UNIFORM CONVERGENCE IN VON NEUMANN'S ERGODIC THEOREM IN ABSENCE OF A SPECTRAL GAP

JONATHAN BEN-ARTZI AND BAPTISTE MORISSE

Abstract. Von Neumann's original proof of the ergodic theorem is revisited. A convergence rate is established under the assumption that one can control the density of the spectrum of the underlying self-adjoint operator when restricted to suitable subspaces. Explicit rates are obtained when the bound is polynomial or logarithmic, with applications to the linear Schrödinger and wave equations. In particular, decay estimates for time-averages of solutions are shown.

MSC (2010): 37A30 (primary); 37A25, 37A10, 37C40, 35Q41, 35L05

Keywords: Von Neumann's ergodic theorem; convergence rates; density of states

Acknowledgements: The authors acknowledge support from Fellowship EP/N020154/1 of the Engineering and Physical Sciences Research Council (EPSRC).

Date: May 23, 2019.

Contents

1. Introduction

Let $\mathcal{H}$ be a separable Hilbert space and let $U_t : \mathcal{H} \to \mathcal{H}$ be a one-parameter group of unitary transformations. Let $H : D(H) \subset \mathcal{H} \to \mathcal{H}$ be its self-adjoint generator: $U_t = e^{itH}$. It is well-known that if H has a spectral gap then von Neumann's ergodic theorem has a polynomial convergence rate. In this note we assume the opposite: that H has continuous spectrum in a neighborhood of 0 (and 0 itself is often an eigenvalue). We show that a bound on the density of the spectrum near 0 also leads to a uniform converge rate, albeit on a suitable subspace $\mathcal{X} \subset \mathcal{H}$. We apply this to the linear Schrödinger and wave equations, to obtain the decay estimates (3.4) and (3.9), respectively.

1.1. **Von Neumann's ergodic theorem.** Von Neumann's ergodic theorem [vN32] is a pillar of modern mathematics. Defining

$$P^T f := \frac{1}{2T}\int_{-T}^{T} U_t f \,\mathrm{d}t, \qquad f \in \mathcal{H},$$

and

$$P := \text{orthogonal projection of } \mathcal{H} \text{ onto } \ker H,$$

we have

Theorem 1.1 (Von Neumann's ergodic theorem). *For any $f \in \mathcal{H}$,*

$$\lim_{T\to+\infty} P^T f = Pf.$$

Sketch of proof. The original proof relies on Stone's theorem (and the spectral theorem, by proxy), i.e. the fact that U_t has a resolution of the identity $\{E(\lambda)\}_{\lambda\in\mathbb{R}}$ for which $U_t = \int_{\mathbb{R}} e^{it\lambda}\mathrm{d}E(\lambda)$. This leads to:

$$\begin{aligned}(P^T - P)f &= \frac{1}{2T}\int_{-T}^{T} U_t f\,\mathrm{d}t - Pf = \frac{1}{2T}\int_{-T}^{T}\int_{\mathbb{R}} e^{it\lambda}\,\mathrm{d}E(\lambda)f\,\mathrm{d}t - Pf \\ &= \frac{1}{2T}\int_{-T}^{T}\int_{\mathbb{R}\setminus\{0\}} e^{it\lambda}\,\mathrm{d}E(\lambda)f\,\mathrm{d}t = \int_{\mathbb{R}\setminus\{0\}} \frac{\sin T\lambda}{T\lambda}\,\mathrm{d}E(\lambda)f. \end{aligned} \tag{1.1}$$

This last expression tends to 0 as $T \to +\infty$. □

The strong convergence $P^T \to P$ can be improved to uniform convergence if H has a spectral gap:

Theorem 1.2 (Ergodic theorem: case of spectral gap). *Assume that $\sigma(H) \subset I_\beta^c \cup \{0\}$ where $I_\beta = (-\beta,\beta)$ and $\beta > 0$. Then*

$$\|P^T - P\|_{\mathcal{H}\to\mathcal{H}} \le \beta^{-1}T^{-1}, \qquad \forall T > 0. \tag{1.2}$$

Proof. See Remark 2.1 below. □

1.2. **Main results.** As mentioned above, we assume the opposite of a spectral gap: we assume that $\sigma(H)$ contains a neighborhood of 0. However, we do not want to have "too much" spectrum near 0. We make this precise as follows. Letting $\{E(\lambda)\}_{\lambda\in\mathbb{R}}$ be the resolution of the identity of H, our main assumption is:

