Diploma Thesis

Convergence Analysis of Time-Splitting Pseudo-Hermite Collocation Methods Applied to Nonlinear Schrödinger Equations

Institute for

Analysis and Scientific Computing

Vienna University of Technology

supervised by

Univ.Prof. Dr.rer.nat. Ansgar Jüngel

by

Nicola Ondracek MSc

Eschenbachstraße 4/7

2100 Korneuburg

Declaration

Unless otherwise indicated in the text or references, this thesis is entirely the product of my own scholarly work. This thesis has not been submitted either in whole or part for a degree at this or any other university or institution.

Vienna, August 2011

Nicola Ondracek MSc

Abstract

The aim of this thesis was to elaborate and extend the results proved in [7]. There, one can find convergence analyses for time and space semi-discretisations and full discretisations applied to the cubic nonlinear Schrödinger equation with a harmonic oscillator potential. The methods used in this treatise include Hermite quadrature and an operator splitting of second order.

The author was able to generalise all these results for Schrödinger equations with a scaled harmonic oscillator potential and a sum of power-nonlinearities up to an arbitrary degree, and was also capable of showing an existence and uniqueness result for equations of this type. Furthermore, using additionally the formal calculus of Lie derivatives [13], the author could prove convergence of arbitrary order of the time semi-discretised equation when using an appropriate higher order splitting scheme.

Proving a higher rate of convergence for the fully discretised scheme, however, turned out to be impossible without additional tools, and remains a challenge for future studies.

Acknowledgements

First of all, I would like to thank my supervisor Univ. Prof. Dr. Ansgar Jüngel for giving me the freedom and all the time I needed to delve into the subject of my thesis and thereby learn so much along the way. He would provide guidance whenever it was needed.

I would also like to express my gratitude to Mag. Herta Wöhrer and Sonja Höllrigl-Binder MSc for proofreading, as well as to DI Stefan Schuchnigg not only for proofreading, but also for discussing problems with me that occurred in the course of writing this thesis.

Finally, I am truely grateful for the support and understanding I received from all of my family and friends, which enabled me not only to overcome difficult periods when writing this work, but also rendered even hard times pleasant.

Contents

1. Introduction

In many areas of applied and theoretical physics, for instance nonlinear optics, Bose-Einstein condensation and plasma physics [4], we encounter a special kind of nonlinear Schrödinger equation, the so-called Gross-Pitaevskii equation (GPE). It reads

$$i\partial_t \psi(x,t) = -\Delta\psi(x,t) + V(x)\psi(x,t) + \gamma|\psi(x,t)|^2\psi(x,t), \tag{1.1}$$

where i is the complex unit, V an external potential and γ a physical parameter. In the case of the modelling of a Bose-Einstein condensate, the wavefuntion ψ solving this equation represents the macroscopic wave function of a condensate if the temperature drops below a critical temperature T_C [21]. Since these condensates are usually prepared in magnetic traps that create a (not necessarily isotropic) harmonic oscillator potential, V takes the form $V(x) = \sum_{k=1}^{3} \xi_k x_k^2$, where ξ_k are physical parameters.
In the derivation of equation (1.1), the cubic term results from considering 2-body interactions. If one also considers 3-body interactions, a cubic-quintic nonlinearity occurs. Taking also into account the interactions of even more particles, we understand that the solution of

$$i\partial_t \psi(x,t) = -\Delta\psi(x,t) + V(x)\psi(x,t) + \sum_{n=1}^{N} \gamma_n |\psi(x,t)|^{2n}\psi(x,t)$$

is of physical relevance as well. If we further generalise the external potential to $V(x) = x^T Ax + U(x)$, for a symmetric, positive matrix A and a bounded, real valued, continuous function U, and consider an arbitrary space dimension, we obtain the equation

$$i\frac{\partial}{\partial t}\psi(x,t) = -\Delta\psi(x,t) + x^T Ax\psi(x,t) + \left(\sum_{n=1}^{N} \gamma_n|\psi(x,t)|^{2n} + U(x)\right)\psi(x,t), \tag{1.2}$$

which will be dealt with throughout this thesis.
This diploma project aims at proving several convergence results for different (semi)-discretisations of this equation, using [7] as the main reference. Hence, the discretisation techniques that we employ are a Hermite spectral collocation method in space and a high-order exponential operator splitting method in time. The latter has become one of the most popular techniques for treating not only Schrödinger equations, but also general (parabolic) evolution equations on Banach spaces. The core of this technique is that we can construct approximate solutions of an equation of the form $\psi_t = (A+B)\psi$ from the solutions of $\psi_t = A\psi$ and $\psi_t = B\psi$. These splitting methods can be of arbitrary order p, with $p = 1$ being the only uneven one [26]. In the same work, Yosida computes a set of coefficients with which one can construct methods up to order $p = 8$, and also provides

leads to how to construct even better schemes.
The method that most works discuss is the Strang-splitting, which is an operator splitting of second order, as originally presented in [23]. For linear operators A and B, we can find convergence proofs for higher order splittings in [24] and [11]. Concerning the Strang-splitting, a rigorous convergence analysis has been performed for the cubic Schrödinger and the Schrödinger-Poisson equation [18], for the multi-configuration time-dependent Hartree equations [17] and the multi-configuration time-dependent Hartree-Fock equations [12]. The most recent advance that we know of in this field is a local error expansion for splitting methods of arbitrary order [13], which will also be employed in this work. All contributions mentioned so far analyse the convergence rate of semi-discretisations in time, only. In [7], the convergence of a full discretisation is shown for the case that a spectral collocation method is combined with a Strang-splitting.
Other possible space discretisation methods that have successfully been employed can be finite element (see for instance [22]), or finite difference (see for instance [19]), methods. However, since the GPE is usually studied in the whole space, pseudo-spectral collocation schemes seem to be the most suitable ones, as the space domain does not have to be truncated (see [3]).
This thesis has been organised as follows.
In chapter 2, we introduce the functional analytic framework in which we have to treat equation (1.2) and prove an existence and uniqueness result for sufficiently regular initial data.
In chapter 3, we prove the convergence of a semi-discretisation in space of (1.2), provided that the solution is adequately regular.
In chapter 4, a convergence result for a splitting-scheme of arbitrary order has been established, again under the assumption of sufficient regularity of the solution.
In chapter 5, the techniques and tools developed in the previous chapters have been combined to investigate the convergence of the full discretisation resulting from the combination of the operator splitting with the collocation method.
However, we have not been able to generalise this result for splittings of arbitrary order, and hence provide a convergence proof only for the Strang-splitting. The further generalisation for higher-order splittings remains a challenge for future work.
Finally, chapter 6 contains numerical studies, illustrating the theoretical results of this thesis.

2. Properties of the Nonlinear Schrödinger Equation

We will present the most important properties of the Schrödinger equation that will be dealt with in this thesis. The respective equation reads

$$
\begin{aligned}
i\frac{\partial}{\partial t}\psi(x,t) &= -\Delta\psi(x,t) + x^T A x\,\psi(x,t) + \left(\sum_{n=1}^{N}\gamma_n|\psi(x,t)|^{2n} + U(x)\right)\psi(x,t), \quad t\in(0,T], \\
\psi(x,0) &= \psi_0(x) \qquad (2.1)
\end{aligned}
$$

where $x \in \mathbb{R}^d$, $\gamma_n \in \mathbb{R}\ \forall n$, U is a bounded, continuous function $\mathbb{R}\to\mathbb{R}$, $T\in\mathbb{R}^+$ and $A \in \mathbb{R}^{d\times d}$ is a symmetric, positive definite matrix with $\operatorname{tr}(A) \geq 1$. Therefore, it allows a decomposition $A = Q^T\tilde{\Lambda}^2 Q$ with an orthogonal matrix Q. For notational simplicity we will most often use the decomposition $A = Q^T\Lambda^4 Q$, which means that Λ is a diagonal matrix containing the (positive) square roots of the eigenvalues of A. These assumptions remain valid throughout the whole thesis. Moreover, we set

$$
L = -\Delta + x^T A x \qquad (2.2)
$$

$$
V(\psi) = \left(\sum_{n=1}^{N}\gamma_n|\psi|^{2n} + U(x)\right)\psi, \qquad (2.3)
$$

which allows us to rewrite (2.1) in the shorter form

$$
i\psi_t = L\psi + V(\psi). \qquad (2.4)
$$

The domain of definitions of L and V will be clarified in sections 2.2 and 2.3, respectively. The final conclusion of this chapter will be an existence and uniqueness result in an appropriate function space. Before proving this, we will state, or partially prove, some results and properties which lead to this theorem and will be of crucial importance for the convergence analysis, at the same time.

2.1. Hilbert Scales and Sobolev Towers

This section contains the most relevant information about Hilbert scales. They will be used for the existence result and for simplifying proofs for the special Hilbert scale which is generated by the operator L. The results presented here can be found in [15], [25] and [6, p. 123ff].

Let H_0 be a Hilbert space and A a strictly positive self-adjoint operator with dense domain $\mathcal{D}(A) \subset H_0$. Additionally, let A fulfil

$$\|u\|_{H_0} \leq \|Au\|_{H_0} \text{ for } u \in \mathcal{D}(A). \tag{2.5}$$

We now define $\mathcal{H} := \bigcap_{k=0}^{\infty} \mathcal{D}(A^k)$ and on this set the s-norms $\|u\|_s = \|A^s u\|_{H_0}$. This leads to the following result [15, §9]:

Lemma 2.1. *For each $s \in \mathbb{R}$, the space $H^s := \overline{\mathcal{H}}^{\|\cdot\|_s}$ is a Hilbert space. Additionally, $H^s = \mathcal{D}(A^s)$. The chain of spaces $(H^s)_{s\in\mathbb{R}}$ is then called a* Hilbert scale *(or, for integer s,* Sobolev tower*) and A its* generating operator*.*

There is another suitable norm on H^s which is generated by the eigenfunctions $(h_n)_{n\in\mathbb{N}}$ and eigenvalues $(\omega_n)_{n\in\mathbb{N}}$ of A. Bearing in mind that the set of eigenfunctions is a complete orthonormal basis (if we choose the norm of the eigenfunctions accordingly), we define for $u_n \in \mathbb{C}$ and $u \in H_0$ $u = \sum_{n\in\mathbb{N}} u_n h_n$

$$\tilde{H}^s := \{u \in H_0 : \|u\|_{\tilde{H}^s} < \infty\}, \ \|u\|^2_{\tilde{H}^s} := \sum_{n\in\mathbb{N}} \omega_n^s |u_n|^2.$$

Some results in the respective literature (for instance in [25]) state that these two definitions lead to the same spaces. Due to the fact that this equivalence is crucial for almost every argument in this thesis, we will give a proof for it.

Lemma 2.2. *The spaces H^s and $\tilde{H}^{2s}$ are identical, or more precisely, they contain the same elements and the norms coincide.*

Proof. Assume that $\psi, \varphi \in \mathcal{D}(A^s)$ and $\varphi = \sum_{k\in\mathbb{N}} \varphi_k h_k$. Then

$$\begin{aligned}
(A^s \sum_{k\in\mathbb{N}} \varphi_k h_k, \psi)_{H_0} &= (\sum_{k\in\mathbb{N}} \varphi_k h_k, A^s \psi)_{H_0} \\
&= \lim_{N\to\infty} (\sum_{k\leq N} \varphi_k h_k, A^s \psi)_{H_0} \\
&= \lim_{N\to\infty} (\sum_{k\leq N} A^s \varphi_k h_k, \psi)_{H_0} \\
&= (\sum_{k\in\mathbb{N}} A^s \varphi_k h_k, \psi)_{H_0},
\end{aligned}$$

where the self-adjointness of A^s and the continuity of the scalar product have been used. Hence,

$$A^s \varphi = A^s \sum_{k\in\mathbb{N}} \varphi_k h_k = \sum_{k\in\mathbb{N}} A^s \varphi_k h_k.$$

The identity of the norms on $H^s \cap \tilde{H}^{2s}$ now follows from

$$\begin{aligned}
\|u\|_{H^s} &= (A^s u, A^s u)_{H_0} \\
&= (A^s \sum_{k\in\mathbb{N}} u_k h_k, A^s \sum_{n\in\mathbb{N}} u_n h_n)_{H_0} \\
&= \sum_{k,n\in\mathbb{N}} \omega_k^s u_k \omega_n^s \overline{u_n} (h_n, h_k)_{H_0} \\
&= \sum_{k\in\mathbb{N}} \omega_k^{2s} |u_k|^2 \\
&= \|u\|_{\tilde{H}^{2s}}.
\end{aligned}$$

For this calculation, we need that A^s can be pulled inside the infinite sum, which is possible $\forall u \in \mathcal{D}(A^s)$, and for such u, the left-hand side of the equation is finite since $A^s : \mathcal{D}(A^s) \to H_0$. Hence, also the right-hand side is bounded and therefore $u \in \tilde{H}^{2s}$. For the other inclusion, we need a different reasoning as we do not know if we may pull A^s outside of the sum. In order to show $H^s \supset \tilde{H}^{2s}$, we take an arbitrary $u \in \tilde{H}^{2s}$. Thus, $\sum_{k\in\mathbb{N}} \omega_k^{2s} |u_k|^2 < \infty$. Furthermore, we set

$$u^n = \sum_{k\leq n} u_k h_k.$$

So, $u^n \to u$ in the sense of H_0 and each $u^n \in \mathcal{D}(A^s)$ since it is a finite linear combination of eigenfunctions of A^s. Furthermore, u^n is also a (strong) Cauchy sequence in the sense of H^s, as there holds

$$\|u^n - u^m\|_{H^s} = \|A^s(u^n - u^m)\|_{H_0} = \sum_{k=m}^{n} \omega_k^{2s} |u_k|^2 \to 0,$$

for $n > m$ and $n, m \to \infty$ because the whole series converges. Since H^s is a Hilbert space, it follows that u^n converges to some $\tilde{u} \in H^s$. Since the H^s-norm dominates the H_0-norm, these limits have to coincide, $u = \tilde{u}$. Therefore, $u \in H^s$, which completes the proof. □

Further properties of Sobolev towers are summarised in

Lemma 2.3. *Let $r, s, \epsilon \in \mathbb{R}^+$ and $s < r$. Then*

- *$H^r \subset H^s$ with a compact and dense inclusion. This in particular means that $\|u\|_s \leq C\|u\|_r$ for some constant $C > 0$(see [15, §9]).*
- *$\|u\|_{\tilde{H}^s} \leq (\min_n \omega_n)^{-2\epsilon} \|u\|_{\tilde{H}^{s+\epsilon}}$ $\forall u \in \tilde{H}^{s+\epsilon}$(see [25]).*
- *The operator A maps H^s into H^{s-1}(see [6, II.5a]).*

2.2. Hermite Functions

As we have seen in the previous section, the knowledge of the eigenbasis of L is essential for working with the Hilbert scale that it generates. So, in this section we are constructing eigenfunctions of L. For the harmonic oscillator potential $\tilde{L}$, the eigenbasis is well known. It consists of the Hermite functions $\tilde{h}_n$. Since the eigenbasis of L will turn out to be a tensor product, a result that shows how to construct these basis functions and the respective eigenvalues from the one-dimensional ones, has been proven first.

Lemma 2.4. *Let I be a finite subset of $\mathbb{N}$, $(X_k)_{k\in I}$ Hilbert spaces and X the product space of the $(X_k)_{k\in I}$. Furthermore, let $(L_k)_{k\in I}$ be a family of operators with $L_k : X_k \to X_k$ having a set of eigenfunctions $(\varphi^k_{n_k})$ with eigenvalues $\lambda^k_{n_k}$. Then the operator $L : X \to X$ defined by $L := \sum_{k\in I}(L_k)$ has eigenfunctions given by the tensor products of the eigenfunctions of L_k and the eigenvalue corresponding to $\bigotimes_{k\in I} \varphi_{n_k}$ is $\sum_{k\in I} \lambda_{n_k}$.*

Proof. We commence with defining $L := \sum_{k\in I}(L_k)$ more precisely. Since each L_k only acts on X_k, its action on the product space X can be described as

$$L_k\psi = L_k \bigotimes_{p\in I} \psi_p = \bigotimes_{p\in I} L_k^{\delta_{pk}} \psi_p,$$

for $\psi \in X$. Hence,

$$L\psi = \sum_{k\in I} L_k \bigotimes_{p\in I} \psi_p = \sum_{k\in I} \bigotimes_{p\in I} L_k^{\delta_{pk}} \psi_p.$$

If we choose $\psi = \bigotimes_{p\in I} \varphi^p_{n_p}$ with $\varphi^p_{n_p}$ being the n_p-th eigenfunction of L_p, we can conclude that

$$L \bigotimes_{p\in I} \varphi^p_{n_p} = \sum_{k\in I} \bigotimes_{p\in I} L_k^{\delta_{pk}} \varphi^p_{n_p} = \sum_{k\in I} \bigotimes_{p\in I} \omega^k_{n_k} \varphi^p_{n_p} = \left(\sum_{k\in I} \omega^k_{n_k} \right) \bigotimes_{p\in I} \varphi^p_{n_p} =: \omega_{n_1,\ldots,n_{|I|}} \psi.$$

□

We now collect some common facts about Hermite functions that will be used in the following computations and can be found, for instance, in [1]. The two most commonly used expressions for the Hermite functions are either based on Rodrigues' formula

$$\tilde{h}_n(z) = c_n e^{\frac{z^2}{2}} \frac{d^n}{dz^n} e^{-z^2}, \quad c_n = (-1)^n / \sqrt{\sqrt{\pi} 2^n n!} \tag{2.6}$$

or define them as a product of a Hermite polynomial multiplied with the according weight function

$$\begin{aligned} &\tilde{h}_n(z) = c_n H_n(z) e^{-\frac{z^2}{2}}, \quad c_n = 1/\sqrt{\sqrt{\pi} 2^n n!}, \\ &H_{n+1}(z) = 2zH_n(z) - 2nH_{n-1}(z), \quad n > 1 \\ &H_0(z) = 1, \quad H_1(z) = 2z, \end{aligned} \tag{2.7}$$

where c_n ensures the normalisation of the $\tilde{h}_n$. Employing these formulas, we can prove the following results.

Lemma 2.5. *Let* $A = Q^T\Lambda^4 Q$, *where* $Q \in \mathbb{R}^{d\times d}$ *is orthogonal and* $\operatorname{diag}(\lambda_1, \ldots, \lambda_d) = \Lambda \in \mathbb{R}^{d\times d}$ *is diagonal. Furthermore, let* $y = \Lambda Q x$, $j \in \mathbb{N}^d$ *a multi-index and* $\tilde{h}_j(\Lambda Q x) = \tilde{h}_j(y) := \prod_{0\le k\le d} \tilde{h}_{j_k}(y_k)$. *Then* $(\prod_{1\le k\le d}\sqrt{\lambda_k} h_j(y))_{j\in\mathbb{N}^d}$ *is a system of orthonormal eigenfunctions of* L *which was defined in (2.2).*

Proof. We start out from expressing Δ_x in the new coordinate y:

$$\begin{aligned}
\frac{\partial}{\partial x_n} &= \sum_{p=1}^{d} \frac{\partial y_p}{\partial x_n}\frac{\partial}{\partial y_p} \\
\frac{\partial}{\partial x_n}\left(\frac{\partial}{\partial x_n}\right) &= \sum_{k=1}^{d} \frac{\partial y_k}{\partial x_n}\frac{\partial}{\partial y_k}\left(\sum_{p=1}^{d}\frac{\partial y_p}{\partial x_n}\frac{\partial}{\partial y_p}\right) \\
&= \sum_{k=1}^{d}\sum_{p=1}^{d}\frac{\partial y_k}{\partial x_n}\left(\frac{\partial^2 y_p}{\partial x_n \partial y_k}\frac{\partial}{\partial y_p} + \frac{\partial y_p}{\partial x_n}\frac{\partial^2}{\partial y_k \partial y_p}\right) \\
&= \sum_{k=1}^{d}\sum_{p=1}^{d}\frac{\partial y_k}{\partial x_n}\frac{\partial y_p}{\partial x_n}\frac{\partial^2}{\partial y_p \partial y_k},
\end{aligned}$$

where we have used the product rule and

$$\frac{\partial^2 y_p}{\partial x_n \partial y_k} = \frac{\partial}{\partial x_n}\delta_{kp} = 0.$$

Since $y_k = \sum_{m=1}^{d}(\Lambda Q)_{km} x_m$, we have $\frac{\partial y_k}{\partial x_p} = (\Lambda Q)_{kp}$. This yields

$$\begin{aligned}
\Delta_x = \sum_{n=1}^{d}\frac{\partial}{\partial x_n}\frac{\partial}{\partial x_n} &= \sum_{n,k,p=1}^{d}(\Lambda Q)_{kn}(\Lambda Q)_{pn}\frac{\partial^2}{\partial y_p \partial y_k} \\
&= \sum_{p,k=1}^{d}\left((\Lambda Q)\cdot(\Lambda Q)^T\right)\frac{\partial^2}{\partial y_p^2} = \sum_{p=1}^{d}\lambda_p^2\frac{\partial^2}{\partial y_p^2},
\end{aligned}$$

due to $\left((\Lambda Q)\cdot(\Lambda Q)^T\right)_{kp} = (\Lambda Q Q^T \Lambda^T)_{kp} = \lambda_p^2\delta_{kp}$.
Therefore, the operator L written in the coordinate y is given by

$$L = -\sum_{p=1}^{d}\lambda_p^2\frac{\partial^2}{\partial y_p^2} + y^T\Lambda^2 y = -\sum_{p=1}^{d}\lambda_p^2\left(\frac{\partial^2}{\partial y_p^2} - y_p^2\right).$$

Hence, lemma 2.4 implies that the eigenfunctions of L are tensor products of the one-dimensional eigenfunctions of $\tilde{L} = -\frac{d^2}{dz^2} + z^2$. These are the one-dimensional standard Hermite functions with according eigenvalues $\tilde{\lambda}_n = 2n+1$. As $\tilde{h}_j(y) = \tilde{h}_j(\Lambda Q x) = h_j(x)$, there holds

$$L h_j(x) = \sum_{k=1}^{d}\lambda_k^2(2j_k+1)h_j(x).$$

Finally, we need to normalise these eigenfunctions. Since

$$\begin{aligned}\int_{\mathbb{R}^d} h_j(x)h_l(x)\,dx &= \int_{\mathbb{R}^d} \tilde{h}_j(\Lambda Qx)\tilde{h}_l(\Lambda Qx)\,dx = \left|\begin{matrix} y = \Lambda Qx \\ dy = (\det\Lambda)dx\end{matrix}\right| \\ &= (\det\Lambda)^{-1}\int_{\mathbb{R}^d}\tilde{h}_j(y)\tilde{h}_l(y)\,dy \\ &= \prod_{1\le k\le d} {\lambda_k}^{-1}\delta_{jl},\end{aligned}$$

for another multi-index $l \in \mathbb{N}^d$, the right scaling factor is $\prod_{1\le k\le d}(\lambda_k)^{\frac{1}{2}}$. □

We can also prove that the momentum and the position operator are ladder operators in the following sense.

Lemma 2.6. *Let $j \in \mathbb{N}^d$ be a multi-index, $<m>\in \mathbb{N}^d$ the m-th unit vector and $y = \Lambda Qx$. Then for the scaled Hermite function $\tilde{h}_j(y)$ there holds*

$$\frac{\partial}{\partial x_l}\tilde{h}_j(y) = \sum_{m=1}^{d}(\Lambda Q)_{ml}\left(\sqrt{\frac{j_m}{2}}\tilde{h}_{j-<m>}(y)(1-\delta_{j_m,0}) - \sqrt{\frac{j_{m+1}}{2}}\tilde{h}_{j+<m>}(y)\right) \tag{2.8}$$

$$x_l h_j(y) = \sum_{m=1}^{d}(Q^T\Lambda^{-1})_{lm}\left(\sqrt{\frac{j_m}{2}}\tilde{h}_{j-<m>}(y)(1-\delta_{j_m,0}) + \sqrt{\frac{j_{m+1}}{2}}\tilde{h}_{j+<m>}(y)\right). \tag{2.9}$$

Proof. Since

$$\frac{\partial}{\partial x_l}\tilde{h}_j(y) = \sum_{m=1}^{d}\frac{\partial y_m}{\partial x_l}\frac{\partial}{\partial y_m}\tilde{h}_j(y) = \sum_{m=1}^{d}(\Lambda Q)_{ml}\frac{\partial}{\partial y_m}\tilde{h}_j(y)$$

$$x_l\tilde{h}_j(y) = ((\Lambda Q)^{-1}y)_l\tilde{h}_j(y) = (Q^T\Lambda^{-1}y)_l\tilde{h}_j(y) = \sum_{k=1}^{d}(Q^T\Lambda^{-1})_{lk}y_k\tilde{h}_j(y),$$

it suffices to prove for $n \ge 1$

$$z\tilde{h}_n(z) = \sqrt{\frac{n}{2}}\tilde{h}_{n-1}(z) - \sqrt{\frac{n+1}{2}}\tilde{h}_{n+1}(z),$$

which follows directly from the recurrence relation (2.7), and

$$\frac{d}{dz}\tilde{h}_n(z) = \sqrt{\frac{n}{2}}\tilde{h}_{n-1}(z) + \sqrt{\frac{n+1}{2}}\tilde{h}_{n+1}(z),$$

for $n \ge 1$, which we will show now. We note that $\frac{c_n}{c_{n-1}} = -\sqrt{\frac{1}{2n}}$ and

$$\begin{aligned}-\frac{1}{2}\frac{d^{n+1}}{dz^{n+1}}e^{-z^2} = \frac{d^n}{dz^n}ze^{-z^2} &= \sum_{k=0}^{n}\binom{n}{k}z^{(k)}(e^{-z^2})^{(n-k)} \\ &= n\frac{d^{n-1}}{dz^{n-1}}e^{-z^2} + z\frac{d^n}{dz^n}e^{-z^2}.\end{aligned}$$

These two formulas imply

$$\begin{aligned}
\frac{d}{dz}\tilde{h}_n(z) &= c_n \frac{d}{dz}\left(e^{\frac{z^2}{2}}\frac{d^n}{dz^n}e^{-z^2}\right) \\
&= c_n e^{\frac{z^2}{2}}\left(\left(z\frac{d^n}{dz^n}e^{-z^2}\right) + \frac{d^{n+1}}{dz^{n+1}}e^{-z^2}\right) \\
&= c_n e^{\frac{z^2}{2}}\left(-n\frac{d^{n-1}}{dz^{n-1}}e^{-z^2} - \frac{1}{2}\frac{d^{n+1}}{dz^{n+1}}e^{-z^2} + \frac{d^{n+1}}{dz^{n+1}}e^{-z^2}\right) \\
&= c_n e^{\frac{z^2}{2}}\left(-n\frac{d^{n-1}}{dz^{n-1}}e^{-z^2} + \frac{1}{2}\frac{d^{n+1}}{dz^{n+1}}e^{-z^2}\right) \\
&= -\frac{nc_n}{c_{n-1}}\tilde{h}_{n-1}(z) + \frac{c_n}{2c_{n+1}}\tilde{h}_{n+1}(z) \\
&= \sqrt{\frac{n}{2}}\tilde{h}_{n-1}(z) - \sqrt{\frac{n+1}{2}}\tilde{h}_{n+1}(z).
\end{aligned}$$

At last, there holds $\tilde{h}_1(z) = \frac{z}{\sqrt{2}\pi^{1/4}}e^{-\frac{z^2}{2}}$ and thus

$$\begin{aligned}
z\tilde{h}_0(z) &= \pi^{-\frac{1}{4}}ze^{-\frac{z^2}{2}} = \sqrt{2}\tilde{h}_1(z) \\
\frac{d}{dz}\tilde{h}_0(z) &= \pi^{-\frac{1}{4}}(-z)e^{-\frac{z^2}{2}} = -\sqrt{2}\tilde{h}_1(z).
\end{aligned}$$

□

2.3. The Scaled Harmonic Oscillator Potential L

We now focus on the operator L that is defined in (2.2). For convenience, we set $\mathcal{D}(L) = \mathcal{S}$, the Schwartz class. Since L is constructed merely by a rescaling of the harmonic oscillator potential $\tilde{L}$, we expect L to inherit all of the properties of $\tilde{L}$. From lemma 2.5 we deduce the existence of an orthonormal eigenbasis of $L^2(\mathbb{R}^d)$ and hence, by proposition A.2, L is essentially self-adjoint, which means that the closure $\overline{L}$ of L is self-adjoint. We will denote this closure by L from now on, instead. From Stone's theorem [20, theorem 1.10.8], we deduce furthermore that $-iL$ is the infinitesimal generator of a strongly continuous semigroup of unitary operators $\mathcal{T}(t)$ on L^2.
Establishing results about the Hilbert scales generated by L is the most elaborate part of this section. Firstly, we have to prove

Lemma 2.7. *L is the generating operator of a Hilbert scale.*

Proof. Since the eigenvalues of L are strictly positive and diverge to infinity, L is strictly positive definite and unbounded. More precisely, the positivity and the normbound (2.5) are implied by

$$\begin{aligned}(L\varphi,\varphi)_{L^2} &= (L\sum_{j\in\mathbb{N}^d}\varphi_j h_j, \sum_{l\in\mathbb{N}^d}\varphi_l h_l)_{L^2}\\ &= (\sum_{j\in\mathbb{N}^d} L\varphi_j h_j, \sum_{l\in\mathbb{N}^d}\varphi_l h_l)_{L^2}\\ &= (\sum_{j\in\mathbb{N}^d} \omega_j\varphi_j h_j, \sum_{l\in\mathbb{N}^d}\varphi_l h_l)_{L^2}\\ &= \sum_{j,l\in\mathbb{N}^d} \omega_j\varphi_j\overline{\varphi}_l\delta_{jl}\\ &= \sum_{j\in\mathbb{N}^d}\omega_j|\varphi_j|^2 \geq \sum_{j\in\mathbb{N}^d}|\varphi_j|^2\\ &= \|\varphi\|_{L^2},\end{aligned}$$

since the smallest eigenvalue $\omega_{0,\dots,0} \geq 1$ by assumption. From the fact that the eigenfunctions h_j form a sequence of unit norm with $\|Lh_j\|_{L^2} = \|\omega_j h_j\|_{L^2} = \omega_j \to \infty$, we deduce the unboundedness of L. □

Before continuing, let us clarify some notations. Firstly, H^s and $\tilde{H}^s$ denote the Hilbert scales introduced in section 2.1 generated by L. Additionally, we write

$$\|\varphi\|_s := \|\varphi\|_{\tilde{H}^s}, \quad \varphi \in \tilde{H}^s.$$

We emphasise that this is not the same as the H^s-norm, as we chose not to square the eigenvalues ω_j in the definition of $\tilde{H}^s$ in accordance with [7]. With this notation, we also have $\|.\|_{L^2} = \|.\|_0$. The following claim is vital for the existence proof and for numerical stability.

