\varphi}_k)_{k\in\mathbb{Z}}$ de ses coefficients de Fourier définie par

$$\forall k \in \mathbb{Z}, \qquad \hat{\varphi}_k = \frac{1}{2\pi}\int_{\mathbb{T}} \varphi(x)\overline{e_k(x)}\mathrm{d}x = \langle e_k, \varphi\rangle.$$

[2]Cette restriction a pour unique but de simplifier la présentation : nos résultats s'étendent *mutatis mutandis* à toute valeur de $d \in \mathbb{N}^*$.

On regroupe les modules des coefficients de Fourier de toute fonction $\varphi \in L^2$ en posant

$$\forall k \in \mathbb{Z} \setminus \{0\}, \qquad |\varphi|_k = \sqrt{|\hat{\varphi}_k|^2 + |\hat{\varphi}_{-k}|^2} \tag{1.10}$$

et $|\varphi|_0 = |\hat{\varphi}_0|$. Tout opérateur linéaire non borné A de L^2 dans L^2 dont le domaine contient les fonctions $(e_k)_{k\in\mathbb{Z}}$ est identifié à la matrice infinie des coefficients $(A_{i,j})_{(i,j)\in\mathbb{Z}^2}$ définis par $A_{i,j} = \langle e_i, Ae_j\rangle$. On définit l'adjoint $A^\star$ d'un tel opérateur A par sa matrice en posant $A^\star_{i,j} = \overline{A_{j,i}}$. Un opérateur dont le domaine contient les $(e_k)_{k\in\mathbb{Z}}$ est dit auto-adjoint lorsque $A = A^\star$, et unitaire lorsque $A^\star A = AA^\star = \mathrm{Id}$ (où Id est l'opérateur identité de L^2). Ainsi, par exemple, le Laplacien est un opérateur auto-adjoint qui s'identifie à la matrice diagonale contenant à la k-ème ligne et à la k-ème colonne le coefficient $-k^2$.

On considère pour $\rho > 0$ le sous-espace U_ρ de L^2 constitué des fonctions $\varphi \in L^2$ telles que $\sup_{k\in\mathbb{Z}} |\hat{\varphi}_k| e^{\rho|k|} < +\infty$, de sorte que les fonctions de U_ρ sont analytiques sur une bande du plan complexe de largeur typique 2ρ. Sur cet espace, on définit la norme analytique par $\|\varphi\|_\rho = \sup_{k\in\mathbb{Z}} |\hat{\varphi}_k| e^{\rho|k|}$. On considère également le sous-espace $\mathcal{A}_\rho$ des opérateurs linéaires non bornés constitué des opérateurs A tels que $\sup_{(i,j)\in\mathbb{Z}^2} |A_{i,j}| e^{\rho|i-j|} < +\infty$, muni de la norme (que l'on appelle également « analytique ») $\|A\|_\rho = \sup_{(i,j)\in\mathbb{Z}^2} |A_{i,j}| e^{\rho|i-j|}$. On considère en outre le $\mathbb{R}$-sous-espace $\mathcal{S}_\rho$ de A_ρ constitué des opérateurs auto-adjoints de $\mathcal{A}_\rho$. Pour tout $K > 0$, on note I_K l'ensemble des couples d'indices $(i,j) \in \mathbb{Z}^2$ tels que $|i| \leq K$ ou $|l| \leq K$. On définit l'espace des opérateurs croix $\mathcal{X}_\rho$ comme le sous-espace de $\mathcal{A}_\rho$ constitué des opérateurs $X \in \mathcal{A}_\rho$ tels que pour tout $(i,j) \in \mathbb{Z}^2$,

$$X_{i,j} \neq 0 \qquad \Longrightarrow \qquad |i| = |j|,$$

et l'espace $\mathcal{X}^K_\rho$ des opérateurs K-quasi-croix comme le sous-espace de $\mathcal{A}_\rho$ constitué des opérateurs $X \in \mathcal{A}_\rho$ tels que

$$X_{i,j} \neq 0 \qquad \Longrightarrow \qquad |i| = |j| \quad \text{ou} \quad (i,j) \notin I_K.$$

Une propriété importante des opérateurs de type $\mathcal{X}$ est qu'ils laissent stable chacun des sous-espaces $\mathrm{Vect}(e_k, e_{-k})$ pour $k \in \mathbb{Z}$. De même les opérateurs de type $\mathcal{X}^K$ laissent stables les sous-espaces $\mathrm{Vect}(e_k, e_{_k})$ pour $|k| \leq K$ ainsi que le sous-espace $\mathrm{Vect}\big((e_k)_{|k|>K}\big)$.

1.3.1.2 Hypothèses et résultats

Nous faisons l'hypothèse de régularité suivante sur le potentiel multiplicatif V de l'équation (1.2) :

Hypothèse 1.3.1 *La fonction réelle V définie sur $\mathbb{T}$ est la restriction d'une fonction analytique au sens où il existe M_V et $\rho_V > 0$ tels que*

$$\forall n \in \mathbb{N}, \forall k \in \mathbb{Z}, \qquad |(\widehat{V^n})_k| \leq M_V^n e^{-\rho_V|k|}.$$

Afin de calculer des solutions numériques de l'équation (1.2), on considère la famille de méthodes numériques indexée par le paramètre réel λ

$$L_\lambda = e^{ih\Delta} e^{-ih\lambda V}, \tag{1.11}$$

où $h > 0$ désigne le pas de temps de la méthode à un pas considérée. De cette manière, on retrouve le propagateur de l'équation de Schrödinger linéaire libre $L_0 = e^{ih\Delta}$ et, pour $\lambda = 1$, on retrouve un propagateur de splitting de Lie $L_1 = e^{ih\Delta} e^{-ihV}$. L'introduction du paramètre réel λ est motivée par l'approche perturbative du problème que nous allons développer : nous allons nous placer dans l'asymptotique des petits potentiels ($\lambda \to 0$) et considérer L_λ comme une perturbation de L_0. On considère, à $h > 0$ et $\lambda \in \mathbb{R}$ fixés, la solution numérique du problème (1.2) issue de la donnée initiale exacte $\varphi^0 = \varphi_0$ et définie par récurrence pour tout $n \in \mathbb{N}$ par $\varphi^{n+1} = L_\lambda \varphi^n$.

Nous faisons l'hypothèse suivante de non résonance sur le pas de temps $h > 0$ de la méthode numérique L_λ :

Hypothèse 1.3.2 *Il existe $\gamma > 0$ et $\nu > 1$ tels que*

$$\forall\, k \in \mathbb{Z}, \quad k \neq 0, \quad \left| \frac{1 - e^{ihk}}{h} \right| \geq \gamma |k|^{-\nu}. \tag{1.12}$$

Remarque 1.3.1 *À γ et ν fixés, l'ensemble des pas de temps $h > 0$ dans l'intervalle $]0, h_0[$ ne vérifiant pas la condition* (1.12) *a une mesure de Lebesgue négligeable devant h_0 lorsque h_0 tend vers* 0 *(voir [HLW06], Chapitre* X*, Lemme 6.3).*

Ces hypothèses nous permettent de démontrer le théorème de forme normale suivant pour le propagateur de splitting L_λ défini par (1.11), dans l'asymptotique des petits potentiels. Ce théorème nous permet de montrer ensuite le Corollaire 1.3.1 de propagation de la régularité de la solution numérique $(\varphi^n)_{n \in \mathbb{N}}$.

Théorème 1.3.1 *On se donne $\gamma > 0$ et $\nu > 1$. Supposons que la fonction V satisfait l'Hypothèse 1.3.1. Pour tout $\lambda \in \mathbb{R}$, on considère le propagateur de splitting L_λ défini par* (1.11)*. Il existe des constantes strictement positives λ_0, σ et c dépendant uniquement de V, M_V, ρ_V, γ et ν telles que pour tout pas de temps $h \in]0, 1[$ satisfaisant la condition* (1.12) *de l'Hypothèse 1.3.2, il existe des familles d'opérateurs unitaires $Q(\lambda)$ et $\Sigma(\lambda)$ analytiques en λ pour $|\lambda| < \lambda_0$, telles que pour tout $\lambda \in]0, \lambda_0[$,*

$$Q(\lambda) \in \mathcal{A}_{\rho_V/4} \quad \textit{et} \quad \Sigma(\lambda) \in \mathcal{X}^K_{\rho_V/4} \quad \textit{avec} \quad K = \lambda^{-\sigma}, \tag{1.13}$$

$$\|Q(\lambda) - \mathrm{Id}\|_{\rho_V/4} \leq \lambda^{1/2} \quad \textit{et} \quad \|\Sigma(\lambda) - e^{ih\Delta}\|_{\rho_V/4} \leq h\lambda^{1/2}, \tag{1.14}$$

et

$$Q(\lambda) L(\lambda) Q(\lambda)^\star = \Sigma(\lambda) + R(\lambda). \tag{1.15}$$

En outre, le reste $R(\lambda)$ satisfait pour tout $\lambda \in (0, \lambda_0)$,

$$\|R(\lambda)\|_{\rho_V/5} \leq \exp(-c\lambda^{-\sigma}). \tag{1.16}$$

Ainsi, il existe un changement de variables unitaire $Q(\lambda)$ analytique à un paramètre, proche de l'identité quand λ tend vers 0 tel que le propagateur de splitting L_λ s'écrit dans la nouvelle base comme la somme d'un terme $\Sigma(\lambda)$ proche du propagateur de Schrödinger libre et d'un terme $R(\lambda)$ exponentiellement petit en norme analytique. Dans les nouvelles variables, l'action du propagateur L_λ quand λ est petit est donc essentiellement guidée par l'action de $\Sigma(\lambda)$. Or, l'opérateur $\Sigma(\lambda)$ est un opérateur K-quasi-croix (avec $K = \lambda^{-\sigma}$) et a donc une dynamique simple sur certains modes dans les nouvelles variables. Plus précisément, pour tout $k \in \{1, \ldots, \lfloor K \rfloor\}$, il existe une matrice unitaire $M_k(\lambda)$ de taille 2×2 telle que pour tout $\theta \in L^2$, si $\Theta = \Sigma(\lambda)\theta$, alors

$$\begin{pmatrix} \widehat{\Theta}_k \\ \widehat{\Theta}_{-k} \end{pmatrix} = M_k(\lambda) \begin{pmatrix} \widehat{\theta}_k \\ \widehat{\theta}_{-k} \end{pmatrix}.$$

La matrice $M_k(\lambda)$ étant unitaire, on a en particulier

$$|\Theta|_k = \sqrt{|\widehat{\Theta}_k|^2 + |\widehat{\Theta}_{-k}|^2} = \sqrt{|\widehat{\theta}_k|^2 + |\widehat{\theta}_{-k}|^2} = |\theta|_k.$$

Cette propriété de conservation dans les nouvelles variables et les différents résultats du théorème précédent permettent de démontrer le corollaire suivant, qui explique la bonne conservation de la régularité de la solution numérique en temps long pour des pas de temps non résonnants :

Corollaire 1.3.1 *Avec les notations du Théorème 1.3.1,*
(i) *Il existe une constante $C > 0$ dépendant uniquement de V, M_V, ρ_V, γ et ν telle que pour tout $h \in]0,1[$ satisfaisant la condition de non résonance* (1.12)*, tout $\lambda \in]0, \lambda_0[$, tout $n \leq \exp(c\lambda^{-\sigma}/2)$, et tout $\varphi^0 \in L^2(\mathbb{T})$, on a*

$$\forall\, k \in \mathbb{N}, \quad k \leq \lambda^{-\sigma}, \quad \big|\, |\varphi^n|_k - |\varphi^0|_k \,\big| \leq C\lambda^{1/2} \|\varphi^0\|_{L^2}. \tag{1.17}$$

(ii) *Pour tout $s > 1/2$ et tout s' tel que $s - s' \geq 1/2$, il existe une constante $c_s > 0$ dépendant uniquement de V, M_V, ρ_V, γ, ν et s telle que pour tout $h \in]0,1[$ satisfaisant la condition de non résonance* (1.12)*, tout $\lambda \in (0, \lambda_0)$, tout $n \leq \exp(c\lambda^{-\sigma}/2)$ et tout $\varphi^0 \in L^2$ tel que $\sup_{k\in\mathbb{N}}(1+k)^s|\varphi^0|_k < +\infty$, on a*

$$\sup_{0\leq k\leq \lambda^{-\sigma}} \Big((1+k)^{s'} \,\big|\, |\varphi^n|_k - |\varphi^0|_k \,\big| \Big) \leq c_s \lambda^{1/2} \sup_{k\in\mathbb{N}} (1+k)^s |\varphi^0|_k. \tag{1.18}$$

(iii) *Pour tout $\rho \in]0, \rho_V/5[$, il existe des constantes strictement positives μ_0 et C dépendant uniquement de V, M_V, ρ_V, γ, ν et ρ telles que pour tout $h \in (0,1)$ satisfaisant la condition de non résonance* (1.12)*, tout $\lambda \in (0, \lambda_0)$, tout $n \leq \exp(c\lambda^{-\sigma}/2)$, tout $\mu \in (0, \mu_0)$ et tout $\varphi^0 \in U_\rho$, on a*

$$\sup_{0\leq k\leq \lambda^{-\sigma}} \Big(e^{\mu k} \,\big|\, |\varphi^n|_k - |\varphi^0|_k \,\big| \Big) \leq C\lambda^{1/2} \|\varphi^0\|_\rho. \tag{1.19}$$

Le premier point du corollaire implique et quantifie la conservation en temps long de la régularité L^2 de la solution numérique, alors que le second traite de la régularité de type $H^{s'}$. Enfin, le dernier point quantifie la conservation de la régularité de type analytique.

En outre, dans chacun des trois cas précédents, l'estimation de la régularité de la solution numérique vaut pour des temps exponentiellement longs par rapport à λ : en effet, la condition sur n implique que le temps de validité des majorations (1.17), (1.18) et (1.19) est au moins $nh \leq n \leq \exp\left(\frac{c}{2\lambda^\sigma}\right)$.

1.3.1.3 Eléments de preuve

Nous présentons ici les principaux arguments de la preuve du Théorème 1.3.1 et du Corollaire 1.3.1. Les détails se trouvent dans leur intégralité au Chapitre 2.

Dans un premier temps, on cherche les opérateurs $Q(\lambda)$ et $\Sigma(\lambda)$ suivant un développement en séries formelles (à coefficients dans des espaces de type $\mathcal{A}$) :

$$Q(\lambda) = \sum_{n\geq 0} \lambda^n Q_n \quad \text{et} \quad \Sigma(\lambda) = \sum_{n\geq 0} \lambda^n \Sigma_n,$$

vérifiant exactement la relation

$$Q(\lambda) L_\lambda Q^\star(\lambda) = \Sigma(\lambda),$$

au sens des séries formelles. On introduit le changement de variables (toujours au sens des séries formelles)

$$S(\lambda) = Q^\star(i\partial_\lambda Q) \quad \text{et} \quad X(\lambda) = \Sigma^\star(i\partial_\lambda \Sigma),$$

où l'on cherche les coefficients (a priori auto-adjoints) respectifs $(X_n)_{n\in\mathbb{N}}$ et $(S_n)_{n\in\mathbb{N}}$ de $X(\lambda)$ et $S(\lambda)$. On est ainsi amené à considérer la suite d'équations homologiques

$$S_n - e^{ih\Delta} S_n e^{-ih\Delta} + X_n = G_n, \tag{1.20}$$

où pour tout $n \in \mathbb{N}$, G_n est auto-adjoint et dépend de V et de $X_0, \ldots, X_{n-1}$ et $S_0, \ldots, S_{n-1}$.

L'équation d'inconnues auto-adjointes X et S

$$S - e^{-ih\Delta} S e^{ih\Delta} + X = G, \tag{1.21}$$

où G est auto-adjoint et connu s'écrit en composantes

$$\forall (k,l) \in \mathbb{Z}^2, \qquad (1 - e^{ih(k^2-\ell^2)}) S_{k\ell} + X_{k\ell} = G_{k\ell}.$$

Il s'agit d'une équation linéaire sous-déterminée. L'estimation de la décroissance des modules des coefficients de la composante S d'une solution (S, X) de cette équation en fonction de ceux de G fait intervenir l'Hypothèse 1.3.2, au moins dans le domaine d'indices $(k,l) \in I_K$. On introduit à cette fin la notion de I_K-solution de l'équation (1.21) dont les détails de la définition constituent la Définition 2.3.1.

On estime ensuite certaines normes analytiques des I_K-solutions successives $(X_n)_{n\in\mathbb{N}}$ et $(S_n)_{n\in\mathbb{N}}$ de l'équation (1.20). On tronque les séries (a priori divergentes) définissant les opérateurs $X(\lambda)$ et $S(\lambda)$ à l'ordre N en considérant des polynômes en λ, $X^N(\lambda)$ et $S^N(\lambda)$. Ensuite, on construit le changement de variables analytique $Q^N(\lambda)$ et le candidat second membre $\Sigma^N(\lambda)$ à partir de $X^N(\lambda)$ et $S^N(\lambda)$, en résolvant des équations différentielles linéaires dans des espaces de Banach analytiques de type $\mathcal{A}$. Enfin, on estime la différence $R_N(\lambda)$ entre $Q^N(\lambda)L_\lambda\big(Q^N\big)^\star(\lambda)$ et $\Sigma^N(\lambda)$ en norme analytique par le terme $(\lambda K^\alpha N^\beta)^N$ (où $\alpha > 0$ et $\beta > 0$ sont explicites en fonction des données), à une constante multiplicative près. Il reste alors à choisir K et N proportionnels à une puissance négative de λ pour obtenir la fin de la preuve du Théorème 1.3.1.

Afin de montrer le Corollaire 1.3.1, on utilise le Théorème 1.3.1 : on écrit la solution numérique de (1.2) obtenue à l'aide du propagateur L_λ défini en (1.11) dans les nouvelles variables en posant pour tout $n \in \mathbb{N}$, $\psi^n = Q(\lambda)\varphi^n$. On considère la projection orthogonale π_K sur le sous-espace de L^2 suivant :

$$E_K = \Big\{\psi \in L^2 \quad | \quad \forall k \in \mathbb{Z}, \quad \big(|k| > K \Rightarrow \hat{\psi}_k = 0\big)\Big\}.$$

La nature K-quasi-croix de l'opérateur $\Sigma(\lambda)$ assure que $\Sigma(\lambda)\pi_K - \pi_K\Sigma(\lambda) = 0$. Ainsi, le sous-espace E_K et son orthogonal sont stables par $\Sigma(\lambda)$. On montre alors qu'il existe une constante $C > 0$ dépendant uniquement de V, M_V, ρ_V, γ et ν telle que pour tout $n \le \exp(c\lambda^{-\sigma}/2)$,

$$\forall\, 0 \le k \le K, \quad \big||\psi^n|_k - |\psi^0|_k\big| \le C\exp(-c\lambda^{-\sigma}/2)\|\varphi_0\|_{L^2}.$$

Ceci montre une bonne conservation, dans les nouvelles variables, des quantités $|\psi^n|_k$, d'une part pour un nombre de modes asymptotiquement grand ($K = \lambda^{-\sigma}$) et d'autre part en temps exponentiellement long ($n \le \exp(c\lambda^{-\sigma}/2)$). Afin d'obtenir des résultats similaires dans les variables initiales, on traite d'abord la question de la conservation dans les grands modes des nouvelles variables en montrant l'existence d'une autre constante $C > 0$ indépendante de λ telle que pour tout $n \le \exp(c\lambda^{-\sigma}/2)$,

$$\Big|\|\pi_K\psi^n\| - \|\pi_K\psi^0\|\Big| \le C\exp(-c\lambda^{-\sigma}/2)\|\varphi_0\|_{L^2}.$$

Enfin, on revient aux variables initiales en utilisant la régularité supposée sur φ_0, les deux inégalités ci-dessus, et les points (1.14), (1.15) et (1.16) du Théorème 1.3.1.

1.3.2 Méthodes de Runge-Kutta exponentielles pour les équations de Schrödinger (1.7) et (1.8)

Nous présentons les résultats obtenus lors de l'analyse des méthodes de Runge-Kutta exponentielles de collocation appliquées à l'équation de Schrödinger dans sa forme linéaire inhomogène (1.7) et dans sa forme non linéaire (1.8) décrites dans la Section 1.2. Les méthodes de Runge-Kutta exponentielles ont été introduites et analysées par exemple dans le cadre d'équations aux dérivées partielles paraboliques dans [HO05b] et [HO05a]. Notre travail consiste en l'adaptation de ces méthodes au contexte de l'équation de Schrödinger et en leur analyse numérique dans ce contexte. En particulier, notre attention se porte sur des conditions suffisantes relatives aux méthodes de collocation à s points pour atteindre l'ordre s, l'ordre $s+1$ et l'ordre $s+2$. Nous comparons dans chaque cas nos résultats avec ceux existant pour les méthodes de Runge-Kutta exponentielles de collocation appliquées à des problèmes paraboliques. Après l'introduction du contexte et des notations utilisées, on présente les résultats obtenus pour les problèmes de Schrödinger linéaires inhomogènes (1.7). Ensuite, on présente les résultats obtenus pour les problèmes de Schrödinger non linéaires (1.8). Les détails de ces résultats, leurs preuves et des expériences numériques les illustrant se trouvent au Chapitre 3.

1.3.2.1 Contexte et notations

Pour $k \in \mathbb{Z}^d$, on note e_k la fonction définie sur $\mathbb{T}^d$ par $e_k(x) = e^{ik.x}$. Un opérateur linaire non borné A de $L^2(\mathbb{T}^d)$ dans lui-même est dit diagonal si son domaine contient tous les $(e_k)_{k\in\mathbb{Z}^d}$ et si pour tout $k \in \mathbb{Z}^d$, il existe $\lambda_k \in \mathbb{C}$ tel que $Ae_k = \lambda_k e_k$. Par exemple, le Laplacien est un opérateur diagonal avec pour tout $k \in \mathbb{Z}^d$, $\Delta e_k = -\|k\|_2^2 e_k$. On peut alors définir un calcul fonctionnel sur de tels opérateurs : si φ désigne une fonction de $\mathbb{C}$ dans $\mathbb{C}$, on note $\varphi(A)$ l'opérateur de domaine l'espace engendré par les $(e_k)_{k\in\mathbb{Z}^d}$ et défini pour tout $k \in \mathbb{Z}^d$ par $\varphi(A)e_k = \varphi(\lambda_k)e_k$. On se fixe un temps final $T > 0$ et un pas de temps $h \in]0, T[$. On note $N = \lfloor \frac{T}{h} \rfloor$. Pour tout $n \in \{0, \dots, N\}$, on pose $t_n = nh$. On se donne $s \in \mathbb{N}^*$ points $(c_1, \dots, c_s)$ distincts du segment $[0, 1]$ et pour $h > 0$ fixé, on note $(l_i)_{1\leq i\leq n}$ la famille de polynômes d'interpolation de Lagrange aux points $c_1h, \dots, c_sh$.

On note u (au lieu de φ) la solution exacte des problèmes (1.7) et (1.8).

1.3.2.2 Méthodes de Runge-Kutta exponentielles de collocation pour l'équation de Schrödinger linéaire inhomogène

Afin de calculer des solutions numériques de l'équation de Schrödinger linéaire inhomogène (1.7), on introduit des méthodes de Runge-Kutta exponentielles de collocation à $s \in \mathbb{N}^*$ points. On donne ensuite des conditions suffisantes sur les coefficients $(c_j)_{1\leq j\leq s}$ de ces méthodes pour que celles-ci soient d'ordre $s+1$ et d'ordre $s+2$. On compare nos résultats avec ceux obtenus dans le cas de la résolution numérique de problèmes paraboliques par ces mêmes méthodes (c.f. [HO05b]).

On définit s opérateurs linéaires sur L^2 en posant pour tout $j \in \{1, \dots, s\}$

$$b_j(h\Delta) = \frac{1}{h}\int_0^h e^{i(h-\tau)\Delta} l_j(\tau)\mathrm{d}\tau. \tag{1.22}$$

On part de la donnée initiale exacte $u_0 = \varphi_0$ du problème de Cauchy (1.7) et, pour tout $n \in \{0, \dots, N-1\}$, à partir d'une approximation u_n de la valeur de la solution exacte

u au temps t_n, on calcule une approximation u_{n+1} de $u(t_{n+1})$ à l'aide de la formule

$$u_{n+1} = e^{ih\Delta}u_n + h\sum_{j=1}^{s} b_j(ih\Delta)f(t_n + c_j h). \tag{1.23}$$

On peut interpréter cette formule comme une discrétisation en temps de la formule de Duhamel pour la solution exacte de l'équation (1.7) sur le segment $[t_n, t_{n+1}]$ qui s'écrit

$$u(t_{n+1}) = e^{ih\Delta}u(t_n) + \int_0^h e^{i(h-\tau)\Delta}f(t_n + \tau)\mathrm{d}\tau.$$

Lorsque l'on ôte le Laplacien dans l'équation (1.2) (formellement, $\Delta := 0$), les opérateurs $(b_j)_{1\leq j\leq s}$ sont des homothéties de L^2 dont les rapports sont les coefficients d'une méthode de Runge-Kutta de collocation classique pour la collection d'équations différentielles ordinaires linéaires inhomogènes indexée par $x \in \mathbb{T}^d$:

$$\frac{\mathrm{d}u(x,.)}{\mathrm{d}t}(t) = f(t,x).$$

On l'appelle méthode de Runge-Kutta sous-jacente.

Le premier résultat que nous obtenons pour les méthodes de Runge-Kutta exponentielles de collocation à s points que nous venons d'introduire pour les problèmes de Schrödinger linéaires inhomogènes (1.7) est qu'elles sont d'ordre au moins s dès que le terme source f a une régularité temporelle suffisante. Précisément,

Théorème 1.3.2 *Soit $r \geq 0$. Appliquons la méthode de Runge-Kutta exponentielle de collocation (1.23) à l'équation de Schrödinger linéaire inhomogène (1.7). Supposons que le terme source $f \in C^s([0,T], H^r(\mathbb{T}^d))$. Alors, la méthode est d'ordre s au sens où il existe une constante strictement positive C dépendant uniquement du choix des points de collocation $(c_1, \dots, c_s)$ telle que pour tout $h \in]0,T[$ et tout $n \in \{0, \dots, N\}$,*

$$\|u_n - u(t_n)\|_{H^r} \leq C\, h^s \int_0^{t_n} \|f^{(s)}(\tau)\|_{H^r}\, \mathrm{d}\tau.$$

La preuve de ce théorème est fondée sur des développements de Taylor en temps du terme source f et sur l'utilisation des propriétés d'interpolation des polynômes de Lagrange $(l_j)_{1\leq j\leq s}$. Remarquons que la constante C du Théorème 1.3.2 est indépendante de T et de r, ce qui constitue une différence avec le résultat similaire pour les problèmes paraboliques.

On montre ensuite que, sous réserve que le terme source f soit suffisamment régulier en temps et en espace, il suffit que la méthode de Runge-Kutta sous-jacente soit d'ordre $s+1$ pour que la méthode (1.23) soit elle aussi d'ordre $s+1$. Précisément,

Théorème 1.3.3 *Soit $r \geq 0$. Supposons que la méthode de Runge-Kutta sous-jacente est d'ordre $s+1$. Supposons que le terme source $f \in C^{s+1}([0,T], L^2(\mathbb{T}^d))$ est tel que $f^{(s)} \in C^1([0,T], H^{r+2})$. Alors, la méthode de Runge-Kutta exponentielle de collocation*

(1.23) *appliquée au problème* (1.7) *est d'ordre* $s+1$ *au sens où il existe une constante strictement positive* C *dépendant uniquement de* $(c_1,\ldots,c_s)$ *telle que pour tout* $h\in]0,T[$ *et tout* $n\in\{0,\ldots,N\}$,

$$\|u(t_n)-u_n\|_{H^r}\le C\,T\,h^{s+1}\Big(\|f^{(s)}(0)\|_{H^{r+2}}+\int_0^{t_n}\|f^{(s+1)}(\tau)\|_{H^{r+2}}\,\mathrm{d}\tau\Big).$$

La preuve de ce résultat est fondée une nouvelle fois sur des développements de Taylor en temps du terme source f ainsi que sur une écriture de l'erreur qui permet d'utiliser une sommation d'Abel. On remarque ici que c'est l'hypothèse de régularité supplémentaire en espace sur le terme source qui différencie le Théorème 1.3.3 du résultat analogue pour les problèmes paraboliques. On note en outre que la constante CT apparaissant dans le Théorème 1.3.3 dépend maintenant explicitement du temps. Toutefois, en supposant une régularité supplémentaire en espace sur le terme source f et en imposant une condition de non résonance sur le pas de temps h, on montre le théorème suivant :

Théorème 1.3.4 *Soit* $r\ge 0$, $\gamma>0$ *et* $\nu>1$. *Supposons que la méthode de Runge-Kutta sous-jacente est d'ordre* $s+1$. *Supposons que le terme source* $f\in C^{s+1}([0,T],L^2(\mathbb{T}^d))$ *est tel que* $f^{(s)}\in C^1([0,T],H^{r+2\nu+2})$. *Alors, la méthode de Runge-Kutta exponentielle de collocation* (1.23) *appliquée à l'équation* (1.7) *est d'ordre* $s+1$ *pour les pas de temps non résonnants au sens où il existe une constante strictement positive* C *dépendant uniquement de* $(c_1,\ldots,c_s)$ *et* γ *telle que pour tout* $h\in]0,T[$ *satisfaisant la condition de non résonance* (1.12), *et tout* $n\in\{0,\ldots,N\}$,

$$\|u(t_n)-u_n\|_{H^r}\le C\,h^{s+1}\Big(\|f^{(s)}(0)\|_{H^{r+2\nu+2}}+\int_0^{t_n}\|f^{(s+1)}(\tau)\|_{H^{r+2\nu+2}}\,\mathrm{d}\tau\Big).$$

Ainsi, φ_0, f et $(c_1,\ldots,c_s)$ étant fixés, on s'attend à avoir un comportement de la courbe d'ordre qui discrimine les pas de temps résonnants des pas de temps non résonnants, même pour un temps final T fixé. C'est ce qu'on observe dans la pratique, comme le montrent les expériences numériques présentées au Chapitre 3.

On montre enfin que, sous l'hypothèse que le terme source f est suffisamment régulier en temps et en espace et que la méthode de Runge-Kutta sous-jacente est d'ordre $s+2$, la méthode de Runge-Kutta exponentielle de collocation (1.23) appliquée au problème (1.7) est elle aussi d'ordre $s+2$. Précisément,

Théorème 1.3.5 *Soit* $r\ge 0$. *Supposons que la méthode de Runge-Kutta sous-jacente est d'ordre* $s+2$. *Supposons que le terme source* $f\in C^{s+2}([0,T],L^2(\mathbb{T}^d))$ *est tel que* $f^{(s)}\in C^2([0,T],H^{r+4})$. *Alors la méthode de Runge-Kutta exponentielle de collocation* (1.23) *est d'ordre* $s+2$ *au sens où il existe une constante strictement positive* C *dépendant uniquement de* $(c_1,\ldots,c_s)$ *telle que pour tout* $h\in]0,T[$ *et tout* $n\in\{0,\ldots,N\}$, *on a*

$$\begin{aligned}\|u(t_n)-u_n\|_{H^r}\le C\,T\,h^{s+2}\Big(&\|f^{(s)}(0)\|_{H^{r+4}}+\|f^{(s+1)}(0)\|_{H^{r+2}}\\&+\int_0^{t_n}\|f^{(s+1)}(\tau)\|_{H^{r+4}}\,\mathrm{d}\tau+\int_0^{t_n}\|f^{(s+2)}(\tau)\|_{H^{r+2}}\,\mathrm{d}\tau\Big).\end{aligned}$$

Ce type de résultat est similaire à celui obtenu dans le cas des problèmes paraboliques dans [HO05b]. Toutefois, il est ici aussi possible de donner une version de ce théorème pour des pas de temps non résonnants :

Théorème 1.3.6 *Soit $r \geq 0$, $\gamma > 0$ et $\nu > 1$. Supposons que la méthode de Runge-Kutta sous-jacente est d'ordre $s+2$. Supposons que le terme source $f \in C^{s+2}([0,T], L^2(\mathbb{T}^d))$ est tel que $f^{(s)} \in C^2([0,T], H^{r+2\nu+4}(\mathbb{T}^d))$. Alors la méthode de Runge-Kutta exponentielle de collocation (1.23) appliquée au problème (1.7) est d'ordre $s+2$ pour les pas de temps non résonnants au sens où il existe une constante strictement positive C dépendant uniquement de $(c_1, \dots, c_s)$ et γ telle que pour tout $h \in]0, T[$ satisfaisant la condition de non-résonance (1.12), et tout $n \in \{0, \dots, N\}$, on a*

$$\begin{aligned}\|u(t_n) - u_n\|_{H^r} \leq C\, h^{s+2}\Big(&\|f^{(s)}(0)\|_{H^{r+2\nu+4}} + \|f^{(s+1)}(0)\|_{H^{r+2\nu+2}} \\ &+ \int_0^{t_n} \|f^{(s+1)}(\tau)\|_{H^{r+2\nu+4}}\, \mathrm{d}\tau + \int_0^{t_n} \|f^{(s+2)}(\tau)\|_{H^{r+2\nu+2}}\, \mathrm{d}\tau\Big).\end{aligned}$$

Les preuves des résultats cités dans cette section sont présentées dans leur intégralité au Chapitre 3. Dans ce même chapitre, on présente également des expériences numériques illustrant ces résultats.

1.3.2.3 Méthodes de Runge-Kutta exponentielles de collocation pour l'équation de Schrödinger non linéaire

Afin de calculer des solutions numériques de l'équation de Schrödinger non linéaire (1.8), on introduit des méthodes de Runge-Kutta exponentielles de collocation à $s \in \mathbb{N}^*$ points. On donne ensuite des conditions suffisantes sur les coefficients $(c_j)_{1 \leq j \leq s}$ de ces méthodes pour que celles-ci soient d'ordre $s+1$ et d'ordre $s+2$. On compare nos résultats avec ceux obtenus dans le cas de la résolution numérique de problèmes paraboliques par ces mêmes méthodes (c.f. [HO05b]).

On initialise la méthode à partir de la donnée initiale exacte $u_0 = \varphi_0$ du problème de Cauchy (1.8). À partir d'une approximation u_n de la valeur $u(t_n)$ de la solution exacte u de (1.8) au temps t_n et de s approximations respectives $u_{n,1}, \dots, u_{n,s}$ de $u(t_n + c_1 h), \dots, u(t_n + c_s h)$, on construit une approximation u_{n+1} de $u(t_{n+1})$ en posant

$$u_{n+1} = e^{ih\Delta} u_n + h \sum_{j=1}^{s} b_j(ih\Delta) g(t_n + c_j h, u_{n,j}), \tag{1.24}$$

où les s opérateurs $(b_j)_{1 \leq j \leq s}$ sont ceux définis en (1.22). Les approximations respectives $u_{n,1}, \dots, u_{n,s}$ de $u(t_n + c_1 h), \dots, u(t_n + c_s h)$ sont calculées auparavant à partir de u_n en résolvant le système de s équations non linaires suivant

$$u_{n,j} = e^{ic_j h\Delta} u_n + h \sum_{k=1}^{s} a_{j,k}(ih\Delta) g(t_n + c_k h, u_{n,k}), \qquad j \in \{1, \dots, s\}, \tag{1.25}$$

où les s^2 opérateurs $(a_{j,k})_{1\leq j,k\leq s}$ sont définis pour tout $(j,k)\in\{1,\dots,s\}^2$ par

$$a_{j,k}(ih\Delta)=\frac{1}{h}\int_0^{c_jh}e^{i(c_jh-\tau)\Delta}l_k(\tau)\mathrm{d}\tau.$$

L'étude des fonctions $b_j(z)$ et $a_{k,l}(z)$ pour $z\in\mathbb{C}$ fait apparaître le fait qu'elles sont bornées sur $i\mathbb{R}$. On en déduit qu'il existe une constante strictement positive $\mathcal{C}$ telle que pour tout $r\geq 0$, tout $(j,k)\in\{1,\dots,s\}^2$, tout $h>0$ et tout $u\in H^r$,

$$\|b_j(ih\Delta)u\|_{H^r}\leq\mathcal{C}\|u\|_{H^r}\qquad\text{et}\qquad\|a_{j,k}(ih\Delta)u\|_{H^r}\leq\mathcal{C}\|u\|_{H^r}.$$

Nous faisons les hypothèses suivantes sur la non linéarité g apparaissant dans l'équation (1.8), pour un certain $r\geq 0$ fixé :

Hypothèse 1.3.3 *La fonction g est telle que*

$$\forall t\in[0,T],\qquad g(t,0)=0.$$

Hypothèse 1.3.4 *La fonction g est Lipschitz-continue au sens où il existe $L>0$ et $\rho>0$ tels que pour tout couple (u,v) de fonctions de H^r tel que $\|u\|_{H^r(\mathbb{T}^d)}\leq\rho$ et $\|v\|_{H^r(\mathbb{T}^d)}\leq\rho$, on a*

$$\forall t\in[0,T],\qquad\|g(t,u)-g(t,v)\|_{H^r(\mathbb{T}^d)}\leq L\|u-v\|_{H^r(\mathbb{T}^d)}.$$

Hypothèse 1.3.5 *Le problème (1.8) admet une solution exacte $u:[0,T]\to H^r(\mathbb{T}^d)$ suffisamment régulière. En particulier, il existe $R>0$ tel que pour tout $t\in[0,T]$, on a $\|u(t)\|_{H^r(\mathbb{T}^d)}\leq R$.*

Hypothèse 1.3.6 *L'application $f:[0,T]\to H^r(\mathbb{T}^d)$ définie par $f(t)=g(t,u(t))$ est suffisamment régulière.*

Remarquons que les hypothèses (1.3.3) et (1.3.4) sont vérifiées par exemple pour la non linéarité cubique $g(u)=|u|^2u$ au moins quand $r>d/2$ et $r\notin\mathbb{N}$.

Sous ces hypothèses, à $\varphi_0\in H^r$ fixé, nous montrons l'existence d'un $h_0>0$ tel que la méthode de Runge-Kutta exponentielle de collocation (1.24)-(1.25) produit une unique solution numérique. En outre, nous montrons que la méthode est d'ordre au moins s. Précisément,

Théorème 1.3.7 *Supposons les hypothèses 1.3.3, 1.3.4, 1.3.5 et 1.3.6 vérifiées. Supposons que $\rho>0$ est suffisamment grand ($\rho\geq 3R\sqrt{s}$) et $h_0>0$ est suffisamment petit au sens où il vérifie $h_0\leq(3\mathcal{C}Ls^2)^{-1}$ et*

$$h_0^s\leq R\Big(\kappa_s e^{2\mathcal{C}sTL}\int_0^T\|f^{(s)}(\tau)\|_{H^r(\mathbb{T}^d)}\mathrm{d}\tau\Big)^{-1}.$$

Alors, pour tout $h\in]0,h_0[$, la méthode de Runge-Kutta exponentielle de collocation (1.24)-(1.25) fournit une unique solution numérique $(u_n)_{0\leq n\leq N}$ telle que pour tout $n\in$

$\{0,\dots,N\}$, $\|u_n\|_{H^s(\mathbb{T}^d)} \leq 2R$ et pour tout $i \in \{1,\dots,s\}$ $\|u_{n,i}\|_{H^r} \leq 3R$. De plus, pour tout $n \in \{0,\dots,N\}$,

$$\|u_n - u(t_n)\|_{H^r(\mathbb{T}^d)} \leq \kappa_s e^{\mathcal{C}_s LT}\, h^s \int_0^{t_n} \|f^{(s)}(\tau)\|_{H^r(\mathbb{T}^d)}\, \mathrm{d}\tau$$

où $\kappa_s = (s+1)\Big(\frac{1}{s!} + \frac{\mathcal{C}_s}{(s-1)!}\Big)$.

La preuve de ce résultat utilise un théorème de point fixe ainsi que des développements de Taylor de f (en temps) et des estimations de type Lipschitz de g (en espace). Ce résultat est similaire à celui obtenu dans le cas des problèmes paraboliques.

Comme pour l'analyse numérique de nos méthodes de Runge-Kutta exponentielles de collocation appliquées à des problèmes de Schrödinger linéaires présentée en 1.3.2.2, on donne ensuite des conditions suffisantes sur les points d'interpolation $(c_1,\dots,c_s)$ des méthodes pour que celles-ci soient d'ordre $s+1$, puis $s+2$. Ces conditions se traduisent en termes de conditions d'ordre pour la méthode de Runge-Kutta classique sous-jacente.

Une condition suffisante pour que la méthode de Runge-Kutta exponentielle de collocation (1.24)-(1.25) soit d'ordre $s+1$ est que la méthode de Runge-Kutta sous-jacente soit d'ordre $s+1$ et que le « second membre » f soit de plus assez régulier à la fois en temps et en espace :

Théorème 1.3.8 *Supposons que la méthode de Runge-Kutta sous-jacente est d'ordre $s+1$. Sous les hypothèses du Théorème 1.3.7, si $f^{(s)} \in C^1([0,T],H^{r+2})$, alors il existe une constante C strictement positive dépendant uniquement des points de collocation $(c_1,\dots,c_s)$, de L et de T telle que pour tout $h \in (0,h_0)$ et tout $n \in \{0,\dots,N\}$, on a*

$$\|u_n - u(t_n)\|_{H^r} \leq C\, h^{s+1} \Big(\|f^{(s)}(0)\|_{H^{r+2}} + \int_0^T \|f^{(s+1)}(\tau)\|_{H^{r+2}}\, \mathrm{d}\tau + \int_0^T \|f^{(s)}(\tau)\|_{H^r}\, \mathrm{d}\tau\Big).$$

La preuve de ce théorème est une nouvelle fois fondée sur des développements de Taylor de f (en temps) ainsi que des estimations de type Lipschitz de g (en espace). En outre, on utilise un lemme de Gronwall discret adapté.

Enfin, afin de prouver une condition suffisante pour que la méthode (1.24)-(1.25) appliquée au problème (1.8) soit d'ordre $s+2$, on renforce les hypothèses sur la non linéarité g ainsi que sur le second membre f.