Assumption A1. There exist

i. a Banach subspace $\mathcal{X} \subset \mathcal{H}$ which is dense in $\mathcal{H}$ in the topology of $\mathcal{H}$, is continuously embedded in $\mathcal{H}$ (and therefore the norm $\|\cdot\|_{\mathcal{X}}$ is stronger than the norm $\|\cdot\|_{\mathcal{H}}$),

ii. a positive number $r > 0$ and a function $\psi \in L^1(I_r)$ that is strictly positive a.e. on I_r (where $I_r = (-r, r)$),

such that the following bound on the Density of States (DoS) of H holds:

$$\left| \frac{\mathrm{d}}{\mathrm{d}\lambda} (E(\lambda)f, g)_{\mathcal{H}} \right| \leq \psi(\lambda) \|f\|_{\mathcal{X}} \|g\|_{\mathcal{X}}, \qquad \forall f, g \in \mathcal{X}, \forall \lambda \in I_r \setminus \{0\}.$$

To state our main theorem we first define the continuous and strictly increasing function $\Psi : (0, r) \to (0, \Psi(r))$ as

$$\Psi(\varepsilon) := \int_{I_\varepsilon} \psi(\lambda) \, \mathrm{d}\lambda, \qquad \forall \varepsilon \in (0, r),$$

and we define the continuous and strictly increasing function $\Xi : (0, r^2\Psi(r)) \to (0, r)$ as the inverse of the function $\varepsilon \mapsto \varepsilon^2 \Psi(\varepsilon)$:

$$\Xi^{-1}(\varepsilon) := \varepsilon^2 \Psi(\varepsilon), \qquad \forall \varepsilon \in (0, r).$$

We note that both functions can be extended to 0 as continuous functions by defining $\Psi(0) = \Xi(0) = 0$.

Theorem 1.3. *Under Assumption A1 the following uniform rate in von Neumann's ergodic theorem holds:*

$$\|P^T - P\|_{\mathcal{X} \to \mathcal{H}} \leq \left(2\Psi\left(\Xi\left(T^{-2}\right)\right)\right)^{1/2}, \qquad \forall T > r^{-1}\Psi^{-1/2}(r). \tag{1.3}$$

Remark 1.4. Considering the definitions of Ψ and Ξ we observe that

$$\lim_{T \to +\infty} \left(2\Psi\left(\Xi\left(T^{-2}\right)\right)\right)^{1/2} = \left(2\Psi\left(\Xi\left(0\right)\right)\right)^{1/2} = 0$$

so that, indeed,

$$\lim_{T \to +\infty} \|P^T - P\|_{\mathcal{X} \to \mathcal{H}} = 0.$$

We obtain explicit rates if given more precise bounds for the DoS:

Corollary 1.5 (Polynomial and logarithmic bounds). *Under Assumption A1, we have:*

1. *Assume that there exist $C, p > 0$ for which*

$$\psi(\lambda) = C|\lambda|^{p-1}, \qquad \lambda \in I_r \setminus \{0\}.$$

Then

$$\|P^T - P\|_{\mathcal{X} \to \mathcal{H}} \leq C(p) T^{-\frac{p}{2+p}}. \tag{1.4}$$

2. *Assume that the following bound holds*

$$\psi(\lambda) = \lambda^{-1} |\log(\lambda)|^{-2}, \qquad \lambda \in I_r \setminus \{0\}.$$

Then for all $T > 0$ large enough,

$$\|P^T - P\|_{\mathcal{X}\to\mathcal{H}} \leq \log\left(T\log(T)^{-1/2}\right)^{-1}. \tag{1.5}$$

Remark 1.6. We note that as $p \to +\infty$ the rate $T^{-\frac{p}{2+p}}$ in (1.4) approaches the rate T^{-1} which holds in the case of a spectral gap. This agrees with the intuition that if p is very large then the function $|\lambda|^{p-1}$ is "nearly" 0 in a neighborhood of $\lambda = 0$.

1.3. Comparison to the RAGE theorem.

It is natural to compare our result to the RAGE theorem (see e.g. [CFKS87]). It states that for any compact operator K and any $\varphi \in \mathcal{H}$,

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T \|KU_tP_{\mathrm{ac}}\varphi\|^2 \,\mathrm{d}t = 0,$$

where P_{ac} is the orthogonal projection onto the absolutely continuous subspace of H. Where Theorem 1.3 proves uniform convergence to the projection onto the kernel of the generator, the RAGE theorem proves a weak convergence to 0 of the time average of the evolution of the continuous part of the spectrum.

1.4. The spectral theorem.

Since the spectral theorem and the resolution of the identity of self-adjoint operators play a central role in this paper, we recall some basic facts.

