CrossMark

A Rigorous Theory of Many-Body Prethermalization for Periodically Driven and Closed Quantum Systems

Dmitry Abanin[1]**, Wojciech De Roeck**[2] **, Wen Wei Ho**[1]**, François Huveneers**[3]

[1] Department of Theoretical Physics, University of Geneva, Geneva, Switzerland
[2] Instituut voor Theoretische Fysica, KU Leuven, Leuven, Belgium. E-mail: wmderoeck@gmail.com
[3] Université Paris-Dauphine, PSL Research University, CNRS, CEREMADE, 75016 Paris, France

Abstract: Prethermalization refers to the transient phenomenon where a system thermalizes according to a Hamiltonian that is not the generator of its evolution. We provide here a rigorous framework for quantum spin systems where prethermalization is exhibited for very long times. First, we consider quantum spin systems under periodic driving at high frequency ν. We prove that up to a quasi-exponential time $\tau_* \sim \mathrm{e}^{c \frac{\nu}{\log^3 \nu}}$, the system barely absorbs energy. Instead, there is an effective local Hamiltonian $\widehat{D}$ that governs the time evolution up to τ_*, and hence this effective Hamiltonian is a conserved quantity up to τ_*. Next, we consider systems without driving, but with a separation of energy scales in the Hamiltonian. A prime example is the Fermi–Hubbard model where the interaction U is much larger than the hopping J. Also here we prove the emergence of an effective conserved quantity, different from the Hamiltonian, up to a time τ_* that is (almost) exponential in U/J.

1. Introduction

1.1. Periodically driven systems. Time-dependent periodic perturbations arise naturally in various physical systems, e.g. when a physical system is irradiated with electromagnetic fields. Such periodically driven quantum systems exhibit rich and often unexpected behaviour [8]. One classic example is the dynamical localization of a kicked quantum rotor [18]. Another well-known (but non-quantum) example is the Kapitza pendulum [22], where a sufficiently fast periodic drive stabilizes the otherwise unstable fixed point in which the pendulum stands on its head.

More recently, it has been suggested that periodic driving can be used as a tool to design interesting and exotic many-body Hamiltonians. The underlying idea here is that the time evolution operator $U(t)$, generated by a periodically modulated Hamiltonian $H(t) = D + V(t)$ with period T, could be recast as

$$U(mT) = \mathrm{e}^{-\mathrm{i}mT\widehat{D}} \quad \text{for } m \in \mathbb{Z}, \tag{1.1}$$

with $\widehat{D}$ the 'effective' Hamiltonian, and $U(T)$ sometimes called 'monodromy' or Floquet' operator. A prime example of (1.1) is where the system consists of non-interacting fermions on the lattice:

$$H(t) = \sum_{x, y \in \Lambda} h(x, y, t) c_x^* c_y,$$

with c_x, c_x^*, fermionic field operators and $h(\cdot, \cdot, t)$ the kernel of a self-adjoint operator on $l^2(\Lambda)$ with the volume Λ a finite subset of $\mathbb{Z}^d$. We write the one-particle unitary evolution $u(t)$, solving $\frac{d}{dt} u(t) = -ih(t)u(t)$ with $u(0) = 1$. Then an algebraic exercise yields (1.1) with

$$\widehat{D} = \sum_{x, y} \widehat{d}(x, y) c_x^* c_y, \qquad \text{with } \widehat{d} \text{ solving } e^{-iT\widehat{d}} = u(T).$$

If $T\|h(t)\| \ll 1$, then the spectrum of $u(T)$ covers only a small patch of the unit circle and one can construct $\widehat{d}$ as a convergent series in $h(t), 0 < t \le T$. Its leading term (as $T \to 0$) is the averaged Hamiltonian $\frac{1}{T} \int_0^T dt h(t)$. Hence, under these conditions $\widehat{D}$ can be chosen as a local many-body Hamiltonian that is moreover locally close to $\frac{1}{T} \int_0^T dt H(t)$. When calling extensive operators 'local', we mean that they can be written as a sum of local terms, the range of which does not grow with volume. The expansion alluded to is in general known as the 'Magnus' expansion', see e.g. [8,29].