Hypothèse 1.3.7 *Supposons qu'il existe deux constantes strictement positives L_1 et $\tilde{L}_1$ telles que*

$$\forall t \in [0,T], \quad \forall v \in H^r(\mathbb{T}^d), \qquad \Big\|\frac{\partial g}{\partial u}(t,u(t))v\Big\|_{H^r(\mathbb{T}^d)} \leq L_1 \|v\|_{H^r(\mathbb{T}^d)}.$$

et

$$\forall t \in [0,T], \quad \forall v \in H^{r+2}(\mathbb{T}^d), \qquad \Big\|\frac{\partial g}{\partial u}(t,u(t))v\Big\|_{H^{r+2}(\mathbb{T}^d)} \leq \tilde{L}_1 \|v\|_{H^{r+2}(\mathbb{T}^d)}.$$

Hypothèse 1.3.8 *Supposons qu'il existe une constante strictement positive L_2 telle que pour tout $t \in [0,T]$, tout $u, v, w \in H^r(\mathbb{T}^d)$ satisfaisant* $\max\{\|u\|_{H^r}, \|v\|_{H^r}, \|w\|_{H^r}\} \leq 2R$*, on a*

$$\left\| g(t,v) - g(t,w) - \frac{\partial g}{\partial u}(t,u)(v-w) \right\|_{H^r} \leq L_2 \Big(\|u-v\|_{H^r} + \|u-w\|_{H^r} \Big) \|v-w\|_{H^r}.$$

Hypothèse 1.3.9 *Supposons qu'il existe deux constantes strictement positives L_3 et $\tilde{L}_3$ telles que pour tout $t \in [0,T]$ et tout $v, w \in H^{r+2}(\mathbb{T}^d)$, on a*

$$\left\| \frac{\partial^2 g}{\partial u^2}(t,u(t))(v,w) \right\|_{H^r} \leq L_3 \|v\|_{H^r} \|w\|_{H^r},$$

$$\left\| \frac{\partial^2 g}{\partial t \partial u}(t,u(t))v \right\|_{H^r} \leq L_3 \|v\|_{H^r},$$

$$\left\| \frac{\partial^2 g}{\partial u^2}(t,u(t))(v,w) \right\|_{H^{r+2}} \leq \tilde{L}_3 \|v\|_{H^{r+2}} \|w\|_{H^{r+2}},$$

et

$$\left\| \frac{\partial^2 g}{\partial t \partial u}(t,u(t))v \right\|_{H^{r+2}} \leq \tilde{L}_3 \|v\|_{H^{r+2}}.$$

On fait en outre l'hypothèse supplémentaire suivante sur la solution exacte u du problème (1.8) :

Hypothèse 1.3.10 *Le problème* (1.8) *admet une solution exacte* $u : [0,T] \to H^{r+2}(\mathbb{T}^d)$ *suffisamment régulière. En particulier, $\varphi_0 = u_0 \in H^{r+2}(\mathbb{T}^d)$.*

Sous ces hypothèses supplémentaires, on montre le théorème suivant

Théorème 1.3.9 *Supposons vérifiées les hypothèses du Théorème 1.3.7 ainsi que l'Hypothèse 1.3.7, l'Hypothèse 1.3.8, l'Hypothèse 1.3.9 et l'Hypothèse 1.3.10. Supposons également que la méthode de Runge-Kutta sous-jacente est d'ordre $s+2$. Si de plus $f \in C^{s+2}([0,T], H^{r+4})$, alors il existe une constante strictement positive C telle que pour tout $h \in (0,h_0)$ et tout $n \in \{0,\ldots,N\}$, on a*

$$\|u(t_n) - u_n\|_{H^r} \leq Ch^{s+2}.$$

La constante C dépend ici de T et est fonction des intégrales des normes des dérivées temporelles de f. Ce théorème est similaire à celui obtenu dans le cas des problèmes paraboliques. La preuve de ce résultat est fondée sur l'utilisation de développements de Taylor en temps de l'erreur (et de f), des hypothèses sur la non linéarité g et d'un lemme de Gronwall discret.

Le Chapitre 3 présente des expériences numériques illustrant les résultats exposés dans cette section et les preuves détaillées de ces résultats.

Chapter 2

Normal form and long time analysis of splitting schemes for the linear Schrödinger equation with small potential

Contents

2.1 Introduction

In this paper, we consider the time-discretization of the linear Schrödinger equation by splitting methods and analyze the long time behavior of the corresponding "numerical" solution. Since no approximation in space is made, the problem is infinite dimensional, and the classical theory used in the case of ordinary differential equations cannot be applied. In particular, the long time behavior of the solution cannot be understood by the use of classical backward error analysis, see [HLW06, LR04]: In the finite dimensional case, a stability argument is invoked by assuming that the numerical solution lies in a compact set of the phase space over very long time. In infinite dimension, the corresponding assumption would require the *a priori* control of the regularity of the numerical solution over long time (see [Can06, OWW04] for the case of the non-linear wave equation).

In the case of splitting methods, exponential methods or standard methods for highly oscillatory equations, it is well known that for some values of the stepsize resonances appear, making the a priori assumption of uniform conservation of regularity irrelevant. In this work, we consider one of the simplest possible situations: The case of a splitting method applied to the linear periodic Schrödinger equation with an analytic and *small* potential in one space dimension. Moreover, we will consider the splitting scheme as a multiplicative symplectic perturbation of the free linear Schrödinger propagator, and show the quasi-persistence of the conservation properties over exponentially long time with respect to the size of the potential.

Before going on, let us mention that another possible way to analyze these problems could be using modulated Fourier expansions: See [CHL06] where the techniques used to analyze the exact solutions of non linear wave equations are clearly aimed at being applied to numerical analysis.

Let us consider the linear Schrödinger equation in one space dimension

$$i\frac{\partial\varphi}{\partial t}(x,t) = -\frac{\partial^2\varphi}{\partial x^2}(x,t) + V(x)\varphi(x,t), \quad \text{with} \quad \varphi(x,0) = \varphi^0(x), \tag{2.1}$$

where $\varphi(x,t)$ is the complex unknown wave function depending on the space variable $x \in \mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$ and the time $t \geq 0$. The potential V is a real function and the function φ^0 is the initial value of the wave function at $t = 0$. In the following, we write $\Delta = \partial_{xx}$ the Laplace operator in x so that the equation (2.1) reads

$$i\partial_t\varphi(t) = H\varphi(t), \quad \varphi(0) = \varphi^0, \quad \text{with} \quad H = -\Delta + V. \tag{2.2}$$

For a given time step $h > 0$, we consider the approximation scheme

$$\varphi(h) \simeq \exp(ih\Delta)\exp(-ihV)\varphi(0) \tag{2.3}$$

where by definition, $\exp(ih\Delta)\varphi$ is the solution $\psi(t)$ at the time $t = h$ of the equation

$$i\partial_t\psi(t) = -\Delta\psi(t), \quad \text{with} \quad \psi(0) = \varphi, \tag{2.4}$$

and similarly $\exp(-ihV)\varphi$ is the solution $\psi(t)$ at the time $t = h$ of the equation

$$i\partial_t \psi(t) = V\psi(t), \quad \text{with} \quad \psi(0) = \varphi. \tag{2.5}$$

If the potential is smooth enough, it can be shown that the approximation (2.3) is a first order approximation of the solution of (2.1), see [JL00] and [BBD02] (where the non-linear case is studied). Moreover, notice that the propagator associated with (2.3) is L^2-unitary and so the scheme conserves the L^2-norm as the exact propagator associated with (2.1) does. Splitting schemes are widely used to approximate the solution of (2.1) (see e.g. [dR87, MQ02] and the references therein) as they are simple to compute using fast Fourier transform: The free Schrödinger part (2.4) can be computed easily in terms of Fourier coefficients, while the solution of (2.5) is treated as an ordinary differential equation on the corresponding grid. Notice also that the Lie splitting scheme (2.3) is conjugated to the Strang splitting so that the long time behavior of these two schemes are equivalent (see Remark 2.2.2 below).

In the finite dimensional case, the behavior of splitting methods for hamiltonian systems is now well understood, see for instance [HLW06]. In particular, the use of the Baker-Campbell-Hausdorff formula shows that for a sufficiently small stepsize depending on the highest eigenvalue of the problem, there exists a modified hamiltonian for the propagator (2.3). The numerical flow can thus be interpreted as the exact solution of a hamiltonian system, at least for exponentially long time with respect to the stepsize. This result holds true for the linear and the non-linear case. When moreover the system is integrable, the techniques of classical perturbation theory apply, and it can be shown that the modified hamiltonian associated with a symplectic numerical method remains integrable, and that invariant tori persist over exponentially long time (see [HLW06], Chapter X, and the classical references therein).

It is worth noticing that in infinite dimension, the persistence of invariant tori for hamiltonian PDEs is a very difficult problem even for the exact solutions, and many progresses have been made very recently. For the case of the non-linear Schrödinger equation, we refer to [BG06] and [EK06] for results in this direction.

In our case, though the initial equation is linear, the splitting propagator can be viewed as a non-linear function of the infinite dimensional operators $-\Delta$ and V, and we use techniques similar to the one used in classical perturbation theory to put the propagator (2.3) under a normal form that will give information on the long time behavior of its solution.

The idea is to consider for a fixed time step h the family of propagators

$$L(\lambda) = \exp(ih\Delta)\exp(-ih\lambda V), \quad \lambda \in \mathbb{R}, \tag{2.6}$$

and to assume that V is analytic. For $\lambda = 0$, we see that $L(0)$ is the free linear Schrödinger propagator. The corresponding solution can be written explicitly in terms of Fourier coefficients. The dynamics is periodic in time and there is no mixing between the different Fourier modes. The regularity of the initial value is preserved.

In the case of the splitting scheme (2.6) when the perturbation parameter λ is small enough, we show that after a linear change of variable realized by a L^2-unitary operator satisfying exponential decay conditions on its coefficients, the propagator $L(\lambda)$ can be put under a normal form and written as an *almost X-shaped* L^2-unitary operator, up to exponentially small terms with respect to the small parameter λ. The coefficients of such an operator vanish, except possibly on the diagonal and the co-diagonal and for asymptotically large modes with respect to λ. This implies the existence of two-dimensional invariant spaces in the new variables, made of functions with zero Fourier coefficients except possibly at the indices k and $-k$ for a given $k \in \mathbb{N}$. This result is valid for modes $k \leq \lambda^{-\sigma}$ where $\sigma > 0$ and for exponentially long time with respect to λ.

To show this result, we use the following non-resonance condition on the stepsize (see [HLW06, Sha00]): There exist $\gamma > 0$ and $\nu > 1$ such that

$$\forall\, k \in \mathbb{Z}, \quad k \neq 0, \qquad \left| \frac{1 - e^{ihk}}{h} \right| \geq \gamma |k|^{-\nu}. \tag{2.7}$$

It can be shown that for a given $h_0 > 0$ close to 0, the set of time steps $h \in (0, h_0)$ that do not satisfy (2.7) has a Lebesgue measure $\mathcal{O}(h_0^{r+1})$ for some $r > 1$ (see [HLW06, Sha00]). The precise results are given in the next section.

Using this almost X-shaped representation, we can analyze the long time behavior of the numerical solution and show that the dynamics can be reduced to two dimensional linear symplectic systems mixing the two modes k and $-k$ for $k \leq \lambda^{-\sigma}$. This implies in particular the quasi-conservation of the regularity of the initial solution for these asymptotically large modes.

2.2 Statement of the results

In this section, we give a precise formulation of our main result and its consequences for the long time behavior of the solution. We then give a sketch of proof and the major ingredients used in the recipe. We conclude this section by showing with numerical experiments the necessity of the non-resonance condition.

2.2.1 The results

For a function $\psi \in L^2(\mathbb{T})$, the associated Fourier coefficients, for $n \in \mathbb{Z}$, are given by the formula

$$\hat{\psi}_n = \frac{1}{2\pi} \int_{\mathbb{T}} e^{-inx} \psi(x) \mathrm{d}x.$$

In all this paper, we identify a function ψ and its Fourier coefficients on $\mathbb{T}$, this means that we write for all $n \in \mathbb{Z}$, ψ_n for $\hat{\psi}_n$ and identify the collection $(\psi_n)_{n\in\mathbb{Z}}$ with the function ψ itself. We denote by $\|\psi\| = (\sum_{n\in\mathbb{Z}} |\psi_n|^2)^{1/2}$ the L^2-norm on $\mathbb{T}$. We also identify operators acting on $L^2(\mathbb{T})$ with operators acting on $l^2(\mathbb{Z})$. Such an operator S can thus be characterized by its complex coefficients $(S_{ij})_{(i,j)\in\mathbb{Z}^2}$. If $\psi = (\psi_n)_{n\in\mathbb{Z}} \in \mathbb{C}^{\mathbb{Z}}$, the product $\varphi = S\psi$ is defined by the sequence $\varphi = (\varphi_n)_{n\in\mathbb{Z}}$ of $\mathbb{C}^{\mathbb{Z}}$ with components $\varphi_n := \sum_{k\in\mathbb{Z}} S_{nk}\psi_k$, provided the summation makes sense. For example, the bounded operator defined as the multiplication by V acting on $L^2(\mathbb{T})$ is identified with the bounded operator (also denoted V) whose coefficients are given for all $(i,j) \in \mathbb{Z}^2$ by $V_{ij} = \hat{V}_{i-j}$.

For two operators A and B, the product AB is the operator whose coefficients are given formally by the definition

$$\forall\,(i,j) \in \mathbb{Z}^2, \quad (AB)_{ij} = \sum_{k\in\mathbb{Z}} A_{ik}B_{kj}. \tag{2.8}$$

For a given operator $S = (S_{ij})_{i,j\in\mathbb{Z}}$, we denote by S^* its adjoint with coefficients

$$\forall\,(i,j) \in \mathbb{Z}^2, \quad S^*_{ij} = \overline{S}_{ji}$$

where the bar denotes the complex conjugation. We say that S is symmetric if it satisfies $S = S^*$, and that S is L^2-unitary if it satisfies $S^*S = \mathrm{Id}$.

For an operator S and $\rho \in \mathbb{R}^+$, we define set

$$\|S\|_\rho = \sup_{k,\ell\in\mathbb{Z}} \left(e^{\rho|k-\ell|}|S_{k\ell}|\right) \tag{2.9}$$

and we define $\mathcal{A}_\rho$ as the space of S with finite analytic operator norm $\|S\|_\rho < \infty$. The space $\mathcal{S}_\rho$ denotes the subspace of symmetric operators of $\mathcal{A}_\rho$.

We define moreover the set of *X-shaped* operators $\mathcal{X}_\rho$ made of the elements of $X \in \mathcal{A}_\rho$ for which we have

$$X_{k\ell} \neq 0 \Longrightarrow |k| = |\ell|.$$

For functions ψ on $\mathbb{T}$, we define

$$\|\psi\|_\rho = \sup_{k\in\mathbb{Z}} \left(e^{\rho|k|}|\psi_k|\right)$$

for a given positive number ρ, and set

$$U_\rho := \{\ \psi \mid \|\psi\|_\rho < \infty\ \}$$

the corresponding function space[1] of functions on $\mathbb{T}$ with finite analytic norm.

Hypothesis 2.2.1 *We assume that there exists a function $\mathcal{V}$ and $\rho_V > 0$ such that $\mathcal{V}$ is analytic on a complex domain containing the closure $\bar{\Omega}$ of $\mathbb{T} + i(-\rho_V, \rho_V)$ and such that for all $x \in \mathbb{T}$, $V(x) = \mathcal{V}(x) \in \mathbb{R}$. Therefore, if M_V denotes the maximum of the function $|\mathcal{V}|$ on $\bar{\Omega}$, then we have*

$$\forall k \in \mathbb{Z}, \quad \forall n \in \mathbb{N}, \qquad |\widehat{(V^n)}_k| \leq M_V^n e^{-\rho_V|k|},$$

so that for all $n \geq 0$, $\|V^n\|_{\rho_V} \leq M_V^n$.

[1]The notation $\|\cdot\|_\rho$ is not ambiguous: If an operator W is induced by a function, i.e. if $W_{ij} = \hat{\psi}_{i-j}$ for $(i,j) \in \mathbb{Z}^2$ and for a function ψ with Fourier coefficients $\hat{\psi}_n$, $n \in \mathbb{Z}$, then we have $\|W\|_\rho = \|\psi\|_\rho$.

For a given $K > 0$ we define the set of indices

$$I_K = \{(k,\ell) \in \mathbb{Z}^2 \quad | \quad |k| \leq K \quad \text{or} \quad |\ell| \leq K\}. \tag{2.10}$$

We then set $\mathcal{X}_\rho^K$ the set of operators that are *almost X-shaped* in the sense where

$$X_{k\ell} \neq 0 \Longrightarrow \Big(|k| = |\ell| \quad \text{or} \quad (k,\ell) \notin I_K \Big)$$

The aim of this paper is to prove the following conjugation result for the propagator:

Theorem 2.2.1 *Assume that V satisfies Hypothesis* 2.2.1 *and let $L(\lambda)$ be defined by* (2.6). *Assume $\gamma > 0$ and $\nu > 1$ are given. There exist positive constants λ_0, σ and c depending only on V, ρ_V, γ and ν, such that for all stepsizes $h \in (0,1)$ satisfying* (2.7)*, there exist families of L^2-unitary operators $Q(\lambda)$ and $\Sigma(\lambda)$ analytic in λ for $|\lambda| < \lambda_0$ such that for $\lambda \in (0, \lambda_0)$,*

$$Q(\lambda) \in \mathcal{A}_{\rho_V/4} \quad \textit{and} \quad \Sigma(\lambda) \in \mathcal{X}_{\rho_V/4}^K \quad \textit{with} \quad K = \lambda^{-\sigma} \tag{2.11}$$

where $\sigma = 1/(32(\nu+1)) > 0$,

$$\|Q(\lambda) - \mathrm{Id}\|_{\rho_V/4} \leq \lambda^{1/2} \quad \textit{and} \quad \|\Sigma(\lambda) - e^{ih\Delta}\|_{\rho_V/4} \leq h\lambda^{1/2}, \tag{2.12}$$

such that the following identity holds:

$$Q(\lambda) L(\lambda) Q(\lambda)^* = \Sigma(\lambda) + R(\lambda). \tag{2.13}$$

Moreover, the remainder term $R(\lambda)$ satisfies, for $\lambda \in (0, \lambda_0)$,

$$\|R(\lambda)\|_{\rho_V/5} \leq \exp(-c\lambda^{-\sigma}). \tag{2.14}$$

Roughly speaking, the preceding result shows that after a unitary change of variables close to the identity in some analytic operator norm, the dynamics can be reduced, up to exponentially small terms, to the action of $\Sigma(\lambda)$ which decouples into 2×2 symplectic systems for each modes $\pm k$. This is valid for asymptotically large modes $|k| \leq \lambda^{-\sigma}$. More precisely, if φ is a function and if $\psi = \Sigma(\lambda)\varphi$, we have for $|k| \leq \lambda^{-\sigma}$,

$$\begin{pmatrix} \psi_k \\ \psi_{-k} \end{pmatrix} = \begin{pmatrix} a_k(\lambda) & b_k(\lambda) \\ c_k(\lambda) & d_k(\lambda) \end{pmatrix} \begin{pmatrix} \varphi_k \\ \varphi_{-k} \end{pmatrix} \tag{2.15}$$

where the 2×2 matrix in this equality is close to the diagonal matrix with entries e^{-ihk^2}, and is unitary. This implies that we have for all k such that $|k| \leq \lambda^{-\sigma}$,

$$|\psi_k|^2 + |\psi_{-k}|^2 = |\varphi_k|^2 + |\varphi_{-k}|^2. \tag{2.16}$$

Combining the conservation law (2.16) of $\Sigma(\lambda)$ with the exponential estimate (2.14) will allow us to derive long time bounds for the iterates of $L(\lambda)$.

In the following, for a function φ, we use the notation

$$|\varphi|_0^2 = |\varphi_0|^2 \quad \text{and} \quad \forall\, k \in \mathbb{Z}\backslash\{0\}, \quad |\varphi|_k^2 = |\varphi_k|^2 + |\varphi_{-k}|^2 \tag{2.17}$$

to denote the energies at the frequency $|k|$.

Moreover, for $s > 0$ we introduce the norm[2]

$$\|\varphi\|_{s,\infty} = \sup_{k\geq 0}\left((1+k)^s|\varphi|_k\right). \tag{2.18}$$

Using the preceding remark and the estimates on the remainder term, we can show the following corollary, which yields long time results for the approximated solution:

Corollary 2.2.1 *We use the notations of the previous Theorem. For $n \in \mathbb{N}$, we set $\varphi^n = L(\lambda)^n\varphi^0$.*
(i) *There exists a constant $C > 0$ depending only on V, ρ_V, γ and ν such that for all $h \in (0,1)$ satisfying* (2.7), *all $\lambda \in (0, \lambda_0)$, $n \leq \exp(c\lambda^{-\sigma}/2)$, and $\varphi^0 \in L^2(\mathbb{T})$,*

$$\forall\, k \in \mathbb{N}, \quad k \leq \lambda^{-\sigma}, \quad \big|\, |\varphi^n|_k - |\varphi^0|_k \,\big| \leq C\lambda^{1/2}\|\varphi^0\|\,. \tag{2.19}$$

(ii) *Assume that $s > 1/2$ is given, and let s' be such that $s - s' \geq 1/2$. Then there exists a constant c_s depending only on V, ρ_V, γ, ν and s, such that for all $h \in (0,1)$ satisfying* (2.7)*, all $\lambda \in (0, \lambda_0)$, $n \leq \exp(c\lambda^{-\sigma}/2)$, and φ^0 such that $\|\varphi^0\|_{s,\infty} < +\infty$, we have*

$$\sup_{0\leq k\leq\lambda^{-\sigma}}\Big((1+k)^{s'}\,\big|\,|\varphi^n|_k - |\varphi^0|_k\,\big|\Big) \leq c_s\lambda^{1/2}\|\varphi^0\|_{s,\infty}\,. \tag{2.20}$$

(iii) *For all $\rho \in (0, \rho_V/5)$, there exist positive constants μ_0 and C (depending only on V, ρ_V, γ, ν and ρ) such that for all $h \in (0,1)$ satisfying* (2.7)*, all $\lambda \in (0, \lambda_0)$, $n \leq \exp(c\lambda^{-\sigma}/2)$, $\mu \in (0, \mu_0)$ and $\varphi^0 \in U_\rho$,*

$$\sup_{0\leq k\leq\lambda^{-\sigma}}\Big(e^{\mu k}\,\big|\,|\varphi^n|_k - |\varphi^0|_k\,\big|\Big) \leq C\lambda^{1/2}\|\varphi^0\|_\rho\,. \tag{2.21}$$

Remark 2.2.1 *We could also diagonalize the 2×2 matrix in* (2.15)*. As the eigenvalues are close to e^{ihk^2}, we could obtain this way a quasiperiodic behavior of the numerical solution in suitable coordinates, with frequencies of the form $k^2 + \mathcal{O}(\lambda^{1/2})$. However, as e^{ihk^2} is a double eigenvalue of the limit matrix, this would not imply the continuity of the transformation in terms of the small parameter λ.*

Remark 2.2.2 *The same results hold true for the family of Strang splitting propagators*

$$\exp(ih\Delta/2)\exp(-ih\lambda V)\exp(ih\Delta/2)$$

which are conjugated to the Lie splitting schemes (2.6) *by the operator* $\exp(ih\Delta/2)$ *that defines an isometry between all $\mathcal{A}_\rho$ spaces.*

Remark 2.2.3 *The choice of $\rho_V/4$ and $\rho_V/5$ in Theorem 2.2.1 is made for convenience in the proof, and could be replaced by any number of the type $\rho_V - \delta$ with $0 < \delta < \rho_V$. This would only change the values of the constants in the statements.*

[2]Note that $\|\varphi\|_{s,\infty} < \infty$ implies that φ is in the Sobolev space $H^{s-1/2-\varepsilon}$ for all $\varepsilon > 0$. If for instance $\|\varphi\|_{s,\infty} < \infty$ for some $s > 1/2$ then $\varphi \in L^2$.

2.2.2 Sketch of proof

The proof of Theorem 2.2.1 is divided into several steps. We describe here the main arguments:
(i) Formal operator series. We first seek the operators $Q(\lambda)$ and $\Sigma(\lambda)$ as formal series of the form

$$Q(\lambda) = \sum_{n\geq 0} \lambda^n Q_n \quad \text{and} \quad \Sigma(\lambda) = \sum_{n\geq 0} \lambda^n \Sigma_n$$

where for all $n \in \mathbb{N}$, the coefficients of the operators Q_n and Σ_n satisfy exponential decay conditions of the form (2.9) and where the equation

$$Q(\lambda)L(\lambda)Q(\lambda)^* = \Sigma(\lambda) \tag{2.22}$$

is satisfied in the sense of formal series. Note that as V is a bounded operator, the formal series $L(\lambda)$ is in fact a power series in λ.

In order to ensure the fact that $Q(\lambda)$ and $\Sigma(\lambda)$ are unitary operators, we introduce the "logarithm" formal series

$$S = Q^*(i\partial_\lambda Q) \quad \text{and} \quad X = \Sigma^*(i\partial_\lambda \Sigma)$$

and we look for symmetric coefficients in the formal series $S(\lambda)$ and $X(\lambda)$. Writing down the equations for the coefficients, we see that for all $n \geq 0$, S_n and X_n have to satisfy an equation that reads

$$S_n - e^{ih\Delta} S_n e^{-ih\Delta} + X_n = G_n \tag{2.23}$$

where G_n is symmetric and depends on V, S_p and X_p for $p = 0, \ldots, n-1$. The study of the homological equation (2.23) is thus the cornerstone for the recursive construction of the operators. The goal of the first part of Section 2.3 is to make explicit the recursive equalities (2.23).
(ii) Solution of the homological equation. In the second part of Section 2.3, we consider the equation

$$S - e^{-ih\Delta} S e^{ih\Delta} + X = G$$

where G is a given symmetric operator. In terms of coefficients, as the Laplace operator is a diagonal operator with entries $-k^2$, $k \in \mathbb{Z}$, this equation can be written

$$(1 - e^{ih(k^2-\ell^2)}) S_{k\ell} + X_{k\ell} = G_{k\ell}$$

for $(k, \ell) \in \mathbb{Z}^2$. We see that in the case where $k^2 = \ell^2$, the value of $X_{k\ell}$ is imposed. This is the reason for the definition of X-shaped operators. Moreover, a problem of small divisors may appear when $h(k^2 - \ell^2)$ is close to a multiple of 2π. The use of the condition (2.7) allows in principle to determine a solution

$$S_{k\ell} = \frac{1}{1 - e^{ih(k^2-\ell^2)}} G_{k\ell}$$

for $k^2 \neq \ell^2$. However, the condition (2.7) and the previous equality do not imply that S is in a space of exponentially decaying matrices $\mathcal{A}_\rho$ for a suitable ρ if G is. This is due to possibly unbounded coefficients $k+\ell$ in the term $|k^2-\ell^2|$. This is the reason for the introduction of the set of indexes I_K in (2.10) and the corresponding I_K-solution of the homological equation (see Definition 2.3.1 below). This explains the final form of the propagator $\Sigma(\lambda)$.

Thanks to the non-resonance condition (2.7), we are now able to state a result for the I_K-solutions of the recursive homological equations, see Proposition 2.3.1. In particular, we obtain the estimates (2.42) which are familiar in backward error analysis and in classical perturbation theory.

(iii) Estimates and optimization of the constants. Sections 2.4 and 2.5 are the most technical ones and show the precise estimates of the Theorem. Using I_K-solutions of the recursive homological equations, we can prove analytic estimates for the coefficients S_n and X_n: see Proposition 2.4.1. We obtain that for large n, the analytical norms of these coefficients are bounded by

$$(K^\alpha n^\beta)^n$$

up to some constants, and for suitable positive coefficients α and β. Though these series diverge, we can define for a fixed index N, the corresponding truncated series and operators $Q^{[N]}(\lambda)$ and $\Sigma^{[N]}(\lambda)$ that satisfy the equation (2.22) up to a remainder term $R^{[N]}(\lambda)$ depending on N. We then use these estimates to give a bound for the remainder term, see Proposition 2.5.2. Roughly speaking, the remainder term is bounded in a suitable $\mathcal{A}_\rho$ space by

$$(\lambda K^\alpha N^\beta)^N$$

for $\lambda \in (0, \lambda_0)$ and up to some constant. We then take K and N proportional to a negative power of λ to obtain an exponential estimate of the form (2.14).

(iv) Proof of Corollary 2.2.1. Section 2.6 is devoted to the proof of this corollary. The idea is to use the almost X-shaped form of $\Sigma(\lambda)$ and the estimate (2.14) of the remainder term $R(\lambda)$. We divide the phase space into the spaces $\{\varphi \,|\, \varphi_k = 0, \quad |k| \leq K\}$ and $\{\varphi \,|\, \varphi_k = 0, \quad |k| > K\}$. If π_K denotes the L^2-orthogonal projection onto the latter space, the almost X-shaped structure of $\Sigma(\lambda)$ implies that $[\Sigma(\lambda), \pi_K] = 0$ so that these spaces are invariant by $\Sigma(\lambda)$. The ingredients to show the results are the following two key estimates: If for all $n \in \mathbb{N}$, we set $\psi^n = Q(\lambda)\varphi^n$, then we prove that for $n \leq \exp(c\lambda^{-\sigma}/2)$ we have

$$\forall\, 0 \leq k \leq K, \quad \big|\, |\psi^n|_k - |\psi^0|_k \big| \leq C \exp(-c\lambda^{-\sigma}/2) \|\varphi^0\|$$

for a constant C depending only on V, ρ_V and the parameters appearing in the non-resonance condition (2.7). This shows the extremely good conservation of the energies at a given mode $k \leq K = \lambda^{-\sigma}$ in the "new" variables, and yields easily the estimates of the corollary for ψ^n. In order to get back to the original variables, we have to control the modes with indexes greater than K by showing that for $n \leq \exp(c\lambda^{-\sigma}/2)$,

$$\Big| \|\pi_K \psi^n\| - \|\pi_K \psi^0\| \Big| \leq C \exp(-c\lambda^{-\sigma}/2) \|\varphi^0\|$$

where C does not depend on λ. The results are obtained by combining these two estimates, and by using the fact that for a smooth function, the term $\|\pi_K \psi^0\|$ is small for big $K = \lambda^{-\sigma}$.

2.2.3 Numerical experiments

We consider the case where the potential function and the initial wave function are given by

$$V(x) = \frac{3}{5 - 4\cos(x)} \quad \text{and} \quad \varphi^0(x) = \frac{2}{2 - \cos(x)}.$$

We use two different time steps:

$$h = 0.2 \quad \text{and} \quad h = \frac{2\pi}{6^2 - 2^2} = 0.196\ldots. \tag{2.24}$$

The first time step satisfies the non-resonance condition[3] (2.7) while the second one is obviously resonant. We use fast Fourier transformations to compute the solution of (2.3).

In Figure 2.1, we take $2^7 + 1 = 129$ Fourier modes. We make 10^6 iterations, and we set $\lambda = 0.1$. We plot the first 5 energies $|\varphi^n|_k$, $k = 0, \ldots, 4$ in logarithmic scale. We see that if the non-resonance condition is not satisfied, the conservation properties are lost.

In Figure 2.2, we use $2^9 + 1 = 513$ Fourier modes and plot all the energies for 10^5 iterations with $\lambda = 0.01$. We see the growth of high order modes (recall that the L^2-norm is conserved).

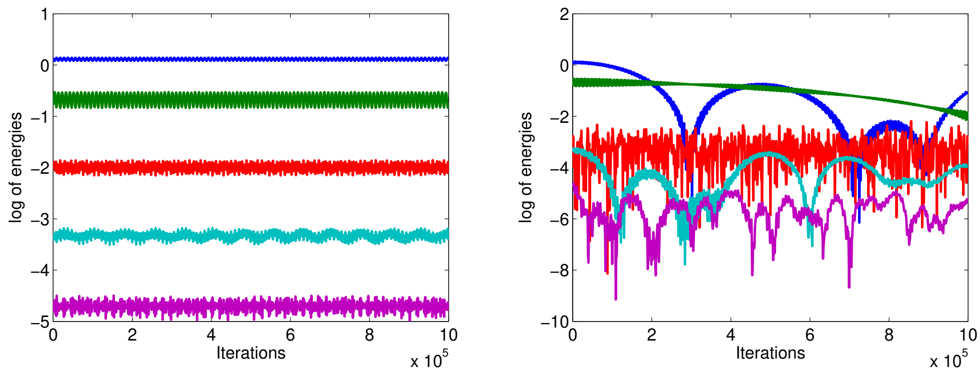

Figure 2.1: Energies of the 5 first modes in logarithmic scale, $\lambda = 0.1$. Non-resonant stepsize (left) and resonant stepsize (right)

[3] We thank Z. Shang who showed us this fact.

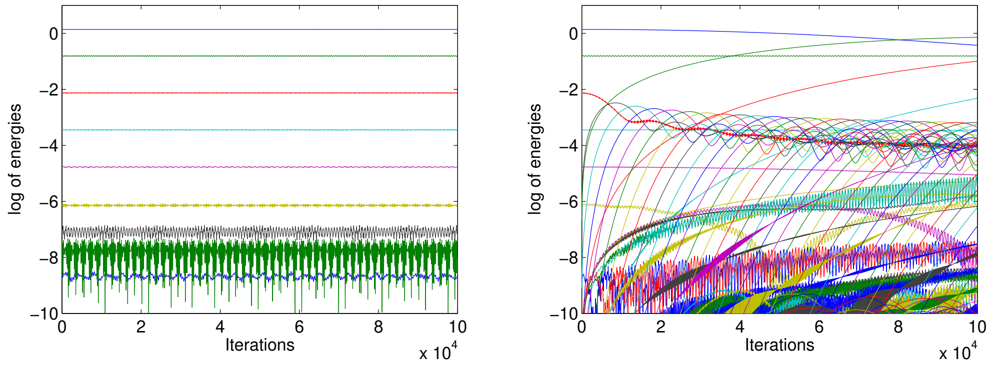

Figure 2.2: Energies in logarithmic scale, $\lambda = 0.01$. Non-resonant stepsize (left) and resonant stepsize (right)

2.3 Formal series and homological equation

2.3.1 Formal operator series

We now start the proof of Theorem 2.2.1. In this Section, we consider all the operators as formal operators series in λ. We show rigorous estimates in the next section. Let $A(\lambda) = \sum_{n\geq 0} \lambda^n A_n$ and $B(\lambda) = \sum_{n\geq 0} \lambda^n B_n$ be two formal series whose coefficients respectively are operators A_n and B_n. The product $A(\lambda)B(\lambda) = C(\lambda)$ is the formal series $C(\lambda) = \sum_{n\geq 0} \lambda^n C_n$ with coefficients

$$C_n = \sum_{k=0}^{n} A_k B_{n-k}, \quad n \geq 0,$$

where the product between A_k and B_{n-k} is the operator product defined in (2.8). Similarly, we can define the action of the derivative in λ on formal series by

$$\partial_\lambda A(\lambda) = \sum_{n\geq 0} \lambda^n (n+1) A_{n+1}$$

for a formal series $A(\lambda) = \sum_{n\geq 0} \lambda^n A_n$.

Recall that $L(\lambda)$ is given by (2.6). As a formal series in λ, we write

$$L(\lambda) = \sum_{n\geq 0} \frac{\lambda^n}{n!} e^{ih\Delta}(-ihV)^n.$$

We thus have the equalities

$$i\partial_\lambda L(\lambda) = e^{ih\Delta} hV e^{-ih\Delta} L(\lambda) \qquad \text{and} \qquad L(0) = e^{ih\Delta}.$$

We now seek formal series $Q(\lambda) = \sum_{n\geq 0} \lambda^n Q_n$ and $\Sigma(\lambda) = \sum_{n\geq 0} \lambda^n \Sigma_n$ satisfying

$$Q(\lambda)L(\lambda)Q(\lambda)^* = \Sigma(\lambda), \tag{2.25}$$

such that

$$Q(\lambda)^*Q(\lambda) = \mathrm{Id} \quad \text{and} \quad \Sigma(\lambda)^*\Sigma(\lambda) = \mathrm{Id}, \tag{2.26}$$

and where $\Sigma(\lambda)$ is an X-shaped operator.

Taking the derivative $i\partial_\lambda$ of the equation (2.25) yields

$$(i\partial_\lambda Q)LQ^* - QL(i\partial_\lambda Q)^* + hQe^{ih\Delta}Ve^{-ih\Delta}LQ^* = i\partial_\lambda \Sigma,$$

whence

$$Q^*(i\partial_\lambda Q)L - L(i\partial_\lambda Q)^*Q + he^{ih\Delta}Ve^{-ih\Delta}L = Q^*(i\partial_\lambda \Sigma)Q. \tag{2.27}$$

We introduce the operators

$$S = Q^*(i\partial_\lambda Q) \quad \text{and} \quad X = \Sigma^*(i\partial_\lambda \Sigma).$$

The formal series Q and Σ can be reconstructed from the series S and X using the initial conditions $Q(0) = \mathrm{Id}$ and $\Sigma(0) = e^{ih\Delta}$.

Note that if the formal series S and X are symmetric, Q and Σ are unitary (i.e. satisfy (2.26)). Moreover, if X an X-shaped operator, Σ will also be X-shaped.

Assuming S is symmetric, the equation (2.27) can be written (we use $\Sigma = QLQ^*$)

$$SL - LS + he^{ih\Delta}Ve^{-ih\Delta}L = Q^*\Sigma XQ = LQ^*XQ$$

whence, as V and $e^{-ih\lambda V}$ commute, and as $L^*L = \mathrm{Id}$,

$$S - L^*SL = hV - Q^*XQ. \tag{2.28}$$

Our formal series problem will hence be the following: Find formal series $S(\lambda) = \sum_{n\geq 0}\lambda^n S_n$ and $X(\lambda) = \sum_{n\geq 0}\lambda^n X_n$ such that for all $n \geq 0$, $S_n^* = S_n$, $X_n^* = X_n$ and X_n is a X-shaped operator, satisfying the equation

$$S(\lambda) - L^*(\lambda)S(\lambda)L(\lambda) = hV - Q^*(\lambda)X(\lambda)Q(\lambda) \tag{2.29}$$

where the operator Q satisfies

$$i\partial_\lambda Q(\lambda) = Q(\lambda)S(\lambda), \quad \text{and} \quad Q(0) = \mathrm{Id}. \tag{2.30}$$

The operator Σ will then be reconstructed using the formulae

$$i\partial_\lambda \Sigma(\lambda) = \Sigma(\lambda)X(\lambda), \quad \text{and} \quad \Sigma(0) = e^{ih\Delta}. \tag{2.31}$$

2.3.2 Coefficients

We now write the previous formal series equations in terms of the coefficients of the operator formal series. The equation (2.30) can be written: For all $n \geq 0$,

$$i(n+1)Q_{n+1} = \sum_{k=0}^{n} Q_k S_{n-k}. \tag{2.32}$$

The equation (2.31) yields similar equalities for Σ_n and X_n. Note that we impose $Q_0 = I$ and $\Sigma_0 = e^{ih\Delta}$.

Let us write the coefficients of the equation (2.29): For all $n \in \mathbb{N}$,

$$S_n - \sum_{p+q+r=n} \frac{(ihV)^p}{p!} e^{-ih\Delta} S_q e^{ih\Delta} \frac{(-ihV)^r}{r!} = \delta_n^0 hV - \sum_{p+q+r=n} Q_p^* X_q Q_r. \tag{2.33}$$

Assuming that S_k and X_k are determined for $k = 0, \cdots, n-1$ and that they are symmetric, the equation (2.33) can be written

$$\begin{aligned} S_n - e^{-ih\Delta} S_n e^{ih\Delta} + X_n &= \delta_n^0 hV \\ &+ \sum_{\{p+q+r=n \mid q \neq n\}} \left(\frac{(ihV)^p}{p!} e^{-ih\Delta} S_q e^{ih\Delta} \frac{(-ihV)^r}{r!} - Q_p^* X_q Q_r \right). \end{aligned} \tag{2.34}$$

Since the summation set is symmetric in r and p, the right hand side of this equation is symmetric.

We thus see that for all n, we have to solve a homological equation

$$S_n - e^{-ih\Delta} S_n e^{ih\Delta} + X_n = G_n, \tag{2.35}$$

where G_n is symmetric, and where we seek S_n and X_n symmetric with X_n a X-shaped operator.

Let us first consider the equation (2.34) for $n = 0$. We have

$$S_0 - e^{-ih\Delta} S_0 e^{ih\Delta} + X_0 = hV. \tag{2.36}$$

The operator Δ is diagonal with entries $-k^2$, $k \in \mathbb{Z}$. Hence for $(k, \ell) \in \mathbb{Z}^2$, the previous equation can be written

$$(S_0)_{k\ell}\big(1 - e^{ih(k^2-\ell^2)}\big) = hV_{k\ell} - (X_0)_{k\ell}$$

so that we can choose for $k^2 \neq \ell^2$, under the condition that $h(k^2 - \ell^2)$ is not a multiple of 2π,

$$(S_0)_{k\ell} = \frac{1}{1 - e^{ih(k^2-\ell^2)}} hV_{k\ell} \quad \text{and} \quad (X_0)_{k\ell} = 0 \tag{2.37}$$

so that X_0 is X-shaped. For $k^2 = \ell^2$ the previous equation degenerates, but we can take

$$(S_0)_{k\ell} = 0 \quad \text{and} \quad (X_0)_{k\ell} = hV_{k\ell}.$$

The previous equations define symmetric operators S_0 and X_0, and X_0 is X-shaped. By induction, the formal series problem of the previous subsection is then solved. However, we cannot obtain estimates in $\mathcal{A}_\rho$-spaces for the operators, even if h satisfies the estimate (2.7). The reason for this is that using (2.7), the equation (2.37) yields the following estimate, for $k \neq \ell$,

$$|(S_0)_{k\ell}| \leq \gamma^{-1} |k^2 - \ell^2|^\nu |V_{k\ell}|$$

and the fact that $V \in \mathcal{A}_{\rho_V}$ does not imply that $S_0 \in \mathcal{A}_\rho$ for a suitable ρ. This is due to the possibly unbounded coefficients $k + \ell$.

2.3.3 Solution of the homological equation

It is worth noticing that the above equation (2.35) is underdetermined. The previous computations motivate the introduction of the following solutions of the homological equation:

Definition 2.3.1 *Let G be a symmetric operator and $K > 0$ be a real number. Assume that $h \in (0,1)$ satisfies* (2.7)*. Let us consider the equation:*

$$S - e^{ih\Delta} S e^{-ih\Delta} + X = G. \tag{2.38}$$

We define as I_K-solution of the homological equation (2.38) *the symmetric operators S and X defined for $(k,l) \in \mathbb{Z}^2$ by*

$$\text{if } (k,\ell) \in I_K \text{ and } |k| \neq |\ell|, \quad \begin{cases} S_{k\ell} &= \dfrac{1}{1 - e^{ih(k^2-\ell^2)}} G_{k\ell}, \\ X_{k\ell} &= 0, \end{cases} \tag{2.39}$$

$$\text{otherwise, if } (k,\ell) \notin I_K \text{ or } |k| = |\ell|, \begin{cases} S_{k\ell} &= 0, \\ X_{k\ell} &= G_{k\ell}, \end{cases} \tag{2.40}$$

where I_K is defined in (2.10)*.*

With this definition, the following theorem yields the cornerstone of the solving process:

Proposition 2.3.1 *Let $G \in \mathcal{S}_\rho$ for a given $\rho > 0$. Assume that h satisfies* (2.7) *with the constants γ and ν. Let S and X be the I_K-solutions of the homological equation*

$$S - e^{-ih\Delta} S e^{ih\Delta} + X = G \tag{2.41}$$

for a given $K \geq 1$. Then for all δ such that $0 < \delta \leq \rho$, we have $S \in \mathcal{S}_{\rho-\delta}$, $X \in \mathcal{S}_\rho \cap \mathcal{X}_\rho^K$, and the following estimates hold:

$$\|S\|_{\rho-\delta} \leq \frac{K^\nu}{\gamma h} \left(\frac{4\nu}{e\delta} \right)^{2\nu} \|G\|_\rho \quad \text{and} \quad \|X\|_\rho \leq \|G\|_\rho . \tag{2.42}$$

Proof. The fact that $\|X\|_\rho \leq \|G\|_\rho$ is clear from (2.40).
Using (2.39) and (2.7), we have for $(k,\ell) \in I_K$ and $|k| \neq |\ell|$,

$$|S_{k\ell}| \leq \frac{1}{\gamma h} |k^2 - \ell^2|^\nu |G_{k\ell}|.$$

But for $(k,\ell) \in I_K$, we easily see that $|k+\ell| \leq 2K + |k-\ell|$. Hence we have

$$\begin{aligned} |k^2 - \ell^2|^2 &\leq |k-\ell|^2 (2K + |k-\ell|)^2 \\ &\leq 2|k-\ell|^2 (4K^2 + |k-\ell|^2) \\ &\leq 8K^2 |k-\ell|^2 + 2|k-\ell|^4 \\ &\leq 16K^2 |k-\ell|^4 \end{aligned}$$

and therefore $|k^2-\ell^2|^\nu \le 2^{2\nu}K^\nu|k-\ell|^{2\nu}$. This implies that for all $(k,\ell)\in I_K$ with $|k|\neq|\ell|$,

$$|S_{k\ell}| \le \frac{K^\nu}{\gamma h}2^{2\nu}|k-\ell|^{2\nu}\|G\|_\rho e^{-\rho|k-\ell|}.$$

Hence, for all $(k,\ell)\in\mathbb{Z}^2$,

$$|S_{k\ell}|e^{(\rho-\delta)|k-\ell|} \le \frac{K^\nu}{\gamma h}2^{2\nu}\|G\|_\rho |k-\ell|^{2\nu}e^{-\delta|k-\ell|}.$$

The fact that for all $(k,\ell)\in\mathbb{Z}^2$,

$$|k-\ell|^{2\nu}e^{-\delta|k-\ell|} \le \left(\frac{2\nu}{e\delta}\right)^{2\nu}$$

allows to conclude. ■

In the next section, we will consider the recursive solutions of the collection of equations (2.34) by induction on n using the I_K-solutions of the homological equation and the previous Proposition.