Lemma 2.8. *Let $\mathcal{T}(t)$ be the semigroup of unitary operators that has $-iL$ as its infinitesimal generator. Then $\mathcal{T}(t)$ is a group of unitary operators on $\tilde{H}^s\,\forall s \in \mathbb{R}^+$.*

Proof. Let $\sum_{j\in\mathbb{N}^d}\varphi_j h_j = \varphi \in \tilde{H}^s$. Since $\mathcal{T}(t)$ is bounded on L^2, we can pull it inside the infinite sum

$$\mathcal{T}(t)\sum_{j\in\mathbb{N}^d}\varphi_j h_j = \sum_{j\in\mathbb{N}^d}\mathcal{T}(t)\varphi_j h_j.$$

Since $\mathcal{T}(t)h_j(x) = e^{-i\omega_j t}h_j$, there holds

$$\sum_{j\in\mathbb{N}^d}\mathcal{T}(t)\varphi_j h_j = \sum_{j\in\mathbb{N}^d} e^{-i\omega_j t}\varphi_j h_j =: \sum_{j\in\mathbb{N}^d}\tilde{\varphi}_j h_j.$$

Since $|\varphi_j| = |\tilde{\varphi}_j|\ \forall j$, there holds

$$\|\varphi\|_s = \sum_{j\in\mathbb{N}^d}\omega_j^s|\varphi_j|^2 = \sum_{j\in\mathbb{N}^d}\omega_j^s|\tilde{\varphi}_j|^2 = \|\tilde{\varphi}\|_s,$$

and hence $\mathcal{T}(t)$ is indeed unitary on every $\tilde{H}^s$. □

A very important feature of this particular Hilbert scale is the boundedness of two operators that are vital in quantum theory. In the following, j is a multi-index and $< l > \in \mathbb{R}^d$ is the l-th unit vector. Hence, $j+ < l >$ adds 1 to the l-th entry of the multi-index j. Furthermore, x_l denotes the l-th coordinate. As the proof of this result is tedious, an auxiliary result is first being proven.

Lemma 2.9. *Let F_l be a family of operators such that for $\varphi \in \mathcal{D}(F_l)$*

$$F_l\varphi = F_l \sum_{j\in\mathbb{N}^d} \varphi_j h_j = \sum_{j\in\mathbb{N}^d} \varphi_j F_l h_j \quad \text{and} \quad F_l h_j = c_{jl} h_{j-<l>} + d_{jl} h_{j+<l>},$$

with real constants c_{jl}, d_{jl} that are monotonously increasing in each element of j. If $\omega^s_{j+<l>} g \leq C\omega^{s+1}_j$ holds $\forall j$ for

$$g = \max\{d^2_{jl}, |c_{j+2<l>,l} \cdot d_{jl}|, c^2_{jl}, |d_{j-2<l>,l} \cdot c_{jl}|\}, \tag{2.10}$$

then $\|F_l\varphi\|_s \leq C\|\varphi\|_{s+1}$, where C may depend on s.

Proof. Obviously, we have to calculate the Hermite coefficients of the new function $F_l\varphi$. In the following computation we drop the l-dependence of the constants:

$$\begin{aligned}
F_l\varphi &= F_l \sum_{j\in\mathbb{N}^d} \varphi_j h_j = \sum_{j\in\mathbb{N}^d} \varphi_j F_l h_j \\
&= \sum_{j\in\mathbb{N}^d} \varphi_j (c_j h_{j-<l>}(1-\delta_{j_l,0}) + d_j h_{j+<l>}) \\
&= \sum_{j\in\mathbb{N}^d} c_j \varphi_j h_{j-<l>}(1-\delta_{j_l,0}) + \sum_{j\in\mathbb{N}^d} d_j \varphi_j h_{j+<l>} \\
&= \sum_{j'\in\mathbb{N}^d} c_{j'+<l>} \varphi_{j'+<l>} h_{j'} + \sum_{j'-<l>\in\mathbb{N}^d} d_{j'-<l>} \varphi_{j'-<l>} h_{j'},
\end{aligned}$$

where we used the substitutions $j' = j- <l>$ in the first and $j' = j+ <l>$ in the second sum. In the first sum, we have to sum over $j' \in \mathbb{N}^d$, because $c_j\varphi_j h_{j-<l>}(1-\delta_{j_l,0}) \neq 0$ if $j_l \geq 1$, which is the same as $j'_l \geq 0$. Rewriting the sums once again, we arrive at

$$\begin{aligned}
F_l\varphi &= \sum_{j'\in\mathbb{N}^d} c_{j'+<l>} \varphi_{j'+<l>} h_{j'} + \sum_{j'\in\mathbb{N}^d} d_{j'-<l>}(1-\delta_{j'_l,0}) \varphi_{j'-<l>} h_{j'} \\
&= \sum_{j'\in\mathbb{N}^d} h_{j'} \left(c_{j'+<l>}\varphi_{j'+<l>} + d_{j'-<l>}\varphi_{j'-<l>}(1-\delta_{j'_l,0})\right) \\
&=: \sum_{j'\in\mathbb{N}^d} h_{j'} \tilde{\varphi}_{j'}.
\end{aligned}$$

Hence, the s-norm of $F_l\varphi$ is given by

$$\|F_l\varphi\|_s = \sum_{j\in\mathbb{N}^d} \omega^s_j |c_{j+<l>}\varphi_{j+<l>} + d_{j-<l>}\varphi_{j-<l>}(1-\delta_{j_l,0})|^2.$$

Employing the triangular inequality and subsequently Young's inequality on the mixed terms of the Hermite coefficient, we arrive at

$$\begin{aligned}|\tilde{\varphi}_j|^2 &\leq c_{j+<l>}^2|\varphi_{j+<l>}|^2 + (1-\delta_{j_l,0})d_{j-<l>}^2|\varphi_{j-<l>}|^2 \\ &\quad + (1-\delta_{j_l,0})c_{j+<l>}d_{j-<l>}\left(\overline{\varphi}_{j+<l>}\varphi_{j-<l>} + \overline{\varphi}_{j-<l>}\varphi_{j+<l>}\right) \\ &\leq |\varphi_{j+<l>}|^2\left(c_{j+<l>}^2 + |c_{j+<l>}d_{j-<l>}|\,(1-\delta_{j_l,0})\right) \\ &\quad + |\varphi_{j-<l>}|^2\left(d_{j-<l>}^2 + |c_{j+<l>}d_{j-<l>}|\right)(1-\delta_{j_l,0})\,.\end{aligned}$$

Changing the index of summation again, we finally obtain

$$\begin{aligned}\|F_l\varphi\|_s &\leq \sum_{j'\in\mathbb{N}^d}\omega_{j'}^s|\varphi_{j'+<l>}|^2\left(c_{j'+<l>}^2 + |c_{j'+<l>}d_{j'-<l>}|\left(1-\delta_{j_l',0}\right)\right) \\ &\quad + \sum_{j'\in\mathbb{N}^d}\omega_{j'}^s|\varphi_{j'-<l>}|^2\left(d_{j'-<l>}^2 + |c_{j'+<l>}d_{j'-<l>}|\right)\left(1-\delta_{j_l',0}\right) \\ &= \sum_{j-<l>\in\mathbb{N}^d}\omega_{j-<l>}^s|\varphi_j|^2\left(c_j^2 + |c_jd_{j-2<l>}|\,(1-\delta_{j_l,1})\right) \\ &\quad + \sum_{j+<l>\in\mathbb{N}^d}\omega_{j+<l>}^s|\varphi_j|^2\left(d_j^2 + |c_{j+2<l>}d_j|\right)(1-\delta_{j_l,0}) \\ &= \sum_{j\in\mathbb{N}^d}\omega_{j-<l>}^s|\varphi_j|^2\left(c_j^2 + |c_jd_{j-2<l>}|\,(1-\delta_{j_l,1})\right)(1-\delta_{j_l,0}) \\ &\quad + \sum_{j\in\mathbb{N}^d}\omega_{j+<l>}^s|\varphi_j|^2\left(d_j^2 + |c_{j+2<l>}d_j|\right).\end{aligned}$$

Therefore, $\|F_l\varphi\|_s \leq C\|\varphi\|_{s+1}$, if there holds

$$\omega_{j+<l>}^s g \leq C'\omega_j^{s+1},$$

because ω_j is monotonously increasing in each j_l. Since this exactly is assumption (2.10), the proof is complete. □

Lemma 2.10. *The momentum operator and the position operator are continuous from* $\tilde{H}^{s+1} \to (\tilde{H}^s, \|.\|_{s+1})$*. More precisely, for* $\varphi \in \tilde{H}^{s+1}$ *and* $l \in \{1,\dots,d\}$ *there holds*

$$\|\frac{\partial}{\partial x_l}\varphi\|_{\tilde{H}^s} \leq C_{\partial x_l}\|\varphi\|_{\tilde{H}^{s+1}} \tag{2.11}$$

$$\|x_l\varphi\|_{\tilde{H}^s} \leq C_{x_l}\|\varphi\|_{\tilde{H}^{s+1}}. \tag{2.12}$$

Proof. Firstly, we have to ensure that we can use lemma 2.9, which means that we can pull x_l and $\frac{\partial}{\partial x_l}$ into the infinite sum, which is not a priori clear for these unbounded operators. It is legitimate to compute the coefficients of the (weak) derivative of $\tilde{H}^{s+1} \cap W^{1,2} \ni \varphi = \sum_{j\in\mathbb{N}^d} \varphi_j h_j(x)$ by differentiating the basis functions h_j because of

$$\begin{aligned}(\frac{\partial}{\partial x_l}\varphi, \psi)_{L^2} &= -(\varphi, \frac{\partial}{\partial x_l}\psi)_{L^2} = -(\sum_{j\in\mathbb{N}^d} \varphi_j h_j, \frac{\partial}{\partial x_l}\psi)_{L^2} \\ &= -\lim_{n\to\infty}(\sum_{j\in\{0,\dots,n\}^d} \varphi_j h_j, \frac{\partial}{\partial x_l}\psi)_{L^2} = \lim_{n\to\infty}(\sum_{j\in\{0,\dots,n\}^d} \varphi_j \frac{\partial}{\partial x_l} h_j, \psi)_{L^2} \\ &= (\sum_{j\in\mathbb{N}^d} \varphi_j \frac{\partial}{\partial x_l} h_j, \psi)_{L^2},\end{aligned}$$

where ψ is a test function, and the continuity of the scalar product was used.
We have to apply a similar procedure to prove that $x_l\varphi = \sum_{j\in\mathbb{N}^d} \varphi_j x_l h_j$. Generally, the position operator Q is defined on

$$\mathcal{D}(Q) = \{f \in L^2 : \, x_l f \in L^2\}.$$

Assuming that $\varphi \in \mathcal{D}(Q)$, there are coefficients $(\psi_j)_{j\in\mathbb{N}^d}$ such that $x_l\varphi = \sum_{j\in\mathbb{N}^d} \psi_j h_j$. We can compute these coefficients by

$$\psi_j = (x_l\varphi, h_j)_{L^2} = (\varphi, x_l h_j)_{L^2}.$$

Therefore, ψ_j is also the coefficient of φ to the basis $x_l h_j$ if these functions actually form a basis of $\mathcal{D}(Q)$. This is indeed the case, since $\operatorname{span}((x_l h_j)_{j\in\mathbb{N}^d})$ is still in L^2 due to lemma 2.6 and all finite linear combinations obviously lie in $\mathcal{D}(Q)$. Hence, $x_l\varphi = \sum_{j\in\mathbb{N}^d} \varphi_j x_l h_j$ and $\frac{\partial}{\partial x_l}\varphi = \sum_{j\in\mathbb{N}^d} \varphi_j \frac{\partial}{\partial x_l} h_j$. Therefore, we can apply lemma 2.9 to estimate the norms of these two terms. Secondly, we see that a bound for $\frac{\partial}{\partial y_l}\tilde{h}_j(y)$ and $y_l\tilde{h}_j(y)$ is sufficient, since due to (2.8) and (2.9) there hold

$$\begin{aligned}\|x_l\varphi(y)\|_{\tilde{H}^s} &\le d \max_{1\le i\le d} \lambda_i \|y_i\varphi\|_{\tilde{H}^s} \\ &\le d \max_{1\le i\le d} \lambda_i C_{y_i} \|\varphi\|_{\tilde{H}^{s+1}} \\ &=: C_{x_l}\|\varphi\|_{\tilde{H}^{s+1}}\end{aligned}$$

and

$$\begin{aligned}\|\frac{\partial}{\partial x_l}\varphi(y)\|_{\tilde{H}^s} &\le d \max_i (Q^T\Lambda^{-1})_{li} \max_{1\le i\le d} \|\frac{\partial}{\partial y_i}\varphi\|_{\tilde{H}^s} \\ &\le d \max_{1\le i\le d} \max_i \frac{1}{\lambda_i} C_{\partial y_i}\|\varphi\|_{\tilde{H}^{s+1}} \\ &=: C_{\partial x_l}\|\varphi\|_{\tilde{H}^{s+1}}.\end{aligned}$$

We now use lemma 2.9 for $c_{jl} = \sqrt{\frac{j_l}{2}}$ and $d_{jl} = \pm\sqrt{\frac{j_l+1}{2}}$. The assumptions of the lemma are fulfilled if we can bound

$$\omega^s_{j+<l>} \frac{j_l+2}{2} \le C\omega^{s+1}_j, \tag{2.13}$$

because $\frac{j_l+2}{2} \geq g$. Since

$$\begin{aligned}
&\omega_{j+<l>}^s \frac{j_l+2}{2} \leq C\omega_j^{s+1} \\
\Leftrightarrow &\left(\sum_{i=1}^d \lambda_i^2(2j_i+1)+2\lambda_l^2\right)^s \frac{j_l+2}{2} \leq C\left(\sum_{i=1}^d \lambda_i^2(2j_i+1)\right)^{s+1} \\
\Leftarrow &\max_i \lambda_i^{2s}\left(\sum_{i=1}^d (2j_i+1)+2\right)^s \frac{j_l+2}{2} \leq C \min_i \lambda_i^{2(s+1)}\left(\sum_{i=1}^d (2j_i+1)\right)^{s+1} \\
\Leftrightarrow &\left(1+\frac{2}{\sum_i(2j_i+1)}\right)^s \frac{j_l+2}{2\sum_i(2j_i+1)} \frac{\max_i \lambda_i^{2s}}{\min_i \lambda_i^{2(s+1)}} \leq C,
\end{aligned}$$

and since the last line holds true with $C=(1+\frac{2}{d})^s \frac{\max_i \lambda_i^{2s}}{\min_i \lambda_i^{2(s+1)}}$, equation (2.13) is valid. We have so far proven that (2.11) and (2.12) hold for $f \in \mathcal{X}_i$, where

$$\mathcal{X}_1 := \{f \in \tilde{H}^{s+1} : \frac{\partial}{\partial x_l} f \in \tilde{H}^s\}, \text{ or } \mathcal{X}_2 := \{f \in \tilde{H}^{s+1} : x_l f \in \tilde{H}^s\},$$

respectively. We will finish the proof by showing that this leads to the desired estimate $\forall f \in \tilde{H}^{s+1}$. Firstly, we note that $\mathcal{X}_i$ is dense in $\tilde{H}^{s+1}$. We can therefore take a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{X}_1$ that converges to $f \in \tilde{H}^{s+1}$ in the sense of $\tilde{H}^{s+1}$. Hence,

$$\|\frac{\partial}{\partial x_l} f_n\|_s \leq C\|f_n\|_{s+1}.$$

Taking the limit $n \to \infty$, we see that

$$\lim_{n\to\infty} \|\frac{\partial}{\partial x_l} f_n\|_s \leq C\|f\|_{s+1},$$

thus there exists a subsequence, denoted again as $(\frac{\partial}{\partial x_l} f_n)_{n\in\mathbb{N}}$ that converges weakly in the sense of $\tilde{H}^s$ to an element $\tilde{f} \in \tilde{H}^s$. Since weak convergence in $\tilde{H}^s$ implies strong convergence in L^2 (and especially also weak convergence in L^2) due to the compact embedding of $\tilde{H}^s \hookrightarrow L^2$, lemma 2.3, we have that $\forall \varphi \in L^2$ and especially for all testfunctions φ

$$\begin{aligned}
&\lim_{n\to\infty} (\frac{\partial}{\partial x_l} f_n, \varphi) = (\tilde{f}, \varphi) \\
\Leftrightarrow \quad &-\lim_{n\to\infty} (f_n, \frac{\partial}{\partial x_l}\varphi) = -(f, \frac{\partial}{\partial x_l}\varphi) = (\frac{\partial}{\partial x_l} f, \varphi).
\end{aligned}$$

Therefore, the weak L^2-limit is $\frac{\partial}{\partial x_l} f$ in the sense of distributions. Since both functions belong to L^2, also the strong L^2-limit coincides and has to be the weak $\tilde{H}^s$-limit at the same time. We conclude that (2.11) holds $\forall f \in \tilde{H}^{s+1}$. Exactly the same limiting procedure yields the validity of (2.12). □

Lemma 2.11. *The space $\tilde{H}^s$ has a continuous embedding into $W^{2,s}$ for integer s. Moreover, assume $\mathbb{N} \ni s > (k + \frac{d}{2})$ and $\varphi \in \tilde{H}^s$. Then $\varphi \in C^k$ with a continuous embedding and*

$$\sup_{x \in \mathbb{R}^d} |\varphi^{(k)}(x)| \leq C\|\varphi\|_{\tilde{H}^s}.$$

Proof. The first claim is immediate since we can bound the k-th Sobolev seminorm by the $\tilde{H}^k$-norm for all $k \leq s$ and all the resulting $\tilde{H}^k$ norms by the $\tilde{H}^s$-norm.
The second claim is a direct consequence of the first one and of the usual Sobolev embedding theorem. □

As mentioned above, the next result will be used to bound products of functions.

Lemma 2.12. *Let $r > \frac{d}{2}$ be an integer and $s > d$ be an even integer.*

- *If $\varphi \in \tilde{H}^r$ and $\psi \in L^2$, then $\|\varphi\psi\|_0 \leq C_{r0}\|\varphi\|_{\tilde{H}^r}\|\psi\|_0$.*
- *If $\psi, \varphi \in \tilde{H}^s$ then $\|\varphi\psi\|_{\tilde{H}^s} \leq C_s\|\varphi\|_{\tilde{H}^s}\|\psi\|_{\tilde{H}^s}$.*

Proof. The first claim follows easily by applying Hölder's inequality and the continuous embedding from lemma 2.11

$$\|\varphi\psi\|_0^2 = \||\varphi\psi|^2\|_{L^1} \leq \||\psi|^2\|_{L^1}\||\varphi|^2\|_{L^\infty} \leq \|\psi\|_0^2\|\varphi\|_{L^\infty}^2 \leq C\|\psi\|_0^2\|\varphi\|_r^2.$$

For the second claim, we perform a change from x to y in the norm which will shorten the computations:

$$\|\varphi(x)\psi(x)\|_s = \|L_x^{\frac{s}{2}}\varphi(x)\psi(x)\|_0 = (\det\Lambda)^{-\frac{1}{2}}\|L_y^{\frac{s}{2}}\varphi(y)\psi(y)\|_0.$$

Hence, we have to estimate

$$\|\left(-\sum_{k=1}^{d}\frac{\partial^2}{\partial y_k^2} + \sum_{n=1}^{d} y_n^2\right)^{\frac{s}{2}}(\varphi\psi)\|_0.$$

As $L^{s/2}$ is a sum of products of $\frac{\partial^2}{\partial x_{n_k}^2}$ and $x_{p_k}^2$ and as the commutator of $\frac{\partial}{\partial x_p}$ and x_k is given by

$$[\frac{\partial}{\partial x_p}, x_k]f = \frac{\partial}{\partial x_p}x_k f - x_k\frac{\partial}{\partial x_p}f = x_k\frac{\partial}{\partial x_p}f + \delta_{kp}f - x_k\frac{\partial}{\partial x_p}f = \delta_{kp}f,$$

we can separate the multiplication operators and derivatives. This means that we can bound $\|L^{\frac{s}{2}}(\varphi\psi)\|_0$ by a sum of terms such as

$$\|\left(\prod_{k=1}^{m} x_{l_k}\right)\left(\prod_{p=m+1}^{n}\frac{\partial}{\partial x_{l_p}}(\varphi\psi)\right)\|_0.$$

For n, which is the number of operators appearing, there holds $n \leq s$, since we started with $\frac{s}{2}$ second order differential operators and $\frac{s}{2}$ multiplication operators x^2. Applying the product rule, we can estimate this term by another sum with summands of the form

$$\|\left(\prod_{k=1}^{m} x_{l_k}\prod_{q=m+1}^{r}\frac{\partial}{\partial x_{l_q}}\varphi\right)\left(\prod_{p=r+1}^{n}\frac{\partial}{\partial x_{l_p}}\psi\right)\|_0$$

or by the same terms, but with φ and ψ exchanged. Due to the first statement of this lemma, each of these terms can again be estimated by

$$\|\left(\prod_{k=1}^{m} x_{l_k} \prod_{q=m+1}^{r} \frac{\partial}{\partial x_{l_q}}\varphi\right)\left(\prod_{p=r+1}^{n} \frac{\partial}{\partial x_{l_p}}\psi\right)\|_0 \le \|\prod_{k=1}^{m} x_{l_k} \prod_{q=m+1}^{r} \frac{\partial}{\partial x_{l_q}}\varphi\|_0 \| \prod_{p=r+1}^{n} \frac{\partial}{\partial x_{l_p}}\psi\|_\sigma, \quad (2.14)$$

for σ being the smallest integer greater than $\frac{d}{2}$, if $\prod_{p=r+1}^{n} \frac{\partial}{\partial x_{l_p}}\psi$ is indeed in $\tilde{H}^\sigma$. As this is an implication of $\psi \in \tilde{H}^{\sigma+n-r}$, by virtue of lemma 2.10, $n-r \le \frac{s}{2}$ and $\sigma < \frac{s}{2}$, there holds $\psi \in \tilde{H}^s \subset \tilde{H}^{\sigma+n-r}$, and hence the estimate is justified. This reasoning also yields

$$\| \prod_{p=r+1}^{n} \frac{\partial}{\partial x_{l_p}}\psi\|_\sigma \le C\|\psi\|_{\sigma+n-r} \le C\|\psi\|_s.$$

In order to estimate the first factor on the right-hand side of (2.14), we note that $r \le n \le s$, which yields

$$\|\prod_{k=1}^{m} x_{l_k} \prod_{q=m+1}^{r} \frac{\partial}{\partial x_{l_q}}\varphi\|_0 \le C\|\varphi\|_r \le C\|\varphi\|_s,$$

due to lemma 2.10. Hence, we conclude

$$\|\psi\varphi\|_s \le C\|\psi\|_s\|\varphi\|_s.$$

□

2.4. The Arbitrary Power-Nonlinearity

If we assign the operator V defined in (2.3) to act on the Schwartz class $\mathcal{S}$, then clearly $V : \mathcal{S} \to \mathcal{S}$. Moreover, we see that for ψ in any normed algebra $\mathcal{X}$ also $V(\psi) \in \mathcal{X}$. This is one of the main reasons why the convergence results and the existence theorem can be established. Furthermore, we can directly compute and bound the Fréchet derivatives of V of arbitrary order, which will be crucial for working with Lie derivatives. To achieve this aim, two auxiliary results are being proven first.

Lemma 2.13. *Let X be a Banach*-algebra with involution operator $x^* =: \overline{x}$, $\mathbf{h}_m = (h_1, \dots, h_m) \in X^m$ and the family of functions $V_{k,n}(x)$ be defined as*

$$V_{k,n}(x) = x^k \overline{x}^n \quad \text{for } n, k \in \mathbb{N} \quad (2.15)$$
$$V_{k,n}(x) \equiv 0 \quad \text{for } -k \in \mathbb{N}^+ \vee -n \in \mathbb{N}^+, \quad (2.16)$$

for $x \in X$.
Then the m-th Fréchet derivative of $V_{kn}(x)$ is given by

$$V_{k,n}^{(m)}(x) = \sum_{j=0}^{m} V_{k-j,n-m+j}(x) \frac{k!}{(k-j)!} \frac{n}{(n-m+j)!} g(\mathbf{h}_m, j).$$

The function g depends on a vector $\mathbf{h}_m$ and a natural j in the following way: g is a sum of products of the elements of $\mathbf{h}_m$. In each product, every h_i appears exactly once, and $m-j$ of these factors are conjugated.