Recent theoretical works [13,25,32] suggest, based on the quantum version of the ergodic hypothesis (Eigenstate Thermalization Hypothesis), that generic periodically driven many-body systems eventually do heat up to an infinite temperature, and so one would *not* expect the strict equality (1.1) to hold for a local $\widehat{D}$. Apart from non-interacting systems, (1.1) can still be true if D describes a disordered many-body localized system, and $T\|V(t)\|$ is sufficiently small. In that case, numerics and theory [3,26,33] suggest that (1.1) holds with $\widehat{D}$ similar to D. But, as said, this is not our prime interest here and we study systems for which (1.1) is not expected to hold strictly.

In this paper, we argue instead that in great generality, the equality (1.1) holds however, approximatively, up to the quasi-exponential time $\tau_* \sim e^{c \frac{\nu}{\log^3 \nu}}$, i.e. for $mT \le \tau_*$, with $\nu = 2\pi/T$ the frequency, for a local Hamiltonian $\widehat{D}$. In particular, we prove that $\widehat{D}$ is approximatively conserved up to time τ_*, and that the evolution of local observables $O(t) = U(t)OU^*(t)$ for stroboscopic times $t \in T\mathbb{N}$ is well-approximated by the effective Heisenberg evolution $e^{it\widehat{D}} O e^{-it\widehat{D}}$.

The physical significance of our work is hence that we rigorously underpin the use of effective Hamiltonians $\widehat{D}$. Part of the appeal of this idea is that $\widehat{D}$ can have different properties from the original static Hamiltonian D, for example, nonzero Chern numbers, see Refs. [4,8,19,23,28].

1.2. Heuristics and connection to non-driven systems. The emergence of an effective conserved quantity is not specific to periodically driven systems, and we now outline how to generalize our results to closed (i.e. non-driven) Hamiltonian dynamics. At the same time, we develop the basic intuition that underlies our results.

1.2.1. Driven systems in Floquet representation. To analyze the evolution equation $i\partial_t \phi(t) = H(t)\phi(t)$, with $H(t) = H(t + T)$ acting on a Hilbert space $\mathcal{H}$, a standard technique is to work in an extended space, see [20] for details. If a sufficiently regular $\psi(\theta, t) \in L^2(\mathbb{R}/T\mathbb{Z}, \mathcal{H})$ solves the equation

$$i\partial_t \psi(\theta, t) = \big(-i\partial_\theta + H(\theta)\big)\psi(\theta, t), \tag{1.2}$$

(with $i\partial_\theta$ defined with periodic boundary conditions) then $\phi(t) = \psi(t, t)$ solves the equation $i\partial_t \phi(t) = H(t)\phi(t)$. Under Fourier transform $L^2(\mathbb{R}/T, \mathcal{H}) \mapsto l^2(\mathbb{Z}, \mathcal{H}) :$ $\psi(\theta) \to \tilde{\psi}(k)$, the equation (1.2) reads

$$i\partial_t \tilde{\psi}(k, t) = k\nu\tilde{\psi}(k, t) + \tilde{H}\tilde{\psi}(k, t) \tag{1.3}$$

where $\nu = 2\pi/T$ and $\tilde{H}$ acts by convolution with $\tilde{H}(k)$. The generator of (1.3), acting on the extended space, is given by

$$G := \nu N + D + V. \tag{1.4}$$

where N is multiplication by k, D is the part of $\tilde{H}$ that acts diagonally in the k-coordinate, by multiplication with

$$\tilde{H}(0) = \int_0^T dt\, H(t)$$

and V is the remaining part, given by

$$V\psi(k) = \sum_{k' \neq k} \tilde{H}(k - k')\psi(k').$$

If $\|V\|, \|D\| \ll \nu$, then spectral perturbation theory would apply and we could construct spectral subspaces for G, corresponding to different k, as those are separated by large gaps ν. In particular, this would lead to the existence of a $\hat{D}$ satisfying (1.1). In a generic many-body system, D, V have norm proportional to the volume and spectral perturbation theory should not be expected to apply. Before continuing this line of thought, we first present the time-independent setup (closed systems), since there the starting point is similar to the setup we arrived at in (1.4)