2.4 Estimates for the coefficients

We now give estimates for the coefficients of the formal series $S(\lambda)$, $X(\lambda)$ and $Q(\lambda)$ given by the recursive I_K-solutions of (2.34) for a fixed $K\ge 1$. Recall that V is assumed to satisfy Hypothesis 2.2.1.

Let us first consider the operators S_0 and X_0 given by the I_K-solution of (2.36). Using Proposition 2.3.1, we have for all $0<\delta\le\rho_V$

$$\|S_0\|_{\rho_V-\delta} \le M_V\frac{K^\nu}{\gamma}\left(\frac{4\nu}{e\delta}\right)^{2\nu} \quad\text{and}\quad \|X_0\|_{\rho_V}\le hM_V. \tag{2.43}$$

In particular, there exists a constant C depending only on V, ρ_V, γ and ν such that

$$\|S_0\|_{\rho_V/3}\le K^\nu C \quad\text{and}\quad \|X_0\|_{\rho_V/3}\le hC. \tag{2.44}$$

Moreover, since $Q_0=\mathrm{Id}$, we have that $\|Q_0\|_{\rho_V/3}=1$.

The goal of this section is to prove the following result:

Proposition 2.4.1 *There exist constants $C_0\ge 1$ and $K_0\ge 1$ depending only on V, ρ_V, γ and ν such that for all $K\ge K_0$, and all $h\in(0,1)$ satisfying (2.7), if we denote by $(S_J)_{J\ge0}$, $(Q_J)_{J\ge0}$ and $(X_J)_{J\ge0}$ the operators obtained by the iterative I_K-solutions of (2.34), we have for all $J\ge1$,*

$$\|S_J\|_{\rho_V/3}+\|Q_J\|_{\rho_V/3} \le \left(C_0K^\alpha J^\beta\right)^J \tag{2.45}$$

$$\text{and}\quad \|X_J\|_{\rho_V/3} \le h\left(C_0K^\alpha J^\beta\right)^J, \tag{2.46}$$

where $\alpha=2\nu$ and $\beta=4\nu+3$.

Before giving the proof of this proposition, we first give the following lemma that provides a bound for the product of two operators with exponential decay of coefficients away from the diagonal:

Lemma 2.4.1 *Let $\rho \geq 0$ and $\delta > 0$. Let $A \in \mathcal{A}_\rho$ and $B \in \mathcal{A}_{\rho+\delta}$. Then the products AB and BA are in the space $\mathcal{A}_\rho$ and we have the estimates*

$$\|AB\|_\rho \leq C_\delta \; \|A\|_\rho \|B\|_{\rho+\delta} \quad \text{and} \quad \|BA\|_\rho \leq C_\delta \; \|A\|_\rho \|B\|_{\rho+\delta}, \tag{2.47}$$

where $C_\delta = (1+e^{-\delta})(1-e^{-\delta})^{-1} > 1$.

Proof. By definition, we have for $(k,\ell,m) \in \mathbb{Z}^3$

$$|A_{k\ell}| \leq \|A\|_\rho \, e^{-\rho|k-\ell|} \quad \text{and} \quad |B_{\ell m}| \leq \|B\|_{\rho+\delta} \, e^{-(\rho+\delta)|\ell-m|}. \tag{2.48}$$

This implies that for all $(k,m) \in \mathbb{Z}^2$,

$$|(AB)_{km}|e^{\rho|k-m|} \leq \|A\|_\rho \|B\|_{\rho+\delta} \, e^{\rho|k-m|} \sum_{\ell\in\mathbb{Z}} e^{-\rho|k-\ell|} e^{-(\rho+\delta)|\ell-m|}.$$

The last sum can also be written after the change of index $\ell \mapsto \ell + m$,

$$\sum_{\ell\in\mathbb{Z}} e^{-\rho|k-\ell-m|} e^{-\rho|\ell|} e^{-\delta|\ell|} \leq \sum_{\ell\in\mathbb{Z}} e^{-\rho|k-m|} e^{+\rho|\ell|} e^{-\rho|\ell|} e^{-\delta|\ell|}$$

where we used $|k-m| - |\ell| \leq |k-m-\ell|$ for all $(k,m,\ell) \in \mathbb{Z}^3$. We thus have

$$\|AB\|_\rho \leq \|A\|_\rho \|B\|_{\rho+\delta} \sum_{\ell\in\mathbb{Z}} e^{-\delta|\ell|}$$

and we easily compute that

$$\sum_{\ell\in\mathbb{Z}} e^{-\delta|\ell|} = \frac{1+e^{-\delta}}{1-e^{-\delta}} \tag{2.49}$$

which yields the first inequality in (2.47). The second is proved similarly. ■

Proof of Proposition 2.4.1. Let $J \geq 1$ be a fixed integer. We set $\delta = \rho_V(2J+1)^{-1}$ and for $j \in \{0,\dots,J+1\}$, $\rho_j = \rho_V - j\delta = (2J+1-j)\delta$.
In the following, for all operator A and for $j \in \{0,\dots,J+1\}$, we set

$$\|A\|_{(j)} := \|A\|_{\rho_j}.$$

Note that if $0 \leq j \leq k \leq J+1$, then $\|A\|_{(k)} \leq \|A\|_{(j)}$. Moreover, $\|A\|_{(0)} = \|A\|_{\rho_V}$ and $\|A\|_{(J+1)} = \|A\|_{\rho_V J/(2J+1)} \geq \|A\|_{\rho_V/3}$.
The inequalities (2.43) and the fact that $Q_0 = \mathrm{Id}$ yield with the previous notations:

$$\|S_0\|_{(1)} \leq M_V \frac{K^\nu}{\gamma}\Big(\frac{4\nu}{e\delta}\Big)^{2\nu}, \quad \|X_0\|_{(0)} \leq hM_V \quad \text{and} \quad \|Q_0\|_{(0)} = 1.$$

We now seek recursive bounds for the norms $\|S_j\|_{(j+1)}$, $\|X_j\|_{(j)}$ and $\|Q_j\|_{(j)}$ for $j = 0,\dots,J$.

The equality (2.32) and Lemma 2.4.1 ensure that, for $n = 0, \dots, J-1$, we have

$$(n+1)\|Q_{n+1}\|_{(n+1)} \le \sum_{k=0}^{n} \|Q_k S_{n-k}\|_{(n+1)} \le C_\delta \sum_{k=0}^{n} \|Q_k\|_{(n)} \|S_{n-k}\|_{(n+1)} .$$

Hence we get the following estimate for all $n = 0, \dots, J-1$:

$$\|Q_{n+1}\|_{(n+1)} \le \frac{C_\delta}{n+1} \sum_{k=0}^{n} \|Q_k\|_{(k)} \|S_{n-k}\|_{(n-k+1)} . \tag{2.50}$$

Let us denote, for $n \in \mathbb{N}^*$,

$$\begin{aligned} G_n^1 &= \sum_{\{p+q+r=n \mid q \neq n\}} \frac{(ihV)^p}{p!} e^{-ih\Delta} S_q e^{ih\Delta} \frac{(-ihV)^r}{r!} \\ \text{and} \qquad G_n^2 &= - \sum_{\{p+q+r=n \mid q \neq n\}} Q_p^* X_q Q_r \end{aligned}$$

the two members of the right hand side of the equation (2.34). Lemma 2.4.1 implies that we have, for $n = 1, \dots, J$,

$$\begin{aligned} \|G_n^1\|_{(n)} &\le \sum_{\{p+q+r=n \mid q \neq n\}} \frac{1}{p!\,r!} \|(ihV)^p e^{-ih\Delta} S_q e^{ih\Delta} (-ihV)^r\|_{(n)} \\ &\le C_\delta \sum_{\{p+q+r=n \mid q \neq n\}} \frac{1}{p!\,r!} \|(ihV)^p\|_{(n-1)} \|e^{-ih\Delta} S_q e^{ih\Delta} (-ihV)^r\|_{(n)} . \end{aligned}$$

Since the coefficients of the diagonal operators $e^{\pm ih\Delta}$ are of modulus one, Lemma 2.4.1 yields

$$\begin{aligned} \|G_n^1\|_{(n)} &\le C_\delta^2 \sum_{\{p+q+r=n \mid q \neq n\}} \frac{1}{p!\,r!} \|(ihV)^p\|_{(n-1)} \|S_q\|_{(n)} \|(-ihV)^r\|_{(n-1)} \\ &\le C_\delta^2 \sum_{\{p+q+r=n \mid q \neq n\}} \frac{1}{p!\,r!} \|(ihV)^p\|_{(0)} \|S_q\|_{(q+1)} \|(-ihV)^r\|_{(0)} . \end{aligned}$$

Using the hypothesis on V, we obtain for all $n = 1, \dots, J$,

$$\|G_n^1\|_{(n)} \le C_\delta^2 \sum_{\{p+q+r=n \mid q \neq n\}} \frac{(hM_V)^{p+r}}{p!r!} \|S_q\|_{(q+1)} . \tag{2.51}$$

Similarly, for $n = 1, \dots, J$, we have

$$\begin{aligned} \|G_n^2\|_{(n)} &\le \sum_{\{p+q+r=n \mid q \neq n\}} \|Q_p^* X_q Q_r\|_{(n)} \\ &\le 2C_\delta \|Q_n\|_{(n)} \|X_0\|_{(n-1)} \\ &\quad + C_\delta^2 \sum_{\{p+q+r=n \mid p,q,r \neq n\}} \|Q_p\|_{(n-1)} \|X_q\|_{(n)} \|Q_r\|_{(n-1)} . \end{aligned}$$

Since $C_\delta > 1$, we can write

$$\|G_n^2\|_{(n)} \leq C_\delta^2 \sum_{\{p+q+r=n|q\neq n\}} \|Q_p\|_{(p)} \|X_q\|_{(q)} \|Q_r\|_{(r)}. \tag{2.52}$$

Now using Proposition 2.3.1 for the I_K-solution of (2.34) we have for $n = 1, \dots, J$,

$$\|S_n\|_{(n+1)} \leq \frac{K^\nu}{\gamma h}\left(\frac{4\nu}{e\delta}\right)^{2\nu} \left(\|G_n^1\|_{(n)} + \|G_n^2\|_{(n)}\right) \tag{2.53}$$

$$\text{and} \qquad \|X_n\|_n \leq \|G_n^1\|_{(n)} + \|G_n^2\|_{(n)}. \tag{2.54}$$

These estimates, together with (2.50), (2.51) and (2.52) give recursive relations between the coefficients $(\|S_j\|_{j+1})_{0\leq j\leq J}$, $(\|Q_j\|_j)_{0\leq j\leq J}$ and $(\|X_j\|_j)_{0\leq j\leq J}$.
Let us define the numbers $(\mathsf{s}_j)_{j\geq 0}$, $(\mathsf{q}_j)_{j\geq 0}$ and $(\mathsf{x}_j)_{j\geq 0}$ as follows:

$$\mathsf{s}_0 = M_V \frac{K^\nu}{\gamma}\left(\frac{4\nu}{e\delta}\right)^{2\nu}, \quad \mathsf{q}_0 = 1 \quad \text{and} \quad \mathsf{x}_0 = hM_V,$$

and for all $n \in \mathbb{N}^*$, with $\kappa = \frac{K^\nu}{\gamma h}\left(\frac{4\nu}{e\delta}\right)^{2\nu}$,

$$\begin{aligned}
\mathsf{s}_n &= \kappa C_\delta^2 \sum_{\{p+q+r=n|q\neq n\}} \left(\frac{(hM_V)^{p+r}}{p!r!}\mathsf{s}_q + \mathsf{q}_p\mathsf{q}_r\mathsf{x}_q\right),\\
\mathsf{q}_n &= \frac{C_\delta}{n}\sum_{k=0}^{n-1} \mathsf{q}_k \mathsf{s}_{n-k-1},\\
\mathsf{x}_n &= C_\delta^2 \sum_{\{p+q+r=n|q\neq n\}} \left(\frac{(hM_V)^{p+r}}{p!r!}\mathsf{s}_q + \mathsf{q}_p\mathsf{q}_r\mathsf{x}_q\right).
\end{aligned}$$

By induction, using the equations (2.50), (2.51), (2.52) and (2.53) we have that for $j = 0, \dots, J$,

$$\|S_j\|_{j+1} \leq \mathsf{s}_j, \quad \|Q_j\|_j \leq \mathsf{q}_j \quad \text{and} \quad \|X_j\|_j \leq \mathsf{x}_j.$$

We then define the corresponding power series

$$\begin{aligned}
\mathsf{s}(t) &= \sum_{n\geq 0} \mathsf{s}_n t^n,\\
\mathsf{q}(t) &= \sum_{n\geq 0} \mathsf{q}_n t^n,\\
\text{and} \qquad \mathsf{x}(t) &= \sum_{n\geq 0} \mathsf{x}_n t^n.
\end{aligned}$$

By definition of the coefficients of these series, we have the following formal identities:

$$\begin{aligned}
\left(1 - \kappa C_\delta^2 (e^{2hM_V t} - 1)\right)\mathsf{s}(t) - \mathsf{s}_0 &= \kappa C_\delta^2 \mathsf{x}(t)(\mathsf{q}(t)^2 - 1),\\
\mathsf{q}'(t) &= C_\delta \mathsf{s}(t)\mathsf{q}(t),\\
\mathsf{x}(t) &= \kappa^{-1}\mathsf{s}(t),
\end{aligned}$$

where q' denotes the derivative of q with respect to t. We then derive that $\mathsf{q}(t)$ is solution of the following ordinary differential equation

$$\mathsf{q}'(t) = \frac{\mathsf{s}_0 C_\delta \mathsf{q}(t)}{1 - C_\delta^2\big(\kappa(e^{2hM_V t} - 1) + (\mathsf{q}(t)^2 - 1)\big)} \quad \text{and} \quad \mathsf{q}(0) = 1. \tag{2.55}$$

The study of this equation will give us bounds on the coefficients q_j. It is clear with the previous equalities that for sufficiently small t, the functions $\mathsf{s}(t)$, $\mathsf{q}(t)$ and $\mathsf{x}(t)$ are analytic in t. Moreover, the following Lemma, whose proof is given in Appendix (Section 2.7, page 62), gives a bound for the radius of convergence of the corresponding power series:

Lemma 2.4.2 *Let us define*

$$K_0^\nu = \max\left(1, \frac{\rho_V^{2(\nu-1)}\gamma}{4}\left(\frac{e}{4\nu}\right)^{2\nu}\right) \quad \textit{and} \quad r^{-1} = 20hM_V\kappa C_\delta^3.$$

For all $K \geq K_0$ and $h \in (0,1)$ satisfying (2.7)*, the functions* x, q *and* s *are analytic in* $(-r, r)$*, and satisfy for all $t \in (-r,r)$ the estimates:*

$$0 < \mathsf{s}(t) \leq \frac{5\sqrt{5}}{4}\kappa h M_V, \quad 0,8 < \mathsf{q}(t) \leq \frac{\sqrt{5}}{2} \quad \textit{and} \quad 0 < \mathsf{x}(t) \leq \frac{5\sqrt{5}}{4}hM_V.$$

End of proof of Proposition 2.4.1: Let us consider the functions $\mathsf{s}(z)$, $\mathsf{q}(z)$ and $\mathsf{x}(z)$ for complex numbers z. As the coefficients of the corresponding power series are non-negative, we have for example for all sufficiently small z, $|\mathsf{s}(z)| \leq \mathsf{s}(|z|)$ and similar inequalities for $\mathsf{q}(z)$ and $\mathsf{x}(z)$. The previous Lemma combined with Cauchy estimates then shows that for $j = J$, we have

$$\mathsf{s}_J \leq \frac{5\sqrt{5}}{4}\kappa h M_V (20hM_V\kappa C_\delta^3)^J = \frac{5\sqrt{5}}{4}20^J C_\delta^{3J}(\kappa h M_V)^{J+1}.$$

By (2.49) and the definition of δ, we have $C_\delta \leq 2e^{\rho_V}\delta^{-1}$. The previous inequality then yields (recall that $\kappa = \frac{K^\nu}{\gamma h}\left(\frac{4\nu}{e\delta}\right)^{2\nu}$):

$$\mathsf{s}_J \leq \frac{5\sqrt{5}}{4}\left(\frac{K^\nu}{\gamma}\left(\frac{4\nu}{e\delta}\right)^{2\nu}\right)^{J+1} M_V^{J+1} 20^J 2^{3J} e^{3J\rho_V}\delta^{-3J}$$

and we easily see that there exists a positive constant C depending on V, ρ_V, γ and ν such that

$$\mathsf{s}_J \leq C^J K^{\nu(J+1)}\delta^{-2\nu(J+1)-3J}.$$

As $\delta^{-1} = (2J+1)/\rho_V$ we can modify C to obtain

$$\mathsf{s}_J \leq C^J K^{\nu(J+1)}(2J+1)^{(2\nu+3)J+2\nu}.$$

Now, since we have $\|S_J\|_{\rho_V/3} \leq \|S_J\|_{(J+1)} \leq \mathsf{s}_J$, we obtain the result with $\alpha = 2\nu$ and $\beta = 4\nu + 3$ after modification of the constant C to become C_0.
The proof of the other estimates for $\|X_J\|_{\rho_V/3}$ and $\|Q_J\|_{\rho_V}$ can be obtained similarly from Lemma 2.4.2. This finishes the proof of Proposition 2.4.1. ■

Remark 2.4.1 *The coefficients α and β given in Proposition 2.4.1 are clearly not optimal. For simplicity, we did not try to optimize them, as well as the constants C_0 and K_0.*

2.5 Error estimates

2.5.1 Construction of the operators

Assume that $K \geq 1$ and $N \geq 1$ are given. Let us define the following polynomials in $\lambda \in \mathbb{R}$:

$$S^{[N]}(\lambda) = \sum_{0 \leq n \leq N} \lambda^n S_n \qquad \text{and} \qquad X^{[N]}(\lambda) = \sum_{0 \leq n \leq N} \lambda^n X_n \tag{2.56}$$

where the coefficients S_n and X_n are defined by the recursive I_K-solutions of (2.32)-(2.34). Proposition 2.4.1 shows that $S^{[N]}(\lambda)$ and $X^{[N]}(\lambda)$ are operators in the space $\mathcal{S}_{\rho_V/3}$ for all $\lambda \in \mathbb{R}$. Lemma 2.4.1 then shows that the multiplication by $S^{[N]}(\lambda)$ or $X^{[N]}(\lambda)$ defines a continuous linear mapping from $\mathcal{A}_{\rho_V/4}$ to itself.

Let us denote by $Q^{[N]}(\lambda)$ and $\Sigma^{[N]}(\lambda)$ the maximal solutions in $\mathcal{A}_{\rho_V/4}$ of the following Cauchy problems

$$\left\{ \begin{array}{l} i\partial_\lambda Q^{[N]}(\lambda) = Q^{[N]}(\lambda) S^{[N]}(\lambda), \\ Q^{[N]}(0) = \mathrm{Id}, \end{array} \right. \qquad \text{and} \qquad \left\{ \begin{array}{l} i\partial_\lambda \Sigma^{[N]}(\lambda) = \Sigma^{[N]}(\lambda) X^{[N]}(\lambda), \\ \Sigma^{[N]}(0) = e^{ih\Delta}. \end{array} \right. \tag{2.57}$$

Since $\lambda \mapsto S^{[N]}(\lambda)$ (respectively $\lambda \mapsto X^{[N]}(\lambda)$) is continuous from $\mathbb{R}$ to $\mathcal{A}_{\rho_V/3}$, $Q^{[N]}(\lambda)$ (resp. $\Sigma^{[N]}(\lambda)$) is defined for all $\lambda \in \mathbb{R}$ and belongs to the space $\mathcal{A}_{\rho_V/4}$.

By construction, the spaces $\mathcal{X}_\rho^K$ are invariant by multiplication. Combining this and Lemma 2.4.1 we see that the multiplication by $X^{[N]}(\lambda)$ defines a continuous linear operator from $\mathcal{X}_{\rho_V/4}^K$ to itself.

As $e^{ih\Delta} \in \mathcal{X}_{\rho_V/4}^M$ and as $\mathcal{X}_{\rho_V/4}^K$ is closed for $\|\cdot\|_{\rho_V/4}$, we have

$$\forall\, \lambda \in \mathbb{R}, \qquad \Sigma^{[N]}(\lambda) \in \mathcal{X}_{\rho_V/4}^K.$$

Since $X^{[N]}(\lambda)$ and $S^{[N]}(\lambda)$ are polynomial functions of λ, $\Sigma^{[N]}(\lambda)$ and $Q^{[N]}(\lambda)$ are convergent power series in λ on a neighborhood of 0.

Moreover, as for all λ the operator $S^{[N]}(\lambda)$ is symmetric, the function

$$\lambda \mapsto Z(\lambda) := Q^{[N]}(\lambda)^* Q^{[N]}(\lambda)$$

is solution of the following Cauchy problem in some suitable $\mathcal{A}$ space:

$$\left\{ \begin{array}{l} i\partial_\lambda Z(\lambda) = [Z(\lambda), S^{[N]}(\lambda)], \\ Z(0) = \mathrm{Id}. \end{array} \right.$$

This implies that for all $\lambda \in \mathbb{R}$, $Z(\lambda) = \mathrm{Id}$ and hence that $Q^{[N]}(\lambda)$ is a unitary operator for all λ. A similar argument using the symmetry of $X^{[N]}(\lambda)$ shows that $\Sigma^{[N]}(\lambda)$ also is a unitary operator.

Let us denote, for λ sufficiently small,

$$Q^{[N]}(\lambda) = \sum_{n=0}^{+\infty} Q_n^{[N]} \lambda^n \qquad \text{and} \qquad \Sigma^{[N]}(\lambda) = \sum_{n=0}^{+\infty} \Sigma_n^{[N]} \lambda^n$$

where for all $n \in \mathbb{N}$, $Q_n^{[N]}$ and $\Sigma_n^{[N]}$ are in $\mathcal{A}_{\rho_V/4}$ (the series are absolutely convergent in this space for λ sufficiently small).

For all $\lambda \in \mathbb{R}$, we can define the following operator in a convenient $\mathcal{A}$ space

$$R^{[N]}(\lambda) = Q^{[N]}(\lambda) L(\lambda) Q^{[N]}(\lambda)^* - \Sigma^{[N]}(\lambda). \tag{2.58}$$

Using Lemma 2.4.1, $R^{[N]}(\lambda) \in \mathcal{A}_{\rho_V/5}$ and in this space, $R^{[N]}(\lambda)$ is the sum of the following absolutely convergent power series in a neighborhood of 0:

$$R^{[N]}(\lambda) = \sum_{n=0}^{+\infty} \lambda^n \left[\sum_{p+q+r=n} \left(Q_p^{[N]} e^{ih\Delta} \frac{(-ih)^q}{q!} V^q (Q_r^{[N]})^* \right) - \Sigma_n^{[N]} \right]. \tag{2.59}$$

Since $Q_0 = Q_0^{[N]} = \mathrm{Id}$, the equality (2.32) and the fact that $Q^{[N]}$ is defined by (2.57) ensure that

$$\forall\, n \in \{1, \ldots, N\}, \qquad Q_n = Q_n^{[N]}.$$

Moreover, the operators $(S_n)_{n\in\mathbb{N}}$ and $(X_n)_{n\in\mathbb{N}}$ satisfy (2.34) for all $n \in \mathbb{N}$ by definition. Hence we have in a suitable $\mathcal{A}$ space,

$$S^{[N]} - L^* S^{[N]} L = hV - (Q^{[N]})^* X^{[N]} Q^{[N]} + \mathcal{O}(\lambda^{N+1}).$$

We deduce that

$$S^{[N]} - L^* S^{[N]} L = hV - (Q^{[N]})^* (\Sigma^{[N]})^* (i\partial_\lambda \Sigma^{[N]}) Q^{[N]} + \mathcal{O}(\lambda^{N+1})$$

by definition (2.57) of $\Sigma^{[N]}$, and the fact that this operator is unitary. Then, since $(\Sigma^{[N]})^* = (Q^{[N]}) L^* (Q^{[N]})^* - (R^{[N]})^*$, we have

$$\begin{aligned} LS^{[N]} - S^{[N]} L = hLV &- (Q^{[N]})^* (i\partial_\lambda \Sigma^{[N]}) Q^{[N]} \\ &+ L (Q^{[N]})^* R^{[N]} (i\partial_\lambda \Sigma^{[N]}) Q^{[N]} + \mathcal{O}(\lambda^{N+1}). \end{aligned}$$

Since $S^{[N]} = (Q^{[N]})^* (i\partial_\lambda Q^{[N]}) = (S^{[N]})^*$, we derive that

$$\begin{aligned} L(i\partial_\lambda Q^{[N]})^* Q^{[N]} - (Q^{[N]})^* (i\partial_\lambda Q^{[N]}) L = hLV &- (Q^{[N]})^* (i\partial_\lambda \Sigma^{[N]}) Q^{[N]} \\ &+ L (Q^{[N]})^* R^{[N]} (i\partial_\lambda \Sigma^{[N]}) Q^{[N]} + \mathcal{O}(\lambda^{N+1}). \end{aligned}$$

After left-multiplying this equality by $Q^{[N]}$ and right-multiplying by $(Q^{[N]})^*$, we get

$$\begin{aligned} &Q^{[N]} L (i\partial_\lambda Q^{[N]})^* - (i\partial_\lambda Q^{[N]}) L (Q^{[N]})^* \\ &- hQ^{[N]} e^{ih\Delta} V e^{-ih\lambda V} (Q^{[N]})^* + i\partial_\lambda \Sigma^{[N]} = Q^{[N]} L (Q^{[N]})^* R^{[N]} (i\partial_\lambda \Sigma^{[N]}) + \mathcal{O}(\lambda^{N+1}). \end{aligned}$$

This equality expresses that

$$i\partial_\lambda R^{[N]} = -Q^{[N]}L(Q^{[N]})^* R^{[N]}(i\partial_\lambda \Sigma^{[N]}) + \mathcal{O}(\lambda^{N+1}).$$

Noticing that $R^{[N]}(0) = e^{ih\Delta} - e^{ih\Delta} = 0$, we derive by induction that

$$\forall\, 0 \le n \le N+1, \qquad \partial_\lambda^n R^{[N]}(0) = 0.$$

Therefore, (2.59) simplifies to

$$R^{[N]}(\lambda) = \sum_{n > N+1}^{+\infty} \lambda^n \left[\sum_{p+q+r=n} \left(Q_p^{[N]} e^{ih\Delta} \frac{(-ih)^q}{q!} V^q (Q_r^{[N]})^* \right) - \Sigma_n^{[N]} \right] \tag{2.60}$$

with absolute convergence in $\mathcal{A}_{\rho_V/5}$.

2.5.2 Estimate for the remainder term

We give now an estimate for the term $R^{[N]}(\lambda)$. We first give estimates for the coefficients of the formal series $\Sigma^{[N]}$ and $Q^{[N]}$:

Lemma 2.5.1 *Using the previous notations, there exists a constant $C_1 \ge 1$ depending only on V, ρ_V, γ and ν such that for all $N \ge 1$, all $n \in \mathbb{N}^*$, all $K \ge K_0$, and all $h \in (0,1)$ satisfying* (2.7),

$$\|\Sigma_n^{[N]}\|_{\rho_V/4} \le h(C_1 K^\alpha N^\beta)^n \tag{2.61}$$

$$\text{and} \qquad \|Q_n^{[N]}\|_{\rho_V/4} \le (C_1 K^\alpha N^\beta)^n \tag{2.62}$$

where $\alpha = 2\nu$ and $\beta = 4\nu + 3$ are given in Proposition 2.4.1.

Proof. Since $\Sigma_0^{[N]} = e^{ih\Delta}$, we have $\|\Sigma_0^{[N]}\|_{\rho_V/4} = 1$. Moreover, since $\Sigma_1^{[N]} = \Sigma_0^{[N]} X_0$, we have with (2.44) that $\|\Sigma_1^{[N]}\|_{\rho_V/4} \le hM_V$. Hence (2.61) holds for $n = 1$ with a constant C_1 depending only on V and ρ_V (notice that $K \ge K_0 \ge 1$).
Assume now that (2.61) holds with some constant C_1 and for all $k \le n \in \mathbb{N}^*$. By definition of $\Sigma^{[N]}$ and the use Lemma 2.4.1, we have

$$\|\Sigma_{n+1}^{[N]}\|_{\rho_V/4} \le \frac{C_{\rho_V/12}}{n+1} \sum_{k=0}^{\min(n,N)} \|\Sigma_{n-k}^{[N]}\|_{\rho_V/4} \|X_k\|_{\rho_V/3}.$$

Using the induction hypothesis and the estimate (2.46), we thus have

$$\|\Sigma_{n+1}^{[N]}\|_{\rho_V/4} \le \frac{C_{\rho_V/12}}{n+1} \sum_{k=0}^{\min(n,N)} h(C_1 K^\alpha N^\beta)^{n-k} h(C_0 K^\alpha k^\beta)^k.$$

If we assume that $C_1 \geq C_0$, we have

$$\begin{aligned}
\|\Sigma_{n+1}^{[N]}\|_{\rho_V/4} &\leq h(C_1 K^\alpha N^\beta)^n \frac{hC_{\rho_V/12}}{n+1} \sum_{k=0}^{\min(n,N)} \left(\frac{k}{N}\right)^{\beta k} \\
&\leq h(C_1 K^\alpha N^\beta)^n (hC_{\rho_V/12}) \frac{1}{n+1} \sum_{k=0}^{\min(n,N)} 1 \\
&\leq h(C_1 K^\alpha N^\beta)^{n+1}
\end{aligned}$$

provided that $C_1 \geq C_{\rho_V/12}$ and since $K \geq 1$. This shows (2.61).
The proof of (2.62) is similar using the estimate (2.45) instead of (2.46). ■

The previous Lemma yields the following estimates:

Proposition 2.5.1 *Using the previous notations, for all* $K \geq K_0$*, all* $N \geq 1$*, all* $h \in (0,1)$ *satisfying* (2.7) *and all* $\lambda \in \mathbb{R}$ *such that* $|\lambda| \leq (2C_1 K^\alpha N^\beta)^{-1}$*, we have*

$$\|Q^{[N]}(\lambda) - \mathrm{Id}\|_{\rho_V/4} \leq 2C_1 K^\alpha N^\beta |\lambda| \tag{2.63}$$

and

$$\|\Sigma^{[N]}(\lambda) - e^{ih\Delta}\|_{\rho_V/4} \leq 2hC_1 K^\alpha N^\beta |\lambda|. \tag{2.64}$$

Proof. We have

$$Q^{[N]}(\lambda) - \mathrm{Id} = \sum_{n\geq 1} \lambda^n Q_n^{[N]}.$$

This implies, using (2.62)

$$\|Q^{[N]}(\lambda) - \mathrm{Id}\|_{\rho_V/4} \leq \sum_{n\geq 1} (|\lambda| C_1 K^\alpha N^\beta)^n$$

and the result follows straightforwardly. The second inequality is proved similarly. ■

We now give an estimate for the remainder term itself. We use the following notation: For a function f defined in a neighborhood of 0 in $\mathbb{R}$ and such that for all $n \in \mathbb{N}$, the derivative $f^{(n)}(0)$ is defined, we set

$$T_N(f)(x) = f(x) - \sum_{p=0}^{\lfloor N \rfloor} x^p \frac{f^{(p)}(0)}{p!} \tag{2.65}$$

where $\lfloor\cdot\rfloor$ denotes the integer part function.

Proposition 2.5.2 *Using the notations of Proposition* 2.4.1*, there exists a constant* $C_2 > 0$ *depending only on* V*,* ρ_V*,* γ *and* ν *such that for all* $h \in (0,1)$ *satisfying* (2.7)*, all* $N \geq 1$*, all* $K \geq K_0$ *and all* $\lambda \in \mathbb{R}$ *such that* $|\lambda| \leq (2C_1 K^\alpha N^\beta)^{-1}$ *we have*

$$\|R^{[N]}(\lambda)\|_{\rho_V/5} \leq (C_2 |\lambda| K^{3\alpha} N^{3(\beta+1)})^N. \tag{2.66}$$

Proof. Using Lemma 2.4.1, the formula (2.60) yields for $\lambda \in \mathbb{R}$ sufficiently small,

$$\begin{aligned}
&\|R^{[N]}(\lambda)\|_{\rho_V/5} \\
&\leq \sum_{n\geq N+1}^{+\infty} |\lambda|^n \left[C_{\frac{\rho_V}{20}} \sum_{p+q+r=n} \left(\|Q_p^{[N]}\|_{\rho_V/5} \, \|\frac{(-ih)^q}{q!} V^q (Q_r^{[N]})^* \|_{\rho_V/4} \right) \right] \\
&\qquad\qquad + \sum_{n\geq N+1}^{+\infty} |\lambda|^n \|\Sigma_n^{[N]}\|_{\rho_V/5} \\
&\leq \sum_{n\geq N+1}^{+\infty} |\lambda|^n \left[C_{\frac{\rho_V}{20}} C_{\frac{\rho_V}{12}} \sum_{p+q+r=n} \left(\|Q_p^{N]}\|_{\rho_V/5} \, \frac{h^q}{q!} \|V^q\|_{\rho_V/3} \, \|(Q_r^{[N]})^* \|_{\rho_V/4} \right) \right] \\
&\qquad\qquad + \sum_{n\geq N+1}^{+\infty} |\lambda|^n \|\Sigma_n^{[N]}\|_{\rho_V/5} .
\end{aligned}$$

Using Lemma 2.5.1 we thus get

$$\begin{aligned}
&\|R^{[N]}(\lambda)\|_{\rho_V/5} \\
&\leq \sum_{n\geq N+1}^{+\infty} |\lambda|^n \left[C_{\frac{\rho_V}{20}} C_{\frac{\rho_V}{12}} \sum_{p+q+r=n} \left((C_1 K^\alpha N^\beta)^p \frac{h^q}{q!} M_V^q (C_1 K^\alpha N^\beta)^r \right) \right] \\
&\qquad\qquad + h \sum_{n\geq N+1}^{+\infty} (|\lambda| C_1 K^\alpha N^\beta)^n.
\end{aligned}$$

For $x \in \mathbb{R}$ such that $|x| < (C_1 K^\alpha N^\beta)^{-1}$, we set

$$f(x) = (1 - C_1 K^\alpha N^\beta x)^{-1} \qquad \text{and} \qquad g(x) = e^{hM_V x}.$$

The previous inequality can then be written

$$\|R_N(\lambda)\|_{\rho_V/5} \leq C_{\frac{\rho_V}{20}} C_{\frac{\rho_V}{12}} T_N(f^2 g)(|\lambda|) + h T_N(f)(|\lambda|)$$

where T_N is defined in (2.65). Now for all $p \in \mathbb{N}$ and all $x \in \mathbb{R}$ such that $|x| < (C_1 K^\alpha N^\beta)^{-1}$, we have

$$\frac{\mathrm{d}^p f}{\mathrm{d}x^p}(x) = \frac{p\,!\,(C_1 K^\alpha N^\beta)^p}{(1 - C_1 K^\alpha N^\beta x)^{p+1}}$$

$$\text{and} \qquad \frac{\mathrm{d}^p (f^2)}{\mathrm{d}x^p}(x) = \frac{(p+1)!\,(C_1 K^\alpha N^\beta)^p}{(1 - C_1 K^\alpha N^\beta x)^{p+2}},$$

and for the function g:

$$\frac{\mathrm{d}^p g}{\mathrm{d}x^p}(x) = (hM_V)^p e^{hM_V x}.$$

By definition of T_N, and as f and g are smooth, we have for $|\lambda| \leq (2C_1K^\alpha N^\beta)^{-1}$,

$$\begin{aligned} T_N(f^2g)(|\lambda|) &\leq \frac{|\lambda|^{\lfloor N\rfloor+1}}{(\lfloor N\rfloor+1)!} \sup_{|t|\leq(2C_1K^\alpha N^\beta)^{-1}} \left|\frac{\mathrm{d}^{\lfloor N\rfloor+1}(f^2g)}{\mathrm{d}x^{\lfloor N\rfloor+1}}(t)\right| \\ \text{and}\quad T_N(f)(|\lambda|) &\leq \frac{|\lambda|^{\lfloor N\rfloor+1}}{(\lfloor N\rfloor+1)!} \sup_{|t|\leq(2C_1K^\alpha N^\beta)^{-1}} \left|\frac{\mathrm{d}^{\lfloor N\rfloor+1}f}{\mathrm{d}x^{\lfloor N\rfloor+1}}(t)\right|. \end{aligned}$$

As for $|t| \leq (2C_1K^\alpha N^\beta)^{-1}$ we have

$$\left|\frac{\mathrm{d}^{\lfloor N\rfloor+1}f}{\mathrm{d}x^{\lfloor N\rfloor+1}}(t)\right| = \frac{(\lfloor N\rfloor+1)!\,(C_1K^\alpha N^\beta)^{\lfloor N\rfloor+1}}{(1-C_1K^\alpha N^\beta t)^{\lfloor N\rfloor+2}} \leq 2(\lfloor N\rfloor+1)!(2C_1K^\alpha N^\beta)^{\lfloor N\rfloor+1},$$

we derive that (notice that $|\lambda| \leq 1/2 \leq 1$)

$$\begin{aligned} T_N(f)(|\lambda|) &\leq 2(2C_1K^\alpha N^\beta|\lambda|)^{\lfloor N\rfloor+1} \\ &\leq (4C_1K^\alpha N^\beta)^{N+1}|\lambda|^N \\ &\leq (16C_1^2K^{2\alpha}N^{2\beta})^N|\lambda|^N. \end{aligned}$$

Similarly, for $|t| \leq (2C_1K^\alpha N^\beta)^{-1}$ we have

$$\frac{\mathrm{d}^{\lfloor N\rfloor+1}(f^2g)}{\mathrm{d}x^{\lfloor N\rfloor+1}}(t) = \sum_{k=0}^{\lfloor N\rfloor+1} \frac{(\lfloor N\rfloor+1)!}{k!(\lfloor N\rfloor+1-k)!}\left(\frac{\mathrm{d}^kf^2}{\mathrm{d}x^k}(t)\right)\left(\frac{\mathrm{d}^{\lfloor N\rfloor+1-k}g}{\mathrm{d}x^{\lfloor N\rfloor+1-k}}(t)\right)$$

and hence

$$\begin{aligned} &\left|\frac{\mathrm{d}^{\lfloor N\rfloor+1}(f^2g)}{\mathrm{d}x^{\lfloor N\rfloor+1}}(t)\right| \\ &\leq \sum_{k=0}^{\lfloor N\rfloor+1} \frac{(\lfloor N\rfloor+1)!}{k!(\lfloor N\rfloor+1-k)!}\left(\frac{(k+1)!\,(C_1K^\alpha N^\beta)^k}{(1-C_1K^\alpha N^\beta t)^{k+2}}\right)(hM_V)^{\lfloor N\rfloor+1-k}e^{hM_Vt}. \end{aligned}$$

As $|t| \leq (2C_1K^\alpha N^\beta)^{-1}$ and as we can always assume that $hM_V \leq M_V \leq C_1$ with $C_1 \geq 1$, the right hand side of the previous equation is bounded by

$$4e^{hM_Vt}C_1^{\lfloor N\rfloor+1}(\lfloor N\rfloor+1)! \sum_{k=0}^{\lfloor N\rfloor+1} \frac{k+1}{(\lfloor N\rfloor+1-k)!}\left(2C_1K^\alpha N^\beta\right)^k$$

and hence by

$$4\sqrt{e}(C_1)^{\lfloor N\rfloor+1}(\lfloor N\rfloor+1)!(\lfloor N\rfloor+1)\sum_{k=0}^{\lfloor N\rfloor+1}\left(2C_1K^\alpha N^\beta\right)^k,$$

using the fact that $|hM_Vt| \leq 1/2$. The last factor is equal to

$$\frac{(2C_1K^\alpha N^\beta)^{\lfloor N\rfloor+2}-1}{2C_1K^\alpha N^\beta-1}.$$

Since $C_1K^\alpha N^\beta \geq 1$, this ratio is bounded by

$$(2C_1K^\alpha N^\beta)^{\lfloor N\rfloor+2}.$$

This yields

$$\left|\frac{\mathrm{d}^{\lfloor N\rfloor+1}(f^2g)}{\mathrm{d}x^{\lfloor N\rfloor+1}}(t)\right| \leq 4\sqrt{e}(C_1)^{\lfloor N\rfloor+1}(\lfloor N\rfloor+1)!(\lfloor N\rfloor+1)(2C_1K^\alpha N^\beta)^{\lfloor N\rfloor+2}$$

and hence for $|\lambda| \leq (2C_1K^\alpha N^\beta)^{-1}$,

$$\begin{aligned} T_N(f^2g)(|\lambda|) &\leq 4\sqrt{e}|\lambda|^{\lfloor N\rfloor+1}\left(2C_1^2K^\alpha(\lfloor N\rfloor+1)^{\beta+1}\right)^{\lfloor N\rfloor+2} \\ &\leq |\lambda|^N(8eC_1^2K^\alpha(2N)^{\beta+1})^{N+2} \\ &\leq |\lambda|^N(2^{3(\beta+4)}e^3C_1^6K^{3\alpha}N^{3(\beta+1)})^N. \end{aligned}$$

Collecting the previous estimates, we get (2.66) for a suitable constant C_2. ∎

We are now able to prove Theorem 2.2.1:

Proof of Theorem 2.2.1. Let $h \in (0,1)$ be a stepsize satisfying (2.7) and consider positive numbers σ_N and σ_K such that

$$\alpha\sigma_K + (\beta+1)\sigma_N \leq 1/4. \tag{2.67}$$

For example, $\sigma_N = \sigma_K = 1/(8\max(\alpha,\beta+1)) = 1/(32(\nu+1)) > 0$ is a possible choice of parameters. We make this specific choice to prove Theorem 2.2.1. For these parameters, we set for all $\lambda \in (0,1)$,

$$K = \lambda^{-\sigma_K} \quad \text{and} \quad N = 1/(2C_1)^{1/\beta}\lambda^{-\sigma_N} \tag{2.68}$$

and we define

$$Q(\lambda) = Q^{[N]}(\lambda), \quad \Sigma(\lambda) = \Sigma^{[N]}(\lambda) \quad \text{and} \quad R(\lambda) = R^{[N]}(\lambda).$$

Since $\sigma_N = \sigma_K > 0$ and C_1 only depends on V, γ and ν, there exists $\lambda_0 \in (0,1)$ depending only on V, ρ_V, γ and ν such that for all $\lambda \in (0,\lambda_0)$, we have $K = \lambda^{-\sigma_N} \geq K_0$ and $N = 1/(2C_1)^{1/\beta}\lambda^{-\sigma_N} \geq 1$. For such a λ, we have

$$(2C_1K^\alpha N^\beta)^{-1} = \lambda^{\alpha\sigma_K+\beta\sigma_N} \geq \lambda$$

since $\alpha\sigma_K + \beta\sigma_N \leq 1$ with (2.67) and $\lambda \leq 1$.
Therefore, Proposition 2.5.1 ensures that

$$\|Q(\lambda) - \mathrm{Id}\|_{\rho_V/4} \leq \lambda^{1-(\alpha\sigma_K+\beta\sigma_N)} \leq \lambda^{1/2}$$

and

$$\|\Sigma(\lambda) - e^{ih\Delta}\|_{\rho_V/4} \leq h\lambda^{1-(\alpha\sigma_K+\beta\sigma_N)} \leq h\lambda^{1/2}.$$

Moreover, Proposition 2.5.2 ensures that

$$\|R(\lambda)\|_{\rho_V/5} \leq (C_2C_1^{-3(\beta+1)/\beta}\lambda^{1-(3\alpha\sigma_K+3(\beta+1)\sigma_N)})^N.$$

As the exponent of λ in the right hand side of this inequality satisfies

$$1 - (3\alpha\sigma_K + 3(\beta+1)\sigma_N) \geq 1/4 > 0$$

by (2.67), after a possible decrease of λ_0 (depending only on V, ρ_V, γ and ν again[4]), we can assume that

$$\forall \lambda \in (0, \lambda_0) \qquad C_2 C_1^{-3(\beta+1)/\beta} \lambda^{1-(3\alpha\sigma_K+3(\beta+1)\sigma_N)} \leq e^{-1}.$$

Therefore, we get eventually that for all $\lambda \in (0, \lambda_0)$,

$$\|R(\lambda)\|_{\rho_V/5} \leq e^{-N} = e^{\frac{-1}{(2C_1)^{1/\beta}}\lambda^{-\sigma_N}},$$

and this finishes the proof of the Theorem. ■

2.6 Proof of the corollary

Before starting the proof of Corollary 2.2.1, we give the following two lemmas:

Lemma 2.6.1 *Let $\mu > 0$. Assume that $R \in \mathcal{A}_\mu$. Let $\varphi \in L^2(\mathbb{T})$ and let $\psi = R\varphi$. Then for all $k \in \mathbb{Z}$, we have*

$$|\psi_k| \leq |\psi|_{|k|} \leq \|\psi\| \leq C_\mu \|R\|_\mu \|\varphi\|$$

where $\|\cdot\|$ is the L^2-norm for functions and where $\|\cdot\|_\mu$ is the norm (2.9) *for operators. Recall that the value of the constant C_μ is given in Lemma* 2.4.1.