Proof. We start with the brief computation

$$
\begin{aligned}
V_{k,n}(x+h) &= \sum_{i=0}^{k} \binom{k}{i} x^i h^{k-i} \sum_{j=0}^{n} \binom{n}{j} \overline{x}^j \overline{h}^{n-j} \\
&= \sum_{i=0}^{k} \sum_{j=0}^{n} \binom{k}{i} \binom{n}{j} x^i h^{k-i} \overline{x}^j \overline{h}^{n-j} \\
&= x^k \overline{x}^n + k x^{k-1} \overline{x}^n h + n x^k \overline{x}^{n-1} \overline{h} \qquad (2.17) \\
&\quad + \underbrace{\sum_{i=0}^{k-1} \sum_{j=0}^{n-1} \binom{k}{i} \binom{n}{j} x^i h^{k-i} \overline{x}^j \overline{h}^{n-j} + \sum_{j=0}^{n-2} \binom{n}{j} x^k \overline{x}^j \overline{h}^{n-j} + \sum_{i=0}^{k-2} \binom{k}{i} x^i h^{k-i} \overline{x}^n}_{\mathrm{o}(h)}, \qquad (2.18)
\end{aligned}
$$

which will be needed later. Furthermore, there holds

$$
\begin{aligned}
h_{m+1}\, g(\mathbf{h}_m, m) &= g(\mathbf{h}_{m+1}, m+1), \\
h_{m+1}\, g(\mathbf{h}_m, j-1) + \overline{h}_{m+1}\, g(\mathbf{h}_m, j) &= g(\mathbf{h}_{m+1}, j), \quad j \in \{1, \dots, m-1\}, \qquad (2.19) \\
\overline{h}_{m+1}\, g(\mathbf{h}_m, 0) &= g(\mathbf{h}_{m+1}, 0).
\end{aligned}
$$

The second line of the equation is valid because in $g(\mathbf{h}_m, j-1)$, $m-j+1$ factors are conjugated, and hence also $h_{m+1}\, g(\mathbf{h}_m, j-1)$ contains $m-j+1$ conjugated elements. The expression $g(\mathbf{h}_m, j)$ on the other hand contains $m-j$ factors that are conjugated, thus $m-j+1$ factors of $\overline{h}_{m+1}\, g(\mathbf{h}_m, j)$ are conjugated.
The claim is proved by induction. First, we are verifying the hypothesis for the seed $m=1$. Since

$$
V_{k,n}(x+h) - V_{k,n}(x) = k x^{k-1} \overline{x}^n h + n x^k \overline{x}^{n-1} \overline{h} + \mathrm{o}(h),
$$

where we employed (2.17), we find that

$$
V'_{k,n}(x) h = k V_{k-1,n}(x) h + n V_{k,n-1}(x) \overline{h}.
$$

For the induction step $m \to m+1$, we compute

$$
\begin{aligned}
V^{(m)}_{k,n}(x+h_{m+1}) =& \sum_{j=0}^{m} V_{k-j,n-m+j}(x+h_{m+1}) \frac{k!n!}{(k-j)!(n-m+j)!} g(\mathbf{h}_m, j) \\
=& \sum_{j=0}^{m} \mathrm{o}(h_{m+1}) + \Big(V_{k-j,n-m+j}(x) + (k-j) V_{k-j-1,n-m+j}(x) h_{m+1} \\
&+ (n-m+j) V_{k-j,n-m+j-1}(x) \overline{h}_{m+1} \Big) \frac{k!n!}{(k-j)!(n-m+j)!} g(\mathbf{h}_m, j)
\end{aligned}
$$

$$\begin{aligned}
\Rightarrow V_{k,n}^{(m+1)}(x)\mathbf{h}_{m+1} =& \sum_{j=0}^{m} \big((k-j)V_{k-j-1,n-m+j}(x)h_{m+1} + (n-m+j)V_{k-j,n-m+j-1}(x)\overline{h}_{m+1}\big) \\
& \cdot \frac{k!n!}{(k-j)!(n-m+j)!} g(\mathbf{h}_m, j) \\
=& \sum_{j=0}^{m} V_{k-j-1,n-m+j} \frac{n!k!}{(k-j-1)!(n-m+j)!} h_{m+1} g(\mathbf{h}_m, j) \\
& + \sum_{j=0}^{m} V_{k-j,n-m+j-1} \frac{n!k!}{(k-j)!(n-m+j-1)!} \overline{h}_{m+1} g(\mathbf{h}_m,j) \\
=& \sum_{j=0}^{m} V_{k-(j+1),n-(m+1)+(j+1)} \frac{n!k!}{(k-(j+1))!(n-m+j))!} h_{m+1} g(\mathbf{h}_m,0) \\
& + \sum_{j=0}^{m} V_{k-j,n-(m+1)+j} \frac{n!k!}{(k-j)!(n-(m+1)+j)!} \overline{h}_{m+1} g(\mathbf{h}_m, j).
\end{aligned}$$

With the substitution $j'=j+1$ in the first sum, this yields

$$\begin{aligned}
V_{k,n}^{(m+1)}(x)\mathbf{h}_{m+1} =& \sum_{j'=1}^{m+1} V_{k-j',n-(m+1)+j'} \frac{n!k!}{(k-j')!(n-(m+1)+j')!} h_{m+1} g(\mathbf{h}_m, j'-1) \\
& + \sum_{j=0}^{m} V_{k-j,n-(m+1)+j} \frac{n!k!}{(k-j)!(n-(m+1)+j)!} \overline{h}_{m+1} g(\mathbf{h}_m, j) \\
=& V_{k-(m+1),n}(x) \frac{n!k!}{(k-(m+1))!n!} h_{m+1} g(\mathbf{h}_m, m) \\
& + \sum_{j'=1}^{m} V_{k-j',n-(m+1)+j'} \frac{n!k!}{(k-j')!(n-(m+1)+j')!} h_{m+1} g(\mathbf{h}_m, j'-1) \\
& + V_{k,n-(m+1)}(x) \frac{n!k!}{k!(n-(m+1))!} \overline{h}_{m+1} g(\mathbf{h}_m, j) \\
& + \sum_{j=1}^{m} V_{k-j,n-(m+1)+j} \frac{n!k!}{(k-j)!(n-(m+1)+j)!} h_{m+1} g(\mathbf{h}_m, j).
\end{aligned}$$

Due to (2.19), we finally obtain

$$
\begin{aligned}
V_{k,n}^{(m+1)}(x)\mathbf{h}_{m+1} =& V_{k-(m+1),n}(x)\frac{n!k!}{(k-(m+1))!n!}h_{m+1}g(\mathbf{h}_m, m+1) \\
&+ V_{k,n-(m+1)}(x)\frac{n!k!}{k!(n-(m+1))!}\overline{h}_{m+1}g(\mathbf{h}_m, 0) \\
&+ \sum_{j=1}^{m}\Bigg(V_{k-j,n-(m+1)+j}\frac{n!k!}{(k-j)!(n-(m+1)+j)!} \\
&\cdot \big(h_{m+1}g(\mathbf{h}_m, j) + \overline{h}_{m+1}g(\mathbf{h}_m, j-1)\big)\Bigg) \\
=& \sum_{j=0}^{m+1} V_{k-j,n-(m+1)+j}\frac{n!k!}{(k-j)!(n-(m+1)+j)!}g(\mathbf{h}_{m+1}, j).
\end{aligned}
$$

□

Corollary 2.1. *Let the assumptions of lemma 2.13 be fulfilled and let X additionally be normed. Then there holds*

$$
\|V_{k,n}^{(m)}\| \leq C\|x\|^{k+n-m}\prod_{i=1}^{m}\|h_i\|,
$$

for a constant $C(k,n,m) \in \mathbb{R}$.

The most important properties of V are summarised in

Proposition 2.1. *Let $(\mathcal{X}, \|.\|_{\mathcal{X}})$ be a normed algebra, $\psi \in \mathcal{X}$ and $V : \mathcal{X} \to \mathcal{X}$. Then*

- *$V(\psi) \leq C(M)M$ for $\psi \in (\mathcal{X}, \|.\|_{\mathcal{X}})$ and $\|\psi\|_{\mathcal{X}} < M$,*
- *$V \in C^\infty(\mathbb{C}, \mathbb{C})$ (in the real sense) and*
- *the m-th Fréchet derivative of V is explicitly given by*

$$
V(\psi)^{(m)} = \sum_{i=1}^{k}\gamma_i\sum_{j=0}^{m}V_{2i+1-j,2i-m+j}(\psi)\frac{(2i+1)!}{(2i+1-j)!}\frac{2i}{(2i-m+j)!}g(\mathbf{h}_m, j) + U\delta_{m1}
$$

for $m \geq 1$ and is therefore bounded on bounded subsets of $\mathcal{X}$.

Proof. We immediately see that in the notation of lemma 2.13, V can be written as

$$
V(\psi) = \sum_{n=1}^{N}\gamma_n V_{2n+1,2n}(\psi) + U(x)\psi.
$$

Hence, all the claims are direct consequences of lemma 2.13 and corollary 2.1. □

2.5. Existence, Uniqueness and Regularity

We are now in the position to prove existence, uniqueness and regularity of our solution given that the initial data are sufficiently regular. Regularity means that the intial data should lie in a space H^m for appropriate m. Our problem now reads

$$\begin{cases} u \in C((-T_{\min}, T_{\max}), H^m) \cap C^1((-T_{\min}, T_{\max}), H^{m-1}) \\ i\frac{du}{dt} = Lu + V(u) \quad t \in (-T_{\min}, T_{\max}) \\ u(0) = u_0. \end{cases} \tag{2.20}$$

In addition, let the energy functional $E(u)$ be defined as

$$E(u) = \frac{1}{2} \int_{\mathbb{R}^d} |\nabla u|^2 + x^T A x |u|^2 - \tilde{V}(|u|^2)\, dx,$$

with $\tilde{V}(y) = \sum_{k=1}^n \gamma_k \frac{y^{k+1}}{k+1} + U(x)y$, where $U(x)$ is the function appearing in the power potential V. Imitating the proof of [5, theorem 4.10.1], we can obtain the following result:

Theorem 2.1. *Assume that $u_0 \in H^m$ and $m > \frac{d}{2}$. Then equation (2.20) has a unique local solution. Additionally, $\|u(t)\|_{H^m} \to \infty$ as $t \nearrow T_{\max}$ if $T_{\max} < \infty$ and $\|u(t)\|_{H^m} \to \infty$ as $t \searrow T_{\min}$ if $T_{\min} < \infty$. Moreover, the total charge (mass) $\|u\|_{L^2}$ and the total energy $E(u)$ are conserved.*

Proof. Step 1: Local existence and uniqueness. Let the semigroup generated by $-iL$ be denoted as $\mathcal{T}(t)$. If we define the operator $\mathcal{B}$ by

$$\begin{aligned} \mathcal{B}(u)(t) &:= \mathcal{T}(t)u_0 + i\mathcal{V}(t) \\ &:= \mathcal{T}(t)u_0 + i\int_0^t \mathcal{T}(t-s)V(u(s))\, ds \end{aligned}$$

then, clearly, a fixed point of $\mathcal{B}$ is the solution of our problem in integral form, which is given in proposition A.3. We want to employ the Banach fixed point theorem which will also establish uniqueness of the solution. Therefore, we need to construct a suitable Banach space on which $\mathcal{B}$ is a contracting self map. We claim that

$$\mathcal{X} := \{u \in L^\infty(I, H^m) : \|u\|_{L^\infty(I,H^m)} \le M\},$$

equipped with the metric

$$d(u, v) := \|u - v\|_{L^\infty(I,L^2)},$$

is indeed a Banach space for $I := (-T, T)$ and arbitrary T, M.
We only need to show that $(\mathcal{X}, d)$ is closed in $L^\infty(I, H^m)$ and are using lemma A.1 for this aim. We choose a sequence $(f_n)_{n\in N} \subset \mathcal{X}$ which is a Cauchy sequence in the sense of $L^\infty(I, H^m)$. As $\|.\|_{L^2} \le C\|.\|_m$, this sequence is also a Cauchy sequence in the sense of $L^\infty(I, L^2)$ and hence converges to some $f \in L^\infty(I, L^2)$. This entails that $(f_n(t))_{n\in\mathbb{N}} \to f(t)$ for almost all $t \in I$. We may choose $X = H^m$ and $Y = L^2$ because each H^m is a Hilbert space and therefore also a reflexive Banach space and the embedding $H^m \hookrightarrow L^2$ is dense

and continuous for all m. From lemma A.1 with $p = q = \infty$, we can conclude that $f \in \mathcal{X}$ and that, therefore, $\mathcal{X}$ is a Banach space.
Since we can estimate

$$\|V(u)\|_{H^m} \leq C'(M)M$$

for $\|u\|_{H^m} \leq M$ due to proposition 2.1, we deduce that $V(u) \in L^\infty(I, H^m)$ if $u \in L^\infty(I, H^m)$. So, $\mathcal{V}(u) \in C(\overline{I}, H^m)$ and hence also $u \in C(\overline{I}, H^m)$. The constant $C'(M)$ additionally depends on the constants γ_n and the highest degree $2N$ appearing in V. Moreover, we already know that the semigroup $\mathcal{T}(t)$ is a semigroup of unitary operators on all H^m. It has to pointed out that this very step is the reason why this results can only be proven in the Hilbert scales generated by L, but not in the usual Sobolev spaces $W^{2,m}$. If, as in the original proof, L is the Laplace operator, then it can be shown that the semigroup generated by this operator is an isometry on every $W^{2,m}$. This is, however, incorrect for the harmonic oscillator potential. Hence,

$$\begin{aligned}\|\mathcal{B}(u)(t)\|_m &\leq \|u_0\|_{H^m} + \int_0^t \|V(u(s))\|_{H^m}\, ds \\ &\leq \|u_0\|_{H^m} + T\|V(u)\|_{L^\infty(I,H^m)} \\ &\leq \|u_0\|_{H^m} + TC'(M)M.\end{aligned}$$

In order to prove the contraction property of $\mathcal{B}$ for accordingly chosen T, M, we compute for $u, v \in \mathcal{X}$

$$\begin{aligned}\|\mathcal{B}(u)(t) - \mathcal{B}(v)(t)\|_{L^2} &= \|\int_0^t \mathcal{T}(t-s)(V(u(s)) - V(v(s)))\, ds\|_{L^2} \\ &\leq \int_0^t \|V(u(s)) - V(v(s))\|_{L^2}\, ds \\ &\leq C(M)\int_0^t \|u - v\|_{L^2}\, ds \\ &\leq C(M)T\|u - v\|_{L^\infty(I,L^2)},\end{aligned}$$

where we have used lemma A.3 for the C^∞-map V. This is possible since functions in H^m are bounded and continuous by lemma 2.11. We note that the constant $C(M)$ already contains the Sobolev imbedding constant. Since the right-hand side does not depend on t, there also holds

$$\|\mathcal{B}(u) - \mathcal{B}(v)\|_{L^\infty(I,L^2)} \leq C(M)T\|u - v\|_{L^\infty(I,L^2)}.$$

Now we set $M = 2\|u_0\|_m$, $K = \max\{C'(M), C(M)\}$, and $KT = KT(M) = \frac{1}{2}$. It follows that $\mathcal{B}$ is a contracting selfmap on $\mathcal{X}$ and thus has a unique fixed point. This fixed point is also the unique solution of (2.20) because $V(u(s)) \in L^\infty(I, H^m) \subset L^1(I, L^2)$, $u \in H^m \subset \mathcal{D}(L)$ and $u \in C(\overline{I}, H^m) \subset L^1(I, \mathcal{D}(L))$. The equation also implies that $u_t \in C(\overline{I}, H^{m-1})$, since L maps $H^m \to H^{m-1}$.

Step 2: Blow-up alternative. Using the local solution at time T, we can start the same procedure to get a solution on a larger interval. It is, however, not possible to guarantee a global solvability, so it is not possible to bound the step size $2T$ from below independently of M. This is specified by the blow-up alternative which is being proven in the following. From the unique local solution $u \in C(\overline{I}, H^m)$, we can define a maximal solution

$$u \in C((-T_{\min}, T_{\max}), H^m) \cap C^1((-T_{\min}, T_{\max}), H^{m-1}),$$

where the differentiability follows from equation (2.20) and $T_{\min}, T_{\max}$ are defined as

$$\begin{aligned} T_{\max} &:= \sup\{T > 0 : \exists \text{ a solution of (2.20) on } [0, T]\} \\ T_{\min} &:= \sup\{T > 0 : \exists \text{ a solution of (2.20) on } [-T, 0]\}. \end{aligned}$$

Assume that $T_{\max} < \infty$ and $\|u(t_n)\|_{H^m} \leq M < \infty$ for a sequence $(t_n)_{n\in\mathbb{N}} \to T_{\max}$ from below. Now we choose k such that $t_k + T(M) > T_{\max}$. From step 1 follows that we can extend the solution to the larger interval $(-T_{\min}, T(M) + t_k)$, which contradicts the maximality of the solution. Hence, either $T_{\max} = \infty$ or $\lim_{t \nearrow T_{\max}} \|u(t)\|_{H^m} = \infty$. Exactly the same reasoning also applies for $T_{\min}$.
Step 3: Conservation laws. Since (2.20) makes sense in L^2, we can take the L^2-scalar product of both sides of the equation with u to arrive at

$$i(u_t, u) + (Lu, u) + (Vu, u) = 0.$$

Since $(Lu, u) + (Vu, u) \in \mathbb{R}$, taking the imaginary part yields

$$\frac{1}{2}\frac{d}{dt}\|u\|_{L^2}^2 = (u_t, u) = 0$$

and thus, the conservation of charge.
Multiplying (2.20) with $\overline{u_t}$ and integrating over $\mathbb{R}^n$ yields

$$i(u_t, u_t) - (Lu, u_t) - (Vu, u_t) = 0.$$

We now take the real part of this equation and integrate by parts to arrive at

$$(\nabla u, \nabla u_t) + (x^T A x\, u, u_t) + (Vu, u_t) = 0,$$

which is justified due to the regularity of u. As this is exactly $\frac{d}{dt}E(u(t)) = 0$, the proof is complete. □

3. Semi-Discretisation in Space

We will now prove a convergence result for the spectral space-discrete version of (2.1) assuming that the initial data and hence the solution by theorem 2.1 are sufficiently regular. The main tools for this proof will be provided in section 3.1. Before we start, let us clarify the notation employed throughout this chapter: For an integer K we set $\mathcal{K} = \{j \in \mathbb{N}^d : 0 \leq j_i \leq K-1\, \forall i\}$ and $\mathbf{K}-\mathbf{1} = (K-1,\dots,K-1) \in \mathbb{N}^d$. Furthermore, we say that a function of the type $p(x)e^{-x^2} = \prod_{n=1}^d p_{j_n}(x_n)e^{-x^2}$ has degree less than $\mathbf{K}$ if every p_{j_i} is of degree less than K. Moreover, let

$$X_{\mathcal{K}} := \operatorname{span}_{j\in\mathcal{K}}(h_j) \subset H^s\, \forall s \in \mathbb{R}^+.$$

We start this treatise with some information on a scaled Hermite quadrature.

3.1. Hermite Quadrature

Due to the scaling of our Hermite basis functions, we cannot apply the usual Gauss-Hermite quadrature, but instead also need to scale the quadrature nodes. Firstly, we recall the standard one-dimensional Hermite quadrature. For a polynomial $p : \mathbb{R} \to \mathbb{R}$ of degree less than $2K-1$, there holds

$$\int_{\mathbb{R}} p(x)e^{-x^2}\,dx = \sum_{k=0}^{K-1} p(\tilde{x}_k)\tilde{w}_k,$$

where the $\tilde{x}_k$ are the K distinct roots of the K-th Hermite polynomial and $\tilde{w}_k = \frac{1}{K\tilde{h}_{K-1}(\tilde{x}_k)}$. Hence, we also have a d-dimensional quadrature at hand. For p being a tensor product of one-dimensional polynomials, each of degree less than $2K-1$, we have

$$\int_{\mathbb{R}^d} p(x)e^{-|x|^2}\,dx = \sum_{j\in\mathcal{K}} p(\tilde{x}_j)\tilde{w}_j,$$

with $\tilde{x}_j = (\tilde{x}_{j_1},\dots,\tilde{x}_{j_d}) \in \mathbb{R}^d$, $\tilde{w}_j = \frac{1}{K^d}\tilde{h}_{\mathbf{K}-\mathbf{1}}(y_k)$ and $\mathbf{K}-\mathbf{1} = (K-1,\dots,K-1) \in \mathbb{N}^d$. We now establish suitable nodes and weights for a quadrature performed with the scaled Hermite functions of section 2.2. Choosing the scaled nodes as $x_j = Q^T\Lambda^{-1}\tilde{x}_j$, we see that

$$h_j(x_k) = \tilde{h}_j(\Lambda Q x_k) = \tilde{h}_j(\tilde{x}_k).$$

Since $\tilde{h}_l(x)$ is integrated exactly, there holds

$$\begin{aligned}\int_{\mathbb{R}^d} h_l(x)\,dx = \int_{\mathbb{R}^d} (\det\Lambda)^{\frac{1}{2}}\tilde{h}_j(\Lambda Qx)\,dx &= \left|\begin{array}{c}\Lambda Qx = y\\ dx = \det(\Lambda)^{-1}dy\end{array}\right|\\ &= \int_{\mathbb{R}^d} (\det\Lambda)^{-\frac{1}{2}}\tilde{h}_l(y)\,dy\\ &= (\det\Lambda)^{-\frac{1}{2}}\sum_{j\in\mathcal{K}}\tilde{h}_l(\tilde{x}_j)\tilde{w}_j\\ &= \sum_{j\in\mathcal{K}} h_l(x_j)(\det\Lambda)^{-1}\tilde{w}_j\\ &=: \sum_{j\in\mathcal{K}} h_l(x_j)w_j,\end{aligned}$$

and thus, the correct quadrature weights for this scheme are $w_j = \det(\Lambda)^{-1}\tilde{w}_j$. Therefore, we can approximate the integral of any function $f \in L^2(\mathbb{R}^d, e^{-|x|^2})$ with $f = \tilde{f}e^{-|x|^2}$ by

$$\begin{aligned}\int_{\mathbb{R}^d} f(x)\,dx &= \int_{\mathbb{R}^d} \tilde{f}(x)e^{-|x|^2}\,dx\\ &\approx \sum_{j\in\mathcal{K}} \tilde{f}(x_j)\frac{1}{(\det\Lambda)K^d h_{\mathbf{K-1}}(x_j)}.\end{aligned}$$

From the orthogonality of the scaled Hermite functions and from proposition A.4, we obtain the two important "orthogonality" properties:

Lemma 3.1. *For the scaled Hermite functions h_j and the scaled Gauss points x_k and weights w_k, there holds*

(1) $\sum_{j\in\mathcal{K}} h_j(x_l)h_{j'}(x_l)\omega_l = \delta_{jj'}$

(2) $\sum_{l\in\mathcal{K}} h_l(x_j)h_l(x_{j'})\omega_j = \delta_{jj'}$

Proof. The first claim is obvious, since the Hermite functions are integrated exactly and are orthonormal up to this scaling factor in L^2. To show the validity of the second one, we note that due to equation (A.9)

$$\sum_{k=0}^{K-1}\frac{1}{k!2^k}H_k(\tilde{x}_q)H_k(\tilde{x}_p) = 0,$$

since $\tilde{x}_q, \tilde{x}_p$ are roots of the K-th Hermite polynomial. Consequently, by multiplying with $e^{-\frac{\tilde{x}_p^2+\tilde{x}_q^2}{2}}$, we get

$$\sum_{k=0}^{K-1}\tilde{h}_k(\tilde{x}_q)\tilde{h}_k(\tilde{x}_p) = 0,$$

the same conclusion for the scaled one-dimensional Hermite functions and, thus, also for the d-dimensional ones. For the last step of the proof, we employ (A.10) to obtain

$$\sum_{k=0}^{K-1} \frac{1}{k!2^k} H_k(\tilde{x}_p) H_k(\tilde{x}_p) = \frac{1}{(K-1)!2^K} H'_K(\tilde{x}_p) H_{K-1}(\tilde{x}_p) = \frac{2K}{(K-1)!2^K} H^2_{K-1}(\tilde{x}_p),$$

where we have once again inferred that $\tilde{x}_p$ is a root of H_K and that $H'_k = 2kH_{k-1}$. Multiplying this equation with $e^{-\tilde{x}_p^2}$ once more, we arrive at

$$\sum_{k=0}^{K-1} \tilde{h}^2_k(\tilde{x}_p) = \frac{2K}{(K-1)!2^K} \tilde{h}^2_n(\tilde{x}) 2^{K-1}(K-1)! = K\tilde{h}^2_{K-1}(\tilde{x}_p) = \frac{1}{\tilde{\omega}_p}.$$

From this, one can easily conclude the same result for the scaled Hermite functions and also the d-dimensional claim. □

3.2. Projections and Interpolation

For a function $\varphi \in \tilde{H}^s$, $s > d/2$ we can think of several ways to restrict it to the span of a subset of the Hermite basis. The most intuitive is the L^2-orthogonal projection $\mathcal{P}$ onto this subspace, which is given by

$$\mathcal{P}_{\mathcal{K}}(\varphi) = \sum_{j \in \mathcal{K}} \varphi_j h_j \Leftrightarrow (\varphi, h_j) = (\mathcal{P}_{\mathcal{K}}(\varphi), h_j) \quad \forall j \in \mathcal{K}. \tag{3.1}$$

From this definition, we see that $\mathcal{P}_{\mathcal{K}}(\varphi)$ is well defined $\forall \varphi \in L^2$, so for this projection we do not need the $\tilde{H}^s$-regularity. In the following, we drop the dependence of $\mathcal{P}_{\mathcal{K}}$ on $\mathcal{K}$, so K is considered fixed for the rest of this chapter. For this projection, we can immediately prove stability:

Lemma 3.2. *For $\varphi \in \tilde{H}^s$, $s \geq \sigma$, there holds*

$$\|\varphi - \mathcal{P}(\varphi)\|_\sigma \leq CK^{\frac{\sigma - s}{2}} \|\varphi\|_s. \tag{3.2}$$

Proof. Since for $j \notin \mathcal{K}$

$$\omega_j = \sum_{i=1}^{d} \lambda_i^2 (2j_i + 1) \geq \lambda^2_{\min} \sum_{i=1}^{d} (j_i + 1) \geq \lambda^2_{\min} K,$$

there holds $\omega_j^{-p} \leq \lambda_{\min}^{-2p} K^{-p} =: CK^{-p}$ for $p \geq 0$. Furthermore, as the sum is absolutely convergent, we have

$$\varphi - \mathcal{P}(\varphi) = \sum_{j \in \mathbb{N}^d} \varphi_j h_j - \sum_{j \in \mathcal{K}} \varphi_j h_j = \sum_{j \notin \mathcal{K}} \varphi_j h_j,$$

which yields the proposed estimate by

$$\begin{aligned}\|\varphi - \mathcal{P}(\varphi)\|_\sigma^2 &= \sum_{j\notin\mathcal{K}} \omega_j^\sigma |\varphi_j|^2 = \sum_{j\notin\mathcal{K}} \omega_j^{\sigma-s+s}|\varphi_j|^2 \\ &\le \sum_{j\notin\mathcal{K}} CK^{\sigma-s}\omega_j^s|\varphi_j|^2 \le CK^{\sigma-s}\sum_{j\in\mathbb{N}^d}\omega_j^\sigma|\varphi_j|^2 \\ &= CK^{\sigma-s}\|\varphi\|_s^2.\end{aligned}$$

□

The second way of restriction that will be used is called the Hermite interpolation $\mathcal{I}_\mathcal{K}$. It is defined by

$$\mathcal{I}_\mathcal{K} : C(\mathbb{R}^d) \to X_\mathcal{K}, \quad \mathcal{I}_\mathcal{K}\varphi(x_j) = \varphi(x_j)\,\forall j\in\mathcal{K}. \tag{3.3}$$

Once again, we drop the dependence of $\mathcal{I}_\mathcal{K}$ on $\mathcal{K}$ and just write $\mathcal{I}$ instead. Since $\mathcal{I}\varphi \in X_\mathcal{K}$, it has the form

$$\mathcal{I}\varphi = \sum_{j\in\mathcal{K}} \hat{\varphi}_j h_j.$$

Bearing in mind the orthogonality of the basis functions on this discrete level, we can compute the interpolation coefficients by

$$\begin{aligned}\mathcal{I}\varphi(x_i) &= \sum_{j\in\mathcal{K}} \hat{\varphi}_j h_j(x_i) \overset{!}{=} \varphi(x_i) \\ \Leftrightarrow \sum_{j\in\mathcal{K}}\hat{\varphi}_j \sum_{i\in\mathcal{K}} h_l(x_i)h_j(x_i)w_i &\overset{!}{=} \sum_{i\in\mathcal{K}} \varphi(x_i)h_l(x_i)w_i \\ \Leftrightarrow \sum_{j\in\mathcal{K}}\hat{\varphi}_j\delta_{jl} = \hat{\varphi}_l &\overset{!}{=} \sum_{i\in\mathcal{K}}\varphi(x_i)h_l(x_i)w_i. \end{aligned}\tag{3.4}$$

Hence, for computing these coefficients we need to be able to evaluate φ in every x_j. This is feasible, if $\varphi \in \tilde{H}^s$ and hence continuous.
In the following lemma some useful information on the interpolation operator $\mathcal{I}$ is summarised

Lemma 3.3. *For $s > \frac{d}{2}$, let $\varphi, \psi \in \tilde{H}^s$. Then*

(1) $\mathcal{I} : C(\mathbb{R}^d) \to X_\mathcal{K}$ is linear,

(2) $\mathcal{I}(\varphi\psi) = \mathcal{I}(\mathcal{I}(\varphi)\mathcal{I}(\psi))$,

(3) $\int_{\mathbb{R}^d} \overline{\mathcal{I}(\varphi)}\mathcal{I}(\psi) = \sum_{j\in\mathcal{K}} \overline{\varphi(x_j)}\psi(x_j)w_j$,

(4) $\|\mathcal{I}(\varphi\psi)\|_{L^2} \le \sup_{j\in\mathcal{K}}|\varphi(x_j)|\|\mathcal{I}(\psi)\|_{L^2}$,

(5) $\mathcal{I}(\mathcal{P}\varphi) = \mathcal{P}\varphi$.