Definition 1.7 (Resolution of the identity). *Let $H : D(H) \subset \mathcal{H} \to \mathcal{H}$ be a self-adjoint operator. Its associated* resolution of the identity *$\{E(\lambda)\}_{\lambda\in\mathbb{R}}$ is a family of projection operators in $\mathcal{H}$ with the property that, for each $\lambda \in \mathbb{R}$, the subspace $\mathcal{H}^\lambda = E(\lambda)\mathcal{H}$ is the largest closed subspace such that*

1. $\mathcal{H}^\lambda$ reduces H, *namely, $HE(\lambda)g = E(\lambda)Hg$ for every $g \in D(H)$. In particular, if $g \in D(H)$ then also $E(\lambda)g \in D(H)$.*
2. *$(Hu, u)_{\mathcal{H}} \leq \lambda(u, u)_{\mathcal{H}}$ for every $u \in \mathcal{H}^\lambda \cap D(H)$.*

Now we are able to state the spectral theorem:

Theorem 1.8 (Spectral theorem). *Let $H : D(H) \subset \mathcal{H} \to \mathcal{H}$ be a self-adjoint operator and let $\{E(\lambda)\}_{\lambda\in\mathbb{R}}$ be the associated resolution of the identity. Then $\{E(\lambda)\}_{\lambda\in\mathbb{R}}$ is unique, and satisfies*

$$H = \int_{\mathbb{R}} \lambda \,\mathrm{d}E(\lambda).$$

In addition to the above, it is useful to state the definition of the spectral measure:

Definition 1.9 (Spectral measure). *Given any $f, g \in \mathcal{H}$ the resolution of the identity defines a complex function of bounded variation on the real line, given by*

$$\mathbb{R} \ni \lambda \mapsto (E(\lambda)f, g)_{\mathcal{H}}.$$

It is well-known that such a function gives rise to a complex measure (depending on f, g) called the spectral measure.

Definition 1.10 (Density of states). *Let $\mathcal{X} \subset \mathcal{H}$ be some closed subspace satisfying Assumption 1. We call the bilinear form*

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}(E(\lambda)\cdot,\cdot)_{\mathcal{H}} : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$$

the density of states of H at λ on the subspace $\mathcal{X}$.

1.5. **Organization of the paper.** Section 2 is devoted to the proofs of the main theorem and its corollaries. In Section 3 we apply these results to the linear Schrödinger and wave equations, to obtain decay estimates for averages of solutions.

2. Proof of the main results

Proof of Theorem 1.3. Our starting point is the observation [Kat95, V-§2.1] that if the bilinear form $\frac{\mathrm{d}}{\mathrm{d}\lambda}(E(\lambda)\cdot,\cdot)_{\mathcal{H}} : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ is bounded at a given $\lambda \in \mathbb{R}$, then there exists a bounded operator $A(\lambda) : \mathcal{X} \to \mathcal{X}^*$ satisfying

$$\langle A(\lambda)f, g\rangle = \frac{\mathrm{d}}{\mathrm{d}\lambda}(E(\lambda)f, g)_{\mathcal{H}}, \qquad \forall f, g \in \mathcal{X},$$

where $\langle\cdot,\cdot\rangle$ is the $(\mathcal{X}^*, \mathcal{X})$ dual space pairing. Moreover, the operator norm of $A(\lambda)$ shares the same bound as the bilinear form. Now, recalling von Neumann's proof as sketched in (1.1) above, we have

$$(P^T - P)f = \int_{\mathbb{R}\setminus\{0\}} \frac{\sin T\lambda}{T\lambda}\,\mathrm{d}E(\lambda)f.$$

We split this integral as follows:

$$\int_{\mathbb{R}\setminus\{0\}} \frac{\sin T\lambda}{T\lambda}\,\mathrm{d}E(\lambda)f = \left(\int_{I_\varepsilon\setminus\{0\}} + \int_{I_\varepsilon^c}\right)\frac{\sin T\lambda}{T\lambda}\,\mathrm{d}E(\lambda)f$$

where $I_\varepsilon = (-\varepsilon, \varepsilon)$ and $0 < \varepsilon < r$ is to be determined later. We start by estimating the second integral:

$$\begin{aligned}
\left\|\int_{I_\varepsilon^c} \frac{\sin T\lambda}{T\lambda}\,\mathrm{d}E(\lambda)f\right\|_{\mathcal{H}}^2 &= \int_{I_\varepsilon^c}\left|\frac{\sin T\lambda}{T\lambda}\right|^2 \mathrm{d}(E(\lambda)f, f)_{\mathcal{H}} \\
&\le \frac{1}{T^2\varepsilon^2}\int_{\mathbb{R}} \mathrm{d}(E(\lambda)f, f)_{\mathcal{H}} \\
&= \frac{1}{T^2\varepsilon^2}\|f\|_{\mathcal{H}}^2 \le \frac{1}{T^2\varepsilon^2}\|f\|_{\mathcal{X}}^2.
\end{aligned}$$