Proof. By definition, we have for all $k \in \mathbb{Z}$,

$$\psi_k = \sum_{\ell\in\mathbb{Z}} R_{k\ell}\varphi_\ell.$$

Hence, we have with the Cauchy-Schwarz inequality

$$\begin{aligned} \|\psi\|^2 &\leq \sum_{k\in\mathbb{Z}} \Big| \sum_{\ell\in\mathbb{Z}} |R_{k\ell}||\varphi_\ell| \Big|^2 \\ &\leq \|R\|_\mu^2 \sum_{k\in\mathbb{Z}} \Big| \sum_{\ell\in\mathbb{Z}} e^{-\frac{\mu}{2}|k-\ell|} |\varphi_\ell| \, e^{-\frac{\mu}{2}|k-\ell|} \Big|^2 \\ &\leq \|R\|_\mu^2 \sum_{k\in\mathbb{Z}} \left[\Big(\sum_{p\in\mathbb{Z}} e^{-\mu|k-p|} \Big) \Big(\sum_{\ell\in\mathbb{Z}} e^{-\mu|k-\ell|} |\varphi_\ell|^2 \Big) \right]. \end{aligned}$$

Using (2.49), we see that

$$\begin{aligned} \|\psi\|^2 &\leq \|R\|_\mu^2 C_\mu \sum_{\ell\in\mathbb{Z}} |\varphi_\ell|^2 \sum_{k\in\mathbb{Z}} e^{-\mu|k-\ell|} \\ &\leq \|R\|_\mu^2 C_\mu^2 \|\varphi\|^2 \end{aligned}$$

and this yields the result using the fact that for all $k \in \mathbb{Z}$,

$$|\psi_k| \leq |\psi|_{|k|} \leq \|\psi\| .$$

■

[4] We recall that C_2 only depends on V, γ and ν by Proposition 2.5.2

Lemma 2.6.2 *Let $\rho > 0$ and $s \geq 0$ be real numbers. There exists a positive constant C depending only on ρ and s such that for all $A \in \mathcal{A}_\rho$ and all φ such that $\|\varphi\|_{s,\infty} < \infty$, we have*

$$\|A\varphi\|_{s,\infty} \leq C\|A\|_\rho \|\varphi\|_{s,\infty}$$

Proof. By definition of $\|\cdot\|_{s,\infty}$ (see (2.18)), we have

$$\forall\, k \in \mathbb{N}, \quad |\varphi|_k \leq \|\varphi\|_{s,\infty} (1+k)^{-s}.$$

We deduce that for all $k \in \mathbb{Z}$

$$\begin{aligned} |(A\varphi)_k| &\leq \sum_{\ell \in \mathbb{Z}} \|A\|_\rho e^{-\rho|k-\ell|}|\varphi_\ell| \\ &\leq \|A\|_\rho \|\varphi\|_{s,\infty} \sum_{\ell\in\mathbb{Z}} e^{-\rho|k-\ell|}(1+|\ell|)^{-s}. \end{aligned}$$

so that

$$\begin{aligned} (1+|k|)^s|(A\varphi)_k| &\leq \|A\|_\rho \|\varphi\|_{s,\infty} \sum_{\ell\in\mathbb{Z}} e^{-\rho|k-\ell|} \left(\frac{1+|k|}{1+|\ell|}\right)^s \\ &\leq \|A\|_\rho \|\varphi\|_{s,\infty} \sum_{\ell\in\mathbb{Z}} e^{-\rho|k-\ell|} \left(\frac{1+|k-\ell|+|\ell|}{1+|\ell|}\right)^s \\ &\leq \|A\|_\rho \|\varphi\|_{s,\infty} \sum_{\ell\in\mathbb{Z}} e^{-\rho|k-\ell|} (1+|k-\ell|)^s \end{aligned}$$

and it is clear that the last sum is finite and independent on k. ∎

We now begin the proof of the Corollary 2.2.1. We first show (i), then (iii) and finally (ii).

Proof of (i). It is clear that we have for all $n \geq 1$,

$$\|\varphi^n\| = \|\varphi^0\|.$$

With the notations of Theorem 2.2.1, let $\psi^n = Q(\lambda)\varphi^n$. As $Q(\lambda)$ is unitary we hence also have

$$\|\psi^n\| = \|\psi^0\| = \|\varphi^0\|.$$

Moreover, by (2.13), we have for all $n \geq 0$,

$$\psi^{n+1} = (\Sigma(\lambda) + R(\lambda))\psi^n$$

Now the fact that $\Sigma(\lambda)$ is an almost X-shaped unitary operator implies that for $0 \leq k \leq K = \lambda^{-\sigma}$, we have

$$|\Sigma(\lambda)\psi^n|_k = |\psi^n|_k.$$

We thus have for all $0 \leq k \leq K$,

$$\begin{aligned} \big|\,|\psi^{n+1}|_k - |\psi^n|_k\big| &= \big|\,|\psi^{n+1}|_k - |\Sigma(\lambda)\psi^n|_k\big| \\ &\leq |\psi^{n+1} - \Sigma(\lambda)\psi^n|_k \\ &\leq |R(\lambda)\psi^n|_k. \end{aligned}$$

Recall that $|R(\lambda)\psi^n|_k^2 = |(R(\lambda)\psi^n)_k|^2 + |(R(\lambda)\psi^n)_{-k}|^2$. Using Lemma 2.6.1 with $\mu = \rho_V/5$, we thus have

$$\begin{aligned} \big| |\psi^{n+1}|_k - |\psi^n|_k \big| &\leq \sqrt{2} C_{\rho_V/5} \|R(\lambda)\|_{\rho_V/5} \|\psi^n\| \\ &\leq \sqrt{2} C_{\rho_V/5} \exp(-c\lambda^{-\sigma}) \|\varphi^0\| \,. \end{aligned}$$

Hence for $n \leq \exp(c\lambda^{-\sigma}/2)$ we have by triangle inequality,

$$\big| |\psi^n|_k - |\psi^0|_k \big| \leq \sqrt{2} C_{\rho_V/5} \exp(-c\lambda^{-\sigma}/2) \|\varphi^0\| \,. \tag{2.69}$$

For all $0 \leq k \leq \lambda^{-\sigma}$, this inequality shows the extremely good conservation of the oscillatory energy $|\psi^n|_k$.

Now using the estimate (2.12) and Lemma 2.6.1, we get for all $(n,k) \in \mathbb{N}^2$,

$$\begin{aligned} |\varphi^n - \psi^n|_k &= |(\mathrm{Id} - Q(\lambda))\varphi^n|_k \\ &\leq \lambda^{1/2} \|\varphi^n\| \\ &\leq \lambda^{1/2} \|\varphi^0\| \,. \end{aligned}$$

Therefore, for all n and k,

$$\begin{aligned} \Big| \big| |\varphi^n|_k - |\varphi^0|_k \big| - \big| |\psi^n|_k - |\psi^0|_k \big| \Big| &\leq \Big| \big(|\varphi^n|_k - |\varphi^0|_k \big) - \big(|\psi^n|_k - |\psi^0|_k \big) \Big| \\ &\leq |\varphi^n - \psi^n|_k + |\varphi^0 - \psi^0|_k \\ &\leq 2\lambda^{1/2} \|\varphi^0\| \,. \end{aligned}$$

We thus deduce that for all $n \leq \exp(c\lambda^{-\sigma}/2)$ and all $k \leq \lambda^{-\sigma}$,

$$\begin{aligned} \big| |\varphi^n|_k - |\varphi^0|_k \big| &\leq \big| |\psi^n|_k - |\psi^0|_k \big| + 2\lambda^{1/2} \|\varphi^0\| \\ &\leq 2\lambda^{1/2} \|\varphi^0\| + \sqrt{2} C_{\rho_V/5} \exp(-c\lambda^{-\sigma}/2) \|\varphi^0\| \,, \end{aligned}$$

so that (2.19) is proved. ∎

Proof of (iii). Recall that λ is chosen in $(0, \lambda_0)$ and that $K = \lambda^{-\sigma}$ with the notations of Theorem 2.2.1. For a function φ, we set

$$(\pi_K \varphi)(x) = \sum_{|k| > K} \varphi_k e^{inkx}. \tag{2.70}$$

The structure of $\Sigma(\lambda)$ implies that $[\pi_K, \Sigma(\lambda)] = 0$. Hence, we have

$$\|\pi_K \Sigma(\lambda) \varphi\| = \|\pi_K \varphi\| \,.$$

Moreover, Theorem 2.2.1 ensures that $\|R(\lambda)\|_{\rho_V/5} \leq \exp(-c\lambda^{-\sigma})$, so using Lemma 2.6.1 we have

$$\|\pi_K R(\lambda) \varphi\| \leq \|R(\lambda) \varphi\| \leq C_{\rho_V/5} \exp(-c\lambda^{-\sigma}) \|\varphi\| \,.$$

Therefore, for all $n \in \mathbb{N}$, we have

$$\begin{aligned}\big|\|\pi_K\psi^{n+1}\| - \|\pi_K\psi^n\|\,\big| &= \big|\|\pi_K\psi^{n+1}\| - \|\pi_K\Sigma(\lambda)\psi^n\|\,\big| \\ &\le \|\pi_K\big(\psi^{n+1} - \Sigma(\lambda)\psi^n\big)\| \\ &\le \|\pi_K R(\lambda)\psi^n\| \\ &\le C_{\rho_V/5}\exp(-c\lambda^{-\sigma})\|\psi^n\| \\ &\le C_{\rho_V/5}\exp(-c\lambda^{-\sigma})\|\varphi^0\|\,.\end{aligned}$$

Hence, we have by triangle inequality for $n \le \exp(c\lambda^{-\sigma}/2)$,

$$\big|\|\pi_K\psi^n\| - \|\pi_K\psi^0\|\,\big| \le C_{\rho_V/5}\exp(-c\lambda^{-\sigma}/2)\|\varphi^0\|\,. \tag{2.71}$$

Let $\rho \in (0, \rho_V/5)$ and assume that $\varphi^0 \in U_\rho$. Notice that we have

$$\|\varphi^0\|^2 = \sum_{k\in\mathbb{Z}} |\varphi^0_k|^2 \le \|\varphi^0\|_\rho^2 C_{2\rho}. \tag{2.72}$$

Since $\psi^0 = Q(\lambda)\varphi^0$, Lemma 2.4.1 ensures that

$$\|\psi^0\|_\rho \le C_{\rho_V/5-\rho}\|Q(\lambda)\|_{\rho_V/5}\|\varphi^0\|_\rho\,.$$

Moreover, since $\|Q(\lambda)\|_{\rho_V/5} \le (1+\lambda_0^{1/2})$ by Theorem 2.2.1, we derive that there exists a constant $C > 0$ depending only on V, ρ_V, γ, ν, and ρ such that

$$\forall k \in \mathbb{Z}, \qquad |\psi^0_k| \le C\|\varphi^0\|_\rho e^{-\rho|k|}.$$

Therefore,

$$\begin{aligned}\|\pi_K\psi^0\|^2 &\le C^2\|\varphi^0\|_\rho^2 \sum_{|k|>K} e^{-2\rho_k} \\ &\le C^2\|\varphi^0\|_\rho^2 e^{-2\rho\lambda^{-\sigma}} 2\sum_{k=0}^{\infty} e^{-2\rho|k|}.\end{aligned}$$

Hence, after a possible increase of C, we get

$$\|\pi_K\psi^0\| \le C\|\varphi^0\|_\rho \exp(-\rho\lambda^{-\sigma}).$$

Then, using (2.71) and (2.72), we get, after another possible increase of C, for all $n \le \exp(c\lambda^{-\sigma}/2)$,

$$\|\pi_K\psi^n\| \le C\exp(-c'\lambda^{-\sigma})\|\varphi^0\|_\rho \tag{2.73}$$

where $c' = \min(c/2, \rho) > 0$ only depends on V, γ, ν, and ρ.
Moreover, the inequality (2.69) shows that for $0 \le k \le \lambda^{-\sigma}$ and all $n \le \exp(c\lambda^{-\sigma}/2)$, we have for all $\mu > 0$,

$$e^{\mu k}\,\big|\,|\psi^n|_k - |\psi^0|_k\big| \le \sqrt{2}C_{\rho_V/5}\exp(\mu K - c\lambda^{-\sigma}/2)\|\varphi^0\|\,,$$

so that if we set $\mu_0 = c'/2 > 0$, we have for $\mu \in (0, \mu_0)$,

$$\sup_{0\le k\le\lambda^{-\sigma}} \Big(e^{\mu k}\,\big|\,|\psi^n|_k - |\psi^0|_k\big|\Big) \le C'\exp(-c'\lambda^{-\sigma}/2)\|\varphi^0\| \tag{2.74}$$

with $C' = \sqrt{2}C_{\rho_V/5}$, since $\mu - c/2 \leq -\mu_0 = -c'/2$.

Recall that for all $n \in \mathbb{N}$, $\varphi^n = Q(\lambda)^*\psi^n$. Using the decomposition

$$\psi^n = (\mathrm{Id} - \pi_K)\psi^n + \pi_K\psi^n$$

we thus have

$$\varphi^n = Q(\lambda)^*(\mathrm{Id} - \pi_K)\psi^n + Q(\lambda)^*\pi_K\psi^n$$

and therefore

$$\varphi^n - \psi^n = (Q(\lambda)^* - \mathrm{Id})(\mathrm{Id} - \pi_K)\psi^n + (Q(\lambda)^* - \mathrm{Id})\pi_K\psi^n. \tag{2.75}$$

Using Lemma 2.6.1 and estimate (2.12), we get for $0 \leq k \leq \lambda^{-\sigma}$,

$$\begin{aligned}
\big| |\varphi^n|_k - |\psi^n|_k \big| &\leq |\varphi^n - \psi^n|_k \\
&\leq |(Q(\lambda)^* - \mathrm{Id})(\mathrm{Id} - \pi_K)\psi^n|_k + |(Q(\lambda)^* - \mathrm{Id})\pi_K\psi^n|_k \\
&\leq |(Q(\lambda)^* - \mathrm{Id})(\mathrm{Id} - \pi_K)\psi^n|_k + C_{\rho_V/4}\lambda^{1/2}\|\pi_K\psi^n\|\,.
\end{aligned}$$

On one hand, we have with (2.73) that for all $\mu \in (0, \mu_0)$, all $n \leq \exp(c\lambda^{-\sigma}/2)$ and all $k \leq \lambda^{-\sigma}$,

$$\|\pi_K\psi^n\|\, e^{+\mu k} \leq C\exp(-c'\lambda^{-\sigma}/2)\|\varphi^0\|_\rho \tag{2.76}$$

since $\mu - c' \leq -c'/2$.
On the other hand, with Lemma 2.4.1 and Theorem 2.2.1, we have

$$\|(Q(\lambda)^* - \mathrm{Id})(\mathrm{Id} - \pi_K)\psi^n\|_\mu \leq C_{\rho_V/4-\mu}\lambda^{1/2}\|(\mathrm{Id} - \pi_K)\psi^n\|_\mu$$

and the inequalities (2.74) and (2.72) ensure that for all $n \leq \exp(c\lambda^{-\sigma}/2)$,

$$\begin{aligned}
\|(\mathrm{Id} - \pi_K)\psi^n\|_\mu &\leq \sqrt{2}\|\psi^0\|_\mu + C'\exp(-c'\lambda^{-\sigma}/2)\|\varphi^0\| \\
&\leq C\|\varphi^0\|_\rho
\end{aligned}$$

after a possible increase of C (still depending only on V, ρ_V, γ, ν and ρ).
As a conclusion, after another possible increase of C, we get for all $n \leq \exp(c\lambda^{-\sigma}/2)$, all $k \leq \lambda^{-\sigma}$ and all $\mu \in (0, \mu_0)$,

$$\big| |\varphi^n|_k - |\psi^n|_k \big| e^{+\mu k} \leq C\lambda^{1/2}\|\varphi^0\|_\rho\,. \tag{2.77}$$

Eventually, we get for all $n \leq \exp(c\lambda^{-\sigma}/2)$, all $k \leq \lambda^{-\sigma}$ and all $\mu \in (0, \mu_0)$

$$\begin{aligned}
&e^{\mu k}\big| |\varphi^n|_k - |\varphi^0|_k \big| \\
&\leq e^{\mu k}\big| |\varphi^n|_k - |\psi^n|_k \big| + e^{\mu k}\big| |\psi^n|_k - |\psi^0|_k \big| + e^{\mu k}\big| |\psi^0|_k - |\varphi^0|_k \big| \\
&\leq C\lambda^{1/2}\|\varphi^0\|_\rho
\end{aligned}$$

after a last increase of C, using (2.77), (2.74) and the fact that

$$\|\psi^0 - \varphi^0\|_\mu \leq \|\psi^0 - \varphi^0\|_\rho \leq C_{\rho_V/4-\rho}\|\varphi^0\|_\rho$$

by Lemma 2.4.1. This proves (2.21) and hence (iii). ∎

Proof of (ii). With the notations of Theorem 2.2.1 and Corollary 2.2.1, we assume that $h \in (0, 1)$ satisfying (2.7), $\lambda \in (0, \lambda_0)$ and $s > 1/2$ are given real numbers. Theorem

2.2.1 and Lemma 2.6.2 ensure that there exists a positive constant c_s depending only on V, ρ_V, γ, ν and s such that

$$\forall k \in \mathbb{N}, \qquad |\psi^0|_k \leq c_s \|\varphi^0\|_{s,\infty} (1+k)^{-s}.$$

Note that, since $s > 1/2$, we have

$$\|\varphi^0\| = \Big(\sum_{k\in\mathbb{N}} |\varphi^0|_k^2\Big)^{1/2} \leq \Big(\sum_{p\in\mathbb{N}} (1+p)^{-2s}\Big)^{1/2} \|\varphi^0\|_{s,\infty} < +\infty$$

and similarly

$$\begin{aligned}
\|\pi_K \psi^0\|^2 &= \sum_{k>K} |\psi^0|_k^2 \\
&\leq c_s^2 \|\varphi^0\|_{s,\infty}^2 \sum_{k>K} (1+k)^{-2s} \\
&\leq c_s^2 \|\varphi^0\|_{s,\infty}^2 \int_{K-1}^{+\infty} (1+t)^{-2s}\,\mathrm{d}t \\
&\leq c_s^2 \|\varphi^0\|_{s,\infty}^2 \frac{K^{-(2s-1)}}{2s-1} \\
&\leq c_s^2 \|\varphi^0\|_{s,\infty}^2 \frac{\lambda^{\sigma(2s-1)}}{2s-1}
\end{aligned}$$

which reads

$$\|\pi_K \psi^0\| \leq c_s \|\varphi^0\|_{s,\infty} \lambda^{\sigma(s-1/2)}$$

after a possible increase of the constant c_s.
Another possible increase of the constant c_s allows us to write, using (2.71), that for $n \leq \exp(c\lambda^{-\sigma}/2)$,

$$\begin{aligned}
\|\pi_K \psi^n\| &\leq C_{\rho_V/5} \exp(-c\lambda^{-\sigma}/2) \|\varphi^0\| + \|\pi_K \psi^0\| \\
&\leq c_s \exp(-c\lambda^{-\sigma}/2) \|\varphi^0\|_{s,\infty} + c_s \|\varphi^0\|_{s,\infty} \lambda^{\sigma(s-1/2)},
\end{aligned}$$

and finally

$$\|\pi_K \psi^n\| \leq c_s \lambda^{\sigma(s-1/2)} \|\varphi^0\|_{s,\infty}. \tag{2.78}$$

Note that c_s only depends on V, γ, ν and s.
Moreover, by (2.69), we have, for $\lambda \in (0, \lambda_0)$, $n \leq \exp(c\lambda^{-\sigma}/2)$ and $0 \leq k \leq K = \lambda^{-\sigma}$,

$$(1+k)^s \big| |\psi^n|_k - |\psi^0|_k \big| \leq C \exp(-c\lambda^{-\sigma}/2)(1+\lambda^{-\sigma})^s \|\varphi^0\|,$$

where C depends only on V. After a possible increase of c_s, we can assume that for all $\lambda \in (0, \lambda_0)$ and all $n \leq \exp(c\lambda^{-\sigma}/2)$,

$$\sup_{0\leq k\leq \lambda^{-\sigma}} \Big((1+k)^s \big| |\psi^n|_k - |\psi^0|_k \big| \Big) \leq c_s \exp(-c\lambda^{-\sigma}/4) \|\varphi^0\|. \tag{2.79}$$

This implies that for all $n \leq \exp(c\lambda^{-\sigma}/2)$ and all k such that $0 \leq k \leq \lambda^{-\sigma}$, we have after a possible increase of c_s,

$$(1+k)^s |\psi^n|_k \leq c_s \|\varphi^0\|_{s,\infty}. \tag{2.80}$$

On one hand, we have with (2.78) and after another increase of the constant c_s, for all $n \leq \exp(c\lambda^{-\sigma})$ and all $k \leq \lambda^{-\sigma}$,

$$\|\pi_K \psi^n\| \, (1+k)^{s'} \leq c_s \lambda^{\sigma(s-s'-1/2)} \|\varphi^0\|_{s,\infty} \, ,$$

and, since $\lambda_0 \leq 1$, for such k and n, we get

$$\|\pi_K \psi^n\| \, (1+k)^{s'} \leq c_s \|\varphi^0\|_{s,\infty} \, . \tag{2.81}$$

On the other hand, with Theorem 2.2.1 and Lemma 2.6.2, we get for all $n \in \mathbb{N}$,

$$\|(Q(\lambda)^* - \mathrm{Id})(\mathrm{Id} - \pi_K)\psi^n\|_{s',\infty} \leq C\lambda^{1/2} \|(\mathrm{Id} - \pi_K)\psi^n\|_{s',\infty} \, ,$$

where C depends only on V and s'. With the inequality (2.80), this yields, after a possible increase of the constant c_s,

$$\|(Q(\lambda)^* - \mathrm{Id})(\mathrm{Id} - \pi_K)\psi^n\|_{s',\infty} \leq c_s \lambda^{1/2} \|\varphi^0\|_{s,\infty} \, . \tag{2.82}$$

Hence, since for all $k \leq \lambda^{-\sigma}$ and all $n \leq \exp(c\lambda^{-\sigma}/2)$,

$$\begin{aligned} \big| |\varphi^n|_k - |\psi^n|_k \big| &\leq |\varphi^n - \psi^n|_k \\ &\leq |(Q(\lambda)^* - \mathrm{Id})(\mathrm{Id} - \pi_K)\psi^n|_k + |(Q(\lambda)^* - \mathrm{Id})\pi_K \psi^n|_k , \end{aligned}$$

we derive, with (2.81) and (2.82), Lemma 2.6.1 and Theorem 2.2.1, that for such n and k,

$$\begin{aligned} &(1+k)^{s'} \big| |\varphi^n|_k - |\psi^n|_k \big| \\ &\leq c_s \lambda^{1/2} \|\varphi^0\|_{s,\infty} + C_{\rho_V/5} \|Q(\lambda)^* - \mathrm{Id}\|_{\rho_V/5} \|\pi_K \psi^n\| \, (1+k)^{s'} \\ &\leq c_s \lambda^{1/2} \|\pi_K \psi^n\| \, , \end{aligned}$$

after a possible increase of the constant c_s.

In addition to this last inequality and (2.79), the fact that for all $k \in \mathbb{N}$,

$$(1+k)^{s'} \Big| |\varphi^0|_k - |\psi^0|_k \Big| \leq c_s \lambda^{1/2} \|\varphi^0\|_{s,\infty}$$

which is a consequence of Theorem 2.2.1 and Lemma 2.6.2, ensures by triangle inequality, that for all $n \leq \exp(c\lambda^{-\sigma}/2)$ and $k \leq \lambda^{-\sigma}$,

$$\begin{aligned} &(1+k)^{s'} \big| |\varphi^n|_k - |\varphi^0|_k \big| \\ &\leq (1+k)^{s'} \big| |\varphi^n|_k - |\psi^n|_k \big| + (1+k)^{s'} \big| |\psi^n|_k - |\psi^0|_k \big| + (1+k)^{s'} \big| |\psi^0|_k - |\varphi^0|_k \big| \\ &\leq c_s \lambda^{1/2} \|\varphi^0\|_{s,\infty} \, , \end{aligned}$$

after a possible increase of the constant c_s, still depending only on V, ρ_V, γ, ν and s. This proves (2.20) and hence (ii). ∎

2.7 Appendix

Proof of Lemma 2.4.2. We consider the differential equation

$$\left\{\begin{array}{lcl} \mathsf{q}'(t) & = & f(t,\mathsf{q}(t)) \\ \mathsf{q}(0) & = & 1 \end{array}\right. ,$$

with

$$f(t,Y) = \frac{\mathsf{s}_0 C_\delta Y}{1 - C_\delta^2\big(\kappa(e^{2hM_V t}-1) + (Y^2-1)\big)},$$

(recall that $\mathsf{s}_0 = M_V \kappa h$). This equation has a unique analytical solution: There exists a number $R > 0$ such that for $t \in (-R, R)$, $\mathsf{q}(t)$ expands in power series of t (see for instance [HNW93]). We can assume that R is maximal in this sense. Let

$$T = \frac{2}{hM_V}\ln\left(1+\frac{1}{4\kappa C_\delta^2}\right) \quad \text{and} \quad D = \left(1+\frac{1}{4C_\delta^2}\right)^{1/2} - 1.$$

For $0 \le t \le T$ we have $\kappa C_\delta^2(e^{2hM_V t}-1) \le \frac{1}{4}$ and if $1 \le Y \le 1+D$, then $C_\delta^2(Y^2-1) \le \frac{1}{4}$. Therefore,

$$\left\{\begin{array}{l} 0 \le t \le T \\ 0 \le Y-1 \le D \end{array}\right. \implies 0 < f(t,Y) \le 2\mathsf{s}_0 C_\delta (D+1). \tag{2.83}$$

This implies that q is an increasing function of t as long as $t \in (0,T)$ and $\mathsf{q}(t) \le D+1$. Notice that, for $t \le 0$, we have $0 < \mathsf{q}(t) \le 1$.

Moreover, the fact that $K \ge K_0 \ge 1$ (see Lemma 2.4.2) implies that $\frac{1}{\kappa} \le 4$ and hence we have

$$T = \frac{2}{hM_V}\ln\left(1+\frac{1}{4\kappa C_\delta^2}\right) \ge \frac{1}{4hM_V\kappa C_\delta^2}. \tag{2.84}$$

If for all $t \in (0,R)$, $1 < \mathsf{q}(t) < 1+D$, we must have in view of (2.83), $T < R$. To the contrary, there exists t_D^* in $(0,R)$ such that $\mathsf{q}(t_D^*) = 1+D$. We can therefore choose $t_D^* \in (0,R)$ in such a way that for all $t \in (0, t_D^*)$, $1 < \mathsf{q}(t) < 1+D$.

Assume that $t_D^* \le T$, then

$$y(t_D^*) - y(0) = D \le \int_0^{t_D^*} f(u, y(u))\mathrm{d}u \le 2\mathsf{s}_0 C_\delta (D+1) t_D^*,$$

using (2.83). Hence we have

$$\begin{aligned} t_D^* &\geq \frac{D}{2\mathsf{s}_0 C_\delta (D+1)} \\ &\geq \frac{(1+\frac{1}{4C_\delta^2})^{\frac{1}{2}}-1}{(1+\frac{1}{4C_\delta^2})^{\frac{1}{2}}} \frac{1}{2\kappa h M_V C_\delta} \\ &\geq \frac{1}{16C_\delta^2\left(1+\frac{1}{4C_\delta^2}\right)} \frac{1}{\kappa h M_V C_\delta} \\ &\geq \frac{1}{4hM_V\kappa C_\delta^2} \frac{1}{5C_\delta}. \end{aligned}$$

Hence, we have

$$R > t_D^* \geq \frac{1}{20hM_V\kappa C_\delta^3}. \tag{2.85}$$

To the contrary, if $t_D^* \geq T$ then we have using (2.84)

$$R > T \geq \frac{1}{4hM_V\kappa C_\delta^2}.$$

As $C_\delta^{-1} \leq 1$, this implies that (2.85) holds in any case. This determines the value of $r < R$ in the Lemma. Now as $t \mapsto \mathsf{q}(t)$ is an increasing positive function, so is $t \mapsto \mathsf{q}'(t) = f(t, \mathsf{q}(t))$.

With the previous notations, in any case, we have with (2.83) that for all $t \in (-r, r)$

$$\mathsf{q}'(t) \leq \frac{\kappa h M_V C_\delta (D+1)}{1-\frac{1}{4}-\frac{1}{4}} \leq \sqrt{5}\kappa h M_V C_\delta.$$

We have already seen that for all $t \in (0, r)$, $1 \leq \mathsf{q}(t) \leq D+1 \leq \frac{\sqrt{5}}{2}$. And it follows from the estimation on q' that

$$\mathsf{q}(-r) = 1 - \int_{-r}^{0} \mathsf{q}'(s)\mathrm{d}s \geq 1 - r\sqrt{5}\kappa h M_V C_\delta > 0.8 \ .$$

This proves the estimates on $\mathsf{q}(t)$. The estimates for $\mathsf{x}(t)$ and $\mathsf{s}(t)$ are then obtained straightforwardly. ∎

Chapter 3

Exponential Runge-Kutta methods for the Schrödinger equation

Contents

3.1 Introduction

Exponential integrators are specific geometric integrators. They have become very popular recently for the numerical integration of first order in time problems (see for example [MW05], where their multiple discoveries from the 60's are recalled). For example, some specific exponential integrators have been used for solving the non-linear Schrödinger equation (see for example [BS05] where the numerical analysis is made in one periodic space dimension).

Exponential Runge-Kutta methods are particular exponential integrators. These methods have been derived and analysed for semi-linear *parabolic* Cauchy problems (see for example [HO05b] for collocation methods and [HO05a] for explicit methods).

This article deals with such exponential Runge-Kutta methods applied to the linear and semi-linear Schrödinger equations, namely

$$\begin{aligned} &\partial_t u(t,x) - i\Delta u(t,x) = f(t,x) && \text{(linear)} \\ \text{and} \quad &\partial_t u(t,x) - i\Delta u(t,x) = f(t,u) && \text{(semi-linear)}, \end{aligned}$$

seen as Cauchy problems in time (no space discretization is made). The analysis in these cases is different from the one performed for parabolic problems since the spectrum of the linear operators are different. We provide a numerical analysis of exponential Runge-Kutta methods of collocation type applied to linear and semi-linear problems on the d-dimensional torus ($d \in \mathbb{N}^*$). We show that s-stage collocation methods are of order s for suitable Sobolev norms. Moreover, when applied to a linear or semi-linear Schrödinger Cauchy problem, we give algebraic sufficient conditions on the collocation points for s-stage methods to be of order $s+1$ and $s+2$. These results need additional assumptions, for example on the non-linear term of semi-linear problems. These assumptions are fulfilled in many cases of interest, for example in the case of the cubic non-linear Schrödinger equation. We also mention, quantify and explain the effects of resonant time steps that appear when solving linear problems over finite time.

This paper is concerned with the numerical integration of hamiltonian PDEs over a finite time interval. For results on numerical integration of hamiltonian PDEs over long time, the reader may refer for example to [CHL06, CHL07, HL07] for results on the non-linear wave equation obtained via modulated Fourier expansions and to [DF07] for results on splitting methods applied to the linear Schrödinger equation with small potential.

Let us mention that many other geometric integrators do exist to integrate numerically Schrödinger Cauchy problems. For example, some of them take advantage of the conservation of phase space properties considering multisymplectic formulations (see for example [BIS07]). This article is not aimed at comparing exponential Runge-Kutta methods of collocation type with these kinds of geometric integrators in any way (if any makes sense).

This article is organized as follows: In Section 3.2, we deal with exponential Runge-Kutta methods of collocation type applied to *linear* Schrödinger Cauchy problems. After the introduction of the notations for these methods for linear problems, we look at order conditions for s-stage methods: We show that these methods are of order at least s and give sufficient conditions to achieve orders $s+1$ and $s+2$. We show the effect of resonant time steps on the dependency of order constants with respect to time. In Section 3.3, we look at exponential Runge-Kutta methods of collocation type applied to *semi-linear* Schrödinger Cauchy problems. After setting notations for these methods applied to non-linear problems, we prove, under suitable assumptions, that s-stage methods provide a unique numerical solution and are of order at least s. Moreover, we give sufficient order conditions for s-stage methods to be of order $s+1$ and $s+2$. We conclude the article with numerical experiments and some comments.

3.2 Linear problems

3.2.1 Notations

We denote by $\mathbb{T}$ the one-dimensional torus $\mathbb{R}/(2\pi\mathbb{Z})$. For a given positive integer d, we denote by $\mathbb{T}^d$ the d-dimensional torus. For $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $y = (y_1, \ldots, y_d) \in \mathbb{R}^d$, we denote by $x.y$ the real number $x_1y_1 + \cdots + x_dy_d$ and by $|x|_2$ the non-negative real number $(x_1^2 + \cdots + x_d^2)^{1/2}$. For $k \in \mathbb{Z}^d$, we denote by p_k the function on $\mathbb{T}^d$ defined for $x \in \mathbb{T}^d$ by $p_k(x) = e^{ik.x}$. $L^2(\mathbb{T}^d)$ (or simply L^2) denotes the set of (classes of) complex functions f on $\mathbb{T}^d$ such that $\int_{\mathbb{T}^d} |f(x)|^2 \mathrm{d}x < +\infty$, endowed with the norm $\|f\|_{L^2(\mathbb{T}^d)} = \left((2\pi)^{-d} \int_{\mathbb{T}^d} |f(x)|^2 \mathrm{d}x\right)^{1/2}$. An unbounded linear operator A acting on L^2 is said to be diagonal if for all $k \in \mathbb{Z}^d$, p_k is in the domain of A and there exists a $\lambda_k \in \mathbb{C}$ such that $Ap_k = \lambda_k p_k$. For example, the Laplace operator Δ acting on $L^2(\mathbb{T}^d)$ is diagonal with for all $k \in \mathbb{Z}^d$, $\Delta p_k = -|k|_2^2 p_k$. Another example is the identity operator on L^2 denoted by Id. For all such A and all function φ from $\mathbb{C}$ to itself, we denote by $\varphi(A)$ the unbounded diagonal linear operator acting on L^2 whose domain is the set of linear combinations of functions $(p_k)_{k\in\mathbb{Z}^d}$ defined for all $k \in \mathbb{Z}^d$ by

$$\varphi(A)p_k = \varphi(\lambda_k)p_k. \tag{3.1}$$

For all function $f \in L^2$ and all $k \in \mathbb{Z}^d$, we denote by $\hat{f}_k$ the Fourier coefficient $(2\pi)^{-d} \int_{\mathbb{T}^d} f(x)e^{-ik.x} \mathrm{d}x$. For $\alpha \in \mathbb{R}^+$, we denote by $H^\alpha(\mathbb{T}^d)$ (or simply H^α) the space of (classes of) complex functions $f \in L^2(\mathbb{T}^d)$ such that $\sum_{k\in\mathbb{Z}^d\setminus\{0\}} |\hat{f}_k|^2 |k|_2^{2\alpha} < +\infty$, endowed with the norm

$$\|f\|_{H^\alpha} = \Big(|\hat{f}_0|^2 + \sum_{k\in\mathbb{Z}^d\setminus\{0\}} |\hat{f}_k|^2 |k|_2^{2\alpha}\Big)^{1/2},$$

Note that $L^2 = H^0$ with the same norm. With these notations, if φ is a function from $\mathbb{C}$ to $\mathbb{C}$ bounded by $M \geq 0$ on $i\mathbb{R}$, then for all $h > 0$ and $\alpha \geq 0$, $\varphi(ih\Delta)$ is a bounded linear operator from H^α to itself with $\|\varphi(ih\Delta)\|_{H^\alpha \to H^\alpha} \leq M$.

Following [HO05b], we set for all $k \in \mathbb{N}^*$

$$\forall z \in \mathbb{C}^*, \qquad \varphi_k(z) = \frac{1}{z^k}\Big(e^z - \sum_{p=0}^{k-1} \frac{z^p}{p!}\Big), \tag{3.2}$$

and $\varphi_k(0) = \frac{1}{k!}$. We also set $\varphi_0 = \exp$. Note that for all $k \in \mathbb{N}$, φ_k is holomorphic on $\mathbb{C}$, and is bounded on $i\mathbb{R}$. For all unbounded linear diagonal operator A and all $h > 0$, we have

$$\forall k \in \mathbb{N}, \qquad \varphi_{k+1}(-hA) = \frac{1}{h^{k+1}} \int_0^h e^{-(h-\tau)A} \frac{\tau^k}{k!} \mathrm{d}\tau. \tag{3.3}$$

In order to compute numerical solutions of the following Cauchy problem

$$\begin{array}{rcll} \partial_t u(t,x) + A\,u(t,x) & = & f(t,x) & (t,x) \in [0,T] \times \mathbb{T}^d \\ u(0,x) & = & u_0(x) & x \in \mathbb{T}^d \end{array} \tag{3.4}$$

where $T > 0$, $A = -i\Delta$, u_0 and f are given, we consider the following numerical methods, called *exponential Runge-Kutta methods of collocation type*; we refer to [HO05b] for a derivation of such methods for semi-linear problems based on the variation-of-constant formula:

$$\forall n \in \{0, \dots, N\}, \qquad \begin{cases} t_n = nh \\ u_{n+1} = e^{-hA}u_n + h\displaystyle\sum_{i=1}^{s} b_i(-hA) f(t_n + c_i h) \end{cases} \tag{3.5}$$

where $h > 0$, $s \in \mathbb{N}^*$, $(c_1, \dots, c_s) \in [0,1]^s$ are given and $N = \lfloor T/h \rfloor$. We assume that for all $(i,j) \in \{1, \dots, s\}^2$, $c_i \neq c_j$ if $i \neq j$. The operators $b_i(-hA)$ are defined by

$$\forall i \in \{1, \dots, n\}, \qquad b_i(-hA) = \frac{1}{h}\int_0^h e^{-(h-\tau)A} l_i(\tau)\mathrm{d}\tau, \tag{3.6}$$

where for all $i \in \{1, \dots, s\}$, l_i is the i-th Lagrange polynomial with respect to the points $(c_j h)_{j\in\{1,\dots,s\}}$:

$$\forall i \in \{1, \dots, s\}, \qquad l_i(\tau) = \prod_{j=1, j\neq i}^{s} \frac{\frac{\tau}{h} - c_j}{c_i - c_j}. \tag{3.7}$$

Note that, for all $i \in \{1, \dots, s\}$, $b_i \in \mathrm{span}(\varphi_0, \dots, \varphi_{s-1})$. We therefore derive that b_i is holomorphic on $\mathbb{C}$ and is bounded on $i\mathbb{R}$.

In the case $A = 0$, the linear PDE (3.4) reduces to the collection of linear ODEs for $x \in \mathbb{T}^d$

$$\frac{\mathrm{d}u}{\mathrm{d}t}(t,x) = f(t,x) \qquad\qquad t \in [0,T],$$

with initial conditions $u(0,x) = u_0(x)$. Moreover, the exponential Runge-Kutta method (3.5) reduces to the classical Runge-Kutta method called *underlying Runge-Kutta method* with coefficients $(b_i(0))_{i\in\{1,\dots,s\}}$ since for all $i \in \{1, \dots, s\}$, $b_i(-hA) = b_i(0)\mathrm{Id}$.

3.2.2 Order of exponential Runge-Kutta methods for linear problems

3.2.2.1 An s-stage method is of order s

Our first result shows that an s-stage exponential Runge-Kutta method of collocation type (3.5) applied to a linear Schrödinger problem (3.4) is of order at least s, provided that the right hand side of (3.4) is sufficiently smooth with respect to the time, when considered as an $L^2(\mathbb{T}^d)$-valued function. This somehow natural result is very similar to the one obtained in the context of parabolic linear equations in [HO05b] (see Theorem 1). However, note that in the case of the linear Schrödinger equation, the order constant appearing does not depend on T.

Theorem 3.2.1 *An s-stage exponential Runge-Kutta method of collocation type* (3.5) *applied to a linear problem* (3.4) *such that $f \in C^s([0,T], L^2(\mathbb{T}^d))$ is of global order s in* *the sense that there exists a positive constant C depending only on $(c_1, \ldots, c_s)$ such that for all $h > 0$ and all $n \in \{0, \ldots, N\}$,*

$$\|u_n - u(t_n)\|_{L^2} \leq C\, h^s \int_0^{t_n} \|f^{(s)}(\tau)\|_{L^2}\, \mathrm{d}\tau. \tag{3.8}$$

Proof. Let us denote for all $n \in \{0, \ldots, N\}$, $e_n = u_n - u(t_n)$. The variation of constant formula for the exact solution u of problem (3.4) reads for $h \in]0, T]$ and $t \in [0, T-h]$,

$$u(t+h) = e^{-hA} u(t) + \int_0^h e^{-(h-s)A} f(s) \mathrm{d}s. \tag{3.9}$$

By Taylor expansion of f in (3.9) and (3.5), and using the properties of the Lagrange collocation polynomials (3.7) involved in the functions $(b_i)_{i\in\{1,\ldots,s\}}$ (see (3.6)), one has

$$\begin{aligned} e_{n+1} = e^{-hA} e_n &+ \int_0^h e^{-(h-\tau)A} \int_0^\tau \frac{(\tau-\sigma)^{s-1}}{(s-1)!} f^{(s)}(t_n+\sigma) \mathrm{d}\sigma \mathrm{d}\tau \\ &- \sum_{i=1}^s \int_0^h e^{-(h-\tau)A} l_i(\tau) \mathrm{d}\tau \int_0^{c_i h} \frac{(c_i h - \sigma)^{s-1}}{(s-1)!} f^{(s)}(t_n+\sigma) \mathrm{d}\sigma. \end{aligned} \tag{3.10}$$

Denoting $\delta_{n+1} = e_{n+1} - e^{-hA} e_n$, and using the properties of the function $\varphi_0 = \exp$, we have that for some constant C depending only on $c_1, \ldots, c_s$,

$$\|\delta_{n+1}\|_{L^2} \leq C\, h^s \int_{t_n}^{t_{n+1}} \|f^{(s)}(t_n+\tau)\|_{L^2}\, \mathrm{d}\tau.$$

This inequality and the relation

$$e_n = \sum_{p=0}^{n-1} e^{-phA} \delta_{n-p}$$

together with the properties of the function $\varphi_0 = \exp$ allow us to conclude.