Proof. (1) is obvious by the definition of the coefficients $\hat{x}_j$.
(3) follows from the fact that $\mathcal{I}\varphi$ and $\mathcal{I}\psi$ each are of degree less than $\mathbf{K}$, and that for functions of degree less than $\mathbf{2K-1}$ the quadrature is exact.
Moreover, we can compute

$$\begin{aligned}
(\widehat{\mathcal{I}\psi\mathcal{I}\varphi})_i &= \sum_{p,l,k\in\mathcal{K}} w_p h_i(x_p)\psi(x_l)\varphi(x_k)\sum_{j\in\mathcal{K}} w_l h_j(x_l)h_j(x_p)\sum_{n\in\mathcal{K}} w_k h_n(x_k)h_n(x_p)\\
&= \sum_{p,l,k} w_p h_i(x_p)\psi(x_l)\varphi(x_k)\delta_{lp}\delta_{kp}\\
&= \sum_{p} w_p h_i(x_p)\psi(x_p)\varphi(x_p)\\
&= \widehat{(\varphi\psi)}_i,
\end{aligned}$$

which shows (2).
Using (3), in order to prove (4) we compute

$$\begin{aligned}
\|\mathcal{I}(\varphi\psi)\|_0^2 &= \int_{\mathbb{R}^d}\mathcal{I}(\varphi\psi)\overline{\mathcal{I}(\varphi\psi)}\,dx\\
&= \sum_{j\in\mathcal{K}}\varphi(x_j)\psi(x_j)\overline{\varphi(x_j)\psi(x_j)}w_j\\
&\leq \sup_{j\in\mathcal{K}}|\varphi(x_j)|^2\sum_{j\in\mathcal{K}}\psi(x_j)\overline{\psi(x_j)}w_j\\
&= \sup_{j\in\mathcal{K}}|\varphi(x_j)|^2\int_{\mathbb{R}^d}|\mathcal{I}(\psi)|^2\,dx.
\end{aligned}$$

Finally, (5) is indeed valid if $\psi = \mathcal{I}\psi\,\forall\psi\in X_{\mathcal{K}}$. This is the case due to

$$\begin{aligned}
\sum_{j\in\mathcal{K}}\hat{\psi}_j h_j &= \sum_{j\in\mathcal{K}}\sum_{l\in\mathcal{K}} h_j(x_l)\psi(x_l)w_l h_j\\
&= \sum_{j\in\mathcal{K}}\sum_{l\in\mathcal{K}} h_j(x_l)\sum_{\iota\in\mathcal{K}}\psi_\iota h_\iota(x_l)w_l h_j\\
&= \sum_{j\in\mathcal{K}} h_j\sum_{\iota\in\mathcal{K}}\psi_\iota\delta_{\iota j}\\
&= \sum_{j\in\mathcal{K}}\psi_j h_j.
\end{aligned}$$

□

We can also prove a stability result for the interpolation operator. Firstly, we show this result for the one-dimensional case according to [8], before we extend it to arbitrary space dimensions.

Lemma 3.4. *Let $\varphi\in W^{1,2}(\mathbb{R})$. Then there holds*

$$\|\mathcal{I}\varphi\|_{L^2}\leq C\left(\sum_{n=0}^{1}K^{-\frac{n}{6}}|\varphi|_{W^{n,2}}\right).$$

Proof. Before we can actually prove this claim, we need to clarify some notations and cite results from various papers. If $d = 1$, then the matrix A is just a constant and hence, the new Hermite Gauss-Points $x_{N,k}$ are $\lambda y_{N,k}$, where $y_{N,k}$ are the original Hermite Gauss-Points for $k = 0, \ldots, 1$. For the new weights $w_{N,k}$, there holds $w_{N,k} = \lambda^{-1}\tilde{w}_{N,k}$. We explicitely write out the dependence of the points and weights on N in this proof only, since we are concerned with bounds that depend on N. Furthermore, we set

$$\tilde{\Lambda}_{N,k} = (y_{N,k-1}, y_{N,k+1}), \quad \tilde{\Delta}_{N,k} = y_{N,k+1} - y_{N,k-1}, \quad a_N = \sqrt{2N}.$$

From [16] we know that there exist constants $C_m > 0$ such that

$$-a_{N+1}(1 - N^{-\frac{2}{3}}) \leq C_1 y_{N,0}, \quad y_{N,N} \leq C_2 a_{N+1}(1 - N^{-\frac{2}{3}}), \tag{3.5}$$

for $k = 1, \ldots, N-1$

$$C_3 \frac{1}{\sqrt{N+1}} \left(1 - \frac{|y_{N,j}|}{a_{N+1}}\right)^{-\frac{1}{2}} \leq \tilde{\Delta}_{N,k} \leq C_4 \frac{1}{\sqrt{N+1}} \left(1 - \frac{|y_{N,j}|}{a_{N+1}}\right)^{-\frac{1}{2}}, \tag{3.6}$$

and from [9] we cite that for $k = 0, \ldots, N$

$$C_5 \frac{1}{\sqrt{N}} \left(1 - \frac{|y_{N,j}|}{a_{N+1}}\right)^{-\frac{1}{2}} \leq \tilde{w}_{N,k} \leq C_6 \frac{1}{\sqrt{N}} \left(1 - \frac{|y_{N,j}|}{a_{N+1}}\right)^{-\frac{1}{2}}. \tag{3.7}$$

It can be observed that it is feasible to choose $C_1 < 1$ and $C_2 > 1$, and that (3.7) and (3.6) also hold for $x_{N,k}$ and $w_{N,k}$, if the constants are chosen accordingly. From these formulas we deduce some estimates that will be needed for the actual proof of our claim. Equation (3.5) yields for $k = 0, N$

$$\left(1 - \frac{|y_{N,k}|}{a_{N+1}}\right)^{-\frac{1}{2}} \leq CN^{1/3}. \tag{3.8}$$

Additionally, we deduce

$$N^{-\frac{1}{2}} \Delta_{N,k}^{-1} \left(1 - \frac{|y_{N,k}|}{a_{N+1}}\right)^{-\frac{1}{2}} \leq C \tag{3.9}$$

from (3.6), and

$$\begin{aligned} N^{-\frac{1}{2}} \Delta_{N,k} \left(1 - \frac{|y_{N,k}|}{a_{N+1}}\right)^{-\frac{1}{2}} &\leq CN^{-1} \left(1 - \frac{|y_{N,k}|}{a_{N+1}}\right)^{-1} \\ &\leq CN^{-1} N^{\frac{2}{3}} = CN^{-\frac{1}{3}}, \end{aligned} \tag{3.10}$$

from combining (3.6) and (3.8). The last result we need to employ reads

$$\sup_{x \in [a,b]} |\varphi(x)|^2 \leq \frac{c}{b-a} \|\varphi\|^2_{L^2(a,b)} + c(a-b)|\varphi|^2_{H^1(a,b)} \tag{3.11}$$

and can be found in [2].
Thus, we can conclude from

$$
\begin{aligned}
\|\mathcal{I}\varphi\|_0 &= \sum_{k=0}^{N} |\varphi(x_k)|^2 w_k \\
&\leq CN^{-\frac{1}{2}} \sum_{k=0}^{N} |\varphi(x_k)|^2 \left(1 - \frac{|y_{N,k}|}{a_{N+1}}\right)^{-\frac{1}{2}} \\
&\leq CN^{-\frac{1}{6}} \Bigg(\|\varphi\|_\infty^2 + \sum_{k=1}^{N-1} \Delta_{N,k}^{-1} \left(1 - \frac{|y_{N,k}|}{a_{N+1}}\right)^{-\frac{1}{2}} \|\varphi\|^2_{L^2(\Lambda_{N,k})} \\
&\quad + \sum_{k=1}^{N-1} \Delta_{N,k} \left(1 - \frac{|y_{N,k}|}{a_{N+1}}\right)^{-\frac{1}{2}} |\varphi|^2_{H^1(\Lambda_{N,k})} \Bigg) \\
&\leq C(N^{-\frac{1}{6}} \|\varphi\|^2_{H^1} + \|\varphi\|^2_{L^2} + N^{-\frac{1}{3}} |\varphi|^2_{H^1}) \\
&\leq C(N^{-\frac{1}{6}} (\|\varphi\|^2_{L^2} + |\varphi|^2_{H^1}) + \|\varphi\|^2_{L^2} + N^{-\frac{1}{3}} |\varphi|^2_{H^1}) \\
&\leq C(\|\varphi\|^2_{L^2} + N^{-\frac{1}{3}} |\varphi|^2_{H^1}),
\end{aligned}
$$

where we used equation (3.11) and the continuous embedding of H^1 into L^∞. □

Lemma 3.5. *Let $\varphi \in W^{d,2}$. Then there holds*

$$
\|\mathcal{I}\varphi\|_{L^2} \leq C \left(\sum_{n=0}^{d} K^{-\frac{n}{6}} |\varphi|_{W^{n,2}} \right).
$$

Proof. Before commencing with the actual proof, let us define some notations. For a function $\varphi : \mathbb{R}^d \to \mathbb{C}$, we denote by $\|\varphi\|_{L^2_j}$ with a multiindex j the norm with respect to the coordinates given in j. For instance, $\|\varphi\|_{L^2_{1,\dots,d-1}}$ denotes the L^2-norm with respect to the first $d-1$ coordinates. Furthermore, for x_j being a $d+1$-dimensional node with according weight w_j, we set $j' = (j_1, \dots, j_{d-1})$ and, thus, $x_{j'} = x_{j_1,\dots,j_{d-1}}$ is a $d-1$-dimensional node with according weight $w_{j'}$. Therefore, $w_{j'} w_{j_d} = w_j$ and, $(x_{j'}, x_{j_d}) = x_j$. Finally, we write $|.|_{H^k}$ for the Sobolev seminorm and $d' = (1, \dots, d)$. We prove the claim by induction and note that we have already established the validity of the hypothesis for the seed $d = 1$ in lemma 3.4. Additionally, we can estimate

$$
\left| \|\varphi\|_{L^2} \frac{\partial}{\partial x_l} \|\varphi\|_{L^2} \right| = \frac{1}{2} \left| \frac{\partial}{\partial x_l} \|\varphi\|_{L^2} \right| \leq \left| \left(\varphi, \frac{\partial}{\partial x_l} \varphi \right)_{L^2} \right| \leq \|\varphi\|_{L^2} \|\frac{\partial}{\partial x_l} \varphi\|_{L^2}.
$$

For the induction step $d \to d+1$, we assume that

$$
\|\mathcal{I}\varphi\|_{L^2} \leq C \left(\sum_{n=0}^{d} K^{-\frac{n}{6}} |\varphi|_{W^{n,2}} \right)
$$

is valid and note that

$$
\begin{aligned}
\|\mathcal{I}\varphi\|_{L^2}^2 &= \sum_{j\in\mathcal{K}_d} w_j|\varphi(x_j)|^2 = \sum_{j'\in\mathcal{K}_{d-1}} w_{j'} \sum_{j_d=0}^{K-1} |\varphi(x_j)|^2 \\
&= \sum_{j'\in\mathcal{K}_{d-1}} w_{j'} |\|\mathcal{I}_d\varphi\|_{L^2(\mathbb{R})}(x_{j'})|^2 = \|\mathcal{I}_{1,\dots,d-1}(\|\mathcal{I}_d(\varphi)\|_{L^2_d})\|_{L^2_{1,\dots,d-1}}.
\end{aligned}
$$

Thus, there holds

$$
\begin{aligned}
\|\mathcal{I}\varphi\|_{L^2(\mathbb{R}^{d+1})} &= \|\mathcal{I}_{d'}(\|\mathcal{I}_{d+1}\varphi\|_{L^2_{d+1}})\|_{L^2_{d'}} \\
&\le C\|\mathcal{I}_{d'}(\|\varphi\|_{L^2_{d+1}} + K^{-\frac{1}{6}}|\varphi|_{H^1_{d+1}})\|_{L^2_{d'}} \\
&\le C\|\mathcal{I}_{d'}\|\varphi\|_{L^2_{d+1}}\|_{L^2_{d'}} + K^{-\frac{1}{6}}\||\varphi|_{H^1_{d+1}}\|_{L^2_{d'}} \\
&\le C\left(\sum_{k=0}^{d} K^{-\frac{k}{6}}\big|\|\varphi\|_{L^2_{d+1}}\big|_{H^k_{d'}} + K^{-\frac{1}{6}}\sum_{k=0}^{d} K^{-\frac{k}{6}}\big||\varphi|_{H^1_{d+1}}\big|_{H^k_{d'}}\right) \\
&\le C\left(\sum_{k=0}^{d} K^{-\frac{k}{6}}|\varphi|_{H^k_{1,\dots,d+1}} + K^{-\frac{1}{6}}\sum_{k=0}^{d} K^{-\frac{k}{6}}|\varphi|_{H^{k+1}_{1,\dots,d+1}}\right) \\
&\le C\sum_{k=0}^{d+1} K^{-\frac{k}{6}}|\varphi|_{H^k_{1,\dots,d+1}},
\end{aligned}
$$

due to

$$
\begin{aligned}
\||\frac{\partial}{\partial x_{d+1}}\varphi\|_{L^2_{d+1}}|_{W^{l,2}_{d'}} &= \max_{|\alpha|=l}\|D^\alpha\|\frac{\partial}{\partial x_{d+1}}\varphi\|_{L^2_{d+1}}\|_{L^2_{d+1}} \le \max_{|\alpha|=l+1}\|D^\alpha\varphi\|_{L^2(\mathbb{R}^{d+1})} \\
&= |\varphi|_{W^{l+1,2}_{1,\dots,d+1}},
\end{aligned}
$$

and

$$
\begin{aligned}
\||\varphi\|_{L^2_{d+1}}|_{W^{l,2}_{d'}} &= \max_{|\alpha|=l}\|D^\alpha\|\varphi\|_{L^2_{d+1}}\|_{L^2_{d'}} \le \max_{|\alpha|=l}\|\|D^\alpha\varphi\|_{L^2_{d+1}}\|_{L^2_{d'}} = \max_{|\alpha|=l}\|D^\alpha\varphi\|_{L^2(\mathbb{R}^{d+1})} \\
&= |\varphi|_{W^{l,2}_{1,\dots,d+1}}.
\end{aligned}
$$

□

Using the stability results for $\mathcal{I}$ and $\mathcal{P}$, we can now prove

Proposition 3.1. *Let $\mathbb{N} \ni s \ge d, \sigma$ and $\varphi \in \tilde{H}^s$. Then there exists $C \in \mathbb{R}^+$ independent of φ and K such that*

$$\|\varphi - \mathcal{I}\varphi\|_\sigma \le CK^{\frac{d}{3}+\frac{\sigma-s}{2}}\|\varphi\|_s$$

Proof. We begin by estimating

$$
\begin{aligned}
\|\mathcal{I}(\mathcal{P}\varphi - \varphi)\|_\sigma &\leq CK^{\frac{\sigma}{2}} \|\mathcal{I}(\mathcal{P}\varphi - \varphi)\|_\sigma \leq CK^{\frac{\sigma}{2}} \left(\sum_{l=1}^{d} K^{-\frac{l}{6}} \left|\mathcal{P}\varphi - \varphi\right|_{W^{l,2}} \right) \\
&\leq CK^{\frac{\sigma}{2}} \left(\sum_{l=1}^{d} K^{-\frac{l}{6}} \|\mathcal{P}\varphi - \varphi\|_{\tilde{H}^l} \right) \leq CK^{\frac{\sigma}{2}} \left(\sum_{l=1}^{d} K^{-\frac{l}{6}} K^{\frac{l-s}{2}} \|\varphi\|_{\tilde{H}^s} \right) \\
&\leq CK^{\frac{\sigma}{2}} \left(\sum_{l=1}^{d} K^{\frac{l}{3}-\frac{s}{2}} \|\varphi\|_{\tilde{H}^s} \right) \leq CdK^{\frac{\sigma}{2}} K^{\frac{d}{3}-\frac{s}{2}} \|\varphi\|_{\tilde{H}^s} \\
&=: C' K^{\frac{d}{3}+\frac{\sigma-s}{2}} \|\varphi\|_{\tilde{H}^s}
\end{aligned}
$$

Since $\mathcal{I}(\mathcal{P}\varphi) = \mathcal{P}\varphi$ according to lemma 3.3(5), we may insert $\mathcal{I}(\mathcal{P}\varphi) - \mathcal{P}\varphi$ on the left-hand side and then use the triangular inequality. We also know that for $j \in \mathcal{K}$

$$
\omega_j = \sum_{i=1}^{d} \lambda_i (2j_i + 1) \leq \max_i(\lambda_i) d(2K+1) \leq \max_i(\lambda_i) d(3K) =: C(\lambda, d)K
$$

and hence, for $\psi \in X_\mathcal{K}$, there holds

$$
\|\psi\|_\sigma^2 = \sum_{j\in\mathcal{K}} \omega_j^\sigma |\psi_j|^2 \leq CK^\sigma \sum_{j\in\mathcal{K}} |\psi_j|^2 = CK^\sigma \|\psi\|_0^2.
$$

Combining all this information, we finally can conclude

$$
\begin{aligned}
\|\varphi - \mathcal{I}\varphi\|_\sigma &\leq \|\mathcal{I}(\mathcal{P}\varphi - \varphi)\|_0 + \|\varphi - \mathcal{P}\varphi\|_\sigma \leq \|\varphi\|_{\tilde{H}^s} \left(C' K^{\frac{d}{3}+\frac{\sigma-s}{2}} + CK^{\frac{\sigma-s}{2}} \right) \\
&\leq \max\{C, C'\} K^{\frac{d}{3}+\frac{\sigma-s}{2}} \|\varphi\|_{\tilde{H}^s}.
\end{aligned}
$$

□

3.3. Convergence Analysis

In the following part of chapter 3, a space semi-discrete version of (2.1) is derived. That means that the space-dependence of the functions is removed by applying a collocation scheme to obtain a system of ordinary differential equations in time. For this purpose, we assume that the solution $\psi(x,t)$ can be written as $\psi(x,t) = \tilde{\psi}(t)g(x)$ with $g \in \tilde{H}^s$ for some s. Then we choose a finite dimensional subspace $X_\mathcal{K}$ of our space solution space $\tilde{H}^s$ and a set of collocation points, which will be the Hermite-Gauss points $(x_j)_{j\in\mathcal{K}}$. Obviously, $(h_j)_{j\in\mathcal{K}}$ is then basis of $X_\mathcal{K}$ and $X_\mathcal{K} \to L^2$ as $K \to \infty$. As usual, the collocation scheme is to expand the numerical solution ψ_K in the basis of $X_\mathcal{K}$ with time-dependent coefficients $\psi_j(t)$ and to demand that the equation is fulfilled in the collocation points. We then arrive at the system

$$
\begin{aligned}
&i\frac{d}{dt}\psi_K(\tilde{x}_k, t) = ((\Delta + x^T A x)\psi_K)(x_k, t) + \left(\sum_{n=1}^{N} \gamma_n |\psi_K|^{2n} \psi_K + U\psi_K \right)(x_k, t) \quad t > 0, k \in \mathcal{K} \\
&\psi_K(x_k, 0) = \psi(x_k, 0).
\end{aligned} \tag{3.12}
$$

Although it is not obvious in this notation, this system is uniquely solvable, which is being proven in

Proposition 3.2. *System (3.12) has a unique local solution.*

Proof. We start with rewriting system (3.12) in the form $i\psi(t) = f(\psi(t))$, with a vector $\psi(t) \in \mathbb{C}^{K^d}$ to clarify how the right-hand side function f operates. For this aim, we multiply (3.12) with $w_k h_p(x_k)$ and then sum over $k \in \mathcal{K}$. This yields

$$\begin{aligned} i\sum_{j\in\mathcal{K}} \psi_j'(t) \sum_{k\in\mathcal{K}} h_j(x_k)h_p(x_k)w_k = &\sum_{j,k\in\mathcal{K}} L\psi_j(t)h_j(x_k)h_p(x_k) + \sum_{j,k\in\mathcal{K}} U(x_k)\psi_j(t)h_j(x_k)h_p(x_k)w_k \\ &+ \left(\sum_{n=1}^{N} \gamma_n \left(\sum_{m,l\in\mathcal{K}} \psi_m(t)\overline{\psi_l(t)}h_m(x_k)h_l(x_k)\right)^n \sum_{j\in\mathcal{K}} \psi_j(t)h_j\right)(x_k). \end{aligned} \tag{3.13}$$

As this leads to very long expressions, we split the computation into three parts. Firstly, the left-hand side of (3.13) reduces to

$$i\sum_{j\in\mathcal{K}} \psi_j'(t) \sum_{k\in\mathcal{K}} h_j(x_k)h_p(x_k)w_k = i\sum_{j\in\mathcal{K}} \psi_j'(t)\delta_{jp} = i\psi_p'(t).$$

Secondly, for the linear part of the right-hand side of (3.13) there holds

$$\begin{aligned} &\sum_{k\in\mathcal{K}} L \sum_{j\in\mathcal{K}} \psi_j(t)h_j(x_k)h_p(x_k)w_k + \sum_{k\in\mathcal{K}} U(x_k)h_p(x_k)w_k \sum_{j\in\mathcal{K}} \psi_j(t)h_j(x_k) = \\ &\quad = \sum_{j,k\in\mathcal{K}} \psi_j(t)\omega_j h_j(x_k)h_p(x_k)w_k + \sum_{j,k\in\mathcal{K}} U(x_k)h_p(x_k)w_k h_j(x_k)\psi_j(t) \\ &\quad =: \omega_p\psi_p(t) + {b_p}^T \cdot M \cdot \psi, \end{aligned}$$

with $M_{kj} = h_j(x_k)$ and $b_p = U(x_k)h_p(x_k)w_k$. Hence, this part is also linear in the finite dimensional case.

Finally, the nonlinear part of (3.13) reads

$$\begin{aligned} &\sum_{k\in\mathcal{K}} h_p(x_k)w_k \sum_{n=1}^{N} \gamma_n \left(\sum_{m,l\in\mathcal{K}} \psi_m(t)\overline{\psi}_l(t)h_m(x_k)h_l(x_k)\right)^n \sum_{j\in\mathcal{K}} \psi_j(t)h_j(x_k) = \\ &\quad = \sum_{n=1}^{N} \gamma_n \sum_{j,k\in\mathcal{K}} \psi_j(t)h_j(x_k)h_p(x_k) \left(\psi(t)^T B_k \overline{\psi}(t)\right), \end{aligned}$$

with $(B_k)_{lm} = h_m(x_k)h_l(x_k)$. As this is a polynomial in $\psi(t)$ and $\overline{\psi}(t)$, we can write system (3.12) as

$$i\psi(t) = C\psi(t) + p(\psi(t), \overline{\psi}(t)) =: f(\psi(t)),$$

where C is a matrix and p a polynomial. Thus, f is Fréchet differentiable (in the real sense) and, therefore, f is also locally Lipschitz continuous. We conclude that there exists a unique local solution of (3.12) due to the Piccard-Lindelöf theorem.

□

Since ψ and $\mathcal{I}\psi$ coincide in x_k , system (3.12) can also be written as

$$\begin{cases} i\frac{d}{dt}\psi_K = ((\Delta + x^T Ax)\psi_K) + \mathcal{I}(\sum_{n=1}^N \gamma_n|\psi_K|^{2n}\psi_K + U\psi_K) & t > 0, \\ \psi_K(.,0) = \mathcal{I}\psi(.,0). \end{cases} \tag{3.14}$$

For this system, we will now state and prove the main result of this chapter.

Theorem 3.1. *Let $\mathbb{N} \ni s > 2 + \lceil \frac{d+1}{2} \rceil + \frac{2d}{3}$ and $T > 0$. Assume that the exact solution $\psi(x,t)$ of (2.20) is in $\tilde{H}^s$ for $t \in [0,T]$ and set $B_s = \sup_{t\in[0,T]} \|\psi(.,t)\|_s$. Then*

$$\|\psi_K(t) - \psi(t)\|_\sigma \leq CK^{1+\frac{d}{3}+\frac{\sigma-s}{2}} \ \forall t \in [0,T],$$

holds for all $K \geq K_0$, where C and K_0 depend on $\sigma, s, d, B_s, A, (|\gamma_n|)_{n\leq N}, N$ and U.