1.2.2. Closed systems. The strategy for closed systems is to find a static Hamiltonian that has the same structure and phenomenology as the extended operator G defined above. Its important features are that there is a term νN with large spectral gaps. However, this is not so easy to achieve in a many body system. If we take $N = \sum_x N_x$ a sum of strictly local and commuting terms, then it is not sufficient that νN_x have large spectral gaps, compared to the strength of other local terms. An obvious possibility is however when the spectrum of all of the νN_x is a subset of $\nu\mathbb{Z}$ and we will exploit this possibility. The Fermi–Hubbard model is a natural example

$$G = J \sum_{x \sim y, \sigma} c_{x,\sigma} c^*_{y,\sigma} + U \sum_x n_{x,\downarrow} n_{x,\uparrow} =: H + UN,$$

with interaction U much larger than J, and U playing the role of ν above. To allow for a unified discussion, we will hence simply write ν instead of U. Note that the spectrum of $N_x = n_{x,\downarrow} n_{x,\uparrow}$ here is 1 (when there are two fermions at x, this is a 'doublon') and 0 (otherwise). Our theory expresses rigorously that doublons and singlons (sites with a single fermion) do not interact up to quasi-exponentially long times in ν/J, see also [9,11,35]. We refer to Sect. 3.4 for more discussion.

1.2.3. Non-convergent perturbation theory. In both of the above setups, the point is that the Hamiltonian $G = \nu N + H$ is dominated by νN, where N is a sum of local terms, and it has spectral gaps that remain open in the thermodynamic limit. As already remarked, it is certainly not true that spectral perturbation theory applies as such, as the norm of the perturbation H grows with volume whereas the gaps do not. However, what is true is that matrix elements of the perturbation are smaller than the gaps of νN. To substantiate this, let $|\eta\rangle = \otimes_x |\eta_x\rangle$ be the product basis of eigenvectors of N. Here $|\eta_x\rangle$ are a basis of the single-site space $\mathbb{C}^4$ in which N_x is diagonal, and we write $N(\eta) \equiv \langle \eta | N | \eta \rangle$. Then, the point is that, whenever $N(\eta) \neq N(\eta')$, then

$$|\langle \eta' | H | \eta \rangle| \leq O(1), \qquad \text{i.e. not growing with volume.} \tag{1.5}$$

This means that

$$\frac{\langle \eta' | H | \eta \rangle}{|\nu N(\eta') - \nu N(\eta)|} = \mathcal{O}(1/\nu) \ll 1 \tag{1.6}$$

That is, perturbation theory seems applicable in first order, and the same analysis can be repeated at higher orders. Even though this is really a weaker property than convergent perturbation theory, it does have a physical consequence, namely that the transitions caused by the perturbation take place at a rate slower than $1/\nu$. As far as we know, the easiest way to see this is by performing a unitary transformation that eliminates the perturbation in first order, e.g. to perform a KAM step, see [21] for a celebrated recent application and an account of the technique. In our case, we can proceed to eliminate further orders and it is clear that the rate of transitions is actually smaller than any power of $1/\nu$. The reason why we cannot proceed *ad infinitum* (and indeed, do not believe it would yield a physically correct result) is that in higher orders the operator H in (1.5) gets replaced with an operator whose local terms grow in range and norm as the order increases. We basically show that only at n_*'th order, with $n_* = \nu$, the condition (1.6) gets violated. Physically, this means that at that order the perturbation connects different subspaces of N resonantly and hence the rate of transition is indeed roughly given by $(1/\nu)^{n_*}$, the strength of the perturbation at that order. For a rather intuitive calculation we refer to [2], where we prove that the linear response heating rate is bounded by $e^{-c\nu}$. The present paper is not restricted to linear response and the final bound on the 'heating rate' is rather $e^{-c\frac{\nu}{\log^3 \nu}}$ due to combinatorial factors. However, in dimension 1, we can obtain $e^{-c\nu}$, see [1], at the cost of restricting H to be a sum of strictly local terms.