■

Remark 3.2.1 *Note that for all $r > 0$, estimation (3.8) holds true with the $L^2(\mathbb{T}^d)$-norm replaced by the $H^r(\mathbb{T}^d)$-norm provided that $f \in C^s([0,T], H^r(\mathbb{T}^d))$.*

3.2.2.2 Achieving order $s+1$

If the underlying Runge-Kutta method is of order $s+1$, then we may expect a better order for the exponential Runge-Kutta method of collocation type, assuming that the right-hand side of (3.4) has higher spatial regularity than just a standard L^2-one. This is a difference with the case of parabolic linear equations studied in [HO05b] (see Theorem 2). Let us first recall (see [HO05b]) that

Lemma 3.2.1 *If the underlying Runge-Kutta method is of order $s+1$, then*

$$\sum_{i=1}^s c_i^s b_i(0) = \frac{1}{s+1}. \tag{3.11}$$

We are now able to state the first result of this section:

Theorem 3.2.2 *Assume $r \geq 0$ is given, the underlying Runge-Kutta method is of order $s+1$ and $f \in C^{s+1}([0,T], L^2(\mathbb{T}^d))$ is such that $f^{(s)} \in C^1([0,T], H^{r+2})$. Then, the exponential Runge-Kutta method of collocation type* (3.5) *applied to the linear problem* (3.4) *is of order $s+1$ in the sense that there exists a positive constant C depending only on $(c_1, \dots, c_s)$ such that we have for all $h > 0$ and all $n \in \{0, \dots, N\}$,*

$$\|u(t_n) - u_n\|_{H^r} \leq C\,T\,h^{s+1}\Big(\|f^{(s)}(0)\|_{H^{r+2}} + \int_0^{t_n} \|f^{(s+1)}(\tau)\|_{H^{r+2}}\,\mathrm{d}\tau\Big). \tag{3.12}$$

Proof. With the notations of the previous proof, using another Taylor series expansion of f in (3.10), we write

$$\delta_{n+1} = \delta_{n+1}^{(1)} + \delta_{n+1}^{(2)},$$

with

$$\begin{aligned}\delta_{n+1}^{(1)} = \Big(&\int_0^h e^{-(h-\tau)A} \int_0^\tau \frac{(\tau-\sigma)^{s-1}}{(s-1)!}\mathrm{d}\sigma\mathrm{d}\tau \\ &- \sum_{i=1}^s \int_0^h e^{-(h-\tau)A} l_i(\tau)\mathrm{d}\tau \int_0^{c_i h} \frac{(c_i h-\sigma)^{s-1}}{(s-1)!}\mathrm{d}\sigma\Big) f^{(s)}(t_n),\end{aligned}$$

and

$$\begin{aligned}\delta_{n+1}^{(2)} = &\int_0^h e^{-(h-\tau)A} \int_0^\tau \frac{(\tau-\sigma)^{s-1}}{(s-1)!} \int_0^\sigma f^{(s+1)}(t_n+\xi)\mathrm{d}\xi\mathrm{d}\sigma\mathrm{d}\tau \\ &- \sum_{i=1}^s \int_0^h e^{-(h-\tau)A} l_i(\tau)\mathrm{d}\tau \int_0^{c_i h} \frac{(c_i h-\sigma)^{s-1}}{(s-1)!} \int_0^\sigma f^{(s+1)}(t_n+\xi)\mathrm{d}\xi\mathrm{d}\sigma.\end{aligned}$$

Accordingly, we set

$$e_n^{(1)} = \sum_{p=0}^{n-1} e^{-phA}\delta_{n-p}^{(1)} \qquad \text{and} \qquad e_n^{(2)} = \sum_{p=0}^{n-1} e^{-phA}\delta_{n-p}^{(2)}.$$

On one hand, there exists a constant $C_1 > 0$ depending only on $(c_1, \dots, c_s)$ such that $\|\delta_{n+1}^{(2)}\|_{H^r(\mathbb{T}^d)} \leq C_1\,h^{s+1} \int_{t_n}^{t_{n+1}} \|f^{(s+1)}(\tau)\|_{H^r(\mathbb{T}^d)}\,\mathrm{d}\tau$. Therefore, we derive that

$$\|e_n^{(2)}\|_{H^r(\mathbb{T}^d)} \leq C_1\,h^{s+1} \int_0^{t_n} \|f^{(s+1)}(\tau)\|_{H^r(\mathbb{T}^d)}\,\mathrm{d}\tau. \tag{3.13}$$

On the other hand, since

$$\delta_{n+1}^{(1)} = h^{s+1}\big(\varphi_{s+1}(-hA) - \frac{1}{s!}\sum_{i=1}^s c_i^s b_i(-hA)\big) f^{(s)}(t_n),$$

an Abel summation process yields

$$\begin{aligned}e_n^{(1)} &= h^{s+1}\psi_{s+1}(-hA) \\ &\Big(\big(\sum_{p=0}^{n-1} e^{-phA}\big) f^{(s)}(0) + \sum_{p=0}^{n-2}\big(\sum_{k=0}^{p} e^{-khA}\big)\big(f^{(s)}(t_{n-p-2}) - f^{(s)}(t_{n-p-1})\big)\Big),\end{aligned}$$

where $\psi_{s+1}(z) = \varphi_{s+1}(z) - \frac{1}{s!}\sum_{i=1}^{s} c_i^s b_i(z)$. Since the underlying Runge-Kutta method is of order $s+1$, Lemma 3.2.1 and the fact that $\varphi_{s+1}(0) = ((s+1)!)^{-1}$ ensure that $\psi_{s+1}(0) = 0$. Moreover, ψ_{s+1} is an holomorphic function on $\mathbb{C}$ and is bounded on $i\mathbb{R}$. We derive that there exists an holomorphic function $\tilde{\psi}_{s+1}$ that is bounded on $i\mathbb{R}$ such that $\forall z \in \mathbb{C}$, $\psi_{s+1}(z) = z\tilde{\psi}_{s+1}(z)$. Hence, there exists a positive constant C_2 depending only on $(c_1, \ldots, c_s)$ such that $\|\tilde{\psi}_{s+1}(-hA)\|_{H^r(\mathbb{T}^d)\to H^r(\mathbb{T}^d)} \leq C_2$. We derive that

$$\|e_n^{(1)}\|_{H^r(\mathbb{T}^d)} \leq C_2 h^{s+1}\Big(\| - hA\big(\sum_{p=0}^{n-1} e^{-phA}\big) f^{(s)}(0)\|_{H^r(\mathbb{T}^d)} + \sum_{p=0}^{n-2} \| - hA\big(\sum_{k=0}^{p} e^{-khA}\big)\int_{t_{n-p-2}}^{t_{n-p-1}} f^{(s+1)}(\tau)\mathrm{d}\tau\|_{H^r(\mathbb{T}^d)}\Big).$$

For all $k \in \mathbb{Z}^d$ and $p \in \{0, \ldots, N-1\}$,

$$\Big| - ih|k|_2^2 \sum_{\ell=0}^{p} e^{-i\ell h|k|_2^2}\Big| \leq (p+1)h|k|_2^2 \leq T|k|_2^2. \tag{3.14}$$

Therefore, for all $p \in \{0, \ldots, N-1\}$ and $u \in H^{r+2}$,

$$\| - hA\sum_{\ell=0}^{p} e^{-\ell hA}u\|_{H^r} \leq T\|u\|_{H^{r+2}}. \tag{3.15}$$

This allows us to bound the H^r-norm of $e_n^{(1)}$. Together with estimation (3.13), this ensures that inequality (3.12) holds true. ∎

Note that the order constant in the previous result depends on T. This is also the case when the exponential Runge-Kutta method (3.5) is applied to a linear parabolic problem (see [HO05b], Theorem 2). In the case of linear Schrödinger problems of type (3.4), if the underlying Runge-Kutta method is of order $s+1$, it is possible to get an order $s+1$ constant that does not depend on T, with a restriction on the values of the time step. The latter has to be non-resonant in the following sense:

Lemma 3.2.2 *Assume the following non-resonance condition on the time step $h > 0$: There exists $\gamma > 0$ and $\nu > 1$ such that*

$$\forall N \in \mathbb{N}^*, \qquad \left|\frac{1 - e^{ihN}}{h}\right| \geq \frac{\gamma}{N^\nu}. \tag{3.16}$$

Then we have for all $n \in \{0, \ldots, N\}$, for all $r \geq 0$ and all $u \in H^{r+2\nu+2}(\mathbb{T}^d)$,

$$\| - hA\sum_{p=0}^{n-1} e^{-phA}u\|_{H^r(\mathbb{T}^d)} \leq \frac{2}{\gamma}\|u\|_{H^{r+2\nu+2}(\mathbb{T}^d)}. \tag{3.17}$$

Proof. With the help of (3.16), we get for $k \in \mathbb{Z}^d \setminus \{0\}$,

$$\Big| - ih|k|_2^2 \sum_{p=0}^{n-1} e^{-iph|k|_2^2}\Big| \leq \Big|\frac{2h|k|_2^2}{1 - e^{-ih|k|_2^2}}\Big| \leq \frac{2}{\gamma}|k|_2^{2\nu+2}.$$

Therefore, estimation (3.17) holds true. ∎

Note that the set of time steps $h \in (0, h_0)$ that do not satisfy (3.16) has a Lebesgue measure in $o(h_0)$ as h_0 tends to 0. See for example [HLW06], Chapter 10, Lemma 6.3. This non-resonance condition has been used in [DF07] to prove results on splitting methods applied to the linear Schrödinger equation in the asymptotics of small analytic potentials.

In addition to the restriction on the values of the time step, a higher spatial regularity of the right-hand side f of (3.4) is assumed to proved the following

Theorem 3.2.3 *Assume that $r \geq 0$, $\gamma > 0$ and $\nu > 1$ are given. Assume that the underlying Runge-Kutta method is of order $s+1$ and that $f \in C^{s+1}([0,T], L^2(\mathbb{T}^d))$ is such that $f^{(s)} \in C^1([0,T], H^{r+2\nu+2})$. Then the exponential Runge-Kutta method of collocation type* (3.5) *applied to the linear problem* (3.4) *is of order $s+1$ for non-resonant time steps in the sense that there exists a positive constant C depending only on $(c_1, \dots, c_s)$ and γ such that if $h > 0$ satisfies* (3.16)*, then we have for all $n \in \{0, \dots, N\}$,*

$$\|u(t_n) - u_n\|_{H^r} \leq C\, h^{s+1}\Big(\|f^{(s)}(0)\|_{H^{r+2\nu+2}} + \int_0^{t_n} \|f^{(s+1)}(\tau)\|_{H^{r+2\nu+2}}\, \mathrm{d}\tau\Big). \tag{3.18}$$

Proof. With the help of Lemma 3.2.2, estimation (3.15) can be replaced by estimation (3.17) in the proof of Theorem 3.2.2 to get

$$\|e_n^{(1)}\|_{H^r} \leq 2\frac{C_2}{\gamma} h^{s+1}\Big(\|f^{(s)}(t_0)\|_{H^{r+2\nu+2}} + \int_{t_0}^{t_{n-1}} \|f^{(s+1)}(\tau)\|_{H^{r+2\nu+2}}\, \mathrm{d}\tau\Big). \tag{3.19}$$

In addition to (3.13), we derive that (3.18) holds true. ∎

The effect of resonant time steps on the global error over the finite time inverval $[0,T]$ is illustrated by numerical experiments in Section 3.4.

3.2.2.3 About order $s+2$

As in the case of exponential Runge-Kutta methods applied to linear parabolic problems studied in [HO05b] (see Theorem 3), if the underlying Runge-Kutta method is of order $s+2$, and one has sufficient spatial regularity in the right-hand side of (3.4), then we may expect the exponential Runge-Kutta method (3.5) applied to the linear problem (3.4) to be of order $s+2$. Before giving precise statements, let us recall the (see [HO05b], Lemma 3)

Lemma 3.2.3 *If the underlying Runge-Kutta method is of order $s+2$, then*

$$\sum_{i=1}^{s} c_i^{s+1} b_i(0) = \frac{1}{s+2}, \tag{3.20}$$

$$\sum_{i=1}^{s} c_i^{s} b_i'(0) = \frac{1}{(s+1)(s+2)}. \tag{3.21}$$

We are now able to state the following

Theorem 3.2.4 *Assume $r \geq 0$ is given, the underlying Runge-Kutta method is of order $s+2$ and $f \in C^{s+2}([0,T], L^2(\mathbb{T}^d))$ is such that $f^{(s)} \in C^2([0,T], H^{r+4})$. Then the exponential Runge-Kutta method of collocation type* (3.5) *applied to the linear Schrödinger equation* (3.4) *is of order $s+2$ in the sense that there exists a positive constant C depending only on $(c_1, \dots, c_s)$ such that for all $h > 0$ and all $n \in \{0, \dots, N\}$, we have*

$$\|u(t_n) - u_n\|_{H^r} \leq C\,T\,h^{s+2}\Big(\|f^{(s)}(0)\|_{H^{r+4}} + \|f^{(s+1)}(0)\|_{H^{r+2}} + \int_0^{t_n} \|f^{(s+1)}(\tau)\|_{H^{r+4}}\,\mathrm{d}\tau + \int_0^{t_n} \|f^{(s+2)}(\tau)\|_{H^{r+2}}\,\mathrm{d}\tau\Big).$$

Proof. With the notations of the previous theorem, we write that $\delta_{n+1}^{(2)} = \delta_{n+1}^{(2,1)} + \delta_{n+1}^{(2,2)}$ with

$$\delta_{n+1}^{(2,1)} = h^{s+2}\Big(\varphi_{s+2}(-hA) - \frac{1}{(s+1)!}\sum_{i=1}^{s} c_i^{s+1} b_i(-hA)\Big) f^{(s+1)}(t_n),$$

and

$$\delta_{n+1}^{(2,2)} = \int_0^h e^{-(h-\tau)A} \int_0^\tau \frac{(\tau-\sigma)^{s+1}}{(s+1)!} f^{(s+2)}(t_n+\sigma)\mathrm{d}\sigma\mathrm{d}\tau - \sum_{i=1}^{s}\int_0^h e^{-(h-\tau)A} l_i(\tau)\mathrm{d}\tau \int_0^{c_i h} \frac{(c_i h-\sigma)^{s+1}}{(s+1)!} f^{(s+2)}(t_n+\sigma)\mathrm{d}\sigma$$

We deduce that there exists $C > 0$ depending only on $(c_1, \dots, c_s)$ such that

$$\|\delta_{n+1}^{(2,2)}\|_{H^r} \leq C\,h^{s+2}\int_{t_n}^{t_{n+1}} \|f^{(s+2)}(\tau)\|_{H^r}\,\mathrm{d}\tau.$$

Therefore,

$$\begin{aligned}\Big\|\sum_{p=0}^{n-1} e^{-phA}\delta_{n-p}^{(2,2)}\Big\|_{H^r} &\leq C\,h^{s+2}\sum_{p=0}^{n-1}\int_{t_{n-p-1}}^{t_{n-p}} \|f^{(s+2)}(\tau)\|_{H^r}\,\mathrm{d}\tau \\ &\leq C\,h^{s+2}\int_{t_0}^{t_n} \|f^{(s+2)}(\tau)\|_{H^r}\,\mathrm{d}\tau. \qquad (3.22)\end{aligned}$$

Since the underlying Runge-Kutta method is of order $s+2$, Lemma 3.2.3 ensures that $\sum_{i=1}^{s} c_i^{s+1} b_i(0) = \frac{1}{s+2}$ (see (3.20)). Moreover, $\varphi_{s+2}(0) = ((s+2)!)^{-1}$. Therefore, there exists an holomorphic function ζ_{s+2} on $\mathbb{C}$ such that

$$\forall z \in \mathbb{C}, \qquad \frac{1}{(s+1)!}\sum_{i=1}^{s} c_i^{s+1} b_i(z) - \varphi_{s+2}(z) = z\zeta_{s+2}(z).$$

Moreover, ζ_{s+2} is bounded on $i\mathbb{R}$. Hence, we have

$$\delta_{n+1}^{(2,1)} = -hAh^{s+2}\zeta_{s+2}(-hA) f^{(s+1)}(t_n),$$

and by an Abel summation process,

$$\sum_{p=0}^{n-1} e^{-phA}\delta_{n-p}^{(2,1)} = h^{s+2}\zeta_{s+2}(-hA)\Big(-hA\big(\sum_{p=0}^{n-1} e^{-phA}\big) f^{(s+1)}(0) + \sum_{p=0}^{n-2} -hA\big(\sum_{k=0}^{p} e^{-khA}\big)\big(f^{(s+1)}(t_{n-p-2}) - f^{(s+1)}(t_{n-p-1})\big)\Big).$$

We derive with estimation (3.14) that there exists a positive constant C_1 depending only on $(c_1, \dots, c_s)$ such that

$$\Big\| \sum_{p=0}^{n-1} e^{-phA}\delta_{n-p}^{(2,1)}\Big\|_{H^r} \leq C_1 T\, h^{s+2}\Big(\|f^{(s+1)}(0)\|_{H^{r+2}} + \int_{t_0}^{t_{n-1}} \|f^{(s+2)}(\tau)\|_{H^{r+2}}\,\mathrm{d}\tau\Big). \tag{3.23}$$

The last point is that, since the underlying Runge-Kutta method is of order $s+2$, we get by Lemmas 3.2.1 and 3.2.3 that $\sum_{i=1}^{s} c_i^s b_i(0) = \frac{1}{s+1}$ and $\sum_{i=1}^{s} c_i^s b_i'(0) = \frac{1}{(s+1)(s+2)}$ (see relation (3.21)). Therefore, there exists an holomorphic function κ_{s+2} on $\mathbb{C}$ such that $\varphi_{s+1}(z) - \frac{1}{s!}\sum_{i=1}^{s} c_i^s b_i(z) = z^2\kappa_{s+2}(z)$. Moreover, κ_{s+2} is bounded on $i\mathbb{R}$. As before, we get

$$\sum_{p=0}^{n-1} e^{-phA}\delta_{n-p}^{(1)} = -hh^{s+1}\kappa_{s+2}(-hA)\Big(-hA\big(\sum_{p=0}^{n-1} e^{-phA}\big) Af^{(s)}(0) + \sum_{p=0}^{n-2} -hA\big(\sum_{k=0}^{p} e^{-khA}\big) A\big(f^{(s)}(t_{n-p-2}) - f^{(s)}(t_{n-p-1})\big)\Big).$$

Therefore, using inequality (3.14), there exists a constant positive C_2 depending only on $(c_1, \dots, c_2)$ such that

$$\Big\| \sum_{p=0}^{n-1} e^{-phA}\delta_{n-p}^{(1)}\Big\|_{H^r} \leq C_2 T\, h^{s+2}\Big(\|f^{(s)}(0)\|_{H^{r+4}} + \int_{t_0}^{t_{n-1}} \|f^{(s+1)}(\tau)\|_{H^{r+4}}\,\mathrm{d}\tau\Big). \tag{3.24}$$

The conclusion follows by adding (3.22), (3.23) and (3.24) together. ■

In the previous theorem, the order constant depends on T. As in the previous section (see Theorem 3.2.3), if the underlying Runge-Kutta method is of order $s+2$, then the exponential Runge-Kutta method (3.5) applied to the linear Schrödinger problem (3.4) is of order $s+2$ with an order constant independant of T provided the time step is non-resonant and the right-hand side f has higher spatial regularity. A precise statement is the following

Theorem 3.2.5 *Assume that $r \geq 0$, $\gamma > 0$ and $\nu > 1$ are given as in Theorem 3.2.3. Assume that the underlying Runge-Kutta method is of order $s+2$ and that*

$f \in C^{s+2}([0,T], L^2(\mathbb{T}^d))$ is such that $f^{(s)} \in C^2([0,T], H^{r+2\nu+4}(\mathbb{T}^d))$. Then the exponential Runge-Kutta method of collocation type (3.5) *applied to the linear Schrödinger equation* (3.4) *is of order $s+2$ for non-resonant time steps in the sense that there exists a positive constant C depending only on $(c_1,\dots,c_s)$ and γ such that if $h>0$ satisfies* (3.16)*, then we have for all $n \in \{0,\dots,N\}$,*

$$\begin{aligned}\|u(t_n)-u_n\|_{H^r} \le C\,h^{s+2}\Big(&\|f^{(s)}(0)\|_{H^{r+2\nu+4}} + \|f^{(s+1)}(0)\|_{H^{r+2\nu+2}} \\ &+ \int_0^{t_n} \|f^{(s+1)}(\tau)\|_{H^{r+2\nu+4}}\,\mathrm{d}\tau + \int_0^{t_n} \|f^{(s+2)}(\tau)\|_{H^{r+2\nu+2}}\,\mathrm{d}\tau\Big).\end{aligned}$$

Proof. With the help of Lemma 3.2.2, estimation (3.23) can be replaced by the following estimation

$$\begin{aligned}&\Big\|\sum_{p=0}^{n-1} e^{-phA}\delta_{n-p}^{(2,1)}\Big\|_{H^r} \\ &\le \frac{C_1}{\gamma}h^{s+2}\Big(\|f^{(s+1)}(0)\|_{H^{r+2\nu+2}} + \int_{t_0}^{t_{n-1}} \|f^{(s+2)}(\tau)\|_{H^{r+2\nu+2}}\,\mathrm{d}\tau\Big), \end{aligned} \tag{3.25}$$

where $C_1>0$ depends only on $(c_1,\dots,c_n)$. The same Lemma yields the existence of a positive constant C_2 depending only on $(c_1,\dots,c_s)$ such that

$$\begin{aligned}&\Big\|\sum_{p=0}^{n-1} e^{-phA}\delta_{n-p}^{(1)}\Big\|_{H^r} \\ &\le \frac{C_2}{\gamma}\,h^{s+2}\Big(\|f^{(s)}(0)\|_{H^{r+2\nu+4}} + \int_{t_0}^{t_{n-1}} \|f^{(s+1)}(\tau)\|_{H^{r+2\nu+4}}\,\mathrm{d}\tau\Big). \end{aligned} \tag{3.26}$$

This estimation is to replace inequality (3.24). As in the proof of the previous theorem, the conclusion follows by adding (3.22), (3.25) and (3.26) together. ∎

3.3 Semi-linear problems

3.3.1 Notations

We consider the following semi-linear Cauchy problem

$$\begin{array}{rcll}\partial_t u(t,x) + A\,u(t,x) &=& g(t,u(t,x)) & (t,x)\in[0,T]\times\mathbb{T}^d \\ u(0,x) &=& u_0(x) & x\in\mathbb{T}^d\end{array} \tag{3.27}$$

where $T>0$, $A=-i\Delta$, $r\ge 0$, $u_0 \in H^r(\mathbb{T}^d)$ and g are given. In the following and in many cases of application, for all $t\in[0,T]$, $g(t)$ comes from a non-linear function from $\mathbb{C}$ to $\mathbb{C}$. For example, for the so-called cubic non-linear Schrödinger equation, $g(t,u) = g(t)(u) = |u|^2u$. In order to solve numerically the problem (3.27), we consider the following numerical methods, also called *exponential Runge-Kutta methods of collocation type*; we refer to [HO05b] for a derivation of such methods for semi-linear problems based on the variation-of-constant formula. We assume $s \in \mathbb{N}^*$, $(c_1, \dots, c_s) \in [0,1]^s$, $h > 0$ are given and set $N = \lfloor T/h \rfloor$. We also assume that if $(i,j) \in \{1, \dots, s\}^2$ are such that $i \neq j$, then $c_i \neq c_j$. We denote for all $n \in \{0, \dots, N\}$, $t_n = nh$. For such n, if $u_n \in H^r(\mathbb{T}^d)$ is a given approximation of the exact solution $u(t_n)$ of the problem (3.27), then we build an approximation u_{n+1} of $u(t_{n+1})$ by first solving the non-linear system consisting of the s following equations

$$u_{n,i} = e^{-c_i hA} u_n + h \sum_{j=1}^{s} a_{i,j}(-hA) g(t_n + c_j h, u_{n,j}), \qquad i \in \{1, \dots, s\}. \tag{3.28}$$

The quantities $u_{n,1}, \dots, u_{n,s} \in H^r(\mathbb{T}^d)$ are s unknown functions and the coefficients $(a_{i,j}(-hA))_{(i,j) \in \{1,\dots,s\}^2}$ are s^2 operators defined by

$$a_{i,j}(-hA) = \frac{1}{h} \int_0^{c_i h} e^{-(c_i h - \tau)A} l_j(\tau) \mathrm{d}\tau.$$

We recall that the $(l_j)_{j \in \{1,\dots,s\}}$ are the Lagrange polynomials defined in (3.7). Then, we define

$$u_{n+1} = e^{-hA} u_n + h \sum_{i=1}^{s} b_i(-hA) g(t_n + c_i h, u_{n,i}), \tag{3.29}$$

where the coefficients $(b_i)_{i \in \{1,\dots,s\}}$ are the ones defined in (3.6). Note that for all $(i,j) \in \{1, \dots, s\}^2$, $a_{i,j}(-hA) \in \mathrm{span}(\varphi_0(c_i hA), \dots, \varphi_{s-1}(c_i hA))$ and the coefficients in each linear combination depend only on the choice of the points $(c_1, \dots, c_s)$. Hence, each coefficient $a_{i,j}(z)$ is bounded on $i\mathbb{R}$ by a constant depending only on the choice of $(c_1, \dots, c_s)$. In the following, $\mathcal{C}$ denotes a bound of the functions $(b_i(z))_{i \in \{1,\dots,s\}}$ and $(a_{i,j}(z))_{(i,j) \in \{1,\dots,s\}^2}$ on $i\mathbb{R}$ depending only on the choice of $(c_1, \dots, c_s)$.

In the case $A = 0$, the non-linear Cauchy problem (3.27) reduces to the collection of non-linear ODEs for $x \in \mathbb{T}^d$

$$\frac{\mathrm{d}u}{\mathrm{d}t}(t,x) = g(t, u(t,x)), \qquad t \in [0,T],$$

with initial values $u(0,x) = u_0(x)$. Moreover, the exponential Runge-Kutta method of collocation type (3.28)-(3.29) reduces to the classical Runge-Kutta method called *underlying Runge-Kutta method* whose coefficients are $(a_{i,j}(0))_{(i,j) \in \{1,\dots,s\}^2}$ and $(b_i(0))_{i \in \{1,\dots,s\}}$, since in that case for all $(i,j) \in \{1, \dots, s\}^2$,

$$a_{i,j}(-hA) = a_{i,j}(0)\mathrm{Id} \qquad \text{and} \qquad b_i(-hA) = b_i(0)\mathrm{Id}.$$

For $t \in [0,T]$, we denote $f(t) = g(t, u(t))$ the right-hand side of (3.27), where u denotes the exact solution of the problem.

The space $H^r(\mathbb{T}^d)^s$ is a Hilbert space when endowed with the norm

$$\|(u_1, \dots, u_s)\|_{r,s} = \Big(\sum_{i=1}^{s} \|u_i\|^2_{H^r(\mathbb{T}^d)} \Big)^{1/2}. \tag{3.30}$$

For $R \geq 0$, we denote by $B(0,R)$ the set of $\boldsymbol{u} \in H^r(\mathbb{T}^d)^s$ such that $\|\boldsymbol{u}\|_{r,s} \leq R$.

3.3.2 Existence and uniqueness of the numerical solution

The section is devoted to the proof of existence and uniqueness of the numerical solution $(u_0, \ldots, u_N)$ provided by the exponential Runge-Kutta method (3.28)-(3.29) when solving the semi-linear problem (3.27) for h sufficiently small and under some suitable assumptions on the exact solution and the right-hand side of (3.27). We will also give a first bound for the numerical error.

Our assumptions on the exact solution u of (3.27) and the non-linearity g are the following:

Hypothesis 3.3.1 *The function g satisfies:*

$$\forall t \in [0, T], \qquad g(t, 0) = 0.$$

Hypothesis 3.3.2 *The non-linearity g in the right-hand side of* (3.27) *is Lipschitz-continuous in the sense that there exists $L > 0$ and $\rho > 0$ such that for all $t \in [0, T]$ and for all $u, v \in H^r(\mathbb{T}^d)$ satisfying $\|u\|_{H^r(\mathbb{T}^d)} \leq \rho$ and $\|v\|_{H^r(\mathbb{T}^d)} \leq \rho$, we have*

$$\|g(t, u) - g(t, v)\|_{H^r(\mathbb{T}^d)} \leq L \|u - v\|_{H^r(\mathbb{T}^d)}. \tag{3.31}$$

Hypothesis 3.3.3 *Equation* (3.27) *admits an exact solution $u : [0, T] \to H^r(\mathbb{T}^d)$ that is sufficiently smooth. In particular, there exists $R > 0$ such that for all $t \in [0, T]$, $\|u(t)\|_{H^r(\mathbb{T}^d)} \leq R$.*

Hypothesis 3.3.4 *The mapping $f : [0, T] \to H^r(\mathbb{T}^d)$ is sufficiently smooth.*

In applications, Hypothesis 3.3.4 will often be a consequence of Hypothesis 3.3.3 and of the regularity of the function g.

Note that these assumptions are fulfilled in many cases of interest, for example in the case of the cubic non-linear Schrödinger equation, at least when $r > d/2$ and $r - d/2 \notin \mathbb{N}$. This can be seen, for example, by adapting classical results of non-linear analysis (see for example [AG91], Chapter II, Proposition 2.2). Precise statements of the adaptations can be found in Appendix A.

Under these assumptions, if h and R are sufficiently small, then for all $n \in \{0, \ldots, N\}$ the non-linear system (3.28) has a unique solution $(u_{n,1}, \ldots, u_{n,s}) \in H^r(\mathbb{T})^s$ such that $\|(u_{n,1}, \ldots, u_{n,s})\|_{r,s} \leq 3R$. Moreover, the approximation u_{n+1} defined by (3.29) is such that $\|u_{n+1}\|_{H^s(\mathbb{T}^d)} \leq 2R$. To prove these assertions, we start with a lemma.

Lemma 3.3.1 *Assume that $h_0 > 0$ is such that*

$$h_0 \leq (3\mathcal{C}Ls^2)^{-1} \tag{3.32}$$

$$\text{and} \qquad \rho \geq 3R\sqrt{s}. \tag{3.33}$$

Then for all $h \in (0, h_0)$, all $t \in [0, T-h]$ and all $y \in H^r(\mathbb{T}^d)$ such that $\|y\|_{H^r(\mathbb{T}^d)} \leq 2R$, the non-linear system

$$\begin{array}{rcl} v_1 & = & e^{-c_1 hA} y + h \sum_{j=1}^{s} a_{1,j}(-hA) g(t + c_j h, v_j) \\ \vdots & & \vdots \\ v_s & = & e^{-c_s hA} y + h \sum_{j=1}^{s} a_{s,j}(-hA) g(t + c_j h, v_j) \end{array}$$

admits a unique solution $\boldsymbol{v} = (v_1, \ldots, v_s) \in B(0, 3R\sqrt{s})$.

Proof. Let us define for $h \in (0, h_0)$, $t \in [0, T-h]$, $y \in H^r(\mathbb{T}^d)$ such that $\|y\|_{H^r(\mathbb{T}^d)} \leq 2R$

$$f_{h,t,y} : \left(\begin{array}{ccc} H^r(\mathbb{T}^d)^s & \longrightarrow & H^r(\mathbb{T}^d)^s \\ (v_i)_{1 \leq i \leq s} & \longmapsto & \left(e^{-c_i hA} y + h \sum_{j=1}^{s} a_{i,j}(-hA) g(t + c_j h, v_j) \right)_{1 \leq i \leq s} \end{array} \right).$$

Hypothesis 3.3.2 and inequality (3.33) ensure that for all $(v_1, \ldots, v_s) \in B(0, 3R\sqrt{s})$, we have

$$\|f_{t,h,y}(v_1, \ldots, v_s) - (e^{-c_1 hA} y, \ldots, e^{-c_s hA} y)\|_{r,s} \leq C h s^{3/2} L(3R\sqrt{s}).$$

Since $\|(e^{-c_1 hA} y, \ldots, e^{-c_s hA} y)\|_{r,s} \leq 2R\sqrt{s}$, the triangle inequality and (3.32) imply that

$$\|f_{t,h,y}(v_1, \ldots, v_s)\|_{r,s} \leq 3R\sqrt{s}.$$

Therefore, $f_{t,h,y}(B(0, 3R\sqrt{s})) \subset B(0, 3R\sqrt{s})$.
On the other hand, for $\boldsymbol{v} = (v_1, \ldots, v_s) \in B(0, 3R\sqrt{s})$ and $\boldsymbol{w} = (w_1, \ldots, w_s) \in B(0, 3R\sqrt{s})$, we have with Hypothesis 3.3.2,

$$\|f_{t,h,y}(\boldsymbol{v}) - f_{t,h,y}(\boldsymbol{w})\|_{r,s} \leq C L h s \|\boldsymbol{v} - \boldsymbol{w}\|_{r,s}.$$

Hence, with (3.32), we derive that

$$\|f_{t,h,y}(\boldsymbol{v}) - f_{t,h,y}(\boldsymbol{w})\|_{r,s} \leq \frac{h}{h_0} \|\boldsymbol{v} - \boldsymbol{w}\|_{r,s}$$

Since $h h_0^{-1} < 1$, equation $\boldsymbol{x} = f_{t,h,y}(\boldsymbol{x})$ exactly has one solution in $B(0, 3R\sqrt{s})$ by the classical Picard fixed-point theorem. ■

We are now able to prove that if h_0 satisfies another smallness condition, then, for all $h \in (0, h_0)$, the exponential Runge-Kutta method of collocation type (3.28)-(3.29) provides a unique numerical solution with for all $n \in \{0, \ldots, N\}$, $\|(u_{n,1}, \ldots, u_{n,s})\|_{r,s} \in B(0, 3R\sqrt{s})$ and $\|u_n\|_{H^r(\mathbb{T}^d)} \leq 2R$. Moreover, such an s-stage method is of order s when applied to the semi-linear problem (3.27).

Theorem 3.3.1 *Assume Hypothesis 3.3.1, 3.3.2, 3.3.3 and 3.3.4 are satisfied with* (3.33). *Assume that $h_0 > 0$ satisfies* (3.32), *and*

$$h_0^s \leq R\Big(\kappa_s e^{2\mathcal{C}sTL} \int_0^T \|f^{(s)}(\tau)\|_{H^r(\mathbb{T}^d)} \mathrm{d}\tau\Big)^{-1}. \tag{3.34}$$

Then, for all $h \in (0, h_0)$, the exponential Runge-Kutta method (3.28)-(3.29) *provides a unique numerical solution such that for all $n \in \{0, \dots, N\}$, $\|u_n\|_{H^s(\mathbb{T}^d)} \leq 2R$ and $\|(u_{n,1}, \dots, u_{n,s})\|_{r,s} \leq 3R$. Moreover, for all $n \in \{0, \dots, N\}$,*

$$\|u_n - u(t_n)\|_{H^r(\mathbb{T}^d)} \leq \kappa_s e^{2\mathcal{C}sLT}\, h^s \int_0^{t_n} \|f^{(s)}(\tau)\|_{H^r(\mathbb{T}^d)} \mathrm{d}\tau \tag{3.35}$$

where $\kappa_s = (s+1)\Big(\frac{1}{s!} + \frac{\mathcal{C}s}{(s-1)!}\Big)$.

Proof. Let $h \in (0, h_0)$. We prove the result by induction. Assume $n \in \{0, \dots, N-1\}$ is such that $\|u_n\|_{H^r(\mathbb{T}^d)} \leq 2R$. Lemma 3.3.1 ensures that the non-linear system (3.28) has a unique solution $(u_{n,1}, \dots, u_{n,s}) \in B(0, 3R\sqrt{s})$. Moreover, Hypothesis 3.3.3 and 3.3.4 ensure that, for the exact solution, we have

$$u(t_{n+1}) = e^{-hA} u(t_n) + h \sum_{i=1}^s b_i(-hA) f(t_n + c_i h) + \delta_{n+1} \tag{3.36}$$

with

$$\begin{aligned}\delta_{n+1} = &\int_0^h e^{-(h-\tau)A} \int_0^\tau \frac{(\tau-\sigma)^{s-1}}{(s-1)!} f^{(s)}(t_n + \sigma) \mathrm{d}\sigma \mathrm{d}\tau \\ &- h \sum_{i=1}^s b_i(-hA) \int_0^{c_i h} \frac{(c_i h - \tau)^{s-1}}{(s-1)!} f^{(s)}(t_n + \tau) \mathrm{d}\tau.\end{aligned}$$

Similarly, for all $i \in \{1, \dots, s\}$, we have

$$u(t_n + c_i h) = e^{-c_i hA} u(t_n) + h \sum_{j=1}^s a_{i,j}(-hA) f(t_n + c_j h) + \Delta_{n,i} \tag{3.37}$$

where

$$\begin{aligned}\Delta_{n,i} = &\int_0^{c_i h} e^{-(c_i h-\tau)A} \int_0^\tau \frac{(\tau-\sigma)^{s-1}}{(s-1)!} f^{(s)}(t_n + \sigma) \mathrm{d}\sigma \mathrm{d}\tau \\ &- h \sum_{j=1}^s a_{i,j}(-hA) \int_0^{c_j h} \frac{(c_j h - \tau)^{s-1}}{(s-1)!} f^{(s)}(t_n + \tau) \mathrm{d}\tau.\end{aligned} \tag{3.38}$$

Hence, if we denote $e_n = u_n - u(t_n)$ and $e_{n,i} = u_{n,i} - u(t_n + c_i h)$, then we have

$$e_{n,i} = e^{-c_i hA} e_n + h \sum_{j=1}^s a_{i,j}(-hA) \big(g(t_n + c_j h, u_{n,j}) - f(t_n + c_j h)\big) - \Delta_{n,i}. \tag{3.39}$$

Therefore, for all $i \in \{1, \ldots, s\}$,

$$\|e_{n,i}\|_{H^r(\mathbb{T}^d)} \leq \|e_n\|_{H^r(\mathbb{T}^d)} + h\mathcal{C}L\Big(\sum_{j=1}^{s} \|e_{n,j}\|_{H^r(\mathbb{T}^d)}\Big) + \|\Delta_{n,i}\|_{H^r(\mathbb{T}^d)}. \tag{3.40}$$

Summing these s inequalities, we get with (3.32) that

$$\sum_{j=1}^{s} \|e_{n,j}\|_{H^r(\mathbb{T}^d)} \leq 2s\|e_n\|_{H^r(\mathbb{T}^d)} + 2\sum_{i=1}^{s} \|\Delta_{n,i}\|_{H^r(\mathbb{T}^d)}. \tag{3.41}$$

On the other hand, substracting (3.36) from (3.29) yields

$$e_{n+1} = e^{-hA}e_n + h\sum_{i=1}^{s} b_i(-hA)\Big(g(t_n + c_i h, u_{n,i}) - f(t_n + c_i h)\Big) - \delta_{n+1}. \tag{3.42}$$

Hence,

$$\|e_{n+1}\|_{H^r(\mathbb{T}^d)} \leq \|e_n\|_{H^r(\mathbb{T}^d)} + h\mathcal{C}L\sum_{i=1}^{s} \|e_{n,i}\|_{H^r(\mathbb{T}^d)} + \|\delta_{n+1}\|_{H^r(\mathbb{T}^d)}. \tag{3.43}$$

With (3.41), we thus get

$$\begin{aligned}
\|e_{n+1}\|_{H^r(\mathbb{T}^d)} &\leq \|e_n\|_{H^r(\mathbb{T}^d)} + 2h\mathcal{C}L\Big(s\|e_n\|_{H^r(\mathbb{T}^d)} + \sum_{j=1}^{s} \|\Delta_{n,j}\|_{H^r(\mathbb{T}^d)}\Big) + \|\delta_{n+1}\|_{H^r(\mathbb{T}^d)} \\
&\leq (1 + 2h\mathcal{C}sL)\|e_n\|_{H^r(\mathbb{T}^d)} + \sum_{j=1}^{s} \|\Delta_{n,j}\|_{H^r(\mathbb{T}^d)} + \|\delta_{n+1}\|_{H^r(\mathbb{T}^d)} \\
&\leq \sum_{p=0}^{n} (1 + 2h\mathcal{C}sL)^p \Big(\sum_{j=1}^{s} \|\Delta_{n-p,j}\|_{H^r(\mathbb{T}^d)} + \|\delta_{n+1-p}\|_{H^r(\mathbb{T}^d)}\Big) \\
&\leq e^{2Nh\mathcal{C}sL} \sum_{p=0}^{n} \Big(\sum_{j=1}^{s} \|\Delta_{n-p,j}\|_{H^r(\mathbb{T}^d)} + \|\delta_{n+1-p}\|_{H^r(\mathbb{T}^d)}\Big),
\end{aligned}$$

using also (3.32). Using relation (3.10) (which still holds for non-linear problems) and relation (3.38), we get that for all $j \in \{1, \ldots, s\}$ and all $p \in \{0, \ldots, n\}$,

$$\|\Delta_{n-p,j}\|_{H^r(\mathbb{T}^d)}, \|\delta_{n+1-p}\|_{H^r(\mathbb{T}^d)} \leq \Big(\frac{1}{s!} + \frac{\mathcal{C}s}{(s-1)!}\Big) h^s \int_{t_{n-p}}^{t_{n-p+1}} \|f^{(s)}(\tau)\|_{H^r(\mathbb{T}^d)} \,\mathrm{d}\tau. \tag{3.44}$$

We derive

$$\|e_{n+1}\|_{H^r(\mathbb{T}^d)} \leq e^{2Nh\mathcal{C}sL}(s+1)\Big(\frac{1}{s!} + \frac{\mathcal{C}s}{(s-1)!}\Big) h^s \int_{t_0}^{t_{n+1}} \|f^{(s)}(\tau)\|_{H^r(\mathbb{T}^d)} \,\mathrm{d}\tau.$$

In view of (3.34), we derive that $\|e_{n+1}\|_{H^r(\mathbb{T}^d)} \leq R$. Therefore,

$$\|u_{n+1}\|_{H^r(\mathbb{T}^d)} \leq \|u(t_{n+1})\|_{H^r(\mathbb{T}^d)} + \|e_{n+1}\|_{H^r(\mathbb{T}^d)} \leq R + R,$$

and estimation (3.35) holds true. ∎

3.3.3 A sufficient condition for order $s + 1$

As in the linear case (see Theorem 3.2.2 and Theorem 3.2.3), the exponential Runge-Kutta method of collocation type (3.28)-(3.29) applied to the non-linear Schrödinger equation (3.27) is of order $s + 1$ provided the underlying Runge-Kutta method is of order $s + 1$. A precise result is stated in Theorem 3.3.2. Note that this result is very similar to the one obtained in [HO05b] in the case of semi-linear parabolic problems (see Theorem 5).