Proof. We start by computing the Hermite interpolation of (2.20). Additionally, we want to differentiate after interpolating and therefore have to show that $\frac{d}{dt}(\mathcal{I}\psi) = \mathcal{I}(\frac{d}{dt}\psi)$. We can only prove this claim if we can write $\frac{d}{dt}\psi = \sum_{j\in\mathbb{N}^d} \tilde{\psi}_j(t)h_j$, which is fulfilled if $\frac{d}{dt}\psi \in L^2(\mathbb{R}^d)\,\forall t \in [0,T]$. If the series converges $\forall t \in [0,T]$ in the sense of $L^2(\mathbb{R}^d)$, it also converges in the sense of $L^2([0,T] \times \mathbb{R}^d)$. Hence, we can choose a testfunction $\varphi \in C_0^\infty([0,T] \times \mathbb{R}^d)$ and compute

$$\begin{aligned}
(\frac{d}{dt}\sum_{j\in\mathbb{N}^d} \psi_j(t)h_j(x), \varphi(x,t))_{L^2([0,T],\mathbb{R}^d)} &= -(\sum_{j\in\mathbb{N}^d} \psi_j(t)h_j(x), \frac{d}{dt}\varphi(x,t))_{L^2([0,T],\mathbb{R}^d)} \\
&= -\lim_{K\to\infty} (\sum_{j\in\mathcal{K}} \psi_j(t)h_j(x), \frac{d}{dt}\varphi(x,t))_{L^2([0,T],\mathbb{R}^d)} \\
&= \lim_{K\to\infty} (\sum_{j\in\mathcal{K}} \frac{d}{dt}\psi_j(t)h_j(x), \varphi(x,t))_{L^2([0,T],\mathbb{R}^d)} \\
&= (\sum_{j\in\mathbb{N}^d} \frac{d}{dt}\psi_j(t)h_j(x), \varphi(x,t))_{L^2([0,T],\mathbb{R}^d)}.
\end{aligned}$$

This yields $\frac{d}{dt}\psi = \frac{d}{dt}\sum_{j\in\mathbb{N}^d} \psi_j(t)h_j = \sum_{j\in\mathbb{N}^d} \frac{d}{dt}\psi_j(t)h_j$. Since we can obviously compute $\frac{d}{dt}\mathcal{I}(\psi)$ by differentiating the coefficients (the sums are finite) and as the computation of the coeffcients of $\mathcal{I}(\frac{d}{dt}\psi)$ is linear, we may pull the differentiation into the interpolation operator. With this reasoning, the Hermite interpolation of (2.20) reads

$$\begin{aligned}
i\frac{\partial}{\partial t}\mathcal{I}(\psi) &= \mathcal{I}(L\psi) + \sum_{n=1}^N \gamma_n \mathcal{I}(|\psi|^{2n}\psi) + \mathcal{I}(U\psi) \\
&= \mathcal{I}(L\psi) + \sum_{n=1}^N \gamma_n \mathcal{I}(\mathcal{I}|\psi|^{2n}\mathcal{I}\psi) + \mathcal{I}(U\psi),
\end{aligned}$$

due to lemma 3.3(2). We subtract this from (3.14) and keep in mind that $\mathcal{I}\psi_K = \psi_K$, due to lemma 3.3(5), to arrive at

$$i\frac{\partial}{\partial t}\delta_K = L\psi_K - \mathcal{I}(L\psi) + \sum_{n=1}^N \gamma_n\big(\mathcal{I}(|\psi_K|^{2n}\psi_K) - \mathcal{I}(|\psi|^{2n}\psi)\big) + U\psi_K - \mathcal{I}(U\psi),$$

for $X_{\mathcal{K}} \ni \delta_K = \psi_K - \mathcal{I}\psi$. Subsequently, we insert the two zeros $L\psi - L\psi$ and $L(\mathcal{I}\psi) - L(\mathcal{I}\psi)$ and rearrange some of the terms. This gives

$$i\frac{\partial}{\partial t}\delta_K = L\delta_K + \sum_{n=1}^{N} \gamma_n \mathcal{I}\big((|\psi_K|^{2n}\psi_K) - \mathcal{I}(|\psi|^{2n}\psi)\big) + U\psi_K - \mathcal{I}(U\psi) + L(\mathcal{I}\psi - \psi) + L\psi - \mathcal{I}(L\psi)$$

and furthermore, by lemma 3.3(2),

$$i\frac{\partial}{\partial t}\delta_K = L\delta_K + \sum_{n=1}^{N} \gamma_n \mathcal{I}\big((|\psi_K|^{2n}\psi_K) - \mathcal{I}(|\mathcal{I}\psi|^{2n}\mathcal{I}\psi)\big) + U\psi_K - \mathcal{I}(U\psi) + L(\mathcal{I}\psi - \psi) + L\psi - \mathcal{I}(L\psi).$$

We now apply lemma A.1 to the term $\sum_{n=1}^{N} \gamma_n \big((|\psi_K|^{2n}\psi_K) - \mathcal{I}(|\mathcal{I}\psi|^{2n}\mathcal{I}\psi)\big)$. In the notation of this lemma, we obtain

$$\begin{aligned}\sum_{n=1}^{N} \gamma_n \big((|\psi_K|^{2n}\psi_K) - \mathcal{I}(|\mathcal{I}\psi|^{2n}\mathcal{I}\psi)\big) = \sum_{n=1}^{N} \gamma_n \big((g(\mathcal{I}\psi, \psi_K, n)\delta_K + f_2(\mathcal{I}\psi, \psi_K, n))\delta_K \\ + f_1(\mathcal{I}\psi, \psi_K, n)\overline{\delta}_K\big).\end{aligned}$$

To shorten the formulas, we now define

$$\eta_K = \sum_{n=1}^{N} \gamma_n \big(f_1(\mathcal{I}\psi, \psi_K, n)\overline{\delta_K} + f_2(\mathcal{I}\psi, \psi_K, n)\delta_K\big) + L(\mathcal{I}\psi - \psi) + (L\psi - \mathcal{I}L\psi). \quad (3.15)$$

Since $\delta_K \in X_{\mathcal{K}}$ and g is a real mapping,

$$\int_{R^d} \overline{\delta_K}\mathcal{I}\left(\sum_{n=1}^{N} \gamma_n g(\mathcal{I}\psi, \psi_K, n)\delta_K\right) dx \in \mathbb{R},$$

by lemma 3.3(3). Furthermore, there holds

$$\begin{aligned}\mathcal{I}(U\psi_K) - \mathcal{I}(U\psi) &= \mathcal{I}(\mathcal{I}(U)\mathcal{I}(\psi_K)) - \mathcal{I}(\mathcal{I}(U)\mathcal{I}(\psi)) \\ &= \mathcal{I}(\mathcal{I}(U)\delta_K)\end{aligned}$$

and thus, due to U being a real map, lemma 3.3(3) yields

$$\int_{\mathbb{R}^d} \overline{\delta_K}\mathcal{I}(U\delta_K)\, dx \in \mathbb{R}.$$

Hence, multiplying (3.15) with $\overline{\delta_K}$, integrating over $\mathbb{R}^d$ and taking the imaginary part of the resulting equation gives, since $(v_t, v) = \frac{1}{2}\frac{\partial}{\partial t}\|v\|^2 = \|v\|\frac{\partial}{\partial t}\|v\| \in \mathbb{R}$,

$$(\frac{\partial}{\partial t}\delta_K, \delta_K)_{L^2} = \mathrm{Im}(\delta_K, \eta_K)_{L^2},$$

and moreover, due to Cauchy-Schwarz inequality and $\mathcal{I}\psi(0) = \psi_K(0)$, there holds

$$\begin{aligned}
&\|\delta_K\|_0 \frac{\partial}{\partial t}\|\delta_K\|_0 \leq \|\delta_K\|_{L^2}\|\eta_K\|_{L^2} \\
\Rightarrow\ &\frac{\partial}{\partial t}\|\delta_K(t)\|_{L^2} \leq \|\eta_K(t)\|_{L^2} \\
&\Rightarrow \|\delta_K(t)\|_{L^2} \leq \int_0^t \|\eta_K(t)\|_{L^2}\,dt - \|\delta_K(0)\|_{L^2} = \int_0^t \|\eta_K(t)\|_{L^2}\,dt,
\end{aligned}$$

since $\psi_K(0) = \mathcal{I}\psi(0)$.
Before continuing, we need estimates for the different terms contained in η_K. Firstly, by propositon 3.1 with $\sigma = 2$

$$\|L(\mathcal{I}\psi - \psi)\|_0 = \|\mathcal{I}\psi - \psi\|_2 \leq CK^{\frac{d}{3}+1-\frac{s}{2}}\|\psi\|_s.$$

Secondly, due to proposition 3.1 with $\sigma = 0$ and $s = s - 2$

$$\|L\psi - \mathcal{I}(L\psi)\|_0 \leq CK^{\frac{d}{3}-\frac{s-2}{2}}\|L\psi\|_{s-2} = CK^{\frac{d}{3}+1-\frac{s}{2}}\|\psi\|_s.$$

Finally, we turn to the most elaborate part,

$$F(\psi_K, \mathcal{I}\psi, N) := \mathcal{I}\left(\sum_{n=1}^{N} \gamma_n\big(f_1(\mathcal{I}\psi, \psi_K, n)\overline{\delta_K} + f_2(\mathcal{I}\psi, \psi_K, n)\delta_K\big)\right).$$

Applying the triangular inequality to $\|F\|_0$ yields

$$\|F(\psi_K, \mathcal{I}\psi, N)\|_0 \leq \sum_{n=1}^{N} |\gamma_n|\|\mathcal{I}(f_1(\psi_K, \mathcal{I}\psi, n)\delta_K)\|_0 + \|\mathcal{I}(f_2(\psi_K, \mathcal{I}\psi, n)\delta_K)\|_0.$$

Each term $\|f_k(\psi_K, I\psi, n)\delta_K\|_0$ for $k = 1, 2$ and $n \in \{1, \dots, N\}$ can be bounded by using 3.3(4), $\psi(x_j) = \mathcal{I}\psi(x_j)\ \forall j \in \mathcal{K}$, and lastly lemma 2.11 for an arbitrary integer $\sigma' > \frac{d}{2}$:

$$\begin{aligned}
\|\mathcal{I}\left(f_k(\psi_K, \mathcal{I}\psi, n)\delta_K\right)\|_0 &\leq f_k(\sup_{j\in\mathcal{K}} |\psi_K(x_j)|, \sup_{j\in\mathcal{K}} |\mathcal{I}(\psi)(x_j)|, n)\|\delta_K\|_0 \\
&= f_k(\sup_{j\in\mathcal{K}} |\psi_K(x_j)|, \sup_{j\in\mathcal{K}} |\psi(x_j)|, n)\|\delta_K\|_0 \\
&\leq f_k(\sup_{x\in\mathbb{R}} |\psi_K(x)|, \sup_{x\in\mathbb{R}} |\psi(x)|, n)\|\delta_K\|_0 \\
&\leq f_k(C\|\psi_K\|_{\sigma'}, C\|\psi(x)\|_{\sigma'}, n)\|\delta_K\|_0.
\end{aligned}$$

Hence,

$$\begin{aligned}
\|F(\psi_K, \mathcal{I}\psi, N)\|_0 &\leq \|\delta_K\|_0 \sum_{n=1}^{N}\big(|\gamma_n| f_1(C\|\psi_K\|_{\sigma'}, C\|\psi(x)\|_{\sigma'}, n) + f_2(C\|\psi_K\|_{\sigma'}, C\|\psi(x)\|_{\sigma'}, n)\big) \\
&=: G(\psi_K, \psi, N)\|\delta_K\|_0.
\end{aligned}$$

We now return to our estimate of δ_K:

$$\begin{aligned}\|\delta_K(t)\|_0 &\leq \int_0^t G(\psi_K,\psi,N)(\tau)\|\delta_K(\tau)\|_0\,d\tau + 2CK^{\frac{d}{3}+1-\frac{s}{2}}\|\psi(t)\|_s\\ &\leq \int_0^t \sup_{\theta\in[0,\tau]} G(\psi_K,\psi,N)(\tau)(\theta)\|\delta_K(\tau)\|_0\,d\tau + 2CK^{\frac{d}{3}+1-\frac{s}{2}}\sup_{\theta\in[0,\tau]}\|\psi(\theta)\|_s.\end{aligned}$$

Our next step is the employment of Gronwall's inequality A.4. We note that $\alpha(\tau) = 2CK^{\frac{d}{3}+1-\frac{s}{2}}\sup_{\theta\in[0,\tau]}\|\psi(\theta)\|_s$ is non-decreasing in τ. Thus,

$$\|\delta_K(t)\|_0 \leq \alpha(t)\exp\left(\int_0^t \beta(\xi)\,d\xi\right),$$

with $\beta(\xi) = \sup_{\theta\in[0,\xi]} G(\psi_K,\mathcal{I}\psi,N)(\theta)$. Since ω_j is monotonously increasing in j, for any $\varphi\in X_{\mathcal{K}}$ we can estimate

$$\|\varphi\|_\sigma^2 = \sum_{j\in\mathcal{K}}\omega_j^\sigma|(\varphi)_j|^2 \leq C_\sigma K^\sigma\sum_{j\in\mathcal{K}}|(\varphi)_j|^2 = C_\sigma K^\sigma\|\varphi\|_0.$$

Hence, for arbitrary $\sigma\leq s$, we obtain

$$\begin{aligned}\|\delta_K(t)\|_\sigma &\leq C_\sigma K^{\frac{\sigma}{2}}\|\delta_K(t)\|_0\\ &\leq 2C_1K^{\frac{d}{3}+1+\frac{\sigma-s}{2}}\exp\left(\int_0^t\sup_{\theta\in[0,\tau]}G(\psi_K,\psi,N)(\theta)\,d\tau\right)\sup_{\theta\in[0,t]}\|\psi(\theta)\|_s\\ &\leq 2C_1K^{\frac{d}{3}+1+\frac{\sigma-s}{2}}\exp\left(\int_0^t\sup_{\theta\in[0,t]}G(\psi_K,\psi,N)(\theta)\,d\tau\right)\sup_{\theta\in[0,t]}\|\psi(\theta)\|_s\\ &\leq 2C_1K^{\frac{d}{3}+1+\frac{\sigma-s}{2}}\exp\left(t\sup_{\theta\in[0,t]}G(\psi_K,\psi,N)(\theta)\right)\sup_{\theta\in[0,t]}\|\psi(\theta)\|_s.\end{aligned}$$

This yields, by applying the triangular inequality and proposition 3.1,

$$\begin{aligned}\|\psi_K(t)-\psi(t)\|_\sigma &\leq \|\psi_K(t)-\mathcal{I}\psi(t)\|_\sigma + \|\mathcal{I}\psi(t)-\psi(t)\|_\sigma\\ &\leq \|\delta_K\|_\sigma + C_2K^{\frac{d}{3}+1+\frac{\sigma-s}{2}}\|\psi\|_s\\ &\leq C_3K^{\frac{d}{3}+1+\frac{\sigma-s}{2}}\sup_{\theta\in[0,t]}\|\psi(\theta)\|_s\left(2\exp\left(t\sup_{\theta\in[0,t]}G(\psi_K,\psi,N)(\theta)\right)+1\right),\end{aligned}$$

where $C_3 = \max\{C_1,C_2\}$.
Now we are almost done. We are only left to show that G can be bounded independently of ψ_K. For this aim, we fix an interval $[0,\Theta]$ in which

$$\sup_{\theta\in[0,\Theta]}\|\psi_K(\theta)\|_\sigma \leq 2C_4\sup_{\theta\in[0,\Theta]}\|\psi(\theta)\|_s,$$

where $C_4 = \max\{C_3, 1\}$. This assumption is justified, since it is valid at $\theta = 0$, and since both functions are continuous in time. We can extend this interval to $\Theta = T$ because for $t \in [0, \Theta]$ (note that G is increasing in both of its arguments) there holds

$$\begin{aligned}\|\psi_K(t)\|_\sigma &\leq \|\psi_K(t) - \psi(t)\|_\sigma + \|\psi(t)\|_\sigma \\ &\leq \|\psi(t)\|_s + C_3 K^{\frac{d}{3}+1+\frac{\sigma-s}{2}} \sup_{\theta\in[0,t]} \|\psi(\theta)\|_s \left(2\exp\left(t \sup_{\theta\in[0,t]} G(\psi,\psi,N)(\theta)\right) + 1\right) \\ &\leq 2C_4 \sup_{\theta\in[0,t]} \|\psi(\theta)\|_s,\end{aligned}$$

for K larger than $\tilde{K}$ with

$$\tilde{K}^{\frac{d}{3}+1+\frac{\sigma-s}{2}} \sup_{\theta\in[0,t]} \|\psi(\theta)\|_s \left(2\exp\left(t \sup_{\theta\in[0,t]} G(\psi,\psi,N)(\theta)\right) + 1\right) = 1.$$

This last result finally yields

$$\|\psi_K(t) - \psi(t)\|_\sigma \leq CK^{\frac{d}{3}+1+\frac{\sigma-s}{2}} B(\sup_{\theta\in[0,t]} \|\psi(\theta)\|_s).$$

$\square$

4. Semi-Discretisation in Time

4.1. The Calculus of Lie Derivatives

In this section, a short introduction to the calculus of Lie derivatives is given. Since we need to employ an extended formalism, which can be used in an infinite dimensional Hilbert space, each definition is presented for the finite dimensional, intuitive case before it is subsequently extended. The results presented in this section can be found in [10, section 5.1] for the finite dimensional case and in the appendix of [13] for the infinite dimensional case.
Let for the rest of this section F, G be continuous vector fields $\mathbb{R}^d \to \mathbb{R}^d$, X a Hilbert space and, furthermore, $\mathcal{F}, \mathcal{G}$ continuous, but possibly unbounded and nonlinear operators with suitable domains $\mathcal{D}(\mathcal{F}) \subset X$ and $\mathcal{D}(\mathcal{G}) \subset X$. Then, for $y \in \mathbb{R}^d$ we have

Definition 4.1. $D_F := \sum_k F^k(y) \frac{\partial}{\partial y_k}$ *is called the* Lie derivative *with respect to* F.

Clearly, D_F is a linear differential operator. The Lie derivative of a differentiable vector field G is hence given by

$$D_F G(y) = G'(y) F(y).$$

Note that this is independent of the dimension of the underlying space. We therefore define for $y \in X$

Definition 4.2. $D_{\mathcal{F}}\mathcal{G}(y) := \mathcal{G}'(y)\mathcal{F}(y)$ *is called the* Lie derivative *of* $\mathcal{G}$ *with respect to* $\mathcal{F}$.

The next step is to extend the wellknown notions of the *flow* and the *flow operator* or *evolution operator* generated by a vector field. In $\mathbb{R}^d$ the flow generated by F is defined as the solution of

$$\dot{y}(t) = F(y(t)), \quad t \in (0, T], \quad y(0) = y_0, \tag{4.1}$$

and commonly denoted as $\varphi_F^t(y_0)$ or $\varphi_F(t, y_0)$. The evolution operator is the operator which maps the initial data and $t \in (0, T]$ to the solution φ_F. If F is linear, then by the theory of semigroups of linear operators, $\varphi_F(t, y_0) = \mathcal{T}(t)(y_0)$, where $\mathcal{T}(t)$ is the semigroup generated by F. Even if F is nonlinear, but Lipschitz continuous and T small enough, the flow operators form a semigroup of linear operators due to the unique solvability of (4.1) on a suitable time interval. This semigroup can be represented with the help of the exponential function and the Lie derivative, which follows directly from

Lemma 4.1.

$$\frac{d}{dt} G(\varphi_F^t(y_0)) = D_F G(\varphi_F^t(y_0)) \tag{4.2}$$

$$G(\varphi_F(t, y_0)) = \sum_{k=0}^{\infty} \frac{t^k}{k!} (D_F^k G)(y_0) = \exp(t D_F) G(y_0). \tag{4.3}$$

Proof. Claim (4.2) is easily proven by applying the chain rule

$$\frac{d}{dt}G(\varphi_F^t(y_0)) = G'(\varphi_F^t(y_0))\dot{\varphi}_F^t(y_0) = G'(\varphi_F^t(y_0))F(\varphi_F^t(y_0)) = D_F G(\varphi_F^t(y_0)), \tag{4.4}$$

and will serve as the induction seed for the hypothesis $\frac{d^k}{dt^k}G(\varphi_F^t(y_0)) = D_F^k G(\varphi_F^t(y_0))$. We compute

$$\begin{aligned}\frac{d^{k+1}}{dt^{k+1}}G(\varphi_F(t,y_0)) &= \frac{d^k}{dt^k}(D_F G(\varphi_F(t,y_0)))\\ &= D_F^k(D_F G(\varphi_F(t,y_0)))\\ &= D_F^{k+1}G(\varphi_F(t,y_0)),\end{aligned}$$

whence (4.3) follows from Taylor expansion. □

Setting $G = \mathrm{Id}$ in (4.3), yields

$$\varphi_F^t(y_0) = \exp(tD_F)y_0$$

as the representation of the flow operator.
Moreover, both

$$\frac{d}{dt}\exp(tD_F)|_{t=0}\, y_0 = D_F \mathrm{Id}\, y_0 \tag{4.5}$$

and

$$\exp(sD_F)\exp(tD_F)y_0 = \exp(tD_F)\exp(sD_F)y_0 = \exp((s+t)D_F)y_0 \tag{4.6}$$

obviously holds true, if s and t are small enough.
A rather curious consequence of (4.3) is Gröbner's Vertauschungssatz, which reveals how the composition of two flow operators has to be defined.

Proposition 4.1 (Vertauschungssatz)**.** *The composition of two flow operators applied to a vector y_0 is given by*

$$e^{tD_F}e^{sD_G}y_0 = \varphi_G(s, \varphi_F(t, y_0))$$

for $s, t \in \mathbb{R}$.

Proof. This statement exactly is equation (4.3) with G substituted by e^{sD_G}. □

In the infinite dimensional setting, the evolution operators are defined analogously and, under suitable conditions at $\mathcal{F}$, they also form a semigroup. However, the exponential cannot be defined without further undesired restrictions on $\mathcal{F}$. Contrariwise, since (4.2) still holds and because of the semigroup property, it is still common to write

$$\varphi_{\mathcal{F}}(t, y_0) = e^{tD_{\mathcal{F}}}y_0,$$

but only as a formal expression. What is more, if we combine two or more solution operators, we have to be aware of proposition 4.1 and reverse the order in which we apply the flow operators to the initial data.
The next perception that needs to be generalised is the commutator of two operators.

Definition 4.3. *Let $\mathcal{F}, \mathcal{G}$ be two operators.*

(1) The Lie commutator *of $\mathcal{F}$ and $\mathcal{G}$ is given by*

$$[\mathcal{F}, \mathcal{G}](y) := \mathcal{F}'(y)\mathcal{G}(y) - \mathcal{G}'(y)\mathcal{F}(y) \tag{4.7}$$

(2) The iterated commutator $\mathrm{ad}^j_{\mathcal{F}}(\mathcal{G})$ *is recursively defined by*

$$\mathrm{ad}^j_{\mathcal{F}}(\mathcal{G}) = [\mathcal{F}, \mathrm{ad}^{j-1}_{\mathcal{F}}(\mathcal{G})], \quad \mathrm{ad}^0_{\mathcal{F}}(\mathcal{G}) = \mathcal{G} \tag{4.8}$$

Clearly, (4.8) makes sense both for linear and nonlinear operators, and (4.7) coincides with the usual commutator for linear operators.
Since one of our goals will be to bound iterated commutators of Lie derivatives of operators, we will rewrite them in terms of the appearing operators themselves.

Lemma 4.2. *For a j-times Fréchet differentiable A there holds*

$$\mathrm{ad}^j_{D_{\mathcal{F}}}(D_{\mathcal{G}})A(v) = (-1)^j D_{\mathrm{ad}^j_{D_{\mathcal{F}}}(\mathcal{G})}A(v) \tag{4.9}$$

Proof. First, we establish the induction seed:

$$\begin{aligned}
[D_{\mathcal{F}}, D_{\mathcal{G}}]Av &= D_{\mathcal{F}}A'(v)\mathcal{G}(v) - D_{\mathcal{G}}A'(v)\mathcal{F}(v) \\
&= A''(v)(\mathcal{G}(v), \mathcal{F}(v)) + A'(v)\mathcal{G}'(v)\mathcal{F}(v) - A''(v)(\mathcal{F}(v), \mathcal{G}(v)) - A'(v)\mathcal{F}'(v)\mathcal{G}(v) \\
&= A'(v)(\mathcal{G}'(v)\mathcal{F}(v) - \mathcal{F}'(v)\mathcal{G}(v)) \\
&= A'(v)[\mathcal{G}, \mathcal{F}](v) \\
&= -D_{[\mathcal{F},\mathcal{G}]}Av.
\end{aligned}$$

For the induction step, we compute

$$\begin{aligned}
\mathrm{ad}^{j+1}_{D_{\mathcal{F}}}(D_{\mathcal{G}})Av &= [D_{\mathcal{F}}, \mathrm{ad}^j_{D_{\mathcal{F}}}(D_{\mathcal{G}})]Av \\
&= (-1)^j [D_{\mathcal{F}}, D_{\mathrm{ad}^j_{\mathcal{F}}(\mathcal{G})}]Av \\
&= (-1)^j D_{-\mathrm{ad}^{j+1}_{\mathcal{F}}(\mathcal{G})}Av \\
&= (-1)^{j+1} D_{\mathrm{ad}^{j+1}_{\mathcal{F}}(\mathcal{G})}Av.
\end{aligned}$$

□

An explicit expression for the n-th iterated Lie commutator, if one of the operators is linear, is of particular interest for our problem.

Lemma 4.3. *Let $\mathcal{F}, \mathcal{G}$ be as above, and, in addition, $\mathcal{F}$ be linear. Then there holds*

$$\mathrm{ad}^n_{\mathcal{F}}(\mathcal{G})(v) = \sum_{\substack{p+|m(n,k)|=n \\ m(n,k)\in\mathbb{N}^k_+}} \mathcal{F}^p \left(c_{n,m(n,k)} \mathcal{G}(v)^{(k)}(\mathcal{F}^{m_1}(v), \ldots, \mathcal{F}^{m_k}(v)) \right),$$

where each $m(n,k)$ is a multi-index (for each k), and each $c_{n,m(n,k)}$ is a real constant.

Proof. The claim is proven by induction. We choose $n = 1$ as the induction seed and see from

$$\begin{aligned}\mathrm{ad}_{\mathcal{F}}(\mathcal{G})(v) &= \mathcal{G}'(v)\mathcal{F}(v) - \mathcal{F}'(v)\mathcal{G}(v)\\ &= \mathcal{F}^0\mathcal{G}^{(1)}(v)\mathcal{F}^1(v) + \mathcal{F}^1(-\mathcal{G}^{(0)}(v))\end{aligned}$$

that the hypothesis is fulfilled. Let us now assume that the claim is true for $n \in \mathbb{N}$. Since $\mathrm{ad}^{n+1}_{\mathcal{F}}(\mathcal{G})(v) = \mathcal{F}'(v)\mathrm{ad}^n_{\mathcal{F}}(\mathcal{G})(v) - (\mathrm{ad}^n_{\mathcal{F}}(\mathcal{G}))'(v)\mathcal{F}(v)$, we see that for the second term we will need an expression for the derivative of the n-th commutator. Again, the first term is of the form demanded, since for any linear operator $\mathcal{F}'(v)u = \mathcal{F}(u)$, so p becomes $p+1$ and hence the summation goes to $n + 1$. Before we compute the second term, we note that for a linear operator $\mathcal{F}$ also $\mathcal{F}^p$ is linear and hence $\mathcal{F}^p(u)'(v) = \mathcal{F}^p(v)$. Thus, we have

$$\begin{aligned}(\mathrm{ad}^n_{\mathcal{F}}(\mathcal{G}))'(v)\mathcal{F}(v) &= \sum_{\substack{p+|m(n,k)|=n\\ m(n,k)\in\mathbb{N}^k_+}} c_{n,m(n,k)} \left((\mathcal{F}(v))^p G^{(k)}(v)((\mathcal{F})^{m(n,k)_1}, \dots, (\mathcal{F}')^{m(n,k)_k})(v)\right)' \mathcal{F}(v)\\
&= \sum_{\substack{p+|m(n,k)|=n\\ m(n,k)\in\mathbb{N}^k_+}} c_{n,m(n,k)} \mathcal{F}^p \Bigg(G^{(k+1)}(v)(\mathcal{F}^{m(n,k)_1}, \dots, \mathcal{F}^{m(n,k)_k}, \mathcal{F})(v)\\
&\quad + \sum_{q=2}^{k+1} \Big(\frac{\partial G^{(k)}(v)}{\partial q}(\mathcal{F}^{m(n,k)_1}, \dots, \mathcal{F}^{m(n,k)_k})(v) \cdot (\mathcal{F}^{m(n,k)_q})'(v)\mathcal{F}(v)\Big)\Bigg)\\
&= \sum_{\substack{p+|m(n,k)|=n\\ m(n,k)\in\mathbb{N}^k_+}} c_{n,m(n,k)} \mathcal{F}^p \Bigg(G^{(k+1)}(v)\big(\mathcal{F}^{m(n,k)_1}, \dots, \mathcal{F}^{m(n,k)_k}, \mathcal{F}\big)(v)\\
&+ \sum_{q=2}^{k+1} \Big(G^{(k)}(v)\big(\mathcal{F}^{m(n,k)_1}, \dots, \mathcal{F}^{m(n,k)_{q-1}}, \mathcal{F}^{m(n,k)_{q+1}}, \dots, \mathcal{F}^{m(n,k)_k}, \mathcal{F}^{m(n,k)_q+1}\big)(v)\Big)\Bigg),\end{aligned}$$

where we have used the chain rule for Fréchet derivatives, the fact that the k-th derivative, seen as a function in $k+1$ variables, is linear in every variable except in the first one and the notation $\frac{\partial}{\partial q}$ for the partial derivative with respect to the q-th variable. This expression is of the form we demanded, since p stays the same, and either k is raised to $k+1$ with

$$p + |m(n, k + 1)| := p + |(m(n, k), 1)| = p + |m(n, k)| + 1 = n + 1,$$

or k also stays the same but one index of $m(n, k)$ is raised by one, and hence

$$p + |m(n + 1, k)| := p + |m(n, k)| + 1 = n + 1.$$

□

It has to be pointed out that in the definition of the summation in lemma 4.3 there is also an implicit restriction on k. Since $m(n, k) \in \mathbb{N}^k_+$, there holds $|m(n, k)| \geq k$ and hence also $p + k \leq n$.