Now we turn to the first integral:

$$\begin{aligned}\left\|\int_{I_\varepsilon\setminus\{0\}} \frac{\sin T\lambda}{T\lambda}\,\mathrm{d}E(\lambda)f\right\|_{\mathcal{H}}^2 &= \int_{I_\varepsilon\setminus\{0\}} \left|\frac{\sin T\lambda}{T\lambda}\right|^2 \mathrm{d}(E(\lambda)f,f)_{\mathcal{H}} \\ &= \int_{I_\varepsilon\setminus\{0\}} \left|\frac{\sin T\lambda}{T\lambda}\right|^2 \langle A(\lambda)f,f\rangle\,\mathrm{d}\lambda \\ &\le \left(\int_{I_\varepsilon\setminus\{0\}} \psi(\lambda)\,\mathrm{d}\lambda\right)\|f\|_{\mathcal{X}}^2 \\ &= \Psi(\varepsilon)\|f\|_{\mathcal{X}}^2.\end{aligned}$$

Altogether, both estimates lead to

$$\|(P^T-P)f\|_{\mathcal{H}}^2 \le \left(\frac{1}{T^2\varepsilon^2}+\Psi(\varepsilon)\right)\|f\|_{\mathcal{X}}^2.$$

We choose the optimal ε, which satisfies

$$\varepsilon^2\Psi(\varepsilon) = \frac{1}{T^2} \qquad \text{(or, equivalently, } \Xi(T^{-2})=\varepsilon).$$

Recalling the definitions of Ψ and Ξ, there holds

$$\|(P^T-P)f\|_{\mathcal{H}} \le \left(2\Psi\left(\Xi\left(T^{-2}\right)\right)\right)^{1/2}\|f\|_{\mathcal{X}}$$

which completes the proof. □

Remark 2.1 (Spectral gap). In the case of a spectral gap (1.2) immediately follows. Indeed, with gap of size β in the above proof one has

$$\|(P^T-P)f\|_{\mathcal{H}} \le \beta^{-1}T^{-1}\|f\|_{\mathcal{H}}.$$

Note that in this case the subspace $\mathcal{X}$ is no longer needed.

Proof of Corollary 1.5. 1. In the case of a polynomial bound $\psi(\lambda)=C|\lambda|^{p-1}$, with some $p,C>0$, the function Ψ is given by $\Psi(\varepsilon)=C\varepsilon^p$. The inverse of the function $\varepsilon\mapsto\varepsilon^2\Psi(\varepsilon)=C\varepsilon^{p+2}$ is simply $\Xi(y)=C^{-1/(2+p)}y^{1/(2+p)}$. Then the rate in (1.3) is

$$\left(2\Psi\left(\Xi\left(T^{-2}\right)\right)\right)^{1/2} = C(p)T^{-\frac{p}{2+p}},$$

which verifies (1.4).

2. In the case of a logarithmic bound $\psi(\lambda)=\lambda^{-1}|\log(\lambda)|^{-2}$ we see that $\Psi(\varepsilon)=|\log(\varepsilon)|^{-1}$. The question is then to compute Ξ, the inverse function of the strictly increasing function $\varepsilon\mapsto\varepsilon^2\Psi(\varepsilon)=\varepsilon^2|\log(\varepsilon)|^{-1}$ near $\varepsilon=0$. It is evident that $\Xi(0)=0$. The definition of Ξ is equivalent to

$$y=\Xi(y)^2\Psi(\Xi(y)) \tag{2.1}$$

so that for all $y > 0$ sufficiently small

$$\begin{aligned} \frac{\Xi(y)^2}{y} &= -\log(\Xi(y)) \\ &= -\frac{1}{2}\log\left(y\frac{\Xi(y)^2}{y}\right) \\ &= -\frac{1}{2}\log(y) - \frac{1}{2}\log\left(\frac{\Xi(y)^2}{y}\right). \end{aligned} \tag{2.2}$$

Since $\Psi(\Xi(y)) \to 0$ as $y \to 0$, the equality (2.1) implies that $\Xi(y)^2/y \to +\infty$ as $y \to 0$. Therefore there exists some $y_0 > 0$ such that $\Xi(y)^2/y > 1$ for all $y < y_0$. Hence the last term in (2.2) is negative, which leads to

$$\frac{\Xi(y)^2}{y} \leq -\frac{1}{2}\log(y).$$

Consequently

$$\Xi(y) \leq y^{1/2}|\log(y^{1/2})|^{1/2}, \qquad \forall y > 0 \text{ sufficiently small.}$$

Recalling that Ψ is an increasing function on $(0, r)$ we have

$$\Psi(\Xi(y)) \leq \left|\log\left(y^{1/2}|\log(y^{1/2})|^{1/2}\right)\right|^{-1}$$

which is (1.5) (recall that $y = T^{-2}$). ☐

3. Examples

3.1. The Laplace operator. Let $\varphi : [0, +\infty) \to [0, +\infty)$ be some continuous and strictly increasing function and define $H = \varphi(-\Delta)$ as a function of the Laplace operator acting in $\mathcal{H} = L^2(\mathbb{R}^d)$ with an appropriate domain for self-adjointness. Note that if $\varphi(x) = x$ is the identity, then $-iH$ is the generator of the *Schrödinger equation*:

$$\begin{cases} \partial_t f(t,x) = i\Delta f(t,x), & t \in \mathbb{R},\, x \in \mathbb{R}^d, \\ f(0,x) = f_0(x), & x \in \mathbb{R}^d. \end{cases}$$