Lemma 3.3.2 *Let X be a subset of $\mathbb{R}^{+1}$. Assume that $T > 0$ and $a \geq 0$ are given. There exists a positive constant C such that for all non-negative function b defined on X and all sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ of non-negative real numbers satisfying for all $h \in X$ and $n \in \mathbb{N}$ such that $0 \leq nh \leq T$, $\varepsilon_n \leq ah\sum_{k=0}^{n-1}\varepsilon_k + b(h)$, we have*

$$\varepsilon_n \leq Cb(h).$$

Proof. Let $h \in X$ be given. By induction, we have for all such function b, all such sequence $(\varepsilon_n)_{n\in\mathbb{N}}$, for all $n \in \mathbb{N}^*$ such that $nh \leq T$,

$$\varepsilon_n \leq [(1 + ah)^n - 1]\varepsilon_0 + (1 + ah)^{n-1}b(h).$$

Since $\varepsilon_0 \leq b(h)$, we derive that

$$\varepsilon_n \leq (2e^{ahn} - 1)b(h).$$

This proves the result with $C = 2e^{aT} - 1$. ■

Theorem 3.3.2 *Assume the underlying Runge-Kutta method is of order $s + 1$. Under the hypothesis of Theorem 3.3.1, if $f^{(s)} \in C^1([0, T], H^{r+2})$, then there exists a positive constant $C > 0$ depending only on $(c_1, \ldots, c_s)$, L, and T such that for all $h \in (0, h_0)$, we have for all $n \in \{0, \ldots, N\}$,*

$$\begin{aligned}\|u_n - u(t_n)\|_{H^r} \leq C\, h^{s+1} \Big(&\|f^{(s)}(0)\|_{H^{r+2}} \\ &+ \int_0^T \|f^{(s+1)}(\tau)\|_{H^{r+2}}\, \mathrm{d}\tau + \int_0^T \|f^{(s)}(\tau)\|_{H^r}\, \mathrm{d}\tau\Big).\end{aligned}$$

Proof. By assumption and with the help of Theorem 3.3.1, the numerical solution provided by the exponential Runge-Kutta method of collocation type (3.28)-(3.29) is well defined for all $n \in \{0, \ldots, N\}$. Identities (3.28) and (3.37) lead again to relation (3.39). Therefore, we derive that the estimation (3.40) still holds. Moreover, with the help of (3.32), we derive that estimation (3.41) also still holds. Similarly, identities (3.29) and (3.36) lead to relation (3.42). By induction, this relation leads to

$$e_n = h\left(\sum_{k=0}^{n-1} e^{-khA}d_{n-k-1}\right) - \sum_{k=0}^{n-1} e^{-khA}\delta_{n-k},$$

[1] Namely, X will be $(0, h_0)$ or the set of $h \in (0, h_0)$ satisfying (3.16).

where $d_k = \sum_{i=1}^{s} b_i(-hA)(g(t_k + c_i h, u_{k,i}) - f(t_k + c_i h))$. As in the proof of Theorem 3.2.2, there exists a constant $C_1 > 0$ depending only on $(c_1, \ldots, c_s)$ such that

$$\Big\| \sum_{k=0}^{n-1} e^{-khA} \delta_{n-k} \Big\|_{H^r} \leq C_1 h^{s+1} \Big(\|f^{(s)}(0)\|_{H^{r+2}} + \int_0^{t_{n-1}} \|f^{(s+1)}(\tau)\|_{H^{r+2}} \,\mathrm{d}\tau \Big). \quad (3.45)$$

On the other hand, with Hypothesis 3.3.2, we have

$$\Big\| \sum_{k=0}^{n-1} e^{-khA} d_{n-k-1} \Big\|_{H^r} \leq \mathcal{C}L \sum_{k=0}^{n-1} \sum_{i=1}^{s} \|e_{k,i}\|_{H^r} .$$

Then, with (3.41), we deduce

$$\Big\| \sum_{k=0}^{n-1} e^{-khA} d_{n-k-1} \Big\|_{H^r} \leq 2s\mathcal{C}L \sum_{k=0}^{n-1} \|e_k\|_{H^r} + 2\mathcal{C}L \sum_{k=0}^{n-1} \sum_{i=1}^{s} \|\Delta_{k,i}\|_{H^r} .$$

Estimation (3.44) provides the existence of a constant $C_2 > 0$ depending only on $(c_1, \ldots, c_s)$ and s such that

$$\sum_{k=0}^{n-1} \sum_{i=1}^{s} \|\Delta_{k,i}\|_{H^r} \leq C_2 h^s \int_0^{t_n} \|f^{(s)}(\tau)\|_{H^r} \,\mathrm{d}\tau.$$

Hence,

$$\begin{aligned} \|e_n\|_{H^r} \leq 2sh\mathcal{C}L \sum_{k=0}^{n-1} \|e_k\|_{H^r} + h^{s+1} \Bigg(& 2\mathcal{C}LC_2 \int_0^{T} \|f^{(s)}(\tau)\|_{H^r} \,\mathrm{d}\tau \\ & + 2C_1 \Big(\|f^{(s)}(0)\|_{H^{r+2\nu+2}} + \int_0^{T} \|f^{(s+1)}(\tau)\|_{H^{r+2\nu+2}} \,\mathrm{d}\tau \Big) \Bigg) . \end{aligned}$$

We conclude with the help of Lemma 3.3.2 with $\varepsilon_n = \|e_n\|_{H^r}$, $a = 2s\mathcal{C}L$ and b is the term multiplied by h^{s+1} in the previous estimation. ∎

Remark 3.3.1 *Note that in this non-linear case, the order constant C appearing in Theorem 3.3.2 depends a priori exponentially on T. This is a difference with the linear case (see Theorem 3.2.2 and Theorem 3.2.3).*

Remark 3.3.2 *We could write a counterpart of Theorem 3.3.2 for non-resonant time steps with suitable assumptions on f essentially by modifying inequality* (3.45). *However, we would still get an order constant depending on T (see Theorem 3.2.3 for the corresponding result for linear problems) and the fact is that we did not manage to observe resonances for non-linear problems (see section 3.4 for numerical experiments).*

3.3.4 Achieving order $s + 2$

In order to give sufficient algebraic conditions on the coefficients of the underlying Runge-Kutta method for the s-stage exponential Runge-Kutta method of collocation type (3.28)-(3.29) to be of order $s + 2$, we reinforce our hypothesis on the smoothness of the non-linearity g. Function g is assumed to be smooth in the sense of functions from $\mathbb{R}^2$ to $\mathbb{R}^2$. For convenience, we denote, for example, for $v \in \mathbb{C} = \mathbb{R}^2$, $\frac{\partial g}{\partial u}(t,v)$ the $\mathbb{R}$-linear mapping from $\mathbb{C} = \mathbb{R}^2$ to $\mathbb{C} = \mathbb{R}^2$ corresponding to the first derivative of $g(t,u)$ as a function of $u \in \mathbb{R}^2$. For $v \in H^r(\mathbb{T}^d)$, we also denote by $\frac{\partial g}{\partial u}(t,v)$ the induced mapping between functions on $\mathbb{T}^d$ defined for a function w on $\mathbb{T}^d$ by

$$\forall x \in \mathbb{T}^d, \qquad \frac{\partial g}{\partial u}(t,v)(w)(x) = \frac{\partial g}{\partial u}(t,v(x))(w(x)).$$

Our additional hypothesis on the non-linearity g are the following:

Hypothesis 3.3.5 *Assume that there exists positive constants L_1 and $\tilde{L}_1$ such that*

$$\forall t \in [0,T], \quad \forall v \in H^r(\mathbb{T}^d), \qquad \left\| \frac{\partial g}{\partial u}(t,u(t))v \right\|_{H^r(\mathbb{T}^d)} \leq L_1 \|v\|_{H^r(\mathbb{T}^d)}.$$

and

$$\forall t \in [0,T], \quad \forall v \in H^{r+2}(\mathbb{T}^d), \qquad \left\| \frac{\partial g}{\partial u}(t,u(t))v \right\|_{H^{r+2}(\mathbb{T}^d)} \leq \tilde{L}_1 \|v\|_{H^{r+2}(\mathbb{T}^d)}.$$

Hypothesis 3.3.6 *Assume that there exists a positive constant L_2 such that for all $t \in [0,T]$, for all $u,v,w \in H^r(\mathbb{T}^d)$ satisfying* $\max\{\|u\|_{H^r}, \|v\|_{H^r}, \|w\|_{H^r}\} \leq 2R$, *one has*

$$\left\| g(t,v) - g(t,w) - \frac{\partial g}{\partial u}(t,u)(v-w) \right\|_{H^r} \leq L_2 \Big(\|u-v\|_{H^r} + \|u-w\|_{H^r} \Big) \|v-w\|_{H^r}.$$

Hypothesis 3.3.7 *Assume that there exists positive constants L_3 and $\tilde{L}_3$ such that for all $t \in [0,T]$ and all $v,w \in H^{r+2}(\mathbb{T}^d)$, one has*

$$\left\| \frac{\partial^2 g}{\partial u^2}(t,u(t))(v,w) \right\|_{H^r} \leq L_3 \|v\|_{H^r} \|w\|_{H^r},$$

$$\left\| \frac{\partial^2 g}{\partial t \partial u}(t,u(t))v \right\|_{H^r} \leq L_3 \|v\|_{H^r},$$

$$\left\| \frac{\partial^2 g}{\partial u^2}(t,u(t))(v,w) \right\|_{H^{r+2}} \leq \tilde{L}_3 \|v\|_{H^{r+2}} \|w\|_{H^{r+2}},$$

and

$$\left\| \frac{\partial^2 g}{\partial t \partial u}(t,u(t))v \right\|_{H^{r+2}} \leq \tilde{L}_3 \|v\|_{H^{r+2}}.$$

Such hypothesis on g are coupled with (and may imply) the following one on the exact solution u of the problem (3.27) that reinforces Hypothesis 3.3.3:

Hypothesis 3.3.8 *Equation* (3.27) *admits an exact solution $u : [0,T] \to H^{r+2}(\mathbb{T}^d)$ that is sufficiently smooth. In particular, $u_0 \in H^{r+2}(\mathbb{T}^d)$.*

Notice that Hypothesis 3.3.5, 3.3.6 and 3.3.7 are satisfied in many cases of interest, for example for the cubic Schrödinger equation when $r > d/2$ with $r - d/2 \notin \mathbb{N}$ provided u is sufficiently smooth (see Appendix A for details).

With these hypotheses, we prove that the exponential Runge-Kutta method (3.28)-(3.29) applied to the semi-linear Schrödinger equation (3.27) is of order $s+2$ provided the underlying Runge-Kutta method is of order $s+2$.

Theorem 3.3.3 *Assume that $r \geq 0$ is given. Assume that the underlying Runge-Kutta method is of order $s+2$, and that the hypotheses of Theorem 3.3.1, Hypothesis 3.3.5, Hypothesis 3.3.6, Hypothesis 3.3.7 and Hypothesis 3.3.8 are fulfilled. If $f \in C^{s+2}([0,T], H^{r+4})$, then there exists a positive constant $C > 0$ such that for all $h \in (0, h_0)$ and all $n \in \{0, \ldots, N\}$, one has*

$$\|u(t_n) - u_n\|_{H^r} \leq Ch^{s+2}.$$

Remark 3.3.3 *Of course, the constant C appearing in Theorem 3.3.3 also involves the final time $T > 0$ and integrals of norms of the time-derivatives of f over $[0,T]$.*

Proof. Let us set for $n \in \{0, \ldots, N\}$

$$J_n = \frac{\partial g}{\partial u}(t_n, u(t_n)),$$

and

$$\tilde{d}_{n,i} = g(t_n + c_i h, u_{n,i}) - g(t_n + c_i h, u(t_n + c_i h)) - J_n e_{n,i}.$$

As before, substracting (3.36) from (3.29) yields

$$e_{n+1} = e^{-hA} e_n - \delta_{n+1} + h \sum_{i=1}^{s} b_i(-hA) J_n e_{n,i} + h \sum_{i=1}^{s} b_i(-hA) \tilde{d}_{n,i}.$$

By induction, we derive that for all $n \in \{0, \ldots, N-1\}$,

$$e_{n+1} = -\sum_{p=0}^{n} e^{-phA} \delta_{n+1-p} \tag{3.46}$$

$$+ \quad h \sum_{i=1}^{s} b_i(-hA) \sum_{p=0}^{n} e^{-phA} J_{n-p} \big(e_{n-p,i} + \Delta_{n-p,i}\big) \tag{3.47}$$

$$- \quad h \sum_{i=1}^{s} b_i(-hA) \sum_{p=0}^{n} e^{-phA} J_{n-p} \Delta_{n-p,i} \tag{3.48}$$

$$+ \quad h \sum_{i=1}^{s} b_i(-hA) \sum_{p=0}^{n} e^{-phA} \tilde{d}_{n-p,i}. \tag{3.49}$$

Since the underlying Runge-Kutta method is of order $s+2$, one deduces, as in the proof of Theorem 3.2.4 that there exists a positive constant C_1 such that for all h and n,

$$\Big\| \sum_{p=0}^{n} e^{-phA} \delta_{n+1-p} \Big\|_{H^r} \leq C_1\, h^{s+2} \Big(\|f^{(s)}(0)\|_{H^{r+4}} + \|f^{(s+1)}(0)\|_{H^{r+2}} + \int_0^T \|f^{(s+1)}(\tau)\|_{H^{r+4}}\, \mathrm{d}\tau + \int_0^T \|f^{(s+2)}(\tau)\|_{H^{r+2}} \Big). \tag{3.50}$$

To bound the error term (3.47), one uses (3.39) and then (3.41) to derive that

$$\begin{aligned} \|e_{n-p,i} + \Delta_{n-p,i}\|_{H^r} &\leq \|e_{n-p}\|_{H^r} + \mathcal{C}hL \sum_{j=1}^{s} \|e_{n-p,j}\|_{H^r} \\ &\leq (1 + 2\mathcal{C}h_0 Ls)\|e_{n-p}\|_{H^r} + 2\mathcal{C}Lh \sum_{j=1}^{s} \|\Delta_{n-p,j}\|_{H^r}\,. \end{aligned}$$

With Hypothesis 3.3.5 and estimation (3.44), we derive that[2]

$$\begin{aligned} &\Big\| h \sum_{i=1}^{s} b_i(-hA) \sum_{p=0}^{n} e^{-phA} J_{n-p,i}\big(e_{n-p,i} + \Delta_{n-p,i}\big) \Big\|_{H^r} \\ &\leq \mathcal{C}hs \left(\sum_{p=0}^{n} L_1 \Big((1 + 2\mathcal{C}h_0 Ls)\|e_{n-p}\|_{H^r} + 2\mathcal{C}Lh \sum_{j=1}^{s} \|\Delta_{n-p,j}\|_{H^r} \Big) \right) \\ &\leq \mathcal{C}hsL_1(1 + 2\mathcal{C}h_0 Ls) \sum_{p=0}^{n} \|e_p\|_{H^r} + 2\mathcal{C}^2 LL_1 \kappa_s h^{s+2} \int_0^T \|f^{(s)}(\tau)\|_{H^r}\, \mathrm{d}\tau. \end{aligned}$$

Hence, there exists positive constants C_2 and C_3 such that for all h and n, the H^r-norm of the term (3.47) is bounded by $C_2 h \sum_{p=0}^{n} \|e_p\|_{H^r} + C_3 h^{s+2}$.

In order to bound the H^r-norm of the error term (3.48), we set

$$\psi_{i,s+1}(z) = \varphi_{s+1}(c_i z) c_i^{s+1} - \frac{1}{s\,!} \sum_{j=1}^{s} a_{i,j}(z) c_j^s,$$

and write

$$\Delta_{n-p,i} = h^{s+1} \psi_{i,s+1}(-hA) f^{(s)}(t_{n-p}) + \widetilde{\Delta}_{n-p,i},$$

[2] See definition of κ_s in Theorem 3.3.1

to get

$$
\begin{aligned}
& h\sum_{i=1}^{s} b_i(-hA)\sum_{p=0}^{n} e^{-phA}J_{n-p}\Delta_{n-p,i} = \\
- \quad & h\sum_{i=1}^{s} b_i(0)\sum_{p=0}^{n} e^{-phA}J_{n-p}h^{s+1}\psi_{i,s+1}(0)f^{(s)}(t_{n-p}) && (3.51)\\
- \quad & h\sum_{i=1}^{s} b_i(-hA)\sum_{p=0}^{n} e^{-phA}J_{n-p}h^{s+1}\big(\psi_{i,s+1}(-hA)-\psi_{i,s+1}(0)\big)f^{(s)}(t_{n-p}) && (3.52)\\
- \quad & h\sum_{i=1}^{s} \big(b_i(-hA)-b_i(0)\big)\sum_{p=0}^{n} e^{-phA}J_{n-p}h^{s+1}\psi_{i,s+1}(0)f^{(s)}(t_{n-p}) && (3.53)\\
+ \quad & h\sum_{i=1}^{s} b_i(-hA)\sum_{p=0}^{n} e^{-phA}J_{n-p}\widetilde{\Delta}_{n-p,i}. && (3.54)
\end{aligned}
$$

Since the underlying Runge-Kutta method is of order at least $s+2$, we derive that

$$\sum_{i=1}^{s} b_i(0)\psi_{i,s+1}(0) = 0,$$

and therefore that the term (3.51) reads

$$h^{s+2}\sum_{p=0}^{n} e^{-phA}J_{n-p}\Big(\sum_{i=1}^{s} b_i(0)\psi_{i,s+1}(0)\Big)f^{(s)}(t_{n-p}) = 0.$$

On the other hand, there exists s holomorphic functions $(\Psi_{i,s+1})_{i\in\{1,\dots,s\}}$ such that

$$\forall i\in\{1,\dots,s\},\quad \forall z\in\mathbb{C},\qquad \psi_{i,s+1}(z)-\psi_{i,s+1}(0) = z\Psi_{i,s+1}(z).$$

Note that the functions $(\Psi_{i,s+1})_{i\in\{1,\dots,s\}}$ are bounded on $i\mathbb{R}$. Let $M>0$ be a bound for these functions. For all $i\in\{i,\dots,s\}$, an Abel summation process yields

$$
\begin{aligned}
& \sum_{p=0}^{n} e^{-phA}J_{n-p}\Psi_{i,s+1}(-hA)\big(-hA\big)f^{(s)}(t_{n-p}) = \\
& \qquad \Big(\sum_{p=0}^{n} e^{-phA}\Big)J_0\Psi_{i,s+1}(-hA)\big(-hA\big)f^{(s)}(0) \\
& \quad + \sum_{p=0}^{n-1}\Big(\sum_{q=0}^{p} e^{-qhA}\Big)\int_{t_{n-p-1}}^{t_{n-p}} \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial g}{\partial u}(t,u(t))\Psi_{i,s+1}(-hA)\big(-hA\big)f^{(s)}(t)\Big]\mathrm{d}t.
\end{aligned}
$$

On one hand,

$$\Big\|\Big(\sum_{p=0}^{n} e^{-phA}\Big)J_0\Psi_{i,s+1}(-hA)\big(-hA\big)f^{(s)}(0)\Big\|_{H^r} \le TL_1M\|f^{(s)}(0)\|_{H^{r+2}}.$$

On the other hand, since after differentiating and with Hypothesis 3.3.5 and 3.3.7, one has

$$
\begin{aligned}
&\Big\| \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial g}{\partial u}(t,u(t))\Psi_{i,s+1}(-hA)\big(-hA\big)f^{(s)}(t)\Big]\Big\|_{H^r} \\
&\quad\leq\ hL_3M\|f^{(s)}(t)\|_{H^{r+2}} + hL_3M\widetilde{R}\|f^{(s)}(t)\|_{H^{r+2}} + hL_1M\|f^{(s+1)}(t)\|_{H^{r+2}},
\end{aligned}
$$

where $\widetilde{R}$ denotes the maximum of $\|u'(t)\|_{H^r}$ on $[0,T]$ (see Hypothesis 3.3.3), we deduce that there exists a positive constant C_4 such that for all h and n, the term (3.52) has an H^r-norm bounded by C_4h^{s+2}.

The same kind of calculations yields, denoting $(\theta_i)_{i\in\{1,\dots,s\}}$ the holomorphic functions such that

$$
\forall i\in\{1,\dots,n\},\quad \forall z\in\mathbb{C},\qquad b_i(z)-b_i(0)=z\theta_i(z)=\theta_i(z)z,
$$

$$
\begin{aligned}
&\sum_{i=1}^{s}\big(b_i(-hA)-b_i(0)\big)\sum_{p=0}^{n}e^{-phA}J_{n-p}\psi_{i,s+1}(0)f^{(s)}(t_{n-p}) = \\
&\qquad \psi_{i,s+1}(0)\Big(\sum_{p=0}^{n}e^{-phA}\Big)\theta_i(-hA)\big(-hA\big)J_0f^{(s)}(0) \\
&\quad +\ \psi_{i,s+1}(0)\sum_{p=0}^{n-1}\Big(\sum_{q=0}^{p}e^{-qhA}\Big)\theta_i(-hA)\big(-hA\big)\int_{t_{n-p-1}}^{t_{n-p}}\frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial g}{\partial u}(t,u(t))f^{(s)}(t)\Big]\mathrm{d}t.
\end{aligned}
$$

If we also denote by M a positive constant bounding the functions $(\theta_i)_{i\in\{1,\dots,s\}}$, then

$$
\begin{aligned}
&\Big\|e^{-qhA}\theta_i(-hA)\big(-hA\big)\int_{t_{n-p-1}}^{t_{n-p}}\frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial g}{\partial u}(t,u(t))f^{(s)}(t)\Big]\mathrm{d}t\Big\|_{H^r} \\
&\quad\leq\ Mh\int_{t_{n-p-1}}^{t_{n-p}}\Big\|A\frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial g}{\partial u}(t,u(t))f^{(s)}(t)\Big]\Big\|_{H^r}\mathrm{d}t \\
&\quad\leq\ Mh\int_{t_{n-p-1}}^{t_{n-p}}\Big\|\frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial g}{\partial u}(t,u(t))f^{(s)}(t)\Big]\Big\|_{H^{r+2}}\mathrm{d}t.
\end{aligned}
$$

Hypothesis 3.3.5 and 3.3.7 ensure that

$$
\begin{aligned}
&\Big\|\frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial g}{\partial u}(t,u(t))f^{(s)}(t)\Big]\Big\|_{H^{r+2}}\mathrm{d}t \\
&\quad\leq\ \tilde{L}_3(1+\widetilde{R})\int_{t_{n-p-1}}^{t_{n-p}}\|f^{(s)}(t)\|_{H^{r+2}}\,\mathrm{d}t+\tilde{L}_1\int_{t_{n-p-1}}^{t_{n-p}}\|f^{(s+1)}(t)\|_{H^{r+2}}\,\mathrm{d}t,
\end{aligned}
$$

if $\widetilde{R}$ stands for a bound of $\|u'(t)\|_{H^{r+2}}$ on $[0,T]$. Hence, the H^r-norm of term (3.53) is bounded by C_5h^{s+2} for some positive constant C_5.

One can easily check by using Hypothesis 3.3.5 that there exists a positive constant C_6 such that for all h and n, the H^r-norm of term (3.54) is bounded by C_6h^{s+2}. Hence, term (3.48) is bounded for all h and n by $(C_4+C_5+C_6)h^{s+2}$.

Eventually, we can estimate term (3.49) by writing, for $i\in\{1,\dots,s\}$, $h\in(0,h_0)$,

$n \in \{0, \dots, N-1\}$, and $p \in \{0, \dots, n\}$, using Hypothesis 3.3.6

$$\|\tilde{d}_{n-p,i}\|_{H^r} \le L_2\Big(\|u(t_{n-p}) - u_{n-p,i}\|_{H^r} + \|u(t_{n-p}) - u(t_{n-p} + c_i h)\|_{H^r}\Big)\|e_{n-p,i}\|_{H^r}.$$

Therefore, we get that

$$\begin{aligned} &\Big\|h\sum_{i=1}^{s} b_i(-hA)\sum_{p=0}^{n} e^{-phA}\tilde{d}_{n-p,i}\Big\|_{H^r} \\ \le\ & \mathcal{C}L_2 h\sum_{p=0}^{n}\sum_{i=1}^{s}\Big(\|e_{n-p,i}\|_{H^r}^2 + \widetilde{R}h\|e_{n-p,i}\|_{H^r}\Big). \end{aligned}$$

Note that we have with (3.41) and (3.44)

$$\begin{aligned} \sum_{i=1}^{s}\|e_{n-p,i}\|_{H^r}^2 &\le \Big(\sum_{i=1}^{s}\|e_{n-p,i}\|_{H^r}\Big)^2 \\ &\le (s+1)\Big(4s^2\|e_{n-p}\|_{H^r}^2 + 4\sum_{i=1}^{s}\|\Delta_{n-p,i}\|_{H^r}^2\Big) \\ &\le (s+1)\Big(4s^2\|e_{n-p}\|_{H^r}^2 + 4\kappa_s^2 h^{2s+1}\int_{t_{n-p-1}}^{t_{n-p}}\|f^{(s)}(\tau)\|_{H^r}^2\,\mathrm{d}\tau\Big) \\ &\le (s+1)\Big(4s^2 R\|e_{n-p}\|_{H^r} + 4\kappa_s^2 h^{2s+1}\int_{t_{n-p-1}}^{t_{n-p}}\|f^{(s)}(\tau)\|_{H^r}^2\,\mathrm{d}\tau\Big). \end{aligned}$$

We derive that there exists a positive constant C_7 such that for all h and n one has the H^r-norm of term (3.49) is bounded by

$$C_7 h\sum_{p=0}^{n}\|e_p\|_{H^r} + C_7 h^{s+2}.$$

We can then apply Lemma 3.3.2 to conclude. ■

3.4 Numerical experiments

We provide two kinds of numerical experiments to illustrate our results. Both of them have been performed with $s = 2$ collocation points.

- The first method we choose is defined by $c_1 = \frac{1}{2}$ and $c_2 = 1$. This method does not satisfy relation (3.11). Hence, its underlying Runge-Kutta method exactly is of order 2.

- The other method we choose is defined by $c_1 = \frac{1}{3}$ and $c_2 = 1$. This method satisfies relation (3.11) but not relation (3.20). Hence, its underlying Runge-Kutta method exactly is of order 3.

3.4.1 Linear problems

First, we consider a linear problem (3.4) in dimension $d = 1$. The functions u_0 and f are chosen in such a way that the exact solution of the problem is $u(t,x) = e^{i\left(\frac{t}{2}\right)^2 \sin(x)}$. We set the final time $T = 2\pi$. We apply Methods 1 and 2 to this problem for different timesteps $h > 0$ such that T/h is an integer. We plot in logarithmic scales the L^2-norm of the final error $e_{\frac{T}{h}}$ in function of h on Figure 3.4.1. Calculations are carried out with Fast Fourier Transforms with 2^8 modes.

One can see the effect of resonant time steps (those for which T/h is close to the square of an integer) in the numerical error at time T.

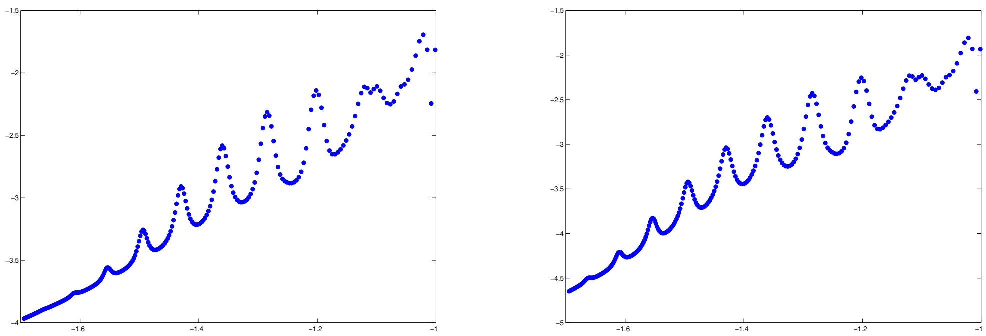

Figure 3.1: Final error in function of the time step for a linear Schrödinger problem. Method 1 (left) and 2 (right); logarithmic scales.

3.4.2 Semi-linear problems

Now, we consider a semi-linear problem (3.27) in dimension $d = 1$. The non-linearity is $g(u) = |u|^2 u$ and the initial data is $u_0(x) = \frac{1+e^{2ix}}{2+\sin(x)}$. We set the final time $T = 2\pi$. We apply Methods 1 and 2 to this problem for different time steps $h > 0$ such that T/h is an integer. We plot in logarithmic scales the L^2-norm of the final error $e_{\frac{T}{h}}$ in function of h on Figure 3.4.2. Calculations are carried out with Fast Fourier Transforms with 2^8 modes.

One can see that no resonance occurs for time steps between 10^{-3} and 10^{-1}, even when T/h is the square of an integer, and that both methods have a numerical order consistent with Theorems 3.3.1 and 3.3.2.

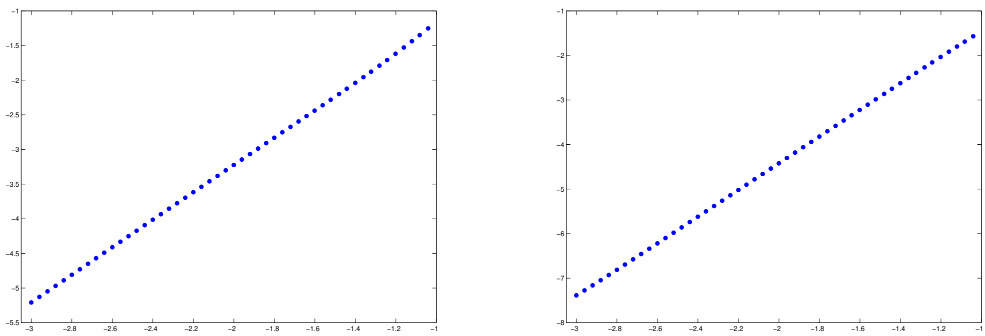

Figure 3.2: Final error in function of the time step for the cubic Schrödinger equation. Method 1 (left, numerical slope 1.9932) and 2 (right, numerical slope 2.9874); logarithmic scales.

Appendix A

Non linear analysis on $\mathbb{T}^d$

Contents

A.1 Introduction

We prove non linear estimates for functions on $\mathbb{T}^d$ using a Littlewood–Paley decomposition. This appendix follows very closely [AG91], where the case of functions on $\mathbb{R}^d$ is studied. As in the case of functions on $\mathbb{R}^d$, the key ingredient of the analysis is the use of dyadic decompositions of functions. Basically, the existing results for functions on $\mathbb{R}^d$ extend to the case of functions on $\mathbb{T}^d$ for similar reasons: Sobolev's embedding theorem (Theorem A.2.1) is still valid and so is Meyer's lemma (Lemma A.3.4). Therefore, this appendix is mostly a transcription of the results and proofs of [AG91].

In Section A.2, we introduce the functional Sobolev and Hölder spaces on $\mathbb{T}^d$ denoted respectively $H^s(\mathbb{T}^d)$ and $C^\alpha(\mathbb{T}^d)$. We give a characterization of these spaces using a Littlewood–Paley decomposition (Propositions A.2.5 and A.2.7). We then give a version of Sobolev's embedding theorem (Theorem A.2.1).

In Section A.3, we present applications of the Littlewood–Paley decomposition for functions on $\mathbb{T}^d$ using the notations and results of Section A.2. We prove that $C^\alpha(\mathbb{T}^d)$ is an algebra for all $\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$ and $H^s(\mathbb{T}^d)$ is an algebra provided $s > d/2$ and $s \notin \mathbb{N}$ (Proposition A.3.1). We then prove a non linear estimate for the composition of functions (Proposition A.3.2) using Meyer's multipliers lemma. As a conclusion, we prove that some non-linearities appearing in the non linear Schrödinger equation (including the cubic one) are locally Lipschitz (Corollary A.3.1) for suitable Sobolev norms.

A.2 Notations and basic properties.

This section is devoted to the introduction of a Littlewood–Paley decomposition for functions on $\mathbb{T}^d$ and a characterization of the Sobolev and Hölder spaces for functions on $\mathbb{T}^d$ that allow us to present a version of Sobolev's embedding theorem for functions on $\mathbb{T}^d$.

A.2.1 Notations

This subsection is aimed at the introduction of a Littlewood–Paley decomposition for complex functions on $\mathbb{T}^d$. We follow [AG91] where the case of complex functions on $\mathbb{R}^d$ is studied.

Notation A.2.1 *Let $\Psi : \mathbb{R} \to \mathbb{R}$ be a smooth function with compact support such that*

- $\forall x \in \mathbb{R}, \quad 0 \le x \le 1/2 \Rightarrow \Psi(x) = 1,$
- $\forall x \in \mathbb{R}, \quad x \ge 1 \Rightarrow \Psi(x) = 0,$
- Ψ *is a non-increasing function over* $\mathbb{R}^+$,

and set

$$\forall \xi \in \mathbb{R}^d, \quad \begin{cases} |\xi| &= \sqrt{\xi_1^2 + \cdots + \xi_d^2} \\ \psi(\xi) &= \Psi(|\xi|) \end{cases} .$$

Note that ψ is a smooth function with compact support on $\mathbb{R}^d$ and that $\psi(\mathbb{R}^d) = [0, 1]$. We then define

Definition A.2.1

$$\forall \xi \in \mathbb{R}^d, \quad \varphi(\xi) = \psi(\xi/2) - \psi(\xi).$$

Proposition A.2.1 *The function φ is smooth on $\mathbb{R}^d$ and its support lies in*

$$\{\xi \in \mathbb{R}^d \quad | \quad 1/2 \le |\xi| \le 2\}.$$

Moreover[1]*,*

$$\forall \xi \in \mathbb{R}^d, \quad \psi(\xi) + \sum_{p=0}^{\infty} \varphi(\xi/2^p) = 1. \tag{A.1}$$

[1]Note that at most 2 terms are not zero in the sum.

Definition A.2.2 *For all $u \in L^1(\mathbb{T}^d)$, we define the Fourier coefficients by*

$$\forall k \in \mathbb{Z}^d, \quad \mathscr{F}(u)(k) = \hat{u}(k) = \hat{u}_k = (2\pi)^{-d} \int_{\mathbb{T}^d} u(x) e^{-ix.k} \mathrm{d}x,$$

and the analytical functions $(u_p)_{p \geq -1}$ by their Fourier coefficients

$$\forall p \in \mathbb{N}, \forall k \in \mathbb{Z}^d, \quad \mathscr{F}(u_p)(k) = \hat{u}(k)\varphi(k/2^p), \tag{A.2}$$

and

$$\forall k \in \mathbb{Z}^d, \quad \mathscr{F}(u_{-1})(k) = \hat{u}(k)\psi(k).$$

Proposition A.2.2 *Relation* (A.1) *ensures that*

$$u = \sum_{p=-1}^{\infty} u_p.$$

The right hand side of this equality is known as the Littlewood–Paley decomposition of u. Note that for all $p \geq 1$, u_p is (the restriction to $\mathbb{T}^d$ of) an holomorphic function on $\mathbb{C}^d$, since it is a trigonometric polynomial.

Definition A.2.3 *We define as the spectrum of $u \in L^1(\mathbb{T}^d)$ the set*

$$\Big\{ k \in \mathbb{Z}^d \quad | \quad \mathscr{F}(u)(k) \neq 0 \Big\}$$

Definition A.2.4 *For all $u \in L^1(\mathbb{T}^d)$, we define*

$$\forall n \in \mathbb{Z} \quad s.t. \quad n \geq -1, \qquad S_n(u) = S_n u = \sum_{p=-1}^{n} u_p \in \mathcal{H}(\mathbb{C}^d).$$

Proposition A.2.3 *For all $u \in L^1(\mathbb{T}^d)$,*

$$\forall n \geq -1, \forall k \in \mathbb{Z}^d, \qquad \mathscr{F}\big(S_n(u)\big)(k) = \psi(k/2^{n+1})\hat{u}(k). \tag{A.3}$$

Proof. For all $\xi \in \mathbb{R}^d$ and all $p \in \mathbb{N}$, we have

$$\begin{aligned} \psi(\xi) + \sum_{q=0}^{p} \varphi(\xi/2^q) &= \psi(\xi) + \sum_{q=0}^{p} \big(\psi(\xi/2^{q+1}) - \psi(\xi/2^q)\big) \\ &= \psi(\xi/2^{p+1}). \end{aligned}$$

∎

Notation A.2.2 *For all function f on $\mathbb{Z}^d$, and all function u on $\mathbb{T}^d$, we denote by $f(D)u$ the function on $\mathbb{T}^d$ whose Fourier coefficients are given by*

$$\forall k \in \mathbb{Z}^d, \qquad \mathscr{F}\big(f(D)u\big)(k) = f(k)\hat{u}(k).$$

Hence, we have

Proposition A.2.4

$$u_{-1} = \psi(D)u \qquad \textit{and} \qquad u_p = \varphi(D/2^p)u.$$

A.2.2 Basic properties

In this subsection, we give a characterization of Sobolev spaces using the Littlewood–Paley decomposition (Proposition A.2.5). We then prove a useful estimate (Proposition A.2.6) that allows us to characterize Hölder spaces using the Littlewood–Paley decomposition (Proposition A.2.7). We start with a lemma dealing with real valued functions.

A.2.2.1 Littlewood–Paley decomposition of real functions

The following lemma ensures that, provided $u \in L^1$ is a real valued function, the u_p's also are real valued functions.

Lemma A.2.1 *If $u \in L^1(\mathbb{T}^d, \mathbb{R})$ then*

$$\forall p \geq -1, \qquad u_p \in C^\infty(\mathbb{T}^d, \mathbb{R}).$$

Proof. The fact that $\psi(\xi)$ and $\varphi(\xi)$ are fuctions of $|\xi|$ implies that

$$\forall k \in \mathbb{Z}^d, \qquad \varphi(-k) = \varphi(k) \qquad \text{and} \qquad \psi(-k) = \psi(k).$$

Moreover, since they are real valued functions, we have for all $k \in \mathbb{Z}^d$ all $p \in \mathbb{N}$, and $u \in L^1(\mathbb{T}^d, \mathbb{R})$,

$$\begin{aligned} \mathcal{F}(u_p)(-k) &= \varphi(-k/2^p)\mathcal{F}(u)(-k) \\ &= \varphi(k/2^p)\overline{\mathcal{F}(u)(k)} \\ &= \overline{\varphi(k/2^p)\mathcal{F}(u)(k)} \\ &= \overline{\mathcal{F}(u_p)(k)} \end{aligned}$$

and hence u_p is real valued. The proof is the same when $p = -1$. ■

A.2.2.2 Littlewood–Paley decomposition and Sobolev spaces

We are going to prove the characterization of the Sobolev spaces using the Littlewood–Paley decomposition (Proposition A.2.5). To do so, we chose a three-steps proof. First of all, we give estimations on the sum of series of functions whose terms have precribed spectrum localizations. This localization implies in our cases a quasi L^2-orthogonality. This is done by Lemma A.2.2 and its generalization Lemma A.2.3. Then, we state 3 lemmas that allow us to bound the L^2-norms of the derivatives of the u_p's by a function of the L^2-norm of u, and to control the size of the H^s-norms of the u_p's by their L^2-norms (and conversely). Once again, the localization of the spectrums of the u_p's is the essential tool of the proofs of these lemmas. As a consequence, we prove the characterization of Sobolev spaces using the Littlewood–Paley decomposition (Proposition A.2.5).

Lemma A.2.2 *For all $\xi \in \mathbb{R}^d$, we have*

$$1/2 \leq \psi^2(\xi) + \sum_{p=0}^{\infty} \varphi^2(\xi/2^p) \leq 1 \tag{A.4}$$

and for all $u \in L^2(\mathbb{T}^d)$,

$$\sum_{p=-1}^{\infty} \|u_p\|^2_{L^2(\mathbb{T}^d)} \leq \|u\|^2_{L^2(\mathbb{T}^d)} \leq 2 \sum_{p=-1}^{\infty} \|u_p\|^2_{L^2(\mathbb{T}^d)}. \tag{A.5}$$

Proof. Assume that $\xi \in \mathbb{R}^d$ is given with $|\xi| \leq 1/2$. Then,

$$\psi^2(\xi) + \sum_{p=0}^{\infty} \varphi^2(\xi/2^p) = \psi^2(\xi) = 1 \leq 1.$$

Assume that $\xi \in \mathbb{R}^d$ is given with $1/2 \leq |\xi| \leq 1$. Then,

$$\psi^2(\xi) + \sum_{p=0}^{\infty} \varphi^2(\xi/2^p) = \psi^2(\xi) + (\psi^2(\xi/2) - \psi(\xi))^2 = 1 + 2\psi(\xi)(\psi(\xi) - 1) \leq 1.$$

because $\psi(\mathbb{R}^d) \subset [0, 1]$.
Assume that $\xi \in \mathbb{R}^d$ is given with $2^q \leq |\xi| \leq 2^{q+1}$ for some $q \in \mathbb{N}$. Then,

$$\begin{aligned} \psi^2(\xi) + \sum_{p=0}^{\infty} \varphi^2(\xi/2^p) &= \varphi^2(\xi/2^q) + \varphi^2(\xi/2^{q+1}) \\ &\leq (\varphi(\xi/2^q) + \varphi(\xi/2^{q+1}))^2 \\ &= \Big(\psi(\xi) + \sum_{p=0}^{\infty} \varphi(\xi/2^p)\Big)^2. \end{aligned}$$

The second line is a consequence of the fact that $\varphi(\xi/2^q) \geq 0$ and $\varphi(\xi/2^{q+1}) \geq 0$, because Ψ is a non-increasing function over $\mathbb{R}^+$: Since we have $|\xi|/2^{q+1} \leq |\xi|/2^q$, we have for example

$$\varphi(\xi/2^q) = \psi(\xi/2^{q+1}) - \psi(\xi/2^q) = \Psi(|\xi|/2^{q+1}) - \Psi(|\xi|/2^q) \geq 0.$$

Now, since the sum (A.1) only has at most 2 non-zero terms, we deduce that for all $\xi \in \mathbb{R}^d$,

$$1 = \Big(\psi(\xi) + \sum_{p=0}^{\infty} \varphi(\xi/2^p)\Big)^2 \leq 2\Big(\psi(\xi)^2 + \sum_{p=0}^{\infty} \varphi^2(\xi/2^p)\Big),$$

because for all $(a, b) \in \mathbb{R}^2$, $(a + b)^2 \leq 2(a^2 + b^2)$. This ensures that (A.4) holds true.
As a conclusion, since we have for all $u \in L^2(\mathbb{T}^d)$,

$$\forall k \in \mathbb{Z}^d, \quad (1/2)|\hat{u}_k|^2 \leq \psi(k)^2|\hat{u}_k|^2 + \sum_{p=0}^{\infty} \varphi^2(k/2^p)|\hat{u}_k|^2 \leq |\hat{u}_k|^2,$$

we derive, summing over $k \in \mathbb{Z}^d$, that

$$(1/2)\|u\|^2_{L^2(\mathbb{T}^d)} \leq \|u_{-1}\|^2_{L^2(\mathbb{T}^d)} + \sum_{p=0}^{\infty} \|u_p\|^2_{L^2(\mathbb{T}^d)} \leq \|u\|^2_{L^2(\mathbb{T}^d)}.$$

This ensures that (A.5) holds true. ∎

A possible generalization is

Lemma A.2.3 *Let $C \geq 1/2$ be given. There exists a positive constant $\mathcal{C}$ such that for all sequence $(a_p)_{p\in\mathbb{N}}$ of functions of $L^2(\mathbb{T}^d)$ satisfying*

$$\forall p \in \mathbb{N}, \qquad \{k \in \mathbb{Z}^d \mid \mathscr{F}(a_p)(k) \neq 0\} \subset \{k \in \mathbb{Z}^d \mid C^{-1}2^{p-1} \leq |k| \leq C2^{p+1}\},$$

we have

$$\|\sum_{p\in\mathbb{N}} a_p\|^2_{L^2(\mathbb{T}^d)} \leq \mathcal{C} \sum_{p\in\mathbb{N}} \|a_p\|^2_{L^2(\mathbb{T}^d)} \in \mathbb{R} \cup \{+\infty\}.$$

Proof. We note that for all $k \in \mathbb{Z}^d$, the set of $p \in \mathbb{N}$ such that $\mathscr{F}(a_p)(k) \neq 0$ is included in the set

$$\mathcal{A}(k) := \Big\{p \in \mathbb{N} \;\Big|\; \frac{\ln|k|}{\ln 2} - \big(1 + \frac{\ln C}{\ln 2}\big)\} \leq p \leq \frac{\ln|k|}{\ln 2} + \big(1 + \frac{\ln C}{\ln 2}\big)\Big\}.$$

We derive that it contains at most $\alpha := 2 + E\big(\big(\frac{\ln C}{\ln 2}\big)\big)^2$ points. Therefore,

$$\begin{aligned}
\|\sum_{p\in\mathbb{N}} a_p\|^2_{L^2\mathbb{T}^d} &= \sum_{k\in\mathbb{Z}^d} \big|\sum_{p\in\mathbb{N}} \mathscr{F}(a_p)(k)\big|^2 \\
&= \sum_{k\in\mathbb{Z}^d} \big|\sum_{p\in\mathcal{A}(k)} \mathscr{F}(a_p)(k)\big|^2 \\
&\leq \sum_{k\in\mathbb{Z}^d} \Big(\sum_{p\in\mathcal{A}(k)} 1^2\Big)\Big(\sum_{p\in\mathcal{A}(k)} \big|\mathscr{F}(a_p)(k)\big|^2\Big) \\
&\leq \alpha \sum_{k\in\mathbb{Z}^d} \sum_{p\in\mathcal{A}(k)} \big|\mathscr{F}(a_p)(k)\big|^2 \\
&\leq \alpha \sum_{k\in\mathbb{Z}^d} \sum_{p\in\mathbb{N}} \big|\mathscr{F}(a_p)(k)\big|^2 \\
&\leq \alpha \sum_{p\in\mathbb{N}} \|a_p\|^2_{L^2(\mathbb{T}^d)}.
\end{aligned}$$

■

We now state and prove 3 lemmas that will be useful for the proof of Proposition A.2.5. The first lemma (Lemma A.2.4) bounds the growth of the L^2-norms of the derivatives u_p and S_p in function of the L^2 norm of u. For an equivalent of Lemma 3 for L^∞-norms, see Lemma A.2.8 and Lemma A.2.9. After the definition of the Sobolev spaces (Definition A.2.6), the second lemma (Lemma A.2.5) ensures that for integer values of s, $H^s(\mathbb{T}^d)$ fits the classical definition of the Sobolev space of index s. The third lemma (Lemma A.2.6) makes explicit the constants in the bound of the L^2-norm of u_p by its H^s-norm, and in the converse bound.