4.2. Operator Splitting

Operator splitting is a very popular technique that enables a numerical treatise of general evolution equations of the form

$$u_t = Au + Bu, \quad u(0) = u_0 \tag{4.10}$$

for general (possibly nonlinear) operators A and B acting on a Hilbert space X that may be finite or infinite dimensional. If A and B were linear, bounded operators (with suitable, compatible domains), then the theory of semigroups of linear operators states that the solution would be given by

$$u(t) = \exp(t(A+B))u_0 = \sum_{k=0}^{\infty} \frac{t^k}{k!}(A+B)^k u_0.$$

The technique of operator splitting of order m is rewriting this expression by a product of $\exp(a_j t D_A)$ and $\exp(d_j t D_B)$ and a remainder term R, where the coefficients a_j and b_j are chosen in a way that the desired order of convergence m is obtained. In other words, we have

$$\begin{aligned} u(t) = \exp(tD_{A+B})u_0 &= \prod_{j=1}^{s} \exp(a_{s+1-j} t D_A) \exp(b_{s+1-j} t D_B) u_0 + R(t, u_0) \\ &= \Phi^n(t) u_0 + R(t, u_0), \end{aligned} \tag{4.11}$$

where $R = \mathcal{O}(t^m)$ as $t \to 0$. We write D_A and D_B instead of A and B as there is no difference for linear operators A, B but (4.11) stays valid also for nonlinear operators. The correct a_j and b_j are computed by equating coefficients of t^n for $n \leq m$ in the Taylor expansions of the left and right-hand sides of (4.11).

For nonlinear and/or unbounded operators, the procedure is exactly the same except that we equate the coefficients in a *formal* Taylor expansion. This is, however, not sufficient for the method to converge. To achieve convergence, we additionally have to be able to bound iterated commutators of D_A and D_B up to a certain order, which is specified in theorem 4.1.

This method is obviously beneficient if the evolution operators $\exp(tD_A)$ and $\exp(tD_B)$ can easily be found and be implented numerically with at least the same order of convergence as the operator splitting itself.

The convergence of this numerical scheme is usually proven by substantiating step-by-step-stability and convergence of the local error, which finally leads to convergence of the global error.

Firstly, we will prove that a splitting method of an arbitrary order is stable if each split step is stable. In the following, the dependence of Φ^n on t will be omitted.

Lemma 4.4. *Let Φ_k be one step of the semi-discrete solution operator $\Phi^n u_0 = \prod_{k=1}^{n} \Phi_k u_0$. If all the Φ_k are stable in the sense that*

$$\|\Phi_k(\varphi) - \Phi_k(\psi)\| \leq e^{C_k h} \|\varphi - \psi\|, \quad k \in \{1, \dots, n\},$$

for an arbitrary norm $\|.\|$, then also Φ is stable in the same sense with the stability constant $C = \sum_{k=1}^{n} C_k$.

Proof. The claim is obviously true for $n = 1$. For the induction step we compute

$$\begin{aligned}\|\Phi^{n+1}(\varphi) - \Phi^{n+1}(\psi)\| &\leq \|\Phi_{n+1}(\Phi^n(\varphi)) - \Phi_{n+1}(\Phi^n(\psi))\| \\ &\leq e^{C_{n+1}h}\|\Phi^n(\varphi) - \Phi^n(\psi)\| \\ &\leq e^{\sum_{k=1}^{n+1} C_k h}\|\varphi - \psi\|\end{aligned}$$

□

A local error expansion for splitting methods of arbitrary order is established in [13]. This result appears to be rather cumbersome, but the important aspect is that we only have to be able to bound iterated Lie-commutators D_A and D_B to employ it. The theorem reads

Theorem 4.1. *Let $T_k = \{\tau = (\tau_1, \tau_2, \ldots, \tau_k) \in \mathbb{R}^k : 0 \leq \tau_k \leq \cdots \leq \tau_1 \leq \tau_0 = t\}$ and $L_k = \{\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{N}^k : 1 \leq \lambda_k \leq \cdots \leq \lambda_1 \leq \lambda_0 = s\}$. Provided that the condition $c_s = 1$ is satisfied, where $c_k = a_1+a_2+\cdots+a_k$ for $1 \leq s$, the defect operator (the difference between the real solution operator and the discrete solution operator) of the exponential splitting method (4.11), when applied to the nonlinear evolution equation (4.10), admits the formal expansion*

$$\mathcal{E}(t, v) = \sum_{k=1}^{p} \sum_{\substack{\mu \in N^k \\ |\mu| \leq p-k}} \frac{1}{\mu!} t^{k+|\mu|} C_{k\mu} \prod_{l=1}^{k} \mathrm{ad}_{D_A}^{\mu_l}(D_B) \exp(tD_A)v + \mathcal{R}_{p+1}(t, v), \quad 0 \leq t \leq T, \tag{4.12}$$

$$C_{k\mu} = \sum_{\lambda \in L_k} \alpha_\lambda \prod_{l=1}^{k} b_{\lambda_l} c_{\lambda_l}^{\mu_l} - \prod_{l=1}^{k} \frac{1}{\mu_l + \cdots + \mu_k + k - l + 1}, \tag{4.13}$$

with a remainder term $\mathcal{R}(t, v) = o(t^{p+1})$.

4.3. Convergence Analysis

We are now well equipped with theoretical tools to prove convergence of an operator splitting scheme applied to equation (2.20). We split it into the two subproblems

$$i\frac{\partial}{\partial t}\psi(x, t) = L\psi = (-\Delta + x^T Ax)\psi(x, t) \tag{4.14}$$

$$i\frac{\partial}{\partial t}\psi(x, t) = V\psi = \left(\sum_{n=1}^{N} \gamma_n |\psi(x, t)|^{2n} + U(x)\right)\psi(x, t). \tag{4.15}$$

In the previous section we demanded that we should have access to the explicit form of the solution operators (or to good numerical approximations of them at least) in order to be able to apply the splitting method advantageously. Since (4.14) is linear, the solution operator $\exp(tD_L)$ is just the semigroup $\mathcal{T}(t)$ that has $-iL$ as its infinitesimal generator. We already know that this semigroup acts on a function $\varphi \in X_{\mathcal{K}}$ by multiplying each Hermite coefficient φ_j with $e^{-i\omega_j t}$ and hence this scheme is very easy to implement. For $\exp(D_V t)$ we have the following result.

Lemma 4.5. *The solution operator* $\exp(D_V t)$ *is given by*

$$\exp(D_V t)u_0(x) = \exp\left(-it\left(\sum_{n=1}^{N} |u_0(x)|^{2n} + U(x)\right)\right) u_0(x)$$

Proof. The claim is clear if along solutions of (4.15), the absolute value of the solution does not change, since the equation then becomes

$$i\frac{\partial}{\partial t}\psi(x,t) = \left(\sum_{n=1}^{N} \gamma_n |u_0|^{2n} + U(x)\right)\psi(x,t), \quad \psi(x,0) = u_0(x),$$

and the only function fulfilling this equation is obviously

$$\exp(D_V t)u_0(x) = \exp\left(-it\left(\sum_{n=1}^{N} |u_0(x)|^{2n} + U(x)\right)\right) u_0(x).$$

To see that $|\psi|$ is constant, we multiply (4.14) with $\overline{\psi}$ and drop the dependency on x and t, which yields

$$i\overline{\psi}\psi_t = \left(\sum_{n=1}^{N} \gamma_n |\psi|^{2n} + U\right)|\psi|^2 \in \mathbb{R}.$$

Conjugating (4.15), and then multiplying it with ψ, we furthermore obtain

$$i\overline{\psi_t}\psi = \left(\sum_{n=1}^{N} \gamma_n |\psi|^{2n} + U\right)|\psi|^2 \in \mathbb{R}.$$

Since $\overline{(\psi_t)} = \frac{\partial}{\partial t}\overline{\psi} =: \overline{\psi_t}$ and the time derivative of a real valued function has to be real as well, there holds

$$\mathbb{R} \ni |\psi|_t = (\psi\overline{\psi})_t^{1/2} = \frac{(\psi\overline{\psi})_t}{2|\psi|} = \frac{\psi\overline{\psi}_t + \psi_t\overline{\psi}}{2|\psi|} \in \mathbb{R}i,$$

and hence $|\psi|_t = 0$ along solutions of (4.15). $\square$

The following proposition provides the stability of one step of the splitting method in both the L^2 and the $\tilde{H}^s$-norms, if the real solution is bounded in $\tilde{H}^s$.

Proposition 4.2. *Let* $s > d$ *be an even integer,* Φ^n *a splitting operator of order* n *with step size* h *and* $\varphi, \psi \in \tilde{H}^s$*. Then*

$$\|\Phi^n(\psi) - \Phi^n(\varphi)\|_\sigma \le e^{Ch}\|\psi - \varphi\|_\sigma, \tag{4.16}$$

where the constant C *depends on* $\sigma, d, A, (|\gamma_k|)_{k\le N}, N, n, U, \|\psi\|_s, \|\varphi\|_s$ *and* σ *is either* s *or* 0*.*

Proof. Due to lemma 4.4 we only have to show the claim for $n = 1$. Therefore, the discrete evolution operator has the form $\Phi^n(t)\varphi = \exp(-itD_V)\exp(-itD_L)\varphi$. If e^{-itD_V} were stable in the sense of (4.16), then we would have

$$\begin{aligned}\|e^{-ihD_T}e^{-ihD_V}\psi - e^{-ihD_T}e^{-ihD_V}\varphi\|_s &\leq e^{Ch}\|\mathcal{T}(t)\psi - \mathcal{T}(t)\varphi\|_s \\ &= e^{Ch}\|\psi - \varphi\|_s,\end{aligned}$$

since $\mathcal{T}(t)$ is linear and unitary on every $\tilde{H}^s$ by lemma 2.8. Hence, the assertion holds if we can establish the stability of $\exp(-itD_V)$. For this aim, let θ and η be the solutions of

$$\begin{aligned} i\frac{\partial}{\partial t}\theta &= \sum_{n=1}^{N}\gamma_n|\psi|^{2n}\theta + U(x)\theta, \quad \theta(0) = \psi \\ i\frac{\partial}{\partial t}\eta &= \sum_{n=1}^{N}\gamma_n|\varphi|^{2n}\eta + U(x)\eta, \quad \eta(0) = \varphi. \end{aligned}$$

Then we have to estimate

$$\begin{aligned}\theta(h) - \eta(h) &= \exp(-ihD_V)\psi - \exp(-ihD_V)\varphi \\ &= \exp\left(-ih\left(\sum_{n=1}^{N}\gamma_n|\psi|^{2n}+U(x)\right)\right)\psi - \exp\left(-ih\left(\sum_{n=1}^{N}\gamma_n|\varphi|^{2n}+U(x)\right)\right)\varphi\end{aligned}$$

We note that θ can be bounded (exactly the same reasoning leads to a bound for η) due to

$$\begin{aligned}\|\theta(t)\|_s &= \|\int_0^t \frac{\partial}{\partial\tau}\theta(\tau)\,d\tau - \theta(0)\|_s \\ &\leq \|\theta(0)\|_s + \int_0^t \|\frac{\partial}{\partial\tau}\theta(\tau)\|_s\,d\tau \\ &= \|\psi\|_s + \int_0^t \|\left(\sum_{n=1}^{N}\gamma_n|\psi|^{2n} + U\right)\theta\|_s\,d\tau \\ &= \|\psi\|_s + \int_0^t \left(\sum_{n=1}^{N} C^{2n}|\gamma_n|\|\psi\|_s^{2n} + C\|U\|_s\right)\|\theta\|_s\,d\tau,\end{aligned}$$

where we have used the assumption that $\tilde{H}^s$ is a normed algebra several times. Hence, we can employ Gronwall's inequality (lemma A.4) for the special case that $\alpha(t) = \|\psi\|_s$ is nondecreasing (since it is constant) to obtain

$$\begin{aligned}\|\theta(t)\|_s &\leq \|\psi\|_s\exp(\int_0^t\sum_{n=1}^{N}|\gamma_n|C_s^{2n}\|\psi\|_s^{2n} + C_s\|U\|_s\,d\tau) \\ &\leq \|\psi\|_s\exp(C't\sum_{n=1}^{N}|\gamma_n|\|\psi\|_s^{2n} + \|U\|_s\,d\tau) \\ &=: \|\psi\|_s\exp(A(\psi)t),\end{aligned}$$

where $C' = \max_{n=1,\dots,2N} C_s^n$.
In the following, we derive a bound for $\|\frac{\partial}{\partial t}(\theta-\eta)\|_s$. Though it would be easy to find a bound in terms of products of ψ and φ, we need to find a bound that can be expressed as a sum of terms containing $(\theta-\eta)$ and $(\psi-\varphi)$ at least once, a form which we need to apply Gronwall's inequality once again. For this purpose, we have proven lemma A.2 which we use to rewrite

$$
\begin{aligned}
-i(\theta-\eta)_t &= \sum_{n=1}^{N} \gamma_n(|\psi|^{2n}\theta - |\varphi|^{2n}\eta) + U(\theta-\eta)\\
&= \sum_{n=1}^{N} \gamma_n(|\psi|^{2n}(\theta-\eta) + (|\psi|^{2n} - |\varphi|^{2n})\eta) + U(\theta-\eta)\\
&= \sum_{n=1}^{N} \gamma_n(|\psi|^{2n}(\theta-\eta) + (h_1(\psi,\varphi,n)(\psi-\varphi) + h_2(\psi,\varphi,n)\overline{(\psi-\varphi)})\eta) + U(\theta-\eta).
\end{aligned}
$$

Therefore, we can estimate

$$
\begin{aligned}
\|(\theta-\eta)_t\|_\sigma \le \|U(\theta-\eta)\|_\sigma &+ \sum_{n=1}^{N} |\gamma_n| \big(\||\psi|^{2n}(\theta-\eta)\|_\sigma \\
&+ \|(h_1(\psi,\varphi,n)(\psi-\varphi) + h_2(\psi,\varphi,n)\overline{(\psi-\varphi)})\eta\|_\sigma \big),
\end{aligned}
$$

for an arbitrary integer $\sigma \ge 0$. Since $\varphi,\psi \in \tilde{H}^s$, also $h_k(\psi,\varphi,n) \in \tilde{H}^s$ since $\tilde{H}^s$ is a normed algebra and $h_k(\psi,\varphi,n)$ is a polynomial in $\psi,\varphi,\overline{\psi},\overline{\varphi}$. Hence, for arbitrary $\rho \in \tilde{H}^s$, due to lemma 2.12, we can estimate

$$
\|\rho h_k(\psi,\varphi,n)\|_0 \le \|\rho\|_0 \|h_k(\psi,\varphi,n)\|_s \quad \text{and} \quad \|\rho h_k(\psi,\varphi,n)\|_s \le \|\rho\|_s \|h_k(\psi,\varphi,n)\|_s.
$$

It is due to this step in the proof, that we need the solution ψ to be in $\tilde{H}^s$, even if we are only interested in L^2-stability. Let us return to our estimate of $\|(\theta-\eta)_t\|_\sigma$. Since $(\theta-\eta)_t(t), \psi-\varphi$ and $\eta(t)$ are elements of $\tilde{H}^s$, this reasoning allows us to estimate the appearing products even further for $\sigma = 0$ or $\sigma = s$:

$$
\begin{aligned}
\|(\theta-\eta)_t\|_\sigma &\le \|\theta-\eta\|_\sigma \left(\sum_{n=1}^{N} |\gamma_n| \left(\|\psi\|_s^{2n} C_{s\sigma}^{2n} \right) + C_{s\sigma}\|U\|_s \right)\\
&\quad + C_{s\sigma}^2 \|\eta\|_s \|\psi-\varphi\|_\sigma \left(\sum_{n=1}^{N} |\gamma_n| \left(\|h_1(\psi,\varphi,n)\|_s + \|h_2(\psi,\varphi,n)\|_s \right) \right)\\
&\le \|\theta-\eta\|_\sigma \left(\sum_{n=1}^{N} |\gamma_n| \left(\|\psi\|_s^{2n} C_{s\sigma}^{2n} \right) + C_{s\sigma}\|U\|_s \right)\\
&\quad + C_{s\sigma}^2 \|\varphi\|_s \exp(A(\varphi)t) \|\psi-\varphi\|_\sigma \left(\sum_{n=1}^{N} |\gamma_n| \left(\|h_1(\psi,\varphi,n)\|_s + \|h_2(\psi,\varphi,n)\|_s \right) \right)\\
&=: \|\theta-\eta\|_\sigma F(\varphi,\psi) + \|\psi-\varphi\|_s \exp(A(\varphi)t) G(\psi,\varphi).
\end{aligned}
$$

The σ-norm of $h_k(\psi, \varphi, n)$ for $k = 1, 2$ and $n = 1, \ldots, N$ can be bounded by a term depending on the σ-norms φ and ψ due to lemma 2.12. We can conclude with the Gronwall inequality, since

$$
\begin{aligned}
\|\theta - \eta\|_\sigma &\leq \|\int_0^t (\theta - \eta)_\tau \, d\tau - (\theta - \eta)(0)\|_\sigma \\
&\leq \int_0^t \|(\theta - \eta)_\tau\|_\sigma \, d\tau + \|(\theta - \eta)(0)\|_\sigma \\
&\leq \int_0^t \|\theta - \eta\|_\sigma F(\varphi, \psi) + \|\psi - \varphi\|_\sigma \exp(A(\varphi)t) G(\psi, \varphi) \, d\tau + \|\psi - \varphi\|_\sigma \\
&\leq \int_0^t \|\theta - \eta\|_\sigma F(\varphi, \psi) \, d\tau + \frac{1}{A(\varphi)} \|\psi - \varphi\|_\sigma \exp(A(\varphi)t) G(\psi, \varphi) + \|\psi - \varphi\|_\sigma \\
&= \int_0^t \|\theta - \eta\|_\sigma F(\varphi, \psi) \, d\tau + (\frac{1}{A(\varphi)} \exp(A(\varphi)t) G(\psi, \varphi) + 1) \|\psi - \varphi\|_\sigma \\
&= \int_0^t \|\theta - \eta\|_\sigma F(\varphi, \psi) \, d\tau + (\frac{1}{A(\varphi)} \exp(A(\varphi)t) G(\psi, \varphi) + 1) \|\psi - \varphi\|_\sigma .
\end{aligned}
$$

As $\alpha(t) = (\frac{1}{A(\varphi)} \exp(A(\varphi)t) G(\psi, \varphi) + 1) \|\psi - \varphi\|_\sigma$ is again nondecreasing, we obtain

$$
\begin{aligned}
\|\theta - \eta\|_\sigma &\leq (\frac{1}{A(\varphi)} \exp(A(\varphi)t) G(\psi, \varphi) + 1) \|\psi - \varphi\|_\sigma \exp\left(\int_0^t F(\varphi, \psi) \, d\tau \right) \\
&\leq \exp(t(F(\varphi, \psi) + A(\varphi) + 1)) \|\psi - \varphi\|_\sigma ,
\end{aligned}
$$

because $1 \leq e^t$ for $t \geq 0$. □

The next step to our convergence result is establishing a bound for the local error.

Proposition 4.3. *Let $s > d$ be an even integer, $B_\sigma = \sup_{t \in (0,T]} \|\psi(t)\|_\sigma$, $\mathcal{E}$ the local error of the discrete splitting scheme and let Φ^n be of classical order μ. If the solution ψ of (2.20) is in $\tilde{H}^{s+2\mu-2}$, then*

$$
\|\mathcal{E}(h)\|_{L^2} \leq C h^{\mu+1},
$$

where C depends on $d, s, A, U, N, (|\gamma_n|)_{n \leq N}, \mu, B_{s+2\mu-2}$.
If $\psi \in \tilde{H}^{s+2\mu}$, then we have the same bound in the $s-norm$:

$$
\|\mathcal{E}(h)\|_s \leq C h^{\mu+1}
$$

where C depends ond $d, s, A, U, N, (|\gamma_n|)_{n \leq N}, \mu, B_{s+2\mu}$.

Proof. We know from the abstract local error expansion (theorem 4.1) that the local error is o$(t^{\mu+1})$ if we can bound all terms of the form

$$
\mathrm{ad}_{D_L}^k (D_V) e^{t D_L} v
$$

for $k \leq \mu$. We can rewrite this expression as

$$\begin{aligned}
\|\mathrm{ad}^n_{D_L}(D_V)e^{tD_L}v\|_s &= \|D_{\mathrm{ad}^n_L(V)e^{tD_L}v}\|_s \\
&= \|(e^{tD_L})'(v)\mathrm{ad}^n_L(V)v\|_s \\
&= \|\mathcal{T}(t)'(v)\mathrm{ad}^n_L(V)v\|_s \\
&= \|\mathcal{T}(t)(\mathrm{ad}^n_L(V)v)\|_s \\
&= \|\mathrm{ad}^n_L(V)v\|_s,
\end{aligned}$$

since $\mathcal{T}(t)$ is a semigroup of linear, unitary operators on every $\tilde{H}^s$. Hence, due to lemma 4.3, we have to estimate

$$\| \sum_{\substack{p+|m(n,k)|=n \\ m(n,k)\in\mathbb{N}^k_+}} L^p\left(c_{n,m(n,k)}V(\psi)^{(k)}(L^{m_1},\ldots,(L^{m_k}))(\psi)\right)\|_\sigma$$

for $\sigma = 0, s$ to obtain an estimate for the local error in the L^2- or in the $\tilde{H}^s$-norm. First, triangular inequality and lemma 2.2 are used to obtain

$$\|\mathrm{ad}^n_{D_L}(D_V)e^{tD_L}\psi\|_\sigma \leq \sum_{\substack{p+|m(n,k)|=n \\ m(n,k)\in\mathbb{N}^k_+}} |c_{n,m(n,k)}|\|V(\psi)^{(k)}(L^{m(n,k)_1}(\psi),\ldots,L^{m(n,k)_k}(\psi))\|_{\sigma+2p}.$$

The next step is estimating the k-th Fréchet derivative of V, which we have already done in corollary 2.1. Hence, for each k and $m(n,k)$, there holds

$$\|V(\psi)^{(k)}(L^{m(n,k)_1},\ldots,L^{m(n,k)_k})(\psi)\|_\sigma \leq C\sum_{n=1}^N|\gamma_n|\|\psi\|_s^{4n+1-m}\prod_{p=1}^m\|L^{m(n,k)_p}(\psi)\|_\sigma.$$

Thus, we finally obtain

$$\begin{aligned}
\|\mathrm{ad}^q_{D_L}(D_V)e^{tD_L}\psi\|_\sigma &\leq \sum_{\substack{p+|m(q,k)|=q \\ m(q,k)\in\mathbb{N}^k_+}} |c_{m(n,k)}|\|V(\psi)^{(k)}(L^{m(q,k)_1}(\psi),\ldots,L^{m(q,k)_k}(\psi))\|_{\sigma+2p} \\
&\leq C_1 \sum_{\substack{p+|m(q,k)|=q \\ m(q,k)\in\mathbb{N}^k_+}} \sum_{n=1}^N |\gamma_n|\|\psi\|_{s+2p}^{2n+1-k}\prod_{\nu=1}^k\|L^{m(q,k)_\nu}\psi\|_{\sigma+2p} \\
&\leq C_2 \sum_{\substack{p+|m(q,k)|=q \\ m(q,k)\in\mathbb{N}^k_+}} \sum_{n=1}^N \|\psi\|_{s+2p}^{2n+1-k}\prod_{\nu=1}^k\|\psi\|_{\sigma+2p+2m(q,k)_\nu} \\
&\leq C_4 \sum_{\substack{p+|m(q,k)|=q \\ m(q,k)\in\mathbb{N}^k_+}} \sum_{n=1}^N \|\psi\|_{s+2p}^{2n+1-k}\|\psi\|^k_{\sigma+2p+2\max(m(q,k))}.
\end{aligned}$$

Since $|m(q,k)| = q - p$ and each element of $m(q,k) \geq 1$, there holds $\max m(q,k) = q - p - k - 1$ and, thus,

$$\begin{aligned}\|\mathrm{ad}^q_{D_L}(D_V)e^{tD_L}\psi\|_\sigma &\leq C_5 \sum_{\substack{p+|m(q,k)|=q\\ m(q,k)\in\mathbb{N}^k_+}} \sum_{n=1}^{N} \|\psi\|^{2n+1-k}_{s+2p}\|\psi\|^k_{\sigma+2q-2(k+1)}\\ &\leq C_6 \sum_{\substack{p+|m(q,k)|=q\\ m(q,k)\in\mathbb{N}^k_+}} \sum_{n=1}^{N} \|\psi\|^{2n+1-k}_{s+2q}\|\psi\|^k_{\sigma+2q-2(k+1)}\end{aligned}$$

holds true. This term is bounded if $\psi \in \tilde{H}^{s+2q}$ and so, the splitting method of order μ induces a local error of order $\mu+1$ in the s-norm, if the true solution lies in $\psi \in \tilde{H}^{s+2\mu}$. For the bound in the L^2-norm we exchange σ and s on the right-hand side (which is possible since for $\varphi, \xi \in \tilde{H}^s$, there hold both $\|\varphi\xi\|_0 \leq \|\varphi\|_0\|\xi\|_s$ and $\|\varphi\xi\|_0 \leq \|\varphi\|_s\|\xi\|_0$). Hence, the splitting method will converge locally with order $\mu + 1$ in the L^2-norm if the true solution ψ lies in $\tilde{H}^{s+2\mu-2}$, since $\max\{2\mu, s + 2\mu - 2\} = s + 2\mu - 2$, due to $s > d \geq 1$. □

We note that the result for the local error coincides with the result of [7], where $\mu = 2$ and hence we need $\psi \in \tilde{H}^{s+2}$. An interesting property of this error expansion is that the degree of the power non-linearity does not influence which regularity we have to assume, but only of what value the constants are.
So, now the proof of the final result of this chapter is only the application of the argument of Lady Windermere's fan.