Let $\{E(\lambda)\}_{\lambda\in\mathbb{R}}$ be the resolution of the identity of H. We use the fact that the Fourier transform is a unitary map relating $-\Delta$ to multiplication by $|\xi|^2$ in order to get:

$$(E(\lambda)f, g)_{\mathcal{H}} = \int_{\varphi(|\xi|^2)\leq\lambda} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\,\mathrm{d}\xi, \qquad \lambda \geq 0. \tag{3.1}$$

Let us show how different choices of subspaces $\mathcal{X}$ can give rise to different results.

3.1.1. *Hilbertian subspace.* Differentiating (3.1) in λ we get

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big|_{\lambda=\lambda_0}(E(\lambda)f,g)_{\mathcal{H}} = \int_{|\xi|=\sqrt{\varphi^{-1}(\lambda_0)}} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\left|\nabla\left(\varphi(|\xi|^2)\right)\right|^{-1}\mathrm{d}\sigma \tag{3.2}$$

where $\mathrm{d}\sigma$ is the Lebesgue (uniform) surface measure on the $d-1$-dimensional sphere of radius $\sqrt{\varphi^{-1}(\lambda_0)}$. The term $\left|\nabla\left(\varphi(|\xi|^2)\right)\right|^{-1} = \frac{1}{2|\xi|\varphi'(|\xi|^2)}$ comes from the coarea formula [Eva10, Appendix C3]. An evaluation of the L^2 functions $\widehat{f}$ and $\widehat{g}$ on the hypersurface $\{|\xi| = \sqrt{\varphi^{-1}(\lambda_0)}\}$ only makes sense if they belong to any Sobolev space $H^s(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$ with $s > 1/2$ by the trace lemma[1]. The functions $\widehat{f}$ and $\widehat{g}$ belong to $H^s(\mathbb{R}^d)$ if and only if f and g belong to $L^{2,s}(\mathbb{R}^d)$, defined as

$$L^{2,s}(\mathbb{R}^d) := \left\{ f \in L^2(\mathbb{R}^d) \ : \ \|f\|^2_{L^{2,s}(\mathbb{R}^d)} := \int_{\mathbb{R}^d} |f(x)|^2(1+|x|^2)^s\,\mathrm{d}x < \infty \right\}.$$

We therefore conclude that we can bound (3.2) using the $L^{2,s}$-norms of f and g, which are stronger than their $\mathcal{H}$-norms:

$$\left|\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big|_{\lambda=\lambda_0}(E(\lambda)f,g)_{\mathcal{H}}\right| \leq \frac{1}{2\sqrt{\varphi^{-1}(\lambda_0)}\varphi'(\varphi^{-1}(\lambda_0))}\|f\|_{L^{2,s}}\|g\|_{L^{2,s}}, \tag{3.3}$$

where we denote $L^{2,s}$ rather than $L^{2,s}(\mathbb{R}^d)$ for brevity. Hence we have that

$$\begin{aligned}
\mathcal{X} &= L^{2,s}(\mathbb{R}^d) \\
\psi(\lambda) &= \frac{1}{2\sqrt{\varphi^{-1}(\lambda)}\varphi'(\varphi^{-1}(\lambda))} \\
\Psi(\varepsilon) &= \int_{I_\varepsilon\setminus\{0\}} \frac{1}{2\sqrt{\varphi^{-1}(\lambda)}\varphi'(\varphi^{-1}(\lambda))}\,\mathrm{d}\lambda.
\end{aligned}$$

In the case of the Schrödinger equation ($\varphi = \mathrm{id}$) we get

$$\Psi(\varepsilon) = \int_{I_\varepsilon\setminus\{0\}} \frac{1}{2\sqrt{\lambda}}\,\mathrm{d}\lambda = \sqrt{\varepsilon}$$

and from (1.4) we get a convergence rate of $T^{-\frac{1}{5}}$. Moreover, since $-\Delta$ has no eigenvalues in this setting (and, in particular, a trivial kernel), we conclude that

$$\left\|\frac{1}{2T}\int_{-T}^{T} e^{-it\Delta} f_0\,\mathrm{d}t\right\|_{L^2} = \|P^T f_0\|_{L^2} \leq CT^{-\frac{1}{5}}\|f_0\|_{L^{2,s}}. \tag{3.4}$$

This also implies that

$$\|P^T f_0\|_{L^q_T L^2_x([0,\infty)\times\mathbb{R}^d)} \leq C(q)\|f_0\|_{L^{2,s}_x(\mathbb{R}^d)}, \qquad \forall q > 5. \tag{3.5}$$

Restricting to any bounded domain $\Omega \subset \mathbb{R}^d$ we may take L^2 norms rather than weighted norms (the weight is uniformly bounded away from 0 and $+\infty$ in Ω) so we have

$$\|P^T f_0\|_{L^q_T L^2_x([0,\infty)\times\Omega)} \leq C(q,\Omega)\|f_0\|_{L^2_x(\Omega)}, \qquad \forall q > 5.$$

[1]This is not entirely optimal, since we are not making use of the fact that this hypersurface is in fact a sphere.