1.3. Previous results. Among our models, the most tractable is certainly the case of non-interacting fermions (strictly speaking, fermionic lattice systems are not covered here, but this could easily be remedied). It is therefore surprising that the phenomenon of 'localization in energy' has barely been rigorously studied in the absence of interaction, i.e., for the one-particle case. In [12], localization was proved for periodically kicked operators. The authors of [37] considered the disordered Anderson model with a local time-periodic perturbation and proved the stability of localization. In [6], the same is achieved for a quasi-periodic perturbation, which can be viewed as the case of multiple frequency dimensions. Very recently, [15] considered the problem from a very similar point of view as in the present paper: [15] proves stability of Anderson localization with a periodic driving term, provided that the driving frequency is not too small. High-frequency asymptotics (instead of strict localization) in periodically driven quantum systems have been investigated in [16,34,39] by techniques similar to ours, but are not applicable to the many-body problem.

We move now to interacting many-body systems. Some recent works proved bounds for linear response heating rate at high frequency, see [2,7]. Moreover, when preparing the first version of this manuscript, a result very similar to the present paper appeared: [24,30] construct an effective Hamiltonian by truncating the Magnus expansion and they prove that it generates the local dynamics up to exponentially long times. In the case where the driving is localized (as opposed to 'a sum of local terms'), [24] contains a much stronger result, namely that the one-cycle unitary of the process is close in operator norm to the unitary generated by an effective Hamiltonian. The main difference between [24,30] and the present paper, apart from the technique, is that our results are valid for arbitrary dimension.

Looking more broadly, we note that there is a large literature on reducibility and quasi-reducibility of Hamiltonian systems, see e.g. [17]. This is in flavour and methodology of course very similar to our work, but the focus is different: reducibility is trivial in the quantum setup as the system is linear, and the only issue is with the locality of the reduced equation, cf. the discussion in the introduction. Finally, our results on closed (time-independent) systems are in a certain sense special cases of a Nekoroshev bound for many-body systems, see [10,14] and references therein.

2. Results for Driven Systems

2.1. Setup. We start from a time-dependent Hermitian Hamiltonian $H(t) = H^*(t)$, acting on $\mathcal{H}_\Lambda = (\mathbb{C}^q)^{\otimes \Lambda}$ with Λ a finite subset of the lattice $\mathbb{Z}^d$, and $q < \infty$.

We are hence dealing with bounded operators throughout, but we will always state bounds that are uniform in the volume Λ. The time-dependence is periodic:

$$H(t) = H(t + T),$$

with period T and the mapping $t \mapsto H(t)$ is measurable and bounded. We split

$$H(t) = D + V(t),$$

by setting

$$D := \langle H \rangle := (1/T) \int_0^T H(t), \qquad V(t) = H(t) - D.$$

The dynamics is given by $U(t)$, the unitary family generated by $H(t)$:

$$U(t) = -\mathrm{i} \int_0^t \mathrm{d}t' H(t') U(t').$$

which can indeed be solved for measurable, bounded $t \mapsto H(t)$. Since our assumptions do not guarantee that one can construct an everywhere differentiable $U(t)$, we have not defined $U(t)$ as the solution of $\partial_t U(t) = -\mathrm{i}H(t)U(t)$, but this will not play any role in what follows.

2.2. Locality and potentials. We need some standard notion of locality. Let $\mathcal{B}_\Lambda$ be the algebra of bounded operators acting on $\mathcal{H}_\Lambda$, equipped with the standard norm $\|O\| = \sup_{\psi \in \mathcal{H}_\Lambda, \|\psi\|=1} \|O\psi\|$. We denote by $\mathcal{B}_S \subset \mathcal{B}_\Lambda$, with $S \subset \Lambda$, the subalgebra of operators of the form $1_{\Lambda \setminus S} \otimes O_S$, which is canonically identified with the operators acting on $(\mathbb{C}^n)^{\otimes S}$: we say that $1_{\Lambda \setminus S} \otimes O_S$ 'acts within S' and in an abuse of notation, also refer to it as O_S. For any operator Z we can decompose (in a nonunique way) $Z = \sum_{S \in \mathcal{P}_c(\Lambda)} Z_S$, where $Z_S \in \mathcal{B}_S$ and $\mathcal{P}_c(\Lambda)$ denotes the set of finite, connected (by adjacency) subsets of Λ. The collection (Z_S) is usually referred to as an (interaction) *potential*, see e.g. [31,36] and in fact, all of our results will be about such potentials. However, to keep the notation simple, we prefer not to make this distinction explicit and to keep the notation Z for operator and potential. In principle, this creates an ambiguity as it is not clear which potential is meant (nonuniqueness of decomposition). In practice, all potentials will be derived in a straightforward way from the potential $(H_S(t))$ of $H(t)$, which is an input to our work. For example, the above definitions specify V, D by linear operations on H, and it is understood that the potentials are given as, for example, $D_S := (1/T) \int_0^T H_S(t)$ and $V_S := H_S - D_S$. We will also define below the potential $e^{\mathrm{ad}_{A_{n_*}}} \ldots e^{\mathrm{ad}_{A_0}} Z$ for potentials A_j, Z. To make this explicit, we expand all exponentials in a power series (recall that all our operators are bounded) and we define the 'commutator of potentials Q, Z' as