[2] $E(x)$ stands for the integer part of the real number x.

Lemma A.2.4 *For all function $u \in L^2(\mathbb{T}^d)$, all $\alpha \in \mathbb{N}^d$, and all $p \geq -1$, we have*

$$\begin{aligned} \|\partial^\alpha u_p\|_{L^2(\mathbb{T}^d)} &\leq 2^{(p+1)|\alpha|}\|u\|_{L^2(\mathbb{T}^d)} \\ \|\partial^\alpha S_p u\|_{L^2(\mathbb{T}^d)} &\leq 2^{(p+1)|\alpha|}\|u\|_{L^2(\mathbb{T}^d)} \end{aligned}$$

Proof. Note that for all $p \in \mathbb{N}$, $k \in \mathbb{Z}^d$,

$$\mathscr{F}(\partial^\alpha u_p)(k) = (ik)^\alpha \varphi(k/2^p)\hat{u}_k.$$

Hence,

$$\begin{aligned} \|\partial^\alpha u_p\|^2_{L^2(\mathbb{T}^d)} &= \sum_{k\in\mathbb{Z}^d} (k)^{2\alpha}\varphi^2(k/2^p)|\hat{u}_k|^2 \\ &\leq 2^{(p+1).2|\alpha|}\sum_{k\in\mathbb{Z}^d} \varphi^2(k/2^p)|\hat{u}_k|^2 \\ &\leq 2^{(p+1).2|\alpha|}\|u\|^2_{L^2(\mathbb{T}^d)} \end{aligned}$$

.

because k^α is bounded by $2^{(p+1)\alpha_1}2^{(p+1)\alpha_2}\dots 2^{(p+1)\alpha_d}$ on the support of $\varphi(./2^p)$, which is included in the set

$$\{\xi \in \mathbb{R}^d \quad | \quad 2^{p-1} \leq |\xi| \leq 2^{p+1}\},$$

and φ is bounded by 1 on $\mathbb{R}^d$. Moreover, for $p = -1$, we have

$$\begin{aligned} \|\partial^\alpha u_{-1}\|^2_{L^2(\mathbb{T}^d)} &= \sum_{k\in\mathbb{Z}^d} (k)^{2\alpha}\psi^2(k)|\hat{u}_k|^2 \\ &\leq \sum_{k\in\mathbb{Z}^d} \psi^2(k)|\hat{u}_k|^2 \\ &\leq 2^{(-1+1).2|\alpha|}\|u\|^2_{L^2(\mathbb{T}^d)}, \end{aligned}$$

because ψ is bounded by 1 on $\mathbb{R}^d$.
A similar proof leads to the result for $S_p(u)$ since for all $p \geq 0$ and all $k \in \mathbb{Z}^d$,

$$\mathscr{F}(\partial^\alpha S_p(u))(k) = (ik)^\alpha \Big(\psi(k) + \sum_{n=0}^{p}\varphi(k/2^n)\Big)\hat{u}_k,$$

and because k^α is bounded by $2^{(p+1)|\alpha|}$ when $|k| \leq 2^{p+1}$, and $\psi + \sum_{n=0}^p \varphi(./2^n)$ is bounded by 1 on $\mathbb{R}^d$. ■

Definition A.2.5 *For $k \in \mathbb{Z}^d$, we denote*

$$< k >= \begin{cases} 1 & if \quad |k| = 0 \\ |k| & if \quad |k| \neq 0 \end{cases}$$

Definition A.2.6 *For $s \in \mathbb{R}^+$, a function $u \in L^1(\mathbb{T}^d)$ is in $H^s(\mathbb{T}^d)$ if*

$$\sum_{k\in\mathbb{Z}^d} < k >^{2s} |\hat{u}_k|^2 < +\infty$$

For such a function u, we set

$$\|u\|_{H^s(\mathbb{R}^d)} = \Big(\sum_{k\in\mathbb{Z}^d} < k >^{2s} |\hat{u}_k|^2\Big)^{\frac{1}{2}}$$

Note that we have $\|u\|_{H^s(\mathbb{T}^d)} = \| < D >^s u\|_{L^2(\mathbb{T}^d)}$.

The following lemma shows that for integer values of s, the definition of the space $H^s(\mathbb{T}^d)$ fits the classical definition of that space.

Lemma A.2.5 *For all $s \in \mathbb{N}$, there exists positive constants c and C such that for all $u \in H^s(\mathbb{T}^d)$,*

$$c\|u\|_{H^s(\mathbb{T}^d)} \leq \Big(\sum_{p=0}^{s}\sum_{|\alpha|=p} \|\partial_x^\alpha u\|^2_{L^2(\mathbb{T}^d)}\Big)^{1/2} \leq C\|u\|_{H^s(\mathbb{T}^d)}\,. \tag{A.6}$$

Proof. Choose $s \in \mathbb{N}$ and $u \in H^s(\mathbb{T}^d)$. Since for all $\alpha \in \mathbb{N}^d$ such that $|\alpha| \leq s$, we have

$$\forall k \in \mathbb{Z}^d, \qquad \mathscr{F}(\partial_x^\alpha u)(k) = (ik)^\alpha \mathscr{F}(u)(k)$$

and

$$\begin{aligned} \forall k \in \mathbb{Z}^d, \qquad |(ik)|^{2\alpha} &= |k_1|^{2\alpha_1}\dots|k_d|^{2\alpha_d} \\ &\leq |k|^{2\alpha_1}\dots|k|^{2\alpha_d} \\ &\leq |k|^{2|\alpha|} \\ &\leq < k >^{2s}, \end{aligned}$$

we derive that the right inequality in (A.6) holds true. Conversely, since there exists a constant C_1 depending only on d and s such that

$$\forall k \in \mathbb{Z}^d, \qquad |k|^{2s} = \big(|k_1| + \dots + |k_d|\big)^{2s} \leq C_1\big(|k_1|^{2s} + \dots + |k_d|^{2s}\big),$$

we derive that the left inequality in (A.6) holds true too. ∎

The following lemma is a link between the L^2 and H^s norms of the u_p's.

Lemma A.2.6 *For all function u, all $s \in \mathbb{R}^+$, and all $p \geq -1$,*

$$2^{(p-1)s}\|u_p\|_{L^2(\mathbb{T}^d)} \leq \|u_p\|_{H^s(\mathbb{T}^d)} \leq 2^{(p+1)s}\|u_p\|_{L^2(\mathbb{T}^d)} \tag{A.7}$$

Proof. Note that for such u, s and for $p \geq 0$,

$$\|u_p\|^2_{H^s(\mathbb{R}^d)} = \sum_{k\in\mathbb{Z}^d} < k >^{2s} \varphi^2(k/2^p)|\hat{u}_k|^2$$

Hence,

$$\begin{array}{rcccl} 2^{2(p-1)s}\displaystyle\sum_{k\in\mathbb{Z}^d}\varphi^2(k/2^p)|\hat{u}_k|^2 & \leq & \|u_p\|^2_{H^s(\mathbb{R}^d)} & \leq & 2^{2(p+1)s}\displaystyle\sum_{k\in\mathbb{Z}^d}\varphi^2(k/2^p)|\hat{u}_k|^2 \\ 2^{2(p-1)s}\|u_p\|^2_{L^2(\mathbb{T}^d)} & \leq & \|u_p\|^2_{H^s(\mathbb{R}^d)} & \leq & 2^{2(p+1)s}\|u_p\|^2_{L^2(\mathbb{T}^d)} \end{array}$$

Moreover, for $p=-1$, we have

$$\begin{array}{rcl} 2^{2(p-1)s}\|u_p\|_{L^2(\mathbb{T}^d)} & = & 2^{-4s}|\hat{u}_0|^2 \\ \|u_p\|_{H^s(\mathbb{R}^d)} & = & |\hat{u}_0|^2 \\ 2^{2(p+1)s}\|u_p\|_{L^2(\mathbb{T}^d)} & = & 2^0|\hat{u}_0|^2 \end{array}$$

This ensures that (A.7) holds true. ■

The following proposition is a characterization of the Sobolev spaces using the smallness of the L^2-norm of the u_p's when p goes to infinity. Note that this characterization does not involve any derivative.

Proposition A.2.5 *For all $s\in\mathbb{R}^+$, if $u\in H^s(\mathbb{T}^d)$, then there exists a sequence of non-negative real numbers $(c_p)_{p\geq -1}$ depending on u such that*

$$\forall p\geq -1, \qquad \|u_p\|_{L^2(\mathbb{T}^d)}\leq 2^s 2^{-ps}c_p(u)\|u\|_{H^s(\mathbb{T}^d)}$$

and $\sum_{p\geq -1}c_p(u)^2\leq 1$.

Conversely, if a function u on $\mathbb{T}^d$ is such that there exists $s\in\mathbb{R}^+$ and a non-negative constant C such that

$$\forall p\geq -1, \qquad \|u_p\|_{L^2(\mathbb{T}^d)}\leq 2^s 2^{-ps}c_p(u)C$$

with $\sum_{p\geq -1}c_p(u)^2\leq 1$, then $u\in H^s(\mathbb{T}^d)$ and $\|u\|_{H^s(\mathbb{T}^d)}\leq 2^{2s+1/2}C$.

Proof. Note that for all $s\in\mathbb{R}^+$ and all function u on $\mathbb{T}^d$, we have

$$\begin{array}{rcl} \displaystyle\sum_{p\geq -1}\|(< D >^s u)_p\|^2_{L^2(\mathbb{T}^d)} & = & \displaystyle\sum_{p\geq -1}\| < D >^s u_p\|^2_{L^2(\mathbb{T}^d)} \\ & = & \displaystyle\sum_{p\geq -1}\|u_p\|^2_{H^s(\mathbb{T}^d)}. \end{array}$$

On one hand, if $u\in H^s(\mathbb{T}^d)$ for some $s\in\mathbb{R}^+$, then $< D >^s u\in L^2(\mathbb{T}^d)$. Hence, with the left-hand side of inequality (A.7), we get

$$\sum_{p\geq -1}2^{2(p-1)s}\|u_p\|^2_{L^2(\mathbb{T}^d)}\leq\sum_{p\geq -1}\|u_p\|^2_{H^s(\mathbb{T}^d)}.$$

Moreover, inequality (A.5) ensures that

$$\sum_{p\geq -1}\|(< D >^s u)_p\|^2_{L^2}\leq\| < D >^s u\|^2_{L^2(\mathbb{T}^d)}.$$

Therefore, we derive

$$\sum_{p\geq -1} 2^{2(p-1)s}\|u_p\|^2_{L^2(\mathbb{T}^d)} \leq \| < D >^s u\|^2_{L^2(\mathbb{T}^d)} .$$

Setting for $u \neq 0$

$$\forall p \geq -1, \qquad c_p(u) := \frac{2^{-s}2^{ps}\|u_p\|_{L^2(\mathbb{T}^d)}}{\|u\|_{H^s(\mathbb{T}^d)}}$$

yields the first part of the result.

On the other hand, if we assume that

$$\forall p \geq -1, \qquad \|u_p\|_{L^2(\mathbb{T}^d)} \leq 2^s C 2^{-ps} c_p(u),$$

then with the right-hand side of inequality (A.7)

$$\begin{aligned}\sum_{p\geq -1} \|(< D >^s u)_p\|^2_{L^2(\mathbb{T}^d)} &\leq 2^{2s} \sum_{p\geq -1} 2^{2ps}\|u_p\|^2_{L^2(\mathbb{T}^d)} \\ &\leq 2^{4s}C^2 \sum_{p\geq -1} 2^{2ps}2^{-2ps}c_p(u)^2.\end{aligned}$$

With inequality (A.5), we get that $< D >^s u \in H^s(\mathbb{T}^d)$ with $\|u\|_{H^s(\mathbb{T}^d)} \leq 2^{1/2}2^{2s}C$. This proves the second part of the result.

■

A.2.2.3 A useful estimate

The goal of this section is to state and prove Proposition A.2.6. This proposition will be useful for the proof of the characterization of Hölder spaces using the Littlewood–Paley decomposition (Proposition A.2.7). We start with the following lemma stating the Poisson summation formula.

Lemma A.2.7 *If $f \in \mathcal{S}(\mathbb{R}^d)$, $\mathcal{F}(f)$ is its Fourier tranform defined by*

$$\forall \xi \in \mathbb{R}^s, \qquad \mathcal{F}(f)(\xi) = \hat{f}(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} f(x)e^{-ix.\xi}\,\mathrm{d}x,$$

then the following relation holds

$$\sum_{k\in\mathbb{Z}^d} f(k) = (2\pi)^{\frac{d}{2}} \sum_{l\in\mathbb{Z}^d} \hat{f}(2\pi l).$$

This relation is know as the Poisson summation formula*.*

Proof. Note that the function

$$S : \begin{pmatrix} \mathbb{R}^d & \longrightarrow & \mathbb{C} \\ t & \longmapsto & \sum_{k\in\mathbb{Z}} f(t+k) \end{pmatrix}$$

is in $C^\infty(\mathbb{R}^d, \mathbb{C})$ and is 1-periodic in each space direction. Hence, its Fourier coefficients defined by

$$\forall l \in \mathbb{Z} \qquad c_l(S) = \int_{[0,1]^d} S(t) e^{-2i\pi l.t} \mathrm{d}t$$

are such that for all $\alpha \in \mathbb{Z}^d$, the set of complex numbers $\left(l^\alpha c_l(S)\right)_{l \in \mathbb{Z}^d}$ is bounded. With this observation in mind, we get for all $l \in \mathbb{Z}^d$,

$$\begin{aligned}
c_l(S) &= \int_{[0,1]^d} \left(\sum_{k \in \mathbb{Z}^d} f(t+k) \right) e^{-2i\pi l.t} \mathrm{d}t \\
&= \sum_{k \in \mathbb{Z}^d} \left(\int_{[0,1]^d} f(t+k) e^{-2i\pi l.t} \mathrm{d}t \right) \\
&= \sum_{k \in \mathbb{Z}^d} \left(\int_{k_1}^{k_1+1} \dots \int_{k_d}^{k_d+1} f(u) e^{-2i\pi l.(u-k)} \mathrm{d}u_d \dots \mathrm{d}u_1 \right) \\
&= \sum_{k \in \mathbb{Z}^d} \left(\int_{k_1}^{k_1+1} \dots \int_{k_d}^{k_d+1} f(u) e^{-2i\pi l.u} \mathrm{d}u_d \dots \mathrm{d}u_1 \right) \\
&= (2\pi)^{\frac{d}{2}} \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} f(u) e^{-i(2\pi l).u} \mathrm{d}u \\
&= (2\pi)^{\frac{d}{2}} \hat{f}(2\pi l).
\end{aligned}$$

Therefore,

$$\sum_{k \in \mathbb{Z}^d} f(k) = S(0) = \sum_{l \in \mathbb{Z}^d} c_l(S),$$

and the result is proved. ■

The previous lemma allows us to prove the following

Proposition A.2.6 *Assume $\chi \in C_0^\infty(\mathbb{R}^d)$ is given. For all $\lambda > 0$, we denote by g_λ the function on $\mathbb{T}^d$ whose Fourier coefficients are given by*

$$\forall k \in \mathbb{Z}^d, \qquad \mathscr{F}(g_\lambda)(k) = \chi(\lambda k).$$

The following estimates hold:

$$\forall \lambda > 0, \qquad \|g_\lambda(x)\|_{L^1_x(\mathbb{T}^d)} \le \|\mathcal{F}(\chi)(\xi)\|_{L^1_\xi(\mathbb{R}^d)}.$$

Note that g_λ is a trigonometric polynomial and that the right-hand side in the previous inequality does not depend on λ.

Proof. First, note that for all positive λ,

$$\|g_\lambda\|_{L^1(\mathbb{T}^d)} = \int_{\mathbb{T}^d} \Big| \sum_{k \in \mathbb{Z}^d} \chi(\lambda k) e^{ik.y} \Big| \mathrm{d}y.$$

Then, let us denote for all positive λ and all $y \in \mathbb{T}^d$,

$$\forall x \in \mathbb{R}^d, \qquad f_\lambda^y(x) = \chi(\lambda x) e^{+ix.y}.$$

Since $\chi \in C_0^\infty(\mathbb{R}^d)$, we have $f_\lambda^y \in S(\mathbb{R}^d)$. Therefore, Lemma A.2.7 ensures that

$$\sum_{k\in\mathbb{Z}^d} f_\lambda^y(k) = (2\pi)^{\frac{d}{2}} \sum_{l\in\mathbb{Z}^d} \mathcal{F}(f_\lambda^y)(2\pi l).$$

Hence, we have for all $\lambda > 0$,

$$\begin{aligned}
\|g_\lambda\|_{L^1(\mathbb{T}^d)} &= \int_{\mathbb{T}^d} \Big| \sum_{l\in\mathbb{Z}^d} \int_{\mathbb{R}^d} \chi(\lambda x) e^{+ix.(y-2\pi l)} \mathrm{d}x \Big| \mathrm{d}y \\
&\leq \sum_{l\in\mathbb{Z}^d} \int_{\mathbb{T}^d} \Big| \int_{\mathbb{R}^d} \chi(\lambda x) e^{+ix.(y-2\pi l)} \mathrm{d}x \Big| \mathrm{d}y \\
&\leq \sum_{l\in\mathbb{Z}^d} \int_{-2\pi l+\mathbb{T}^d} \Big| \int_{\mathbb{R}^d} \chi(\lambda x) e^{+ix.u} \mathrm{d}x \Big| \mathrm{d}u \\
&\leq \int_{\mathbb{R}^d} \Big| \int_{\mathbb{R}^d} \chi(\lambda x) e^{+ix.u} \mathrm{d}x \Big| \mathrm{d}u \\
&\leq \int_{\mathbb{R}^d} \Big| \int_{\mathbb{R}^d} \chi(\lambda x) e^{+i(\lambda x).(u\lambda)} \mathrm{d}(\lambda x) \Big| \mathrm{d}(u/\lambda) \\
&\leq \int_{\mathbb{R}^d} \Big| \int_{\mathbb{R}^d} \chi(x) e^{+ix.u} \mathrm{d}x \Big| \mathrm{d}u.
\end{aligned}$$

This proves the result. ■

Proposition A.2.6 will be useful for the proofs of the following L^∞-type estimates.

A.2.2.4 Littlewood–Paley decomposition and Hölder spaces

We want to prove Proposition A.2.7 which is a characterization of the Hölder spaces using the smallness of the L^∞-norm of the u_p's when p tends to infinity. We start with 3 lemmas. The first one (Lemma A.2.8) bounds the L^∞-norm of the derivatives of the u_p's in function of the L^∞-norm of u. The second one (LemmaA.2.9) bounds the L^∞-norm of the derivatives of the $S_p(u)$'s in function of the L^∞-norm of u. These two lemmas are the equivalent of Lemma A.2.4 for the L^∞-norm. The last lemma (Lemma A.2.10) makes explicit the constants in the bounds of the L^∞-norms of u_p by its derivatives of a given order.

Lemma A.2.8 *For all $\alpha \in \mathbb{N}^d$, there exists a positive constant C such that*

$$\forall p \geq -1, \forall u \in L^\infty(\mathbb{T}^d), \qquad \|\partial_x^\alpha u_p\|_{L^\infty(\mathbb{T}^d)} \leq C\, 2^{p|\alpha|} \|u\|_{L^\infty(\mathbb{T}^d)}.$$

Proof. Let us denote for $p \in \mathbb{N}$, g_p the function on $\mathbb{T}^d$ such that

$$\forall k \in \mathbb{Z}^d, \qquad \mathscr{F}(g_p)(k) = \varphi(k/2^p).$$

Relation (A.2) ensures that

$$\forall p \in \mathbb{N}, \forall \alpha \in \mathbb{Z}^d, \forall x \in \mathbb{T}^d, \qquad \partial_x^\alpha u_p(x) = \Big((\partial_x^\alpha g_p) \star u \Big)(x).$$

Hence,

$$\forall p \in \mathbb{N}, \forall \alpha \in \mathbb{Z}^d, \qquad \|\partial_x^\alpha u_p\|_{L^\infty(\mathbb{T}^d)} \leq \|\partial_x^\alpha g_p\|_{L^1(\mathbb{T}^d)} \|u\|_{L^\infty(\mathbb{T}^d)}.$$

Since for all $\alpha \in \mathbb{N}^d$ we have

$$\forall p \in \mathbb{N}, \forall k \in \mathbb{Z}^d, \qquad \mathscr{F}(\partial_x^\alpha g_p)(k) = (ik)^\alpha \varphi(k/2^p) = 2^{p|\alpha|} \Big(\frac{ik}{2^p} \Big)^\alpha \varphi(k/2^p),$$

Proposition A.2.6 ensures that

$$\forall p \in \mathbb{N}, \forall \alpha \in \mathbb{N}^d, \qquad \|\partial_x^\alpha g_p\|_{L^1(\mathbb{T}^d)} \leq 2^{p|\alpha|} \|\mathcal{F}((ix)^\alpha \varphi(x))(\xi)\|_{L^1_\xi(\mathbb{R}^d)}.$$

Hence, the result is proved for $p \in \mathbb{N}$ with $C = \|\mathcal{F}((ix)^\alpha \varphi(x))(\xi)\|_{L^1_\xi(\mathbb{R}^d)}$.

In the case $p = -1$, we have for all function $u \in L^\infty(\mathbb{T}^d)$,

$$\forall x \in \mathbb{T}^d, \qquad u_{-1}(x) = (2\pi)^{-d} \int_{\mathbb{T}^d} u(y) \mathrm{d}y \leq \|u\|_{L^\infty(\mathbb{T}^d)}.$$

Hence, if $\alpha = 0$, then we can choose $C = 1$. If $\alpha \neq 0$, then $\partial_x^\alpha u_{-1} = 0$, and we can choose any positive C.

∎

Similarly, we get the following

Lemma A.2.9 *For all $\alpha \in \mathbb{N}^d$, there exists a positive constant $C > 0$ such that*

$$\forall p \geq -1, \forall u \in L^\infty(\mathbb{T}^d), \qquad \|\partial_x^\alpha S_p(u)\|_{L^\infty(\mathbb{T}^d)} \leq C\, 2^{p|\alpha|} \|u\|_{L^\infty(\mathbb{T}^d)}.$$

Proof. As in the proof of the previous lemma, for all $p \geq -1$, we denote by h_p the function on $\mathbb{T}^d$ such that

$$\forall k \in \mathbb{Z}^d, \qquad \mathscr{F}(h_p)(k) = \psi(k/2^{p+1}).$$

Relation (A.3) ensures that

$$\forall \alpha \in \mathbb{N}^d, \forall x \in \mathbb{T}^d, \qquad \partial_x^\alpha S_p(u)(x) = \big((\partial_x^\alpha h_p) \star u \big)(x).$$

Note that

$$\forall p \geq -1, \forall k \in \mathbb{Z}^d, \qquad \mathscr{F}(\partial_x^\alpha h_p)(k) = (ik)^\alpha \psi(k/2^{p+1}) e^{ik.y} = 2^{p|\alpha|} \Big(\frac{ik}{2^p} \Big)^\alpha \psi\Big(\frac{k/2}{2^p} \Big)$$

Hence, Proposition A.2.6 yields

$$\forall p \geq -1, \forall \alpha \in \mathbb{N}^d, \qquad \|\partial_x^\alpha h_p\|_{L^1(\mathbb{T}^d)} \leq 2^{p|\alpha|} \| \big[\mathcal{F}\big((ix)^\alpha \psi(x/2) \big) \big](\xi) \|_{L^1_\xi(\mathbb{R}^d)}.$$

and the lemma holds true with $C = \| \big[\mathcal{F}\big((ix)^\alpha \psi(x/2) \big) \big](\xi) \|_{L^1_\xi(\mathbb{R}^d)}$. ∎

The last estimation lemma is the following

Lemma A.2.10 *For all $l \in \mathbb{N}$, there exists a positive constant C such that for all function u on $\mathbb{T}^d$,*

$$\forall p \geq 0, \qquad \frac{1}{C}2^{pl}\|u_p\|_{L^\infty(\mathbb{T}^d)} \leq \sum_{|\alpha|=l} \|\partial_x^\alpha u_p\|_{L^\infty(\mathbb{T}^d)} \leq C2^{pl}\|u_p\|_{L^\infty(\mathbb{T}^d)}. \tag{A.8}$$

Proof. Assume $l \in \mathbb{N}$ is given. To prove the right inequality in (A.8), we note that if $\chi \in C_0^\infty(\mathbb{R}^d)$ is such that

$$\forall \xi \in \operatorname{supp}\varphi, \qquad \chi(x) = 1,$$

then we have for all $\alpha \in \mathbb{N}^d$ such that $|\alpha| = l$,

$$\begin{aligned} \forall p \in \mathbb{N}, \forall k \in \mathbb{Z}^d, \qquad \mathscr{F}(\partial_x^\alpha u_p) &= (ik)^\alpha \varphi(k/2^p)\hat{u}(k) \\ &= 2^{pl}\Big(\frac{ik}{2^p}\Big)^\alpha \chi(k/2^p)\varphi(k/2^p)\hat{u}(k). \end{aligned}$$

Therefore, if we denote by $g_{\alpha,p}$ the function whose Fourier coefficients are given by

$$\forall k \in \mathbb{Z}^d, \qquad \mathscr{F}(g_{\alpha,p})(k) = \Big(\frac{ik}{2^p}\Big)^\alpha \chi(k/2^p),$$

then we derive that

$$\forall x \in \mathbb{T}^d, \qquad \partial_x^\alpha u_p(x) = 2^{pl}\big(g_{\alpha,p} \star u_p\big)(x).$$

Proposition A.2.6 yields

$$\|g_{\alpha,p}\|_{L^1(\mathbb{T}^d)} \leq \|\mathcal{F}((ix)^\alpha \chi(x))(\xi)\|_{L^1_\xi(\mathbb{R}^d)}.$$

The right inequality in (A.8) is hence proved with $C = \sum_{|\alpha|=l} \|\mathcal{F}((ix)^\alpha\chi(x))(\xi)\|_{L^1_\xi(\mathbb{R}^d)}$.

To prove the left inequality in (A.8), we impose to the function χ to satisfy moreover $\chi \equiv 0$ on a neighborhood of 0. For all $\alpha \in \mathbb{N}^d$ such that $|\alpha| = l$, we define

$$\forall \xi \in \mathbb{R}^d, \qquad \chi_\alpha(\xi) = \frac{\xi^\alpha \chi(\xi)}{\sum_{|\beta|=l}(\xi^\beta)^2}.$$

With this definition, we have that $\chi_\alpha \in C_0^\infty(\mathbb{R}^d)$ and

$$\forall \xi \in \mathbb{R}^d, \qquad \varphi(\xi) = \Big(\sum_{|\alpha|=l} \xi^\alpha \chi_\alpha(\xi)\Big)\varphi(\xi).$$

Hence, for all $p \in \mathbb{N}$,

$$\begin{aligned} \forall k \in \mathbb{Z}^d, \quad \mathscr{F}(u_p)(k) &= \varphi(k/2^p)\hat{u}(k) \\ &= \left(\sum_{|\alpha|=l}\Big(\frac{k}{2^p}\Big)^\alpha \chi_\alpha(k/2^p)\right)\varphi(k/2^p)\hat{u}(k) \\ &= 2^{-pl}i^{-l}\sum_{|\alpha|=l}\chi_\alpha(k/2^p)(ik)^\alpha\hat{u}(k). \end{aligned}$$

This yields

$$\forall p \in \mathbb{N}, \forall k \in \mathbb{Z}^d, \qquad 2^{pl}\mathscr{F}(u_p)(k) = i^{-l}\sum_{|\alpha|=l}\chi_\alpha(k/2^p)\mathscr{F}(\partial_x^\alpha u_p)(k).$$

For all $\alpha \in \mathbb{N}^d$ such that $|\alpha| = l$, we denote by $g_{\alpha,p}$ the function on $\mathbb{T}^d$ whose Fourier coefficients are the $(\chi_\alpha(k/2^p))_{k\in\mathbb{Z}^d}$. Proposition A.2.6 yields

$$\sum_{|\alpha|=l}\|g_{\alpha,p}\|_{L^1(\mathbb{T}^d)} \leq \sum_{|\alpha|=l}\|\mathcal{F}(\chi_\alpha)\|_{L^1(\mathbb{R}^d)}.$$

Since

$$\forall l \in \mathbb{N}, \forall p \in \mathbb{N}, \forall x \in \mathbb{T}^d, \qquad 2^{pl}u_p(x) = i^{-l}\Big(\big(\sum_{|\alpha|=l} g_{\alpha,p}\big)\star u_p\Big)(x),$$

the left inequality in (A.8) is proved with $C = \sum_{|\alpha|=l}\|\mathcal{F}(\chi_\alpha)\|_{L^1(\mathbb{R}^d)}$.
This proves the result. ∎

Let us set the following

Definition A.2.7 *For $\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$, $C^\alpha(\mathbb{T}^d)$ is the subset of complex functions on $\mathbb{T}^d$ whose periodic extension to $\mathbb{R}^d$ is in $C^k(\mathbb{R}^d)$ (where $k := E(\alpha) \in \mathbb{N}$ is the integer part of α) such that*

$$\sup\left\{\frac{|\partial_x^\beta u(x) - \partial_x^\beta u(y)|}{|x-y|^{\alpha-k}}\ \middle|\ x, y \in \mathbb{R}^d,\ \beta \in \mathbb{N}^d \text{ with } x \neq y \text{ and } |\beta| = k\right\} < +\infty.$$

For $\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$ and $u \in C^\alpha(\mathbb{T}^d)$, we set

$$\|u\|_{C^\alpha(\mathbb{T}^d)} := \|u\|_{L^\infty(\mathbb{T}^d)} + \text{the above supremum.}$$

From now on, we identify any function u on $\mathbb{T}^d$ with its periodic extension to $\mathbb{R}^d$. We have the following characterization of the Hölder spaces.

Proposition A.2.7 *For all $\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$, there exists a positive constant $C > 0$ such that*

$$\forall u \in C^\alpha(\mathbb{T}^d), \forall p \geq -1, \qquad \|u_p\|_{L^\infty(\mathbb{T}^d)} \leq C\, 2^{-p\alpha}\|u\|_{C^\alpha(\mathbb{T}^d)}.$$

Conversely, for all $\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$, there exists a positive constant $C > 0$ such that for all function u on $\mathbb{T}^d$, if

$$\exists \mathcal{C} > 0, \forall p \geq -1, \qquad \|u_p\|_{L^\infty(\mathbb{T}^d)} \leq \mathcal{C}2^{-p\alpha},$$

then

$$u \in C^\alpha(\mathbb{T}^d) \qquad \text{and} \qquad \|u\|_{C^\alpha(\mathbb{T}^d)} \leq C\mathcal{C}.$$

Proof. Assume $\alpha \in (0,1)$. Since for all $x \in \mathbb{T}^d$,

$$u_{-1}(x) = \psi(D)u(x),$$

we derive that

$$\|u_{-1}\|_{L^\infty(\mathbb{T}^d)} \leq \|g\|_{L^1(\mathbb{T}^d)} \|u\|_{L^\infty(\mathbb{T}^d)},$$

where g is the function on $\mathbb{T}^d$ whose Fourier coefficients are given by

$$\forall k \in \mathbb{Z}^d, \qquad \mathscr{F}(g)(k) = \psi(k).$$

Moreover, if we denote for all $p \in \mathbb{N}$ by g_p the function on $\mathbb{T}^d$ whose Fourier coefficients are given by

$$\forall k \in \mathbb{Z}^d, \qquad \mathscr{F}(g_p)(k) = \varphi(k/2^p),$$

then for all $x \in \mathbb{T}^d$, we have

$$u_p(x) = (2\pi)^{-d} \int_{\mathbb{T}^d} g_p(x-y)u(y)\mathrm{d}y = (2\pi)^{-d} \int_{\mathbb{T}^d} g_p(x-y)(u(y)-u(x))\mathrm{d}y$$

since $\int_{\mathbb{T}^d} g_p(x)\mathrm{d}x = 0$. Following the proof of Proposition A.2.6, we write for $x \in \mathbb{T}^d$,

$$\begin{aligned} u_p(x) &= (2\pi)^{-d} \int_{\mathbb{T}^d} \Big(\sum_{k\in\mathbb{Z}^d} \varphi(k/2^p) e^{ik.(x-y)}(u(y)-u(x)) \Big) \mathrm{d}y \\ &= (2\pi)^{-d} \int_{\mathbb{T}^d} (u(y)-u(x)) \Big(\sum_{l\in\mathbb{Z}^d} \int_{\mathbb{R}^d} \varphi(z/2^p) e^{iz.(x-y-2\pi l)} \mathrm{d}z \Big) \mathrm{d}y \end{aligned}$$

with the help of Lemma A.2.7. This yields, for all $x \in \mathbb{T}^d$,

$$\begin{aligned} u_p(x) &= (2\pi)^{-d} \sum_{l\in\mathbb{Z}^d} \int_{\mathbb{T}^d} (u(y)-u(x)) \Big(\int_{\mathbb{R}^d} \varphi(z/2^p) e^{iz.(x-y-2\pi l)} \mathrm{d}z \Big) \mathrm{d}y \\ &= (2\pi)^{-d} \sum_{l\in\mathbb{Z}^d} \int_{2\pi l - x + \mathbb{T}^d} (u(t+x-2\pi l)-u(x)) \Big(\int_{\mathbb{R}^d} \varphi(z/2^p) e^{iz.t} \mathrm{d}z \Big) \mathrm{d}t \\ &\leq (2\pi)^{-d} \int_{\mathbb{R}^d} |u(t+x)-u(x)| \Big| \int_{\mathbb{R}^d} \varphi(z/2^p) e^{iz.t} \mathrm{d}z \Big| \mathrm{d}t, \end{aligned}$$

and eventually

$$\begin{aligned} |u_p(x)| &\leq (2\pi)^{-d/2} \|u\|_{C^\alpha(\mathbb{T}^d)} \int_{\mathbb{R}^d} |t|^\alpha \left|\mathcal{F}(\varphi(./2^p))(t)\right| \mathrm{d}t \\ &\leq (2\pi)^{-d/2} \|u\|_{C^\alpha(\mathbb{T}^d)} 2^{-p\alpha} \int_{\mathbb{R}^d} |2^p t|^\alpha \left|\mathcal{F}(\varphi(.))(2^p t)\right| \mathrm{d}(2^p t) \\ &\leq (2\pi)^{-d/2} \|u\|_{C^\alpha(\mathbb{T}^d)} 2^{-p\alpha} \int_{\mathbb{R}^d} |t|^\alpha \left|\mathcal{F}(\varphi(.))(t)\right| \mathrm{d}(t). \end{aligned}$$

The first part of the result is then proved with $C = (2\pi)^{-d/2} \int_{\mathbb{R}^d} |t|^\alpha |\mathcal{F}(\varphi(.))(t)| \mathrm{d}(t)$ in the case $\alpha \in (0,1)$. The proof of this first part extends to the general case $\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$ by the use of inequalities (A.8) and the fact that $(\partial_x^\alpha u)_p = \partial_x^\alpha u_p$.

To prove the second part of the result, we assume that $\alpha \in (0,1)$ and we note that if there exists $\mathcal{C} > 0$ such that

$$\forall p \geq -1, \qquad \|u_p\|_{L^\infty(\mathbb{T}^d)} \leq \mathcal{C}2^{-p\alpha},$$

then

$$\forall p \geq -1, \qquad \sum_{q \geq p} \|u_q\|_{L^\infty(\mathbb{T}^d)} \leq \mathcal{C} \sum_{q\geq p} 2^{-\alpha q} \leq \big(\mathcal{C} \sum_{k \geq 0} 2^{-k\alpha}\big) 2^{-p\alpha},$$

and $u \in C^0(\mathbb{T}^d)$.

Moreover, for all $p \geq -1$, forall $x, y \in \mathbb{R}^d$,

$$|S_p(x) - S_p(y)| \leq |x-y| \sum_{q=-1}^{p} \|\nabla u_q\|_{L^\infty(\mathbb{T}^d)}.$$

and, with inequality (A.8), there exists $C > 0$ such that for all $q \in \mathbb{N}$,

$$\|\nabla u_q\|_{L^\infty(\mathbb{T}^d)} \leq Cd2^q \|u_q\|_{L^\infty(\mathbb{T}^d)} \leq C\mathcal{C}d2^{q(1-\alpha)}.$$

Hence, after a possible increase of the constant to become $C_1 > 0$, we have for all $x, y \in \mathbb{R}^d$,

$$\forall q \geq -1, \qquad |S_q u(x) - S_q u(y)| \leq C_1 \mathcal{C} |x-y| 2^{q(1-\alpha)}.$$

Therefore, we get for all $x, y \in \mathbb{R}^d$,

$$\forall p \geq -1, \qquad |u(x) - u(y)| \leq C_1 \mathcal{C}|x-y|2^{p(1-\alpha)} + \big(C \sum_{k\geq 0} 2^{-k\alpha}\big)2^{-p\alpha}$$

and taking $p \geq -1$ as the largest integer such that $2^p \leq |x-y|^{-1}$ yields the result in the case $\alpha \in (0,1)$. The proof of the second part of the proposition extends to the general case $\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$ as the proof of the first part did. ∎

A.2.3 Sobolev's embedding theorem

We are now able to prove the following version of Sobolev's embedding theorem for functions on $\mathbb{T}^d$.

Theorem A.2.1 *For all $s \in \mathbb{R}$ such that $s > d/2$ and $s - d/2 \notin \mathbb{N}$, for all $u \in H^s(\mathbb{T}^d)$, we have $u \in C^{s-d/2}(\mathbb{T}^d)$*[3]*.*

Moreover, for all such s, there exists a positive constant $C > 0$ such that

$$\forall u \in H^s(\mathbb{T}^d), \qquad \|u\|_{C^{s-d/2}(\mathbb{T}^d)} \leq C\|u\|_{H^s(\mathbb{T}^d)}.$$

Proof. For $p \in \mathbb{N}$ and $x \in \mathbb{T}^d$, we write

$$u_p(x) = \sum_{k \in \mathbb{Z}^d} \varphi(k/2^p)\hat{u}(k)e^{ik.x}$$

[3] See Definition A.2.7.

Hence, for all $p \in \mathbb{N}$,

$$\begin{aligned}
\|u_p\|_{L^\infty(\mathbb{T}^d)} &\leq \sum_{|k|\leq 2^{p+1}} |\varphi(k/2^p)\hat{u}(k)| \\
&\leq \Big(\sum_{|k|\leq 2^{p+1}} 1^2\Big)^{1/2}\Big(\sum_{k\in\mathbb{Z}^d} \varphi(k/2^p)^2|\hat{u}(k)|^2\Big)^{1/2} \\
&\leq 2^{d/2}2^{pd/2}\|u_p\|_{L^2} \\
&\leq 2^{d/2+s}2^{-p(s-d/2)}c_p(u)\|u\|_{H^s(\mathbb{T}^d)},
\end{aligned}$$

using Proposition A.2.5 and the fact that $u \in H^s(\mathbb{T}^d)$. Such an estimate also holds for $p = -1$. Since for all $p \geq -1$, $c_p(u) \leq 1$, Proposition A.2.7 ensures that $u \in C^{s-d/2}(\mathbb{T}^d)$ and the constant positive constant $C > 0$ (depending on d and s) of the theorem exists.

■

A.3 Applications

This section is devoted to the proof of applications of the Littlewood–Paley decomposition introduced in Section A.2. First of all, we prove that $C^\alpha(\mathbb{T}^d)$ is an algebra and $H^s(\mathbb{T}^d)$ is an algebra provided $s > d/2$ and $s \notin \mathbb{N}$ (Proposition A.3.1). We then prove a version of Meyer's multipliers lemma (Lemma A.3.4) that allows us to prove a non linear composition estimation theorem for functions on $\mathbb{T}^d$ (Proposition A.3.2). As a conclusion, we prove that some non linearities appearing in the non linear Schrödinger equation on $\mathbb{T}^d$ (including the cubic one) are locally Lipschitz (Corollary A.3.1).

A.3.1 Products of functions on $\mathbb{T}^d$

The goal of this subsection is the proof of Proposition A.3.1. We first state and prove two lemmas. The first lemma (Lemma A.3.1) deals with the summation of functions in $C^\alpha(\mathbb{T}^d)$ whose spectrum's localization is known. The second lemma (Lemma A.3.2) deals with the summation of functions in $H^s(\mathbb{T}^d)$ whose spectrum localization is known.