3.1.2. *Non-Hilbertian subspace.* Considering (3.2) again, we may change variables so that the integration takes place on the unit sphere in $\mathbb{R}^d$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big|_{\lambda=\lambda_0}(E(\lambda)f,g)_{\mathcal{H}} &= \int_{|\xi|=\sqrt{\varphi^{-1}(\lambda_0)}} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\left|\nabla\left(\varphi(|\xi|^2)\right)\right|^{-1}\mathrm{d}\sigma\\ &= \sqrt{\varphi^{-1}(\lambda_0)}^{d-1}\int_{\mathbb{S}^{d-1}}\frac{\widehat{f}\left(\tau\sqrt{\varphi^{-1}(\lambda_0)}\right)\overline{\widehat{g}\left(\tau\sqrt{\varphi^{-1}(\lambda_0)}\right)}}{2\sqrt{\varphi^{-1}(\lambda_0)}\varphi'(\varphi^{-1}(\lambda_0))}\,\mathrm{d}\tau\\ &= \frac{\sqrt{\varphi^{-1}(\lambda_0)}^{d-2}}{2\varphi'(\varphi^{-1}(\lambda_0))}\int_{\mathbb{S}^{d-1}}\widehat{f}\left(\tau\sqrt{\varphi^{-1}(\lambda_0)}\right)\overline{\widehat{g}\left(\tau\sqrt{\varphi^{-1}(\lambda_0)}\right)}\,\mathrm{d}\tau\end{aligned}$$

where $\mathrm{d}\tau$ is the uniform measure on the unit sphere in $\mathbb{R}^d$. Another way to make sense of the restriction of L^2 functions to a hypersurface is if they are bounded, i.e. one can bound:

$$\begin{aligned}\left|\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big|_{\lambda=\lambda_0}(E(\lambda)f,g)_{\mathcal{H}}\right| &\le \frac{|\mathbb{S}^{d-1}|\sqrt{\varphi^{-1}(\lambda_0)}^{d-2}}{2\varphi'(\varphi^{-1}(\lambda_0))}\|\widehat{f}\|_{L^\infty(\mathbb{R}^d)}\|\widehat{g}\|_{L^\infty(\mathbb{R}^d)}\\ &\le \frac{|\mathbb{S}^{d-1}|\sqrt{\varphi^{-1}(\lambda_0)}^{d-2}}{2\varphi'(\varphi^{-1}(\lambda_0))}\|f\|_{L^1(\mathbb{R}^d)}\|g\|_{L^1(\mathbb{R}^d)}.\end{aligned}$$

Thus we obtain

$$\begin{aligned}\mathcal{X} &= L^1(\mathbb{R}^d)\cap L^2(\mathbb{R}^d)\\ \psi(\lambda) &= \frac{|\mathbb{S}^{d-1}|\sqrt{\varphi^{-1}(\lambda_0)}^{d-2}}{2\varphi'(\varphi^{-1}(\lambda_0))}\\ \Psi(\varepsilon) &= \int_{I_\varepsilon\setminus\{0\}}\frac{|\mathbb{S}^{d-1}|\sqrt{\varphi^{-1}(\lambda)}^{d-2}}{2\varphi'(\varphi^{-1}(\lambda))}\,\mathrm{d}\lambda.\end{aligned}$$

We again consider the Schrödinger case where we obtain $\psi(\lambda)=\frac{1}{2}|\mathbb{S}^{d-1}|\lambda^{\frac{d}{2}-1}$ so that

$$\Psi(\varepsilon)=\int_{I_\varepsilon\setminus\{0\}}\frac{1}{2}|\mathbb{S}^{d-1}|\lambda^{\frac{d}{2}-1}\,\mathrm{d}\lambda=\frac{1}{d}|\mathbb{S}^{d-1}|\varepsilon^{d/2}$$

From (1.4) we get a convergence rate of $T^{-\frac{d}{4+d}}$. For any $f_0\in L^2\cap L^1$, there holds

$$\left\|\frac{1}{2T}\int_{-T}^{T}e^{-it\Delta}f_0\,\mathrm{d}t\right\|_{L^2}=\|P^Tf_0\|_{L^2}\le CT^{-\frac{d}{4+d}}\|f_0\|_{L^1}.$$