Theorem 4.2. *Let the assumptions of proposition 4.3 be satisfied and* $B_\sigma = \sup_{t\in[0,T]} \|\psi\|_\sigma$. *We then have the global error bound*

$$\begin{aligned}&\|\psi^n - \psi(t_n)\|_0 \leq C_1 h^\mu \quad \text{for } 0 \leq t_n = nh \leq T \quad \text{and } \psi \in \tilde{H}^{s+2\mu-2}\\ &\|\psi^n - \psi(t_n)\|_s \leq C_2 h^\mu \quad \text{for } 0 \leq t_n = nh \leq T \quad \text{and } \psi \in \tilde{H}^{s+2\mu},\end{aligned}$$

where C_1 *and* C_2 *both depend on* $d, s, A, U, N, (|\gamma_n|)_{n\leq N}, T$, C_1 *additionally on* $B_{s+2\mu-2}$ *and* C_2 *additionally on* $B_{s+2\mu}$.

Proof. The argument of Lady Windermere's fan is that we can estimate the norm of the global error $\|E\|$ by summing $\|E_n\|$ which are norms of the local errors that are induced of the n-th step of the scheme, for $n = 1, \dots, M$, and are transported until the final time T. In order to estimate these norms, we use the stability estimates from propositon 4.2 and the bounds on the local errors from proposition 4.3 to compute for $\sigma = 0, s$

$$\begin{aligned}\|E\|_\sigma &\leq \sum_{n=1}^{N} \|E_n\|_\sigma \leq \sum_{n=1}^{M} e^{C(T-t_n)}\|e_i\|_\sigma\\ &\leq \sum_{n=1}^{M} e^{C(T-t_n)}C'h^{\mu+1} \leq \sum_{n=1}^{M}(he^{C(T-t_n)})C'h^\mu\\ &\leq \int_0^T e^{C(T-t)}dtC'h^\mu =: \tilde{C}h^\mu,\end{aligned}$$

because $t \to e^{C(T-t)}$ is decreasing and hence $\sum_{n=1}^{M}(he^{C(T-t_n)})$ is the Riemannian lower sum of this function. $\square$

5. Full Discretisation

In this chapter we will finally prove a convergence result for the full discretisation of (2.20), for which we use an operator splitting in time and a Hermite pseudo-spectral collocation in space. As mentioned in the introduction, we will provide a convergence proof for the second-order Strang splitting only, which is a scheme with coefficients $c_1 = c_2 = \frac{1}{2}$ and $d_1 = 0, d_2 = 1$. In contrast to that, a detailed solution algorithm for arbitrary order p will also be given.

5.1. Presentation of the Fully Discretised Scheme

This section follows the presentation of a fourth order algorithm in [3]. We denote the fully discrete evolution operator with time step size h, space discretisation parameter K and order p by Φ^p_{hK}. Furthermore, we denote the solution operators that provide the intermediate results by Θ^q_{hk}, $q = 1, \ldots, p$, hence

$$\Phi^p_{hK} = \prod_{q=1}^{p} \Theta^q_{hK} = \prod_{q=1}^{p} e^{-ic_q t D_T} e^{-id_q t D_{\tilde{V}}}.$$

The change from the usual potential V to $\tilde{V} = \mathcal{I}(V)$ has to be employed as this scheme is defined on $X_{\mathcal{K}}$ for some K and the solution after each time step must still lie in $X_{\mathcal{K}}$. Thus, we have to use the space discretised version of (2.20). The numerical solution at time step m is stored in a tensor ψ^m of order d, such that $\psi^m_j \approx \psi(x_j, mh)$. The operator Φ^p_{hK} approximates the real solution $\psi(x,t)$ in the Hermite points x_j at times mh by alternatingly solving the two subsystems

$$i\frac{\partial}{\partial t}\psi_K = -\Delta\psi_K + x^T A x \psi_K = L\psi_K \tag{5.1}$$

$$i\frac{\partial}{\partial t}\psi_K = \sum_{n=1}^{N} \gamma_n \mathcal{I}(|\psi_K|^{2n}\psi_K) + \mathcal{I}(U\psi_K) \tag{5.2}$$

$$\psi_K(0) = \mathcal{I}(\psi_0)$$

In section 4.2, it was demanded that these subproblems need to be numerically solvable with an error which is smaller than the one introduced by the splitting itself. The next lemma shows that this is indeed the case.

Lemma 5.1. *Equations (5.1) and (5.2) are uniquely solvable and exactly integrable in time.*

Proof. We begin with (5.2). Since ψ_K coincides with $\mathcal{I}\psi_K$ in the collocation points x_j, the equation evaluated in these nodes becomes

$$\begin{aligned} i\frac{\partial}{\partial t}\psi_K(x_j,t) &= \sum_{n=1}^{N} \gamma_n \mathcal{I}(|\psi_K|^{2n}\psi_K)(x_j,t) + \mathcal{I}(U\psi_K)(x_j,t) \\ &= \sum_{n=1}^{N} \gamma_n (|\psi_K(x_j,t)|^{2n}\psi_K(x_j,t)) + U(x_j)\psi_K(x_j,0) \\ &= \sum_{n=1}^{N} \gamma_n (|\psi_K(x_j,0)|^{2n})\psi_K)(x_j,t) + U(x_j)\psi_K(x_j,0), \end{aligned}$$

and is hence solved by

$$\psi_K(x_j,t) = \exp\left(-it\sum_{n=1}^{N}\gamma_n(|\psi_K|^{2n}\psi_K)(x_j,0) + U(x_j)\right)\psi_K(x_j,0),$$

as has already been proven in lemma 4.5.
Concerning equation (5.1), we recall that the semigroup generated by $-iL$, $\mathcal{T}(t)$, acts on a function ψ_K by multiplying its coefficients ψ_j with respect to the Hermite basis h_j with $e^{-i\omega_j t}$. □

For a practical implementation, we have to figure out how we can acquire the Hermite coefficients from knowing the values of ψ only in the Gauss-Hermite points. This can be done using the Hermite interpolation, because, as we have already shown in lemma 3.3(5), that for $\varphi \in X_K$, there holds $\hat{\varphi}_j = \varphi_j$, and hence

$$\varphi_j(t) = \sum_{k\in\mathcal{K}} w_k \varphi(x_k,t) h_j(x_k).$$

Rediscovering the values at the collocation points is simply done by

$$\varphi(x_k,t) = \sum_{j\in\mathcal{K}} \varphi_j(t) h_j(x_k).$$

Hence, the scheme in an implementable form reads as follows. Let ψ^{qk} be a tensor storing the numerical solution after k intermediate steps between the time steps m and $m+1$ and the tensor U be the function $U(x)$ evaluated in the collocation points. Then for $k = 1, \dots, p$ and for each time step we have to compute

$$\psi_r^{m,k} = \sum_{j\in\mathcal{K}} \exp\left(-ic_{k+1}h\sum_{n=1}^{N}\gamma_n(|\psi_r^{m,k-1,+}|^{2n}\psi_r^{m,k-1,+}) + U_r\right)\psi_r^{m,k-1,+} \tag{5.3}$$

$$\psi_r^{m,k,+} = \sum_{j\in\mathcal{K}} e^{-ihd_{k+1}\omega_j} h_j(x_r)\hat{\psi}^{m,k}, \quad \hat{\psi}^{m,k} = \sum_{l\in\mathcal{K}} w_l \psi_l^{m,k} h_j(x_l). \tag{5.4}$$

By applying the two orthogonality relations that the Hermite functions fulfil (lemma 3.1), we can prove

Lemma 5.2. Φ^p_{hK} *conserves the L^2-norm of the discretised initial datum, which means*

$$\|\varphi_{hK}\|_{L^2} = \|\mathcal{I}(\psi_0)\|_{L^2}.$$

Proof. The claim is true if it holds for one step of the splitting. Since the first step (5.3) obviously does not change the norm, only the norm conservation of the second step has to be ensured. We drop the dependencies of the solution on the time step and simply denote it as ψ^m. The L^2-norm of a function $\varphi \in X_{\mathcal{K}}$ can be computed by

$$\|\varphi\|_{L^2} = \int_{\mathbb{R}^d} |\psi|^2\, dx = \sum_{j\in\mathcal{K}} |\psi(x_j)|^2 w_j,$$

since the degree of φ is smaller than $\mathbf{2K-1}$. Hence, we can compute $\|\psi(mh)\|_{L^2}$ via the weighted l^2-norm of ψ^m, and so the claim follows from the calculation

$$\begin{aligned}
\|\psi^{k+1}\|{=}\sum_{j\in\mathcal{K}} w_j|\psi^k|^2 &= \sum_{j\in\mathcal{K}} w_j \left|\sum_{l\in\mathcal{K}} e^{-itc_{k+1}\omega_l}\hat{\psi}^k_l h_l(x_j)\right| \\
&= \sum_{l,\iota\in\mathcal{K}} e^{-itc_{k+1}\omega_\iota} e^{itc_{k+1}\omega_l}\hat{\psi}^k_\iota \overline{\hat{\psi}^k_l} \sum_{j\in\mathcal{K}} w_j h_l(x_j) h_\iota(x_j) \\
&= \sum_{l,\iota\in\mathcal{K}} e^{-itc_{k+1}\omega_\iota} e^{itc_{k+1}\omega_l}\hat{\psi}^k_\iota \overline{\hat{\psi}^k_l}\delta_{l\iota} \\
&= \sum_{l\in\mathcal{K}} \left|\hat{\psi}^k_l\right|^2 = \sum_{l\in\mathcal{K}} \left|\sum_{j\in\mathcal{K}} w_j\psi^k_j h_l(x_j)\right|^2 \\
&= \sum_{\iota,j\in\mathcal{K}} w_j\psi^k_j\psi^k_\iota \sum_{l\in\mathcal{K}} w_\iota h_l(x_\iota) h_l(x_j) \\
&= \sum_{\iota,j\in\mathcal{K}} w_j\psi^k_j\psi^k_\iota\delta_{\iota,j} \\
&= \sum_{j\in\mathcal{K}} w_j|\psi^k_j|^2,
\end{aligned}$$

where the orthogonality relations lemma 3.1(1) and subsequently lemma 3.1(2) have been used. □

Therefore, a fully discrete operator splitting inherits the property of norm conservation from the continuous equation (2.20). What is more, the scheme is obviously explicit and, since we can substitute h for $-h$, also time reversible. Chapter 6 contains numerical data gathered from experiments wherein the scheme we have just presented was used. But before dealing with that, we will look at the convergence properties of a special Φ^p_{hK} in the following section.

5.2. Convergence Analysis

Employing the triangular inequality, we can estimate the error of the full discretisation by

$$\|\psi^n_{hK} - \psi(x,t_n)\|_{L^2} \le \|\psi^n_{hK} - \psi^n_h\|_{L^2} + \|\psi^n_h - \psi(t_n)\|_{L^2}.$$

The second part is the global error of the semidiscretisation in time, which we already bounded in theorem 4.2. The first part will be studied employing methods that were introduced in the two previous chapters. The final proof uses the fan of Lady Windermere and hence we need to find a local error bound and show step-by-step stability. As we will rely heavily on an intermediate result which is a bound for $\|\Phi^p_{hK}(\mathcal{I}\varphi) - \mathcal{I}(\Phi^p_h\varphi)\|_{L^2}$, we will prove stability and boundedness of the local error for this expression, starting with the stability.

Proposition 5.1. *Let $s > \frac{d}{2}$ be an integer. If $\varphi, \psi \in X_\mathcal{K}$, then*

$$\|\Phi^p_{hK}(\psi) - \Phi^p_{hK}(\varphi)\|_0 \leq e^{CF(\|\psi\|_s, \|\varphi\|_s)h}\|\psi - \varphi\|_0,$$

with $C > 0$ depending on $s, d, A, N, (\gamma_n)_{n\leq N}$ and U.

Proof. The proof is very similar to the one of proposition 4.2. By lemma 4.4, it is again sufficient to show stability of one step and once again, since the semigroup $\mathcal{T}(t)$ which is generated by $-iL$ is unitary, we only have to estimate

$$\|e^{-ihdD_{\tilde{V}}}\psi - e^{-ihdD_{\tilde{V}}}\varphi\|_0,$$

which means that we have to bound the norm of $\theta(h) - \eta(h)$, where θ and η are the solutions of

$$i\partial_t\theta = \mathcal{I}(\sum_{n=1}^N \gamma_n|\theta|^{2n}\theta + U(x)\theta), \quad \theta(0) = \psi$$
$$i\partial_t\eta = \mathcal{I}(\sum_{n=1}^N \gamma_n|\eta|^{2n}\eta + U(x)\eta), \quad \eta(0) = \varphi.$$

Subtracting the second equation from the first one, we obtain

$$\begin{aligned} i\partial_t(\theta - \eta) &= \mathcal{I}\left(\sum_{n=1}^N(\gamma_n|\theta|^{2n}\theta - |\eta|^{2n}\eta) + U(x)(\theta - \eta)\right) \\ &= \sum_{n=1}^N \gamma_n\mathcal{I}\Big(g(\eta,\theta,n)(\theta-\eta) + f_1(\eta,\theta,n)(\overline{(\theta-\eta)}) \\ &\qquad + f_2(\eta,\theta,n)(\theta-\eta)\Big) + \mathcal{I}(U(\theta-\eta)). \end{aligned}$$

Substituting $\delta = (\theta - \eta)$, multiplying with $\overline{\delta}$ and integrating over $\mathbb{R}^d$ yields

$$i(\partial_t\delta,\delta)_0 = \sum_{n=1}^N \gamma_n(\mathcal{I}\big(g(\eta,\theta,n)\delta + f_1(\eta,\theta,n)\overline{\delta} + f_2(\eta,\theta,n)\delta\big) + \mathcal{I}(U\delta),\delta)_0.$$

Taking the imaginary part of this equation yields

$$(\partial_t,\delta)_0 = \sum_{n=1}^N |\gamma_n| \left(\mathcal{I}(f_1(\eta,\theta,n)\delta + f_2(\eta,\theta,n)\overline{\delta}),\delta\right)_0$$

since $(\partial_t\delta,\delta)_0 = \frac{1}{2}\partial_t\|\delta\|_0^2 = \|\delta\|_0\|\partial_t\delta\|_0 \in \mathbb{R}$ and g is real valued, and furthermore, due to Cauchy-Schwarz' inequality

$$\|\partial_t\delta\|_0 \leq \sum_{n=1}^{n} |\gamma_n| \|\mathcal{I}(f_1(\eta,\theta,n)\delta + f_2(\eta,\theta,n)\overline{\delta})\|_0.$$

The next step is to apply lemma 3.3 several times, to see that there holds

$$\begin{aligned}\|\mathcal{I}(f_k(|\eta|,|\theta|,n)\delta)\|_0 &\leq |f_k(\sup_{j\in\mathcal{K}}|\eta(x_j,t)|, \sup_{j\in\mathcal{K}}|\theta(x_j,t)|,n)|\|\delta\|_0 \\ &= |f_k(\sup_{j\in\mathcal{K}}|\psi(x_j)|, \sup_{j\in\mathcal{K}}|\varphi(x_j)|,n)|\|\delta\|_0, \\ &\leq f_k(\|\psi\|_s,\|\varphi\|_s,n)\|\delta\|_0,\end{aligned}$$

since the absolute value of the solution is conserved by $e^{-idhD_{\tilde{V}}}$ at the points x_j. Hence, we obtain

$$\begin{aligned}\partial_t\|\delta\|_0 &\leq \|\delta\|_0 \sum_{n=1}^{N} |\gamma_n|(f_1(\|\psi\|_s,\|\varphi\|_s,n) + f_2(\|\psi\|_s,\|\varphi\|_s,n)) \\ &=: \|\delta\|_0 \tilde{F}_1(\psi,\varphi).\end{aligned}$$

Finally, the differential form of the Gronwall inequality (lemma A.5) yields

$$\|\delta(t)\|_0 \leq e^{\tilde{F}_1(\psi,\varphi)t}\|\psi-\varphi\|_0,$$

and thus, the proof is complete if we set $F = \sum_{k=1}^{p} \tilde{F}_k$ due to lemma 4.4. □

This was the last step that allowed a useful generalisation of the techniques of [7] for a splitting of higher order. For the rest of this section we will restrict our analysis to the second-order Strang splitting with a solution operator that we denote as $\Phi_{hK} = e^{-i\frac{h}{2}L}e^{-ihD_{\tilde{V}}}e^{-i\frac{h}{2}L}$ for the fully discrete and as Φ_h for the time semi-discrete scheme. For this splitting a bound for the local error is being deduced first.

Proposition 5.2. *Let φ be an element of $\tilde{H}^s$ for some integer $s > \frac{d}{2}$. Then there exists a function F that is monotonously increasing in both of its arguments such that*

$$\|\Phi_{hK}(\mathcal{I}\varphi) - \mathcal{I}(\Phi_h(\varphi))\|_0 \leq ChK^{1+\frac{d}{3}-\frac{s}{2}}\left(e^{F(\|\mathcal{I}(\varphi)\|_\sigma,\|\varphi\|_\sigma)h}\|\varphi\|_s + \|\Phi_h\varphi\|_0\right), \tag{5.5}$$

with $C > 0$ depending on $s, d, A, N, (\gamma_n)_{n\leq N}$ and U.

Proof. We start with analysing the error by establishing bounds for the two components

$$\|\theta_1 - \mathcal{I}(\eta_1)\|_0 \quad \text{and} \quad \|\theta_2 - \mathcal{I}(\eta_2)\|_0,$$

where η_k and θ_k are solutions of the systems

$$i\partial_t\theta_1 = L\theta_1, \quad \theta_1(0) = \psi_1 \in X_{\mathcal{K}} \tag{5.6}$$
$$i\partial_t\eta_1 = L\eta_1, \quad \eta_1(0) = \varphi_1 \tag{5.7}$$

and

$$i\partial_t\theta_2 = \sum_{n=1}^{N} \gamma_n \mathcal{I}(|\theta_2|^{2n}\theta_2 + U(x)\theta_2), \quad \theta_2(0) = \psi_2 \in X_{\mathcal{K}} \tag{5.8}$$

$$i\partial_t\eta_2 = \sum_{n=1}^{N} \gamma_n \mathcal{I}(|\eta_2|^{2n}\eta_2 + U(x)\eta_2), \quad \eta_2(0) = \varphi_2. \tag{5.9}$$

Subtracting the Hermite interpolation of (5.7) from (5.6), we obtain

$$\begin{aligned} i\partial_t(\theta_1 - \mathcal{I}(\eta_1)) &= L\theta_1 - \mathcal{I}(L\eta_1) + L\eta_1 - L\eta_1 + L(\mathcal{I}\eta_1) - L(\mathcal{I}\eta_1) \\ &= L(\theta_1 - \mathcal{I}(\eta_1)) + L(\mathcal{I}(\eta_1) - \eta_1) + L\eta_1 - \mathcal{I}(L\eta_1). \end{aligned}$$

We multiply this equation with $\overline{(\theta_1 - \mathcal{I}(\eta_1))} =: \overline{\delta}$ and integrate over $\mathbb{R}^d$ to attain

$$i(\partial_t\delta, \delta)_0 = (L(\delta), \delta)_0 + (L(\mathcal{I}(\eta_1) - \eta_1), \delta)_0 + (L\eta_1 - \mathcal{I}(L\eta_1), \delta)_0.$$

As L is selfadjoint, taking the imaginary part yields

$$\|\delta\|_0\partial_t\|\delta\|_0 = \mathrm{Im}(L(\mathcal{I}(\eta_1) - \eta_1), \delta)_0 + (L\eta_1 - \mathcal{I}(L\eta_1), \delta)_0,$$

which gives, by Cauchy-Schwarz inequality and lemmas 3.5 and 2.12

$$\begin{aligned} |\partial_t\|\delta\|_0| &\le \|L(\mathcal{I}(\eta_1) - \eta_1)\|_0 + \|L\eta_1 - \mathcal{I}(L\eta_1)\|_0 \\ &\le \|\mathcal{I}(\eta_1) - \eta_1\|_2 + C_1K^{\frac{d}{3} - \frac{s-2}{2}}\|L\eta_1\|_{s-2} \\ &\le C_2K^{\frac{d}{3} - \frac{s-2}{2}}\|\eta_1\|_s + C_1K^{\frac{d}{3} - \frac{s-2}{2}}\|\eta_1\|_s \\ &=: CK^{1+\frac{d}{3} - \frac{s}{2}}\|\eta_1\|_s. \end{aligned}$$

Integrating over $(0,t)$ results in the first bound

$$\|\theta_1(t) - \mathcal{I}(\eta_1)(t)\|_0 \le \|\psi_1 - \mathcal{I}(\varphi_1)\|_0 + \|\eta_1\|_s.$$

To estimate $\|\theta_2(t) - \mathcal{I}(\eta_2)(t)\|_0$, we subtract the Hermite interpolation of (5.9) from (5.8) and notice that we can proceed exactly as in the proof of proposition 5.1. Thus,

$$\|\theta_2(t) - \mathcal{I}(\eta_2)(t)\|_0 \le e^{tF(\|\psi_2\|_\sigma, \|\varphi_2\|_\sigma)}\|\psi_2 - \mathcal{I}(\varphi_2)\|_0.$$

From combining these two estimates we obtain for $\psi_2 \in X_{\mathcal{K}}$

$$\begin{aligned} \|e^{-iL\frac{h}{2}}e^{-ihD_{\tilde{V}}}\psi_2 - e^{-iL\frac{h}{2}}e^{-ihD_V}\varphi_2\|_0 &\le CK^{1+\frac{d}{3}-\frac{s}{2}}\|e^{-ihD_V}\varphi_2\|_0 + \|e^{-ihD_{\tilde{V}}}\psi_2 - e^{-ihD_V}\varphi_2\|_0 \\ &\le CK^{1+\frac{d}{3}-\frac{s}{2}}\|e^{-ihD_V}\varphi_2\|_0 + e^{ChF(\|\psi_2\|_\sigma, \|\varphi_2\|_\sigma)}\|\psi_2 - \varphi_2\|_0 \end{aligned}$$

Substituting $e^{-iL\frac{h}{2}}\varphi$ for ψ_2, and $e^{-iL\frac{h}{2}}\mathcal{I}(\varphi)$ for φ_2, we can estimate the last term by

$$\begin{aligned} e^{ChF(\|\psi_2\|_\sigma, \|\varphi_2\|_\sigma)}\|\psi_2 - \varphi_2\|_0 &\le e^{ChF(\|\mathcal{I}\varphi\|_\sigma, \|\varphi\|_\sigma)}\left(\|\mathcal{I}\varphi - \mathcal{I}\varphi\|_0 + CK^{1+\frac{d}{3}-\frac{s}{2}}\|\varphi\|_s\right) \\ &= e^{ChF(\|\mathcal{I}\varphi\|_\sigma, \|\varphi\|_\sigma)}CK^{1+\frac{d}{3}-\frac{s}{2}}\|\varphi\|_s. \end{aligned}$$

Furthermore, since e^{-iLt} is unitary on every $\tilde{H}^s$, we can substitute $\|e^{-iLt}\varphi\|_\sigma = \|\varphi\|_\sigma$ and

$$\|e^{-ihD_V}e^{-iL\frac{h}{2}}\varphi\|_s = \|e^{-iL\frac{h}{2}}e^{-ihD_V}e^{-iL\frac{h}{2}}\varphi\|_s = \|\Phi_h\varphi\|_s,$$

to finally obtain

$$\|\Phi_{hk}(\mathcal{I}\varphi) - \mathcal{I}(\Phi_h\varphi)\|_0 \leq ChK^{1+\frac{d}{3}-\frac{s}{2}}\left(e^{F(\|\mathcal{I}(\varphi)\|_\sigma,\|\varphi\|_\sigma)h}\|\varphi\|_s + \|\Phi^h\varphi\|_0\right).$$

□

With propositions 5.1 and 5.2 we are now in the position to prove the final result of this thesis.