$$(\mathrm{ad}_Q(Z))_S := \sum_{S_1, S_2 : S_1 \cup S_2 = S} [Q_{S_1}, Z_{S_2}]$$

where we note that only $S_1 \cap S_2 \neq \emptyset$ contribute. This defines inductively $e^{\mathrm{ad}_Q} Z$.

We define a family of norms on (time-dependent) potentials, parametrized by a spatial decay rate κ

$$\|Z\|_\kappa := \sup_{x \in \Lambda} \sum_{S \in \mathcal{P}_c(\Lambda) : S \ni x} e^{\kappa |S|} \sup_t \|Z_S(t)\|, \quad \kappa > 0$$

Note that these norms are tailor-made for operators (potentials) that are sums of local terms that themselves are independent of the global volume Λ, in particular for many-body Hamiltonians. In principle, one could take Λ infinite, i.e. $\Lambda = \mathbb{Z}^d$. In that case Z still makes sense as a potential (but not as an operator) and $e^{\mathrm{ad}_{A_j}}$, in case A_j is anti-Hermitian, still makes sense as an automorphism of the C^*-algebra of quasilocal operators (but e^{A_j} does not make sense as a unitary). We will not adopt this point of view and we prefer to have a finite Λ throughout, so that also the operators remain well-defined.

2.3. General results. We will always assume that the frequency ν is large compared to some local energy scales, namely that there is a decay rate $\kappa_0 > 0$ such that

$$\nu \geq \frac{9\pi \|V\|_{\kappa_0}}{\kappa_0}, \quad n_* \geq 1, \tag{2.1}$$

where

$$n_* := \left\lfloor \frac{\nu/\nu_0}{(1 + \ln \nu/\nu_0)^3} \right\rfloor - 2, \quad \text{with } \nu_0 := \frac{54\pi}{\kappa_0^2} \left(\|D\|_{\kappa_0} + 2\|V\|_{\kappa_0} \right).$$

In the theorem below (and further in the text), C refer to numerical constants that can be chosen independent of all model parameters. Most importantly, they are independent of the volume Λ. By K we denote numbers that can depend on all model parameter, but not on the frequency ν and the volume Λ.

Theorem 2.1. *Assume that* (2.1) *holds, then there are Hermitian operators (potentials) $\widehat{D}$, $\widehat{V}(t)$ and a unitary $Y(t)$ such that the unitary*

$$\widehat{U}(t) := Y(t)U(t)$$

solves

$$\widehat{U}(t) = -\mathrm{i}\int_0^t \mathrm{d}t'\,(\widehat{D} + \widehat{V}(t'))\widehat{U}(t'),$$

and the following are satisfied:

1. *$\widehat{D}$ is time-independent and $Y(t)$, $\widehat{V}(t)$ are T-periodic. $Y(t) = 1$ for stroboscopic times $t \in T\mathbb{N}$.*
2. *Set the decay rate $\kappa_{n_*} := \kappa_0(1 + \log(n_* + 1))^{-1}$, then*

$$\|\widehat{D} - D\|_{\kappa_{n_*}} \le C(\nu_0/\nu), \qquad \|\widehat{V}\|_{\kappa_{n_*}} \le C(2/3)^{n_*}\|V\|_{\kappa_0}. \tag{2.2}$$