This leads to the global-in-time estimate:

$$\|P^Tf_0\|_{L^q_TL^2_x([0,\infty)\times\mathbb{R}^d)}\le C(q)\|f_0\|_{L^1_x(\mathbb{R}^d)},\qquad \forall f_0\in L^2\cap L^1,\ \forall q>\frac{4+d}{d}. \tag{3.6}$$

Remark 3.1. It is natural to compare the estimates (3.5) and (3.6) with:

1) The well-known *Strichartz estimates*

$$\left\|e^{it\Delta/2}f_0\right\|_{L^q_tL^r_x(\mathbb{R}\times\mathbb{R}^d)}\le C(d,q,r)\|f_0\|_{L^2_x(\mathbb{R}^d)}$$

where $2\le q,r\le\infty$, $\frac{2}{q}+\frac{d}{r}=\frac{d}{2}$ and $(q,r,d)\ne(2,\infty,2)$ (see [Tao06]).

2) *Smoothing estimates*, such as

$$\left\||D_x|^{1/2}e^{it\Delta}f_0\right\|_{L^2_tL^{2,-s}_x(\mathbb{R}\times\mathbb{R}^d)}\le C(d)\|f_0\|_{L^2_x(\mathbb{R}^d)}$$

where $s > 1/2$, see [BAK92]. A detailed comparison between these estimates is elusive at the present time, and is the subject of ongoing research.

3.2. **The wave operator.** We now consider the linear, homogeneous wave equation

$$\begin{cases} \partial_t^2 f(t,x) - \Delta f(t,x) = 0, & t \in \mathbb{R},\ x \in \mathbb{R}^d, \\ f(0,x) = f_0(x),\ \partial_t f(0,x) = g_0(x), & x \in \mathbb{R}^d. \end{cases}$$

We let $\mathcal{H} = L^2(\mathbb{R}^d)$ and consider the self-adjoint operator (with an appropriate domain) $H = -\Delta$. We first need to convert the above problem into a first order system. We follow a well-known procedure: define

$$f_\pm := \frac{1}{2}\left(\sqrt{H} f \pm i\partial_t f\right). \tag{3.7}$$

Then we compute

$$\begin{aligned} \partial_t f_\pm &= \frac{1}{2}\left(\sqrt{H}\partial_t f \pm i\partial_t^2 f\right) = \frac{1}{2}\left(\sqrt{H}\partial_t f \mp iHf\right) = \frac{1}{2}\sqrt{H}\left(\partial_t f \mp i\sqrt{H} f\right) \\ &= \frac{i}{2}\sqrt{H}\left(-i\partial_t f \mp \sqrt{H} f\right) = \mp i\sqrt{H} f_\pm. \end{aligned}$$

It follows that the vector

$$F(t,x) := \begin{pmatrix} f_+(t,x) \\ f_-(t,x) \end{pmatrix}$$

satisfies the equation

$$F'(t) = -iKF \qquad \text{where} \qquad K = \begin{pmatrix} \sqrt{H} & 0 \\ 0 & -\sqrt{H} \end{pmatrix}.$$

Denoting $\{E_{\sqrt{H}}(\lambda)\}_{\lambda\in\mathbb{R}}$ and $\{E_K(\lambda)\}_{\lambda\in\mathbb{R}}$ the resolutions of the identity of $\sqrt{H}$ and K, respectively, we first observe that $E_{-\sqrt{H}}(\lambda) = I - E_{\sqrt{H}}(-\lambda)$ so that

$$E_K(\lambda) = E_{\sqrt{H}}(\lambda) \oplus (I - E_{\sqrt{H}}(-\lambda)), \qquad \forall \lambda \in \mathbb{R}.$$

For $\sqrt{H}$, we know from (3.3) that all constants in the estimate of the density of states cancel, so that for $s > 1/2$,

$$\left|\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(E_{\sqrt{H}}(\lambda) f, g\right)_{\mathcal{H}}\right| \le \|f\|_{L^{2,s}} \|g\|_{L^{2,s}}, \qquad \forall \lambda \in \mathbb{R}.$$

This implies that

$$\left|\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(E_K(\lambda) F, G\right)_{\mathcal{H}\oplus\mathcal{H}}\right| \le \|F\|_{L^{2,s}\oplus L^{2,s}} \|G\|_{L^{2,s}\oplus L^{2,s}}, \qquad \forall \lambda \in \mathbb{R}.$$

The bound on the DoS is therefore simply $\psi(\lambda) = 1$, so that from (1.4) we get a convergence rate of $T^{-\frac{1}{3}}$ (noting that the kernel is empty):

$$\left\| \frac{1}{2T}\int_{-T}^{T} e^{-itK} F_0 \,\mathrm{d}t \right\|_{\mathcal{H}\oplus\mathcal{H}} \le C T^{-\frac{1}{3}} \|F_0\|_{L^{2,s}\oplus L^{2,s}}. \tag{3.8}$$