Theorem 5.1. *Let $s > \left\lceil\frac{d+1}{2}\right\rceil + 2 + \frac{2d}{3}$ be an even integer. Let, furthermore, the exact solution $\psi(x,t) = \psi$ of equation (2.20) be an element of $\tilde{H}^{s+2}$ for $t \in [0,T]$ and let $\sup_{t\in[0,T]}\|\psi\|_s = B_s$. Then there exist a time discretisation parameter h_0, a space discretisation parameter K_0 and a constant C that all depend on $A, d, s, T, (|\gamma_n|)_{n\leq N}, N, B_{s+2}$ and a bound of $U(x)$, such that $\forall h \leq h_0$ and $\forall K \geq K_0$ there holds*

$$\|\psi^n_{hK} - \psi(t_n)\|_0 \leq C(K^{1+\frac{d}{3}-\frac{s}{2}} + h^2), \quad 0 \leq t_n = nh \leq T.$$

Proof. We begin by splitting

$$\begin{aligned}\|\psi^n_{hK} - \psi(t_n)\|_0 &\leq \|\psi^n_{hK} - \psi^n_h\|_0 + \|\psi^n_h - \psi(t_n)\|_0\\ &\leq \|\psi^n_{hK} - \mathcal{I}(\psi^n_h)\|_0 + \|\mathcal{I}(\psi^n_h) - \psi^n_h\|_0 + \|\psi^n_h - \psi(t_n)\|_0,\end{aligned}$$

and see that the first term is the only one we have not bounded up to now. We set $\sigma' = \left\lceil\frac{d+1}{2}\right\rceil$ and

$$\alpha(k) = \max_{p\in\{0,\dots,k-1\},q\in\{0,\dots,k-p-1\}} \|(\Phi_{hK})^q(\mathcal{I}(\psi^p_h))\|_{\sigma'}.$$

The next step is to apply the argument of Lady Windermere's fan to $\|\psi^n_{hK} - \mathcal{I}(\psi^n_h)\|_0$ using propositions 5.1 and 5.2. As we can thus estimate the global error by summing the transported local errors, we obtain

$$\begin{aligned}\|\psi^n_{hK} - \mathcal{I}(\psi^n_h)\|_0 &= \|(\Phi_{hK})^n(\mathcal{I}\psi_0) - \mathcal{I}((\Phi_h)^n\psi_0)\|_0\\ &\leq \sum_{p=0}^{n-1} e^{CF(\|\alpha(n)\|_{\sigma'},\|\alpha(n)\|_{\sigma'})h(n-p-1)}\|\Phi_{hK}(\mathcal{I}(\psi^p_h)) - \mathcal{I}(\Phi_h(\psi^p_h))\|_0\\ &\leq \sum_{p=0}^{n-1} e^{CF(\|\alpha(n)\|_{\sigma'},\|\alpha(n)\|_{\sigma'})h(n-p-1)}ChK^{1+\frac{d}{3}-\frac{s}{2}}\left(e^{C\|\mathcal{I}\varphi^n_h\|_{\sigma'}\|\varphi^n_h\|_{\sigma'}h}\|\psi^p_h\|_s + \|\psi^{p+1}_h\|_s\right),\end{aligned}$$

since F is increasing in both of its arguments. Now we need to elimate the dependency of the error on the numerical solution ψ^p_h. Due to theorem 4.2 and the boundedness of ψ, there holds

$$\|\psi^p_h\|_{\sigma'} \leq \|\psi^p_h\|_s \leq \|\psi(t_p)\|_s + \|\psi^p_h - \psi(t_p)\|_s \leq C',$$

where C depends on $d, s, T, (|\gamma_n|)_{n\leq N}, N, B_{s+2}, A$ and U. Furthermore, due to lemma 3.5, there also holds

$$\begin{aligned}\|\mathcal{I}(\psi_h^p)\|_{\sigma'} &\leq \|\psi_h^p\|_{\sigma'} + \|\psi_h^p - \mathcal{I}(\psi_h^p)\|_{\sigma'} \\ &\leq \|\psi_h^p\|_s + C_1 K^{\frac{d}{3}+\frac{\sigma'}{2}-\frac{\sigma'}{2}-\frac{2d}{6}}\|\psi_h^p)\|_{\sigma'+\frac{2d}{3}} \\ &\leq (1+C_1)\|\psi_h^p\|_{\sigma'} \leq (1+C_1)C' =: C.\end{aligned}$$

Hence, we can estimate the error by

$$\|\psi_{hK}^n - \mathcal{I}(\psi_h^n)\|_0 \leq \sum_{p=0}^{n-1} e^{CF(\alpha(n),\alpha(n))h(n-p-1)} ChK^{1+\frac{d}{3}-\frac{s}{2}}.$$

Since $e^{C(T-t)}$ is decreasing, the right-hand side constitutes a Riemannian lower sum and is hence estimated by

$$\begin{aligned}\|\psi_{hK}^n - \mathcal{I}(\psi_h^n)\|_0 &\leq \int_0^T e^{CtF(\alpha(n),\alpha(n))} CK^{1+\frac{d}{3}-\frac{s}{2}}\, dt \\ &= CK^{1+\frac{d}{3}-\frac{s}{2}} \left(\frac{e^{CTF(\alpha(n),\alpha(n))}-1}{e^{CF(\alpha(n),\alpha(n))}}\right),\end{aligned}$$

for $CF(\|\alpha(n)\|_{\sigma'}, \|\alpha(n)\|_{\sigma'}) \neq 0$. If this expression is 0, the lower sum can be bounded by T, because then we merely have to integrate $\int_0^{nh} 1\, dt$. This additionally yields

$$\|\psi_{hK}^n - \mathcal{I}(\psi_h^n)\|_{\sigma'} \leq CK^{1+\frac{d}{3}+\frac{\sigma'-s}{2}} \left(\frac{e^{CTF(\alpha(n),\alpha(n))}-1}{e^{CF(\alpha(n),\alpha(n))}}\right)$$

and due to lemma 3.5

$$\begin{aligned}\|\psi_{hK}^n - \psi_h^n\|_{\sigma'} &\leq \|\psi_{hK}^n - \mathcal{I}(\psi_h^n)\|_{\sigma'} + \|\psi_h^n - \mathcal{I}(\psi_h^n)\|_{\sigma'} \\ &\leq K^{1+\frac{d}{3}+\frac{\sigma'-s}{2}} \left(\frac{e^{CTF(\alpha(n),\alpha(n))}-1}{e^{CF(\alpha(n),\alpha(n))}} + 1\right).\end{aligned}$$

As $\alpha(n)$ depends on the numerical solution, we have to eliminate it from the error estimate. To achieve this aim, we assume $\alpha(n) < 2C$ and perform an additional time step to $(n+1)h \leq T$. Since $(\Phi_{hK})^{n-p}(\mathcal{I}\psi_h^p)$ is a part of Lady Windermere's fan for $p = 0, \ldots, n$ the estimate above has to hold for these points as well. Hence, there holds

$$\|(\Phi_{hK})^{n-p}(\mathcal{I}(\varphi_h^p)) - \psi_h^n\|_{\sigma'} \leq CK^{1+\frac{d}{3}+\frac{\sigma'-s}{2}} \left(\frac{e^{CTF(\alpha(n),\alpha(n))}-1}{e^{CF(\alpha(n),\alpha(n))}} + 1\right),$$

and thus also

$$\|(\Phi_{hK})^{n-p}(\mathcal{I}(\varphi_h^p))\|_{\sigma'} \leq \|\psi_h^n\|_{\sigma'} + CK^{1+\frac{d}{3}+\frac{\sigma'-s}{2}} \left(\frac{e^{CTF(2C,2C)}-1}{e^{CF(2C,2C)}} + 1\right).$$

Therefore, if K is large enough such that

$$CK^{1+\frac{d}{3}+\frac{\sigma'-s}{2}}\left(\frac{e^{CTF(2C,2C)}-1}{e^{CF(2C,2C)}}\right)<1,$$

also $\alpha(n+1)\leq 2C$, whence it follows by induction that $\alpha(n)\leq 2C\ \forall n$, and so we can conclude that

$$\|\psi_{hk}^n-\mathcal{I}(\psi_h^n)\|_0\leq CK^{1+\frac{d}{3}-\frac{s}{2}},$$

where C depends on $A,d,s,N,(\gamma_n)_{n\leq N},U,T$ and B_{s+2}.
Hence, we conclude

$$\begin{aligned}\|\psi_{hK}^n-\psi(t_n)\|_0&\leq\|\psi_{hK}^n-\psi_h^n\|_0+\|\psi_h^n-\psi(t_n)\|_0\\&\leq\|\psi_{hK}^n-\mathcal{I}(\psi_h^n)\|_0+\|\mathcal{I}(\psi_h^n)-\psi_h^n\|_0+\|\psi_h^n-\psi(t_n)\|_0\\&\leq C_1K^{1+\frac{d}{3}-\frac{s}{2}}+C_2K^{1+\frac{d}{3}-\frac{s}{2}}+C_3h^2\\&\leq C(K^{1+\frac{d}{3}-\frac{s}{2}}+h^2),\end{aligned}$$

where the bound of the second term follows from lemma 3.5 and the estimate for the third term follows from theorem 4.2. □

Corollary 5.1. *Let the assumptions of theorem 5.1 be fulfilled and let additionally,* $\psi(x,t)\in\tilde{H}^\rho$ *with* $\rho=\max\{s+2,\sigma+4\}$ *for an even integer* σ *with* $s\geq\sigma>d$*. Then there also holds*

$$\|\psi_{hK}^n-\psi(t_n)\|_\sigma\leq C(K^{1+\frac{d}{3}+\frac{\sigma-s}{2}}+h^2),$$

with a constant $C(d,s,\sigma,T,B_\rho,A,(|\gamma_n|)_{n\in\mathbb{N}},N,U)$*.*

Proof. We employ the same splitting as above to obtain

$$\begin{aligned}\|\psi_{hK}^n-\psi(t_n)\|_\sigma&\leq\|\psi_{hK}^n-\psi_h^n\|_\sigma+\|\psi_h^n-\psi(t_n)\|_\sigma\\&\leq\|\psi_{hK}^n-\mathcal{I}(\psi_h^n)\|_\sigma+\|\mathcal{I}(\psi_h^n)-\psi_h^n\|_\sigma+\|\psi_h^n-\psi(t_n)\|_\sigma\\&\leq C_1K^{1+\frac{d}{3}+\frac{\sigma-s}{2}}+C_2K^{1+\frac{d}{3}+\frac{\sigma-s}{2}}+C_3h^2\\&\leq C(K^{1+\frac{d}{3}+\frac{\sigma-s}{2}}+h^2),\end{aligned}$$

where the bound on the first term follows from

$$\|\psi_{hk}^n-\mathcal{I}(\psi_h^n)\|_\sigma\leq CK^{\frac{\sigma}{2}}\|\psi_{hk}^n-\mathcal{I}(\psi_h^n)\|_0\leq CK^{1+\frac{d}{3}+\frac{\sigma-s}{2}},$$

and the bound on the last term from theorem 4.2 under the additional regularity assumptions on ψ. □

6. Numerical Results

We end this thesis with some numerical results. In chapter 5, we introduced the fully discrete solution operator Φ_{hK}^p which (for different p) was used to compute the errors and solution plots that are contained in this chapter. The first plot, figure 6.1, illustrates the behaviour of the solution of a cubic nonlinear Schrödinger equation in two space dimensions for the initial datum

$$\psi_0(x,y) = e^{-\frac{x^2+y^2}{2}}$$

and the matrix $A = \mathrm{diag}(2,4)$. The space-discretisation parameter K was set to 40 and the time-step size to 0.005, and the simulation ended at $T = 16$.
The second plot, figure 6.2, shows the convergence rate of the error in the L^2-norm, which is induced by operator splittings of order $2, 4$ and 8 that were applied to the one-dimensional cubic nonlinear Schrödinger equation with $A = 2$. The initial datum was $\psi_0(x) = e^{-\frac{x^2}{2}}$ and the "real" solution was chosen as the numerical solution with time-step size $h = 2^{-14}$. The space discretisation parameter of this experiment was $K = 120$. After a start-up phase which is of different length for each order, the expected order of convergence is clearly visible.

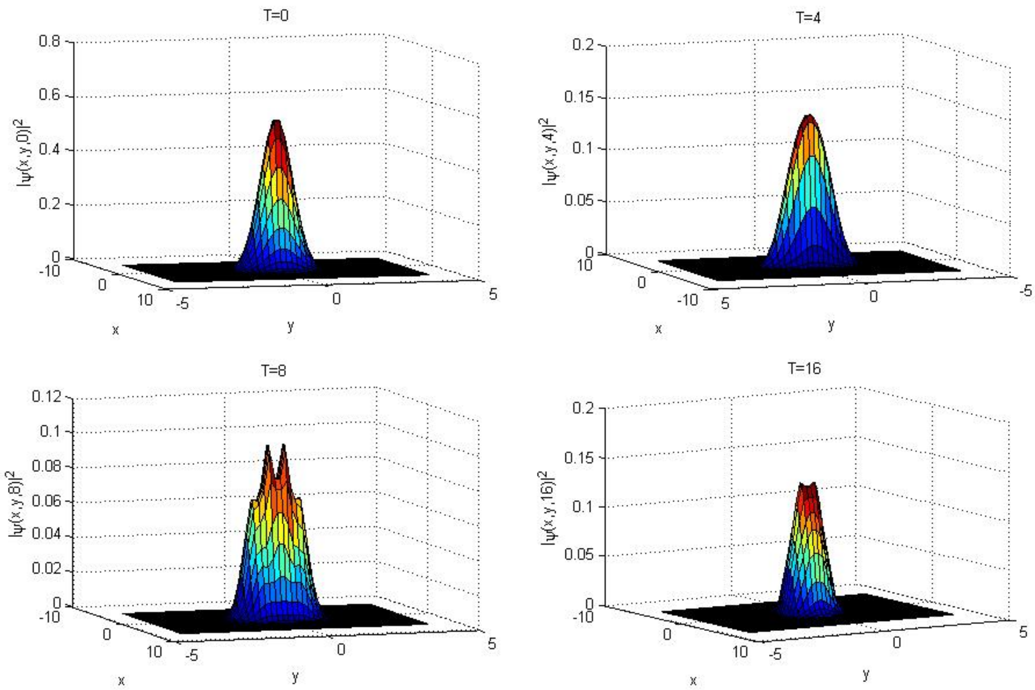

Figure 6.1.: Time evolution of $|\psi(x, y)|^2$

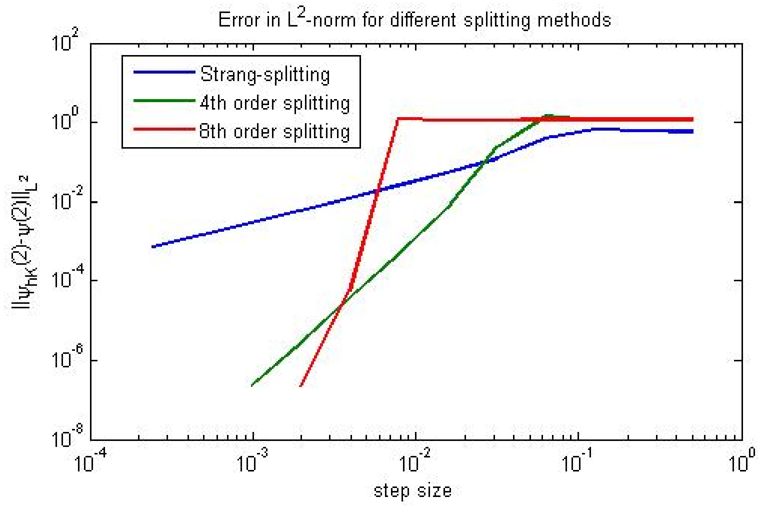

Figure 6.2.: Error of operator splittings of different order

A. Appendix

A.1. Technical Results

In this section, two lemmas are presented that we needed to prove in order to obtain our desired convergence results. We chose to store them in the appendix so that the proofs of those results did not have to interrupted in the respective sections.

Lemma A.1. *For $x, y \in \mathbb{C}$ and $\forall i \in \mathbb{N}$ there are functions $g : \mathbb{C}^2 \times \mathbb{N} \to \mathbb{R}$, $f_1 : \mathbb{C}^2 \times \mathbb{N} \to \mathbb{C}$ and $f_2 : \mathbb{C}^2 \times \mathbb{N} \to \mathbb{C}$, all of which are polynomials in $x, y, \overline{x}, \overline{y}$, such that*

$$|y|^{2i}y - |x|^{2i}x = g(x,y,i)(y-x) + f_1(x,y,i)(\overline{y-x}) + f_2(x,y,i)(y-x).$$

A triple of functions fulfilling this condition is given by

$$g(x,y,i) = |y|^{2i} + |x|^{2i}, \tag{A.1}$$

$$f_1(x,y,i) = \begin{cases} \sum_{j=0}^{\frac{i-2}{2}} |xy|^{2j} xy \left(|x|^{2i-2-4j} + |y|^{2i-2-4j}\right) & i = 2k \\ \sum_{j=0}^{\frac{i-3}{2}} |xy|^{2j} xy \left(|x|^{2i-2-4j} + |y|^{2i-2-4j}\right) + |xy|^{i-1}xy & i = 2k+1 \end{cases}, \tag{A.2}$$

$$f_2(x,y,i) = \begin{cases} \sum_{j=0}^{\frac{i-4}{2}} |xy|^{2j+2} \left(|x|^{2i-2-4j} + |y|^{2i-2-4j}\right) & i = 2k \\ \sum_{j=0}^{\frac{i-5}{2}} |xy|^{2j+2} \left(|x|^{2i-2-4j} + |y|^{2i-2-4j}\right) + |xy|^{i+1} & i = 2k+1 \end{cases}. \tag{A.3}$$

Proof. We will first show the existence of such a triple by induction which will also provide us with a recurrence relation for these functions. During this proof we will drop the dependence of all appearing functions on x, y. For $i = 0$ the claim is obviously fulfilled for $g \equiv 1, f_m \equiv 0$. For $i = 1$ we have

$$\begin{aligned} |y|^2y - |x|^2x &= |y|^2y - |x|^2x + |y|^2x + |x|^2y - |y|^2x - |x|^2y \\ &= (|y|^2 + |x|^2)(y-x) + xy\overline{y} - x\overline{x}y \\ &= (|y|^2 + |x|^2)(y-x) + xy(\overline{y-x}), \end{aligned}$$

whence $g(1) = |y|^2 + |x|^2, f_1(1) = xy, f_2(1) = 0$.
For the induction step from i to $i+1$ we rewrite the left hand side in the following way:

$$\begin{aligned} |y|^{2i+2}y - |x|^{2i+2}x &= |y|^{2i+2}y - |x|^{2i+2}x + |y|^{2i+2}x - |y|^{2i+2}x + |x|^{2i+2}y - |x|^{2i+2}y \\ &= (|y|^{2i+2} + |x|^{2i+2})(y-x) + |y|^{2i+2}x - |x|^{2i+2}y \\ &= (|y|^{2i+2} + |x|^{2i+2})(y-x) + \overline{y}^{i+1}y^{i+1}x - \overline{x}^{i+1}x^{i+1}y \\ &= (|y|^{2i+2} + |x|^{2i+2})(y-x) + xy(|y|^{2i}\overline{y} - |x|^{2i}\overline{x}). \end{aligned}$$

Thus, the induction hypothesis yields

$$\begin{aligned}|y|^{2i+2}y-|x|^{2i+2}x &= (|y|^{2i+2}+|x|^{2i+2})(y-x) + xy\overline{\left(g(i)(y-x) + f_1(i)(\overline{y-x}) + f_2(i)(y-x)\right)} \\ &= (|y|^{2i+2}+|x|^{2i+2})(y-x) + xy((g(i) + \overline{f_2(i)})(\overline{y-x}) + \overline{f_1(i)}(y-x)).\end{aligned}$$

This proves that (A.1) is a possible choice for g, if f_1 and f_2 fulfil

$$xy(g(i) + \overline{f_2(i)}) = f_1(i+1) \tag{A.4}$$

$$xy\overline{f_1(i)} = f_2(i+1). \tag{A.5}$$

Plugging (A.5) into (A.4) and recalling the starting values results in the recurrence relation

$$\begin{aligned}f_1(i) &= xy(\overline{xy}f_1(i-2) + g(i-1)) \\ &= xy(|x|^{2i-2} + |y|^{2i-2} + \overline{xy}f_i(i-2)), \quad i > 1 \\ f_1(0) &= 0 \\ f_1(1) &= xy,\end{aligned}$$

from which f_1, and subsequently also f_2, can be calculated $\forall i$. A direct computation shows that f_1 defined in (A.2) satisfies the recurrence relation and that (A.3) satisfies (A.5). □

Lemma A.2. *For $x, y \in \mathbb{C}$ there are functions h_1, h_2 with $h_i : \mathbb{C}^2 \times \mathbb{N}$ that are polynomials in $x, y, \overline{x}, \overline{y}$, which satisfy*

$$|x|^{2k} - |y|^{2k} = h_1(x, y, k)(x - y) + h_2(x, y, k)(\overline{x - y}).$$

A possible choice for h_i is

$$h_1(x, y, k) = \sum_{i=0}^{k-1} |x|^{2i}|y|^{2(i-1-i)}\overline{y} \tag{A.6}$$

$$h_2(x, y, k) = \sum_{i=0}^{k-1} |x|^{2i}|y|^{2(i-1-i)}x. \tag{A.7}$$

Proof. This proof is performed analogously to the one of lemma A.1. Again, an induction argument yields a recurrence relation. The dependence of h_i on x, y is dropped once again. For $k = 1$, one computes

$$\begin{aligned}|x|^2 - |y|^2 &= x\overline{x} - y\overline{y} + x\overline{y} - x\overline{y} \\ &= \overline{y}(x - y) + x(\overline{x - y}),\end{aligned}$$

resulting in $h_1(1) = \overline{y}, h_2(1) = x$.
For the induction step k to $k + 1$, we rewrite

$$\begin{aligned}|x|^{2(k+1)} - |y|^{2(k+1)} &= |x|^{2k}(|x|^2 + |y|^2 - |y|^2) - |y|^{2k}|y|^2 \\ &= |y|^2(|x|^{2k} - |y|^{2k}) + |x|^{2k}(|x|^2 - |y|^2) \\ &= |y|^2(h_1(k)(x-y) + h_2(k)(\overline{x-y})) + |x|^{2k}(h_1(1)(x-y) + h_2(1)(\overline{x-y})),\end{aligned}$$

due to the induction hypothesis. For that matter, the following recurrence relation for $h_i(k)$ has been deduced:

$$\begin{aligned} h_i(k+1) &= |y|^2 h_i(k) + |x|^{2k} h_i(1), \quad i = 1, 2 \\ h_1(1) &= \overline{y} \\ h_2(1) &= x. \end{aligned}$$

We can see by a direct computation that (A.6) and (A.7), respectively, fulfil these relations.

□

A.2. Results from the Literature

For the convenience of the reader, this section contains results that are either not well known or appear in various forms. These lemmas and propositions are only cited, their proofs can be found in the respective reference.

Proposition A.1. *Consider two Banach spaces $X \hookrightarrow Y$ and $1 > p, q \leq \infty$. Let $(f_n)_{n\in\mathbb{N}}$ be a bounded sequence in $L^q(I,Y)$ and let $f : I \to Y$ be such that $f_n(t) \rightharpoonup f(t)$ in Y as $n \to \infty$ for a.a. $t \in I$. If $(f_n)_{n\in\mathbb{N}}$ is bounded in $L^p(I,X)$ and if X is reflexive then $f \in L^p(I,X)$ and $\|f\|_{L^p(I,X)} \leq \liminf_{n\to\infty} \|f_n\|_{L^p(I,X)}$.[5, theorem 1.2.5]*

Lemma A.3. *Let $m > d/2$ be an integer and let $g \in C^m(\mathbb{C},\mathbb{C})$ satisfy $g(0) = 0$. Then $\forall M > 0$ there exists a constant $C(M) > 0$ such that*

$$\|g(u) - g(v)\|_{L^2} \leq C(M)\|u - v\|_{L^2}$$

for all $u, v \in \tilde{H}^m$ with $\|u\|_{L^\infty}, \|v\|_{L^\infty} \leq M$. [5, lemma 4.10.2]

Proposition A.2. *Let H be a separable Hilbert space and $(\mathcal{D}(T), T)$ a positive symmetric unbounded operator on H on the domain of definition $\mathcal{D}(T)$. Assume that there exists an orthonormal basis (e_j) of H such that $e_j \in \mathcal{D}(T)$ $\forall j$ are eigenfunctions of T with corresponding eigenvalues λ_j.*
Then T is essentially self-adjoint. [14, lemma 5.10]

Proposition A.3 (Duhamel's Formula)**.** *Let $(A, \mathcal{D}(A))$ be the self adjoint infinitesimal generator of the semigroup of linear operators $\mathcal{T}(t)$. If $x \in \mathcal{A}$ and $f \in L^1([0,T],X)$ and if $u \in W^{1,1}((0,T),X)$ or if $u \in L^1((0,T),\mathcal{D}(A))$, then u satisfies*

$$u(t) = \mathcal{T}(t)x + i\int_0^t \mathcal{T}(t-s)f(s)\,ds$$

if and only if

$$\begin{cases} u \in L^1((0,T),\mathcal{D}(A)) \cap W^{1,1}((0,T),X), \\ i\frac{du}{dt} + Au + f = 0 \quad a.e.\ on\ [0,T], \\ u(0) = x. \end{cases} \tag{A.8}$$

[5, remark 1.6.1(v)]

Lemma A.4 (Gronwall inequality, integral form)**.** *Let $I := [a,b]$ a real interval, u, α, β be continuous, real valued functions and β additionally nonnegative.*
If

$$u(t) \leq \alpha(t) + \int_a^t \beta(s)u(s)\,ds$$

holds $\forall t \in I$, then

$$u(t) \leq \alpha(t) + \int_a^t \alpha(s)\beta(s)\exp(\int_s^t \beta(\tau)d\tau)\,ds, \quad \forall t \in I.$$

If α is additionally nondecreasing on I then

$$u(t) \leq \alpha(t) \exp(\int_a^t \beta(s)\, ds), \quad \forall t \in I.$$

Lemma A.5 (Gronwall inequality, differential form)**.** *Let $I := [a,b]$ a real interval, u, α be continuous, real valued functions and u additionally differentiable on (a,b).*
If u satisfies

$$u'(t) \leq \alpha(t) u(t), \quad \forall t \in (a,b),$$

then there holds

$$u(t) \leq u(a) \exp(\int_a^t \alpha(s)\, ds), \quad \forall t \in (a,b).$$

Proposition A.4 (Christoffel-Darboux-formula)**.** *[1] Let $(H_k(x))_{k\in\mathbb{N}}$ be the Hermite polynomials and $x \neq y$. Then we have*

$$\sum_{k=0}^{n} \frac{1}{k!2^k} H_k(x) H_k(y) = \frac{1}{n!2^{n+1}} \frac{H_n(y)H_{n+1}(x) - H_n(x)H_{n+1}(y)}{x-y}. \tag{A.9}$$

For $x = y$, there holds

$$\sum_{k=0}^{n} \frac{1}{k!2^k} |H_k(x)|^2 = \frac{1}{n!2^{n+1}} (H_n(x)H'_{n+1}(x) - H'_n(x)H_{n+1}(x). \tag{A.10}$$

Bibliography

[1] M. Abramowitz and I. Stegun. *Handbook of mathematical functions with formulas, graphs and mathematical tables*, volume 55. National Buereo of Standard Applied Mathematics Series, 1964.

[2] A. Alvares and B. Carreras. Interaction dynamics for the solitary waves of a nonlinear Dirac model. *Physical Letters A*, 86:327–332, 1981.

[3] W. Bao and J. Shen. A fourth-order time-splitting Laguerre-Hermite pseudo-spectral method for Bose-Einstein condensates. *SIAM Journal of Scientific Computing*, 26:2010–2028, 2005.

[4] J. Belmonte-Beitia and J. Cuevas. Solitons for the cubic-quintic nonlinear schrödinger equation with time- and space-modulated coefficients. *Journal of physics A: mathematical and theoretical*, 42:262–268, 2009.

[5] T. Cazenave. *Semilinear Schroedinger Equations.* Courant Institute of Mathemical Sciences, 2003.

[6] K.-J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations.* Springer-Verlag, New York, 2000.

[7] L. Gauckler. Convergence of a split-step Hermite method for the Gross-Pitaevskii equation. *IMA Journal of Numerical Analysis*, 31:396–414, 2010.

[8] B. Guo, J. Shen, and C. Xu. Spectral and pseudospectral approximations using Hermite functions: application to the Dirac equation. *Advances in Computational Mathematics*, 19:35–55, 2003.

[9] B. Guo and C. Xu. Hermite pseudospectral method for nonlinear partial differential equations. *Mathematical Modelling and Numerical Analysis*, 34:859–872, 2000.

[10] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration.* Springer-Verlag Berlin Heidelberg, 2 edition, 2005.

[11] E. Hansen and A. Ostermann. Exponential splitting for unbounded operators. *Mathematics of Computation*, 78:1485–1496, 2009.

[12] O. Koch and C. Lubich. Variational splitting time integration of the MCTDHF equations in electron dynamics. *IMA Journal of Numerical Analysis*, 31:379–395, 2011.

[13] O. Koch and M. Thalhammer. High-Order splitting methods for the MCTDHF equations in electron dynamics. *Siam Journal on Numerical Analysis*, 2011. To appear.

[14] Kowalski. Spectral theory in Hilbert spaces. Lecture script, ETH Zuerich, 2009.

[15] S. Krein and J. Petunin. Scales of Banach spaces. *Russian Math. Surveys*, 21:85–160, 1966.

[16] A. Levin and D. Lubinsky. Christoffel functions, orthogonal polynomials and Nevai's conjecture for Freud weights. *Constructive Approximation*, 8:461–533, 1992.

[17] C. Lubich. A variational splitting integrator for quantum molecular dynamics. *Applied Numerical Mathematics*, 48:355–368, 2004.

[18] C. Lubich. On splitting methods for Schrödinger-Poisson and cubic nonlinear Schrödinger equations. *Mathematics of Computation*, 77:2141–2153, 2008.

[19] K. B. P. Ruprecht, M. Holland and M. Edwards. Time-dependent solution of the nonlinear Schrödinger equation for trapped neutral atoms. *Physical Reviews A*, 51:4704–4711, 1995.

[20] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*. Springer Verlag, New York, 1983.

[21] L. Pitaevskii and S. Stringari. *Bose-Einstein condensation*. Oxford University Press Inc., 2003.

[22] B. Schneider. Parallel solver for the time-dependent linear and nonlinear Schrödinger equation. *Physical Reviews E*, 73:036708–1–036708–9, 2006.

[23] G. Strang. On the construction and comparison of difference schemes. *SIAM Journal of Numerical Analysis*, 5:506–517, 1968.

[24] M. Thalhammer. High-order exponential operator splitting methods for time-dependent Schrödinger equations. *SIAM Journal of Numerical Analysis*, 46:2022–2038, 2008.

[25] N. H. V. Dominguez and F. Sayas. Hilbert scales and Sobolev spaces defined by associated Legendre functions. *Journal of Computational and Applied Mathematics*, 235:3481–3501, 2011.

[26] H. Yosida. Construction of higher order symplectic integrators. *Physics Letters A*, 150:262–268, 1990.