A.3.1.1 Two lemmas

For functions in Hölder spaces, we have the following

Lemma A.3.1 *For all $\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$, there exists a positive constant $\mathcal{C}$ such that for each sequence $(a_q)_{q\geq -1}$ of functions on $\mathbb{T}^d$ such that there exists positive constants C_1, C_2 satisfying*

$$\forall q \geq -1, \qquad \{k \in \mathbb{Z}^d \mid \mathscr{F}(a_q)(k) \neq 0\} \subset \{k \in \mathbb{Z}^d \mid |k| \leq C_1 2^q\},$$

and

$$\forall q \geq -1, \qquad \|a_q\|_{L^\infty(\mathbb{T}^d)} \leq C_2 2^{-\alpha q},$$

we have $u := \sum_{q\geq -1} a_q \in C^\alpha(\mathbb{T}^d)$ *and*

$$\|u\|_{C^\alpha(\mathbb{T}^d)} \leq \mathcal{C}C_2.$$

Proof. If $(a_q)_{q\geq -1}$ is as in the above statement, and if $p, q \in \mathbb{N}$ are such that $C_1 2^q \leq 2^{p-1}$, then we have $\mathscr{F}(a_q)(.)\varphi(./2^p) \equiv 0$. Hence, if we set $N = 2 + E\big(\frac{\ln(C_1)}{\ln 2}\big)$, then

$$\forall p \in \mathbb{N}, \qquad \Big(\sum_{q\geq -1} a_q\Big)_p = \sum_{q\geq p-N} (a_q)_p.$$

Hence, if we set $u = \sum_{q\geq -1} a_q$, then for all $p \in \mathbb{N}$,

$$\begin{aligned} \|u_p\|_{L^\infty(\mathbb{T}^d)} &\leq \sum_{q\geq p-N} \|(a_q)_p\|_{L^\infty(\mathbb{T}^d)} \\ &\leq C \sum_{q\geq p-N} \|a_q\|_{L^\infty(\mathbb{T}^d)} \end{aligned}$$

where $C > 0$ is given by Lemma A.2.8. We deduce for all $p \in \mathbb{N}$,

$$\|u_p\|_{L^\infty(\mathbb{T}^d)} \leq CC_2 \sum_{q\geq p-N} 2^{-\alpha q} \leq CC^3 2^{-p\alpha},$$

where $C_3 > 0$ depends on both C_1 and C_2. Proposition A.2.7 and the fact that

$$\begin{aligned} \|u_{-1}\|_{L^\infty(\mathbb{T}^d)} &\leq \sum_{q\geq -1} \|(a_q)_{-1}\|_{L^\infty(\mathbb{T}^d)} \\ &\leq C \sum_{q\geq -1} \|(a_q)\|_{L^\infty(\mathbb{T}^d)} \\ &\leq CC_2 \sum_{q\geq -1} 2^{-\alpha q} \\ &\leq CC_3, \end{aligned}$$

after a possible increase of C_3, ensure that $u \in C^\alpha$ and there exists a positive constant $\mathcal{C}$ satisfying the conclusion of the result. ■

For functions in Sobolev spaces, one has the following lemma, which is the counterpart of Lemma A.3.1.

Lemma A.3.2 *For all* $s \in \mathbb{R}^+$, *there exists a positive constant* $\mathcal{C}$ *such that for each sequence* $(a_q)_{q\geq -1}$ *of functions on* $\mathbb{T}^d$ *such that there exists positive constants* C_1, C_2 *and a positive sequence* $(c_q)_{q\geq -1}$ *satisfying*

$$\forall q \geq -1, \qquad \{k \in \mathbb{Z}^d \mid \mathscr{F}(a_q)(k) \neq 0\} \subset \{k \in \mathbb{Z}^d \mid |k| \leq C_1 2^q\},$$

$$\forall q \geq -1, \qquad \|a_q\|_{L^2(\mathbb{T}^d)} \leq C_2 c_q 2^{-sq},$$

and $\sum_{q\geq -1} c_q^2 \leq 1$, *we have* $u := \sum_{q\geq -1} a_q \in H^s(\mathbb{T}^d)$ *and*

$$\|u\|_{H^s(\mathbb{T}^d)} \leq \mathcal{C}C_2.$$

Proof. As in the proof of the previous lemma, we have

$$\|u_p\|_{L^\infty(\mathbb{T}^d)} \le C \sum_{q\ge p-N} \|a_q\|_{L^\infty(\mathbb{T}^d)}$$

where C is given by Lemma A.2.8. Hence, for all $p \in \mathbb{N}$,

$$\begin{aligned}\|u_p\|_{L^2(\mathbb{T}^d)} &\le CC_2 \sum_{q\ge p-N} c_q 2^{-sq}\\ &\le CC_2 2^{-ps} \sum_{q\ge p-N} c_q 2^{-s(q-p)/2} 2^{-s(q-p)/2}\\ &\le CC_2 2^{-ps}\Big(\sum_{q\ge p-N} 2^{-s(q-p)}\Big)^{1/2}\Big(\sum_{q\ge p-N} c_q^2 2^{-s(q-p)}\Big)^{1/2}.\end{aligned}$$

Note that

$$\sum_{p\ge -1}\Big(\sum_{q\ge p-N} c_q^2 2^{-s(q-p)}\Big)^{2/2} < +\infty.$$

Since we also have

$$\|u_{-1}\|_{L^2(\mathbb{T}^d)} \le CC_2 \sum_{q\ge -1} c_q 2^{-qs},$$

Proposition A.2.5 ensures that the result holds true. ∎

These two lemmas are interesting since they provide, in a rather basic framework, the idea that the spectrum localization of the terms of a series of functions can provide useful information on the sum of the series. Note that Lemma A.3.3 is a generalization of Lemma A.3.2.

A.3.1.2 Products of functions on $\mathbb{T}^d$

We are now able to state and prove the main result of this subsection, dealing with products of functions in Hölder and Sobolev spaces on $\mathbb{T}^d$.

Proposition A.3.1 *For all $\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$, there exists a positive constant C such that for all $u, v \in C^\alpha(\mathbb{T}^d)$,*

$$\|uv\|_{C^\alpha(\mathbb{T}^d)} \le C\big(\|u\|_{C^\alpha(\mathbb{T}^d)}\|v\|_{L^\infty(\mathbb{T}^d)} + \|u\|_{L^\infty(\mathbb{T}^d)}\|v\|_{C^\alpha(\mathbb{T}^d)}\big).$$

Moreover, for all $s \in \mathbb{R}^+$, there exists a positive constant C such that for all $u, v \in L^\infty(\mathbb{T}^d) \cap H^s(\mathbb{T}^d)$, $uv \in L^\infty(\mathbb{T}^d) \cap H^s(\mathbb{T}^d)$ and

$$\|uv\|_{H^s} \le C\big(\|u\|_{L^\infty(\mathbb{T}^d)}\|v\|_{H^s(\mathbb{T}^d)} + \|u\|_{H^s(\mathbb{T}^d)}\|v\|_{L^\infty(\mathbb{T}^d)}\big).$$

Remark A.3.1 *The two estimates of Proposition A.3.1 are tame estimates since the power of the highest derivative involved in the right hand sides is one in both cases.*

Proof. Note that in both cases, we have

$$\begin{aligned} uv &= \sum_{p,q\geq -1} u_p v_q \\ &= \sum_{q\geq -1}\sum_{p=-1}^{q} u_p v_q + \sum_{q\geq -1}\sum_{p\geq q+1} u_p v_q \\ &= \sum_{q\geq -1} S_p(u) v_q + \sum_{p\geq 0}\sum_{q\leq p-1} v_q u_p \\ &= \sum_{q\geq -1} S_p(u) v_q + \sum_{p\geq 0} S_{p-1}(v) u_p. \end{aligned}$$

For all $q \geq -1$, we have

$$\|S_q(u)v_q\|_{L^\infty(\mathbb{T}^d)} \leq \|S_q(u)\|_{L^\infty(\mathbb{T}^d)} \|v_q\|_{L^\infty(\mathbb{T}^d)} \leq C\|u\|_{L^\infty(\mathbb{T}^d)} \|v\|_{C^\alpha(\mathbb{T}^d)} 2^{-\alpha p},$$

with the help of Lemma A.2.9 and Proposition A.2.7. A symmetric estimate holds for $u_q S_{q-1}(v)$. Proposition A.2.7 ensures that the first part of the result holds true. Similarly, we have for all $q \geq -1$,

$$\|S_q(u)v_q\|_{l^2(\mathbb{T}^d)} \leq \|S_q(u)\|_{L^\infty(\mathbb{T}^d)} \|v_q\|_{L^2(\mathbb{T}^d)} \leq C\|u\|_{L^\infty(\mathbb{T}^d)} \|v\|_{L^2(\mathbb{T}^d)} c_p(v) 2^{-\alpha s},$$

and a symmetric result for $u_q S_{q-1}(v)$, with the help of Lemma A.2.8 and Proposition A.2.5. Proposition A.2.5 allows us to conclude. ∎

A.3.2 Estimates for the composition of functions

This subsection is devoted to the statement and proof of the non linear estimate for the composition of functions on $\mathbb{T}^d$ of Proposition A.3.2 and its application to some non linearities of the Schrödinger equation on $\mathbb{T}^d$ (Corollary A.3.1). First of all, we generalize Lemma A.3.2 in order to prove a version of Meyer's multiplier lemma (Lemma A.3.4). We then apply this result to prove Proposition A.3.2.

A.3.2.1 A generalization of Lemma A.3.2

Lemma A.3.3 *For all $s \in \mathbb{R}^+$, there exists a positive constant $\mathcal{C}$ such that for all double sequence $(a_{p,k})_{p,k\geq -1}$ of functions on $\mathbb{T}^d$ such that there exists a positive constant C_1 and two positive sequences $(C_2(k))_{k\geq -1}$ and $(c_p)_{p\geq -1}$ satisfying*

$$\forall k,p \geq -1, \qquad \{t \in \mathbb{Z}^d \mid \mathscr{F}(a_{p,k})(t) \neq 0\} \subset \{t \in \mathbb{Z}^d \mid |t| \leq C_1 2^{p+k}\}$$

$$\forall k,p \geq -1, \qquad \|a_{p,k}\|_{L^2(\mathbb{T}^d)} \leq C_2(k) 2^{-s(p+k)} c_p,$$

and $\sum_{p\geq -1} c_p^2 \leq 1$, we have for all $k \geq -1$, $\sum_{p\geq -1} a_{p,k} \in H^s(\mathbb{T}^d)$ and

$$\forall k \geq -1, \qquad \|\sum_{p\geq -1} a_{p,k}\|_{L^2(\mathbb{T}^d)} \leq \mathcal{C} C(k).$$

Proof. As in the proof of Lemma A.3.2, note that if $p,k \geq -1$ and $q \in \mathbb{N}$ are such that

$$C_1 2^{p+k} \leq 2^{q-1},$$

then

$$\forall t \in \mathbb{Z}^d, \qquad \mathscr{F}\big((a_{p,k})_q\big)(t) = 0.$$

Hence, if we denote $N = 1 + \frac{\ln C_1}{\ln 2}$ and set

$$\forall k \geq -1, \qquad M_k(a) := \sum_{p \geq -1} a_{p,k},$$

then for all $k \geq -1$ and all $q \in \mathbb{N}$, we have

$$(M_k(a))_q = \sum_{p \geq q-k-N} (a_{p,k})_q.$$

As in the proof of Lemma A.3.2, we derive that

$$\|(M_k(a))_q\|_{L^2(\mathbb{T}^d)} \leq \sum_{p \geq q-k-N} \|(a_{p,k})_q\|_{L^2(\mathbb{T}^d)} \leq \sum_{p \geq q-k-N} \|a_{p,k}\|_{L^2(\mathbb{T}^d)},$$

with the help of Lemma A.2.4. Hence, using the hypothesis, we derive

$$\begin{aligned}
\|(a_{p,k})_q\|_{L^2(\mathbb{T}^d)} &\leq C_2(k) \sum_{p \geq q-k-N} 2^{-s(p+k)} c_p \\
&\leq C_2(k) \sum_{p \geq q-N} 2^{-sp} c_{p-k} \\
&\leq C_2(k) 2^{-sq} \sum_{p \geq q-N} 2^{-s(p-q)} c_{p_k} \\
&\leq C_2(k) 2^{-sq} \Big(\sum_{p \geq q-N} 2^{-s(p-q)} \Big)^{1/2} \Big(\sum_{p \geq q-N} 2^{-s(p-q)} c_{p-k}^2 \Big)^{1/2}.
\end{aligned}$$

Note that

$$\sum_{q \geq -1} \sum_{p \geq q-N} 2^{-s(p-q)} c_{p_k}^2 \leq 2^{Ns} \sum_p c_{p-k}^2 \sum_{q \geq 0} 2^{-sq}$$

is bounded with bound independent of k. Moreover, a similar calculus yields a constant C_s satisfying

$$\forall p,k \geq -1, \qquad \|(M_k(a))_{-1}\|_{L^2(\mathbb{T}^d)} \leq C_2(k) C_s.$$

Proposition A.2.5 allows us to conclude. ∎

A.3.2.2 Meyer's multipliers lemma

Lemma A.3.4 *Let $\delta \in \mathbb{R}$ and $(m_p)_{p \geq -1}$ be a sequence of functions of $C^\infty(\mathbb{T}^d)$ such that for some sequence $(C_k)_{k \geq -1}$ of non-negative numbers we have*

$$\forall p \geq -1, \forall k \in \mathbb{N}, \qquad \sum_{|\alpha|=k} \|\partial_x^\alpha m_p\|_{L^\infty(\mathbb{T}^d)} \leq C_k 2^{p(k+\delta)}.$$

There exists a positive constant C depending only on $(C_k)_{k\leq E(s-\delta)+1}$ such that for all $u \in H^s(\mathbb{T}^d)$, $Mu := \sum_{p\geq -1} m_p u_p \in H^{s-\delta}(\mathbb{T}^d)$ and

$$\|Mu\|_{H^{s-\delta}(\mathbb{T}^d)} \leq C\|u\|_{H^s(\mathbb{T}^d)}.$$

Proof. We set $C > 4$ and define for all $p \geq -1$ and all $k \in \mathbb{N}$,

$$\forall l \in \mathbb{Z}^d, \qquad \mathscr{F}(m_{p,k})(l) = \varphi(2^{-p}l/(C2^k))\mathscr{F}(m_p)(l),$$

and $\mathscr{F}(m_{p,-1})(l) = \psi(2^{-p}l/C)\mathscr{F}(m_p)(l)$. Then, we set

$$\forall k \geq -1, \qquad M_k(u) := \sum_{p\geq -1} m_{p,k}u_p.$$

Note that for all $p \geq 0$,

$$\{k \in \mathbb{Z}^d \mid \mathscr{F}(m_{p,-1}u_p)(k) \neq 0\} \subset \{k \in \mathbb{Z}^d \mid |k| \leq C2^p + (2^{p+1} - 2^{p-1}) = (C+3)2^p\},$$

and this is also true for $p = -1$. Moreover, for all $p \geq -1$,

$$\begin{aligned} \|m_{p,-1}u_p\|_{L^2(\mathbb{T}^d)} &\leq \|m_{p,-1}\|_{L^\infty(\mathbb{T}^d)} \|u_p\|_{L^2(\mathbb{T}^d)} \\ &\leq \|m_{p,-1}\|_{L^\infty(\mathbb{T}^d)} 2^s c_p 2^{-ps}\|u\|_{H^s(\mathbb{T}^d)}, \end{aligned}$$

with the help of Proposition A.2.5. Lemma A.2.8 with $\alpha = 0$ ensures that there exists $C > 0$ such that

$$\forall p \geq -1, \qquad \|m_{p,-1}\|_{L^\infty(\mathbb{T}^d)} \leq C\|m_p\|_{L^\infty(\mathbb{T}^d)}.$$

The hypothesis on the derivatives of the functions m_p ensures that

$$\forall p \geq -1, \qquad \|m_{p,-1}u_p\|_{L^\infty(\mathbb{T}^d)} \leq CC_0c_p2^s2^{-p(s-\delta)}\|u\|_{H^s(\mathbb{T}^d)}.$$

Proposition A.2.5 now ensures that $M_{-1}u \in H^{s-\delta}$ since $s > \delta$.
For all $k \in \mathbb{N}$ and all $p \geq -1$,

$$\{t \in \mathbb{Z}^d \mid \mathscr{F}(m_{p,k}u_p)(t) \neq 0\} \subset \{t \in \mathbb{Z}^d \mid |t| \leq 4C2^{k+p}\}.$$

Moreover, we have, as previously,

$$\|m_{p,k}u_p\|_{L^2(\mathbb{T}^d)} \leq \|m_{p,k}\|_{L^\infty(\mathbb{T}^d)} 2^s 2^{-ps} c_p \|u\|_{H^s(\mathbb{T}^d)},$$

with Proposition A.2.5.
A straightforward modification of Lemma A.2.10 ensures that for all l, there exists a positive constant $C(l)$ such that

$$\forall p \geq -1, \forall k \in \mathbb{N}, \qquad \|m_{p,k}\|_{L^\infty(\mathbb{T}^d)} \leq C(l) \sum_{|\alpha|=l} \|\partial_x^\alpha m_{p,k}\|_{L^\infty(\mathbb{T}^d)} 2^{-l(p+k)}.$$

Using the hypothesis, this yieds for all $n \in \mathbb{N}$,

$$\forall p \geq -1, \forall k \in \mathbb{N}, \qquad \|m_{p,k}\|_{L^\infty(\mathbb{T}^d)} \leq C(l)C_l 2^{-kl}2^{p\delta}.$$

Plugging these estimations one in another yields for all $p \geq -1$ and all $k, l \in \mathbb{N}$

$$\begin{aligned} \|m_{p,k}u_p\|_{L^2(\mathbb{T}^d)} &\leq C(l)C_l 2^s \|u\|_{H^s(\mathbb{T}^d)} 2^{-kl} c_p 2^{-p(s-\delta)} \\ &\leq C(l)C_l 2^s \|u\|_{H^s(\mathbb{T}^d)} 2^{k(s-l-\delta)} c_p 2^{-(p+k)(s-\delta)}. \end{aligned}$$

This estimation and Lemma A.3.3 ensure that there exists a positive constant A independant on k such that for all $k \in \mathbb{N}$, $M_k(u) \in H^{s-\delta}(\mathbb{T}^d)$ and

$$\forall k \in \mathbb{N}, \qquad \|M_k(u)\|_{H^{s-\delta}(\mathbb{T}^d)} \leq AC_l 2^{k(s-l-\delta)} \|u\|_{H^s(\mathbb{T}^d)}.$$

It is now clear that the sequence $(M_k)_{k\geq -1}$ converges in norm to M in the space of linear continuous operators from $H^s(\mathbb{T}^d)$ to $H^{s-\delta}(\mathbb{T}^d)$ and that its norm in this space depends only on $C_0 + C_l$ (and not on the C_k's with bigger k's). ■

A.3.2.3 Application to the composition of functions

Using Lemma A.3.3 and Lemma A.3.4, we are able to prove the following non linear estimate for functions on $\mathbb{T}^d$.

Proposition A.3.2 *For all function $F \in C^\infty(\mathbb{R}^2, \mathbb{C})$ with $F(0,0) = 0$ and all $R, s > 0$, there exists a positive constant C_R such that for all $u, v \in L^\infty(\mathbb{T}^d, \mathbb{R}) \cap H^s(\mathbb{T}^d)$ with $\|u\|_{L^\infty(\mathbb{T}^d)} \leq R$ and $\|v\|_{L^\infty(\mathbb{T}^d)} \leq R$, we have $F(u,v) \in L^\infty(\mathbb{T}^d) \cap H^s(\mathbb{T}^d)$, with*

$$\|F(u,v)\|_{H^s(\mathbb{T}^d)} \leq C_R\big(\|u\|_{H^s(\mathbb{T}^d)} + \|v\|_{H^s(\mathbb{T}^d)}\big).$$

Remark A.3.2 *Note that the estimation of Proposition A.3.2 also is a tame estimate.*

Proof. First of all, we write $F(u,v)$ as the following series

$$F(u,v) = F(S_0u, S_0v) + \sum_{p\geq 0} \Big(F(S_{p+1}u, S_{p+1}v) - F(S_pu, S_pv)\Big). \tag{A.9}$$

Let us set $r = \lfloor s \rfloor + 1$. For the first in that sum, let us consider the partial derivatives $\partial_x^\alpha F(S_0u, S_0v)$ when $\alpha \in \mathbb{N}^d$ satisfies $|\alpha| \leq r$. If $\alpha = 0$, since $F(0,0) = 0$ and we have for some positive constant C_1

$$\|S_0u\|_{L^\infty(\mathbb{T}^d)} \leq C_1\|u\|_{L^\infty(\mathbb{T}^d)} \qquad \text{and} \qquad \|S_0v\|_{L^\infty(\mathbb{T}^d)} \leq C_1\|v\|_{L^\infty(\mathbb{T}^d)}$$

with the help of Lemma A.2.9, we derive that

$$\forall x \in \mathbb{T}^d, \qquad |F(S_0u, S_0v)(x)| \leq C_2\big(|S_0u(x)| + |S_0v(x)|\big)$$

for some positive constant C_2 depending only on F and R[4]. We derive that for all $u, v \in L^\infty(\mathbb{T}^d) \cap H^s(\mathbb{T}^d)$ such that $\|u\|_{L^\infty(\mathbb{T}^d)} \leq R$ and $\|v\|_{L^\infty(\mathbb{T}^d)} \leq R$, we have

$$\begin{aligned} \|F(S_0u, S_0v)\|_{L^2(\mathbb{T}^d)} &\leq C_2\sqrt{2}\big(\|S_0u\|_{L^2(\mathbb{T}^d)} + \|S_0v\|_{L^2(\mathbb{T}^d)}\big) \\ &\leq C_2\sqrt{2}\big(\|u\|_{L^2(\mathbb{T}^d)} + \|v\|_{L^2(\mathbb{T}^d)}\big) \end{aligned}$$

[4]Note that C_2 also depends on C_1 and hence on the choice of the function Ψ for the Littlewood-Paley decomposition.

with the help of Lemma A.2.4.
If $\alpha \in \mathbb{N}^d$ is such that $|\alpha| \geq 1$, then $\partial_x^\alpha F(S_0u, S_0v)$ is a sum of the form:

$$\sum_{\substack{\beta\in\mathbb{N}^2 \\ 1\leq|\beta|\leq|\alpha|}} (*)\frac{\partial^{\beta_1+\beta_2}F}{(\partial u)^{\beta_1}(\partial v)^{\beta_2}}(S_0u, S_0v)\,(\partial_x^{\gamma_1}S_0u)\dots(\partial_x^{\gamma_{\beta_1}}S_0u)\,(\partial_x^{\eta_1}S_0v)\dots(\partial_x^{\eta_{\beta_2}}S_0v),$$

with $\gamma_1 + \cdots + \gamma_{\beta_1} + \eta_1 + \cdots + \eta_{\beta_1} = \alpha$ and for all i, $|\gamma_i| \geq 1$ and $|\eta_i| \geq 1$. The symbol $(*)$ in the previous sum stands for some numerical coefficients independant on u, v and F. Note that for all $\beta \in \mathbb{N}^2$ and $x \in \mathbb{T}^d$, $\big|\big(\partial_u^{\beta_1}\partial_v^{\beta_2}F\big)(S_0u, S_0v)(x)\big|$ is bounded by

$$\sup\Big\{\Big|\frac{\partial^{\beta_1+\beta_2}F}{(\partial u)^{\beta_1}(\partial v)^{\beta_2}}(U,V)\Big| \quad , \quad U,V\in\mathbb{R} \text{ s.t. } |U|,|V| \leq C_1R\Big\}.$$

We derive that there exists a positive constant C_3 depending only on F, R and s[5] such that for all $\beta \in \mathbb{N}^2$ satisfying $1 \leq |\beta| \leq r$ and all $u, v \in L^\infty(\mathbb{T}^d) \cap H^s(\mathbb{T}^d)$ satisfying $\|u\|_{L^\infty(\mathbb{T}^d)} \leq R$ and $\|v\|_{L^\infty(\mathbb{T}^d)} \leq R$, we have

$$\|\big(\partial_u^{\beta_1}\partial_v^{\beta_2}F\big)(S_0u, S_0v)\|_{L^\infty(\mathbb{T}^d)} \leq C_3.$$

Moreover, with the help of Lemma A.2.4 and Lemma A.2.9, there exists a positive constant C_4 depending only on s such that for all $\gamma, \eta \in \mathbb{N}^d$ satisfying $1 \leq |\gamma| \leq r$ and $1 \leq |\eta| \leq r$, we have

$$\|\partial_x^\gamma S_0u\|_{L^\infty(\mathbb{T}^d)} \leq C_4R \qquad \text{and} \qquad \|\partial_x^\eta S_0v\|_{L^\infty(\mathbb{T}^d)} \leq C_4R,$$

and

$$\|\partial_x^\gamma S_0u\|_{L^2(\mathbb{T}^d)} \leq C_4\|u\|_{L^2(\mathbb{T}^d)} \qquad \text{and} \qquad \|\partial_x^\eta S_0v\|_{L^2(\mathbb{T}^d)} \leq C_4\|u\|_{L^2(\mathbb{T}^d)}.$$

As a consequence, there exists a positive constant C_5 depending only on F, R and s such that for all $\alpha \in \mathbb{N}^d$ satisfying $1 \leq |\alpha| \leq r$ and all $u, v \in L^\infty(\mathbb{T}^d) \cap H^s(\mathbb{T}^d)$ satisfying $\|u\|_{L^\infty(\mathbb{T}^d)} \leq R$ and $\|v\|_{L^\infty(\mathbb{T}^d)} \leq R$, we have

$$\|\partial_x^\alpha F(S_0u, S_0v)\|_{L^2(\mathbb{T}^d)} \leq C_5\big(\|u\|_{L^2(\mathbb{T}^d)} + \|v\|_{L^2(\mathbb{T}^d)}\big)$$

Eventually, with the help of Lemma (A.2.5), we derive that there exists a positive constant C_6 depending only on F, R and s such that

$$\begin{aligned}\|F(S_0u, S_0v)\|_{H^s(\mathbb{T}^d)} &\leq \|F(S_0u, S_0v)\|_{H^r(\mathbb{T}^d)}\\ &\leq C_6\big(\|u\|_{L^2(\mathbb{T}^d)} + \|v\|_{L^2(\mathbb{T}^d)}\big)\\ &\leq C_6\big(\|u\|_{H^s(\mathbb{T}^d)} + \|v\|_{H(\mathbb{T}^d)}\big).\end{aligned}$$

For the second term in (A.9), we write for all $p \in \mathbb{N}$, that $F(S_{p+1}u, S_{p+1}v) - F(S_pu, S_pv)$ is equal to

$$\int_0^1 \Big(\frac{\partial F}{\partial u}(S_pu + tu_{p+1}, S_pv + tv_{p+1})u_{p+1} + \frac{\partial F}{\partial v}(S_pu + tu_{p+1}, S_pv + tv_{p+1})v_{p+1}\Big)\mathrm{d}t.$$

[5] The same remark holds for C_3 and the others constants to appear in the proof.

In order to use Lemma A.3.4, we prove that the derivatives of the functions on $\mathbb{T}^d$

$$\int_0^1 \frac{\partial F}{\partial u}(S_p u + t u_{p+1}, S_p v + t v_{p+1})\mathrm{d}t \quad \text{and} \quad \int_0^1 \frac{\partial F}{\partial v}(S_p u + t u_{p+1}, S_p v + t v_{p+1})\mathrm{d}t$$

satisfy suitable estimates in L^∞-norm. For brevity, we only show it for the functions $G(S_p u, S_p v) := \frac{\partial F}{\partial u}(S_p u, S_p v)$. Assume $\alpha \in \mathbb{N}^d$ is given. If $\alpha = 0$, then the estimate is obtained as below:

$$\forall p \geq -1, \qquad \|G(S_p u, S_p v)\|_{L^\infty(\mathbb{T}^d)} \leq C_7$$

for some positive constant C_7 depending on F, and R, since $(S_p(u), S_p(v))(x)$ stays in a compact subset of $\mathbb{R}^2$ independant on p by Lemma A.2.9 and G is a smooth function on $\mathbb{R}^2$ (since F is).
Is $|\alpha| \geq 1$, then, as below, $\partial_x^\alpha G(S_p u, S_p v)$ is a sum of the form

$$\sum_{\substack{\beta \in \mathbb{N}^2 \\ 1 \leq |\beta| \leq |\alpha|}} (*) \frac{\partial^{\beta_1+\beta_2} G}{(\partial u)^{\beta_1} (\partial v)^{\beta_2}} (S_p u, S_p v)\, (\partial_x^{\gamma_1} S_p u) \ldots (\partial_x^{\gamma_{\beta_1}} S_p u)\, (\partial_x^{\eta_1} S_p v) \ldots (\partial_x^{\eta_{\beta_2}} S_p v).$$

with $\gamma_1 + \cdots + \gamma_{\beta_1} + \eta_1 + \cdots + \eta_{\beta_1} = \alpha$ and for all i, $|\gamma_i| \geq 1$ and $|\eta_i| \geq 1$. On one hand, Lemma A.2.9 ensures that for all $p \in \mathbb{N}$, and all i,

$$\|\partial_x^{\gamma_i} S_p u\|_{L^\infty(\mathbb{T}^d)} \leq C(\alpha) 2^{p|\gamma_i|} \|u\|_{L^\infty(\mathbb{T}^d)}$$

and

$$\|\partial_x^{\eta_i} S_p v\|_{L^\infty(\mathbb{T}^d)} \leq C(\alpha) 2^{p|\eta_i|} \|v\|_{L^\infty(\mathbb{T}^d)}$$

since $|\gamma_i|, |\eta_i| \leq |\alpha|$. On the other hand, as below, there exists a positive constant C_8 depending only on F, R and α such that

$$\forall \beta \in \mathbb{N}^2 \text{ s.t. } |\beta| \leq |\alpha|, \|\big(\partial_u^{\beta_1} \partial_v^{\beta_2} G\big)(S_p u, S_p v)\|_{L^\infty(\mathbb{T}^d)} \leq C_8.$$

Since the equality $\gamma_1 + \cdots + \gamma_{\beta_1} + \eta_1 + \cdots + \eta_{\beta_1} = \alpha$ implies that $|\gamma_1| + \cdots + |\gamma_{\beta_1}| + |\eta_1| + \cdots + |\eta_{\beta_1}| = |\alpha|$, we derive that every term in the above decomposition of $\partial_x^\alpha G(S_p u, S_p v)$ is bounded in L^∞-norm by a constant depending on F, R and α times $2^{p|\gamma_1|+\cdots+p|\gamma_{\beta_1}|+p|\eta_1|+\cdots+p|\eta_{\beta_2}|} = 2^{p|\alpha|}$. As a conclusion, for all $l \in \mathbb{N}^*$, there exists a positive constant C_l depending only on F, R and l such that

$$\forall p \in \mathbb{N}, \qquad \sum_{|\alpha|=l} \|\partial_x^\alpha G(S_p u, S_p v)\|_{L^\infty(\mathbb{T}^d)} \leq C_l 2^{pl}.$$

so that Lemma A.3.4 allows us to conclude. ∎

A.3.2.4 A Lipschitz-type estimate

Using the non linear estimate for Proposition A.3.2, we are able to prove that some non linearities appearing in the non linear Schrödinger equation on $\mathbb{T}^d$ are locally Lipschitz for suitable Sobolev norms. We start with the following

Lemma A.3.5 *Let $s > d/2$ be given such that $s - d/2 \notin \mathbb{N}$ and let $f : \mathbb{R}^2 \to \mathbb{C}$ be a* *smooth function with*

$$f(0,0) = \partial_{x_1} f(0,0) = \partial_{x_2} f(0,0) = 0.$$

For all $R > 0$, there exists a positive constant L_R satisfying for all $u_1, u_2, v_1, v_2 \in H^s(\mathbb{T}^d, \mathbb{R})$ such that for all $i \in \{1,2\}$, $\|u_i\|_{H^s(\mathbb{T}^d)}, \|v_i\|_{H^s(\mathbb{T}^d)} \le R$,

$$\|f(u_1,u_2) - f(v_1,v_2)\|_{H^s(\mathbb{T}^d)} \le L_R \Big(\|u_1 - v_1\|_{H^s(\mathbb{T}^d)} + \|u_2 - v_2\|_{H^s(\mathbb{T}^d)} \Big).$$

Proof. Let us denote $u = (u_1, u_2)$ and $v = (v_1, v_2)$. Since f is smooth, one has the following expression for the difference $f(u) - f(v)$:

$$\Big(\int_0^1 \partial_{x_1} f(v + t(u-v)) \mathrm{d}t \Big)(u_1 - v_1) + \Big(\int_0^1 \partial_{x_2} f(v + t(u-v)) \mathrm{d}t \Big)(u_2 - v_2).$$

Theorem A.2.1 ensures that for all $t \in [0,1]$ and all $i \in \{1,2\}$,

$$\|v_1 + t(u_1 - v_1)\|_{L^\infty(\mathbb{T}^d)} \le C\|v_1 + t(u_1 - v_1)\|_{H^s(\mathbb{T}^d)} \le 2CR$$

for some positive constant C (depending on s). Theorem A.3.2 applied to the two first derivatives of f leads to the existence of a positive constant M such that for all u, v satisfying the hypothesis of the lemma and all $t \in [0,1]$, one has

$$\|\partial_{x_1} f(v + t(u-v))\|_{H^s(\mathbb{T}^d)} \le M(CR + CR)$$

and

$$\|\partial_{x_2} f(v + t(u-v))\|_{H^s(\mathbb{T}^d)} \le M(CR + CR).$$

After integration, and using Proposition A.3.1 and then Theorem A.2.1, we derive that there exists a positive constant C_1 depending on f, s and R such that

$$\Big\| \Big(\int_0^1 \partial_{x_1} f(v + t(u-v)) \mathrm{d}t \Big)(u_1 - v_1) \Big\|_{H^s(\mathbb{T}^d)} \le C_1 \|u_1 - v_1\|_{H^s(\mathbb{T}^d)}$$

and

$$\Big\| \Big(\int_0^1 \partial_{x_2} f(v + t(u-v)) \mathrm{d}t \Big)(u_2 - v_2) \Big\|_{H^s(\mathbb{T}^d)} \le C_1 \|u_2 - v_2\|_{H^s(\mathbb{T}^d)}.$$

The result is then proved with $L_R = 2C_1$. ■

This lemma allows us to prove the following

Corollary A.3.1 *For all $R > 0$, all $\sigma \in \mathbb{N}^\star$ and all $s > d/2$ such that $s - d/2 \notin \mathbb{N}$, there exists a positive constant M such that for all $u, v \in H^s(\mathbb{T}^d)$ satisfying $\|u\|_{H^s(\mathbb{T}^d)} \le R$ and $\|v\|_{H^s(\mathbb{T}^d)} \le R$, one has*

$$\||u|^{2\sigma} u - |v|^{2\sigma} v\|_{H^s(\mathbb{T}^d)} \le M \|u - v\|_{H^s(\mathbb{T}^d)}.$$

Proof. Consider the function defined for $x, y \in \mathbb{R}$ by

$$f(x,y) = (x^2 + y^2)^\sigma x + i(x^2 + y^2)^\sigma y.$$

The previous lemma ensures that there exists a positive constant L such that

$$\begin{aligned} \|f(u)-f(v)\|^2_{H^s(\mathbb{T}^d)} &\leq L^2\Big(\|\Re(u-v)\|_{H^s(\mathbb{T}^d)}+\|\Im(u-v)\|_{H^s(\mathbb{T}^d)}\Big)^2 \\ &\leq 2L^2\Big(\|\Re(u-v)\|^2_{H^s(\mathbb{T}^d)}+\|\Im(u-v)\|^2_{H^s(\mathbb{T}^d)}\Big) \end{aligned}$$

Hence the result is proved with $M = 2L$ with the help of Lemma A.3.6. ■

In dimension $d = 1$ for example, Corollary A.3.1 ensures that the cubic non linearity ($\sigma = 1$) of the Schrödinger equation on $\mathbb{T}^d$ is locally Lipschitz on any of the $H^s(\mathbb{T}^d)$ with $s \in \mathbb{N}^\star$.

Lemma A.3.6 *For all $f \in H^s(\mathbb{T}^d)$, one has*

$$\|\Re(f)\|_{H^s(\mathbb{T}^d)} \leq \|f\|_{H^s(\mathbb{T}^d)} \qquad \textit{and} \qquad \|\Im(f)\|_{H^s(\mathbb{T}^d)} \leq \|f\|_{H^s(\mathbb{T}^d)}\,.$$

Proof. Let us denote

$$2u = 2\Re(f) = f + \bar{f} \qquad \text{and} \qquad 2iv = 2i\Im(f) = f - \bar{f}.$$

For all $k \in \mathbb{Z}^d$, one has

$$\begin{aligned} 4|\mathscr{F}(u)(k)|^2 &= \big(|\mathscr{F}(f)(k)| + |\mathscr{F}(\bar{f})(k)|\big)^2 \\ &\leq 2|\mathscr{F}(f)(k)|^2 + 2|\mathscr{F}(\bar{f})(k)|^2 \\ &\leq 2|\mathscr{F}(f)(k)|^2 + 2|\mathscr{F}(f)(-k)|^2. \end{aligned}$$

Hence, for all $k \in \mathbb{Z}$,

$$|\mathscr{F}(u)(k)|^2 + |\mathscr{F}(u)(-k)|^2 \leq |\mathscr{F}(f)(k)|^2 + |\mathscr{F}(f)(-k)|^2.$$

This proves the left inequality. The same kind of calculations for v yields for all $k \in \mathbb{Z}^d$,

$$|\mathscr{F}(v)(k)|^2 + |\mathscr{F}(v)(-k)|^2 \leq |\mathscr{F}(f)(k)|^2 + |\mathscr{F}(f)(-k)|^2,$$

so that the right inequality also holds true. ■

Bibliography

[AG91] A. Alinhac and P. Gérard. *Opérateurs pseudo-différentiels et théorie de Nash-Moser*. InterÉditions/Éditions du CNRS, 1991.

[BBD02] C. Besse, B. Bidégaray, and S. Descombes. Order estimates in time of splitting methods for the nonlinear Schrödinger equation. *SIAM J. Numer. Anal. 40, no. 5*, pages 26–40, 2002.

[BG06] D. Bambusi and B. Grébert. Birkhoff normal form for PDEs with tame modulus. *Duke Math. J. 135, (3)*, pages 507–567, 2006.

[BIS07] H. Berland, A.L. Islas, and C.M. Schober. Conservation of phase space properties using exponential integrators on the cubic Schrödinger equation. *Journal of Computational Physics, 255(1)*, pages 284–299, 2007.

[Bou93a] J. Bourgain. Exponential sums and nonlinear Schrödinger equations. *Geom. and Funct. Anal. 3*, pages 157–178, 1993.

[Bou93b] J. Bourgain. Fourier restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations I, Schrödinger equations. *Geom. and Funct. Anal. 3*, pages 107–156, 1993.

[Bou94] J. Bourgain. Periodic nonlinear Schrödinger equation and invariant measures. *Commun. Math. Phys. 166*, pages 1–26, 1994.

[BS05] H. Berland and B. Skaflestad. Solving the nonlinear Schrödinger equation using exponential integrators. In *Proceedings SIMS 2005 46th Conference on Simulation and Modeling, Trondheim, Norway*, 2005.

[Can06] B. Cano. Conserved quantities of some Hamiltonian wave equations after full discretization. *Numer. Math. 103*, pages 197–223, 2006.

[CHL06] D. Cohen, E. Hairer, and C. Lubich. Long-time analysis of nonlinearly perturbed wave equations via modulated Fourier expansions. *to appear in Arch. Ration. Mech. Anal.*, 2006.

[CHL07] D. Cohen, E. Hairer, and C. Lubich. Conservation of energy, momentum and actions in numerical discretizations of nonlinear wave equations. *Preprint*, 2007.

[DF07] G. Dujardin and E. Faou. Normal form and long time analysis of splitting schemes for the linear Schrödinger equation with small potential. *Numerische Mathematik 106, Number 2*, pages 223–262, 2007.

[dR87] H. de Raedt. Product formula algorithms for solving the time dependent Schrödinger equation. *Comput. Phys. Rep. 7*, pages 1–72, 1987.

[Duj08] G. Dujardin. Exponential Runge-Kutta methods for the Schrödinger equation. *Submitted*, 2008.

[EK06] H. L. Eliasson and S. B. Kuksin. KAM for non-linear Schrödinger equation. *to appear in Annals of Mathematics.*, 2006.

[Gin95] J. Ginibre. Le problème de Cauchy pour des EDP semi-linéaires périodiques en variables d'espace. *Séminaire N. Bourbaki 796*, pages 163–187, 1994-1995.

[HL07] E. Hairer and C. Lubich. Spectral semi-discretisations of weakly nonlinear wave equations over long times. *to appear in Foundations of Comput. Math*, 2007.

[HLW06] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration. Structure-preserving Algorithms for Ordinary Differential Equations, Second Edition*. Springer, 2006.

[HNW93] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I. Nonstiff Problems.* Springer Series in Computational Mathematics 8. Springer, 1993.

[HO05a] M. Hochbruck and A. Ostermann. Explicit exponential Runge-Kutta methods for semilinear parabolic problems. *SIAM J. Numer. Anal. 43, No. 3*, pages 1069–1090, 2005.

[HO05b] M. Hochbruck and A. Ostermann. Exponential Runge-Kutta methods for parabolic problems. *Appl. Numer. Math. 53, Issues 2-4*, pages 323–339, 2005.

[JL00] T. Jahnke and C. Lubich. Error bounds for exponential operator splittings. *BIT 40*, pages 735–744, 2000.

[LR04] B. Leimkuhler and S. Reich. *Simulating Hamiltonian dynamics.* Cambridge Monographs on Applied and Computational Mathematics, 14. Cambridge University Press, Cambridge, 2004.

[MQ02] R. I. McLachlan and G. R. W. Quispel. Splitting methods. *Acta Numerica 11*, pages 341–434, 2002.

[MW05] B. Minchev and W.M. Wright. A review of exponential integrators for semilinear problems. Technical Report 2, Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway, 2005.

[OWW04] M. Oliver, M. West, and C. Wulff. Approximate momentum conservation for spatial semidiscretizations of semilinear wave equations. *Numer. Math. 97*, pages 493–535, 2004.

[Sha00] Z. Shang. Resonant and Diophantine step sizes in computing invariant tori of Hamiltonian systems. *Nonlinearity 13*, pages 299–308, 2000.

List of Figures

Résumé

Cette thèse consiste en l'analyse numérique de méthodes de résolution d'équations aux dérivées partielles de type Schrödinger : sur le tore de dimension $d \in \mathbb{N}^\star$, on s'intéresse à la résolution numérique de l'équation de Schrödinger linaire avec potentiel multiplicatif, de l'équation de Schrödinger linéaire inhomogène et de l'équation de Schrödinger non linéaire.

Dans une première partie, on étudie des méthodes de splitting en temps, symplectiques, pour l'équation de Schrödinger linéaire avec potentiel multiplicatif. Dans l'asymptotique des petits potentiels, on démontre par une méthode perturbative un théorème de forme normale pour le propagateur de ces méthodes. Ce théorème permet ensuite de démontrer des propriétés de conservation en temps long de la régularité de la solution numérique pour des pas de temps non résonnants.

La seconde partie est consacrée à l'analyse numérique de méthodes de Runge-Kutta exponentielles pour l'équation de Schrödinger linéaire inhomogène et pour l'équation de Schrödinger non linéaire. Dans une perspective d'ordre élevé et en temps fini, on donne des conditions suffisantes pour que les méthodes de collocation à s points soient d'ordre s, $s+1$ et $s+2$ pour les deux types de problèmes envisagés. On illustre, quantifie et explique en outre l'effet des résonances numériques qui apparaissent lors de la résolution des problèmes linéaires inhomogènes par de telles méthodes.

Abstract

This thesis provides a numerical analysis of numerical methods for partial differential equations of Schrödinger type on the d-dimensional torus, namely the linear Schrödinger equation with potential, the inhomogeneous linear Schrödinger equation and the non linear Schrödinger equation.

The first part of this thesis deals with symplectic time-splitting methods for the linear Schrödinger equation with potential. Under a non resonance condition, we prove a normal form theorem for the numerical propagator. This theorem allows us to derive properties of preservation of the regularity of the numerical solution for non resonant time steps.

The second part of this thesis presents a numerical analysis of exponential Runge-Kutta methods for the inhomogeneous linear Schrödinger equation and for the non linear Schrödinger equation. Over a finite time interval, we give sufficient order conditions for s-stage collocation methods to be of order s, $s+1$ and $s+2$ when applied to any of these two problems. Moreover, we illustrate and explain the effect of the numerical resonances that may occur when solving inhomogeneous linear problems with such methods.