To obtain direct bounds for the average of the solution $f(t)$ of the wave equation, we use that $f = \sqrt{H}^{-1}(f_+ + f_-)$ to write

$$\begin{aligned}\frac{1}{2T}\int_{-T}^{T} f(t)\,\mathrm{d}t &= \frac{1}{2T}\int_{-T}^{T}\sqrt{H}^{-1}(f_+ + f_-)(t)\,\mathrm{d}t\\ &= \frac{1}{2T}\int_{-T}^{T}\sqrt{H}^{-1}\Big(\big(e^{-itK}F_0\big)_1 + \big(e^{-itK}F_0\big)_2\Big)\,\mathrm{d}t.\end{aligned}$$

Estimate (3.8) leads then to

$$\begin{aligned}\left\|\frac{1}{2T}\int_{-T}^{T} f(t)\,\mathrm{d}t\right\|_{\mathcal{H}} &\le CT^{-\frac{1}{3}}\left(\left\|\sqrt{H}^{-1}(F_0)_1\right\|_{L^{2,s}} + \left\|\sqrt{H}^{-1}(F_0)_2\right\|_{L^{2,s}}\right)\\ &\le CT^{-\frac{1}{3}}\left(\|f_0\|_{L^{2,s}} + \left\|\sqrt{H}^{-1}g_0\right\|_{L^{2,s}}\right), \qquad (3.9)\end{aligned}$$

where in the second inequality we have used (3.7):

$$2\left\|\sqrt{H}^{-1}f_\pm(t=0)\right\|_{L^{2,s}} \le \|f_0\|_{L^{2,s}} + \left\|\sqrt{H}^{-1}g_0\right\|_{L^{2,s}}.$$

Denoting $P^T(f_0, g_0) := \frac{1}{2T}\int_{-T}^{T} f(t)\,\mathrm{d}t$ for brevity, we deduce the global-in-time estimate

$$\left\|P^T(f_0,g_0)\right\|_{L^q_T L^2_x([0,\infty)\times\mathbb{R}^d)} \le C(q)\left(\|f_0\|_{L^{2,s}(\mathbb{R}^d)} + \left\|\sqrt{H}^{-1}g_0\right\|_{L^{2,s}(\mathbb{R}^d)}\right), \qquad \forall q > 3. \tag{3.10}$$

Remark 3.2. Here we compare our estimates to *Strichartz estimates* for the wave equation (see [Tat01]):

$$\|f\|_{L^q_t L^p_x(\mathbb{R}\times\mathbb{R}^d)} \le C(d,q,p,s)\left(\|f_0\|_{H^s(\mathbb{R}^d)} + \|g_0\|_{H^{s-1}(\mathbb{R}^d)}\right)$$

for triplets satisfying $2 \le p, q \le \infty$ and

$$\frac{1}{q} + \frac{d}{p} = \frac{d}{2} - s, \qquad \frac{2}{q} + \frac{d-1}{p} \le \frac{d-1}{2}.$$

In particular, we can compare (3.10) with the Strichartz estimate for $(q,p,s) = (\infty, 2, 0)$:

$$\|f\|_{L^\infty_t L^2_x(\mathbb{R}\times\mathbb{R}^d)} \le C(d)\left(\|f_0\|_{L^2_x(\mathbb{R}^d)} + \|g_0\|_{H^{-1}_x(\mathbb{R}^d)}\right).$$

References

[BAK92] Matania Ben-Artzi and Sergiu Klainerman. Decay and regularity for the Schrödinger equation. *J. d'Analyse Mathématique*, 58(1):25–37, dec 1992.

[CFKS87] H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon. *Schrödinger operators with application to quantum mechanics and global geometry.* Texts and Monographs in Physics. Springer-Verlag, Berlin, study edition, 1987.

[Eva10] Lawrence C. Evans. *Partial Differential Equations (Graduate Studies in Mathematics).* American Mathematical Society, 2010.

[Kat95] Tosio Kato. *Perturbation Theory for Linear Operators.* Springer-Verlag, 1995.

[Tao06] Terence Tao. *Nonlinear Dispersive Equations: Local and Global Analysis.* 2006.

[Tat01] Daniel Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. II. *Amer. J. Math.*, 123(3):385–423, 2001.

[vN32] John von Neumann. Proof of the quasi-ergodic hypothesis. *Proc. Natl. Acad. Sci.*, 18(2):70–82, 1932.

School of Mathematics, Cardiff University, Cardiff CF24 4AG, Wales, UK

E-mail address: `Ben-ArtziJ@cardiff.ac.uk`

E-mail address: `MorisseB@cardiff.ac.uk`