3. *The unitaries $Y(t)$ are defined by $Y(t) = \mathrm{e}^{A_{n_*}(t)}\ldots\mathrm{e}^{A_0(t)}$ with operators $A_j(t)$ to be specified later. They are close to identity and quasilocal in the sense that*

$$\|Y(t)ZY^*(t) - Z\|_{\kappa_{n_*}} \le C(\nu_0/\nu)\|Z\|_{\kappa_0}, \qquad \text{for any operator (potential) } Z, \text{ for all } t \in \mathbb{R}.$$

(we write $Y(t)ZY^*(t) = \mathrm{e}^{\mathrm{ad}_{A_{n_*}}}\ldots\mathrm{e}^{\mathrm{ad}_{A_0}}Z$ to interpret it as a potential, see Sect. 2.2*)*

For a Z that is local, and not merely a sum of local terms, the last bound of the theorem above is obviously useless; in this case, it actually follows easily from our construction that $Y(t)ZY^*(t) - Z$ has small operator norm. Note that the times $t \in T\mathbb{Z}$ are, artificially, singled out to play a distinguished role, namely $Y(T\mathbb{Z}) = 1$. A more intuitive way to phrase this result is by going to a periodically rotating frame, given by the unitary $Y(t)$. If $\psi(t)$ solves $\partial_t\psi(t) = -\mathrm{i}H(t)\psi(t)$, then $\widehat{\psi}(t) = Y(t)\psi(t)$ solves $\widehat{\psi}(t) = -\mathrm{i}\widehat{H}(t)\widehat{\psi}(t)$ with

$$\widehat{H}(t) = Y(t)H(t)Y^*(t) - \mathrm{i}Y(t)\partial_t Y^*(t).$$

Theorem 2.1 asserts then that the Hamiltonian $\widehat{H}(t)$ generating this evolution is almost time-independent for large ν, since the time-dependent part $\widehat{V}$ has local terms of order $(2/3)^{\frac{\nu/\nu_0}{(1+\ln(\nu/\nu_0))^3}}$.

2.4. Physical consequences.

As explained above, our technical result suggests that the evolution is well-described by an effective Hamiltonian, at least for stroboscopic times. This implies that the effective Hamiltonian $\widehat{D}$ is a quasi-conserved quantity, see (2.3) in the upcoming theorem. Since D is close to $\widehat{D}$, this also implies that D itself is well-conserved, (2.4). In other words, apart from a quantity of order $1/\nu$, the energy density grows very slowly.

Theorem 2.2 (Slow heating).

$$\frac{1}{|\Lambda|} \|U^*(t)\hat{D}U(t) - \hat{D}\| \leq t\, K_0\, (2/3)^{n_*} \quad \text{for } t \in T\mathbb{N}, \tag{2.3}$$

$$\frac{1}{|\Lambda|} \|U^*(t)DU(t) - D\| \leq t\, K_0\, (2/3)^{n_*} + C(\nu_0/\nu), \quad \text{for any } t \geq 0 \tag{2.4}$$

with $K_0 = C\|D\|_{\kappa_0}\|V\|_{\kappa_0}$ and n_* as in Theorem 2.1.

Note that in this theorem, we use the standard operator norm $\|\cdot\|$, but we divide by volume. Next, we state how the evolution of local operators is approximated by the evolution with the time-independent Hamiltonian $\hat{D}$. For a local observable O, the difference

$$\hat{U}^*(t)O\hat{U}(t) - e^{it\hat{D}}Oe^{-it\hat{D}}$$

grows very slowly as it is due to the term $\hat{V}$. However, to make this precise, we need to control spatial spreading of $e^{it\hat{D}}Oe^{-it\hat{D}}$, which is done by invoking a Lieb-Robinson bound. Since $\hat{U}(t) = U(t)$ for stroboscopic times, we get then

Theorem 2.3 (Approximation of local observables). *For any rate* $0 < r < \ln(3/2)$, *there are numbers* $K(O)$, $K'(O) < \infty$, *depending on model parameters and the observable* O, *but not* ν, *such that*

$$\|U^*(t)OU(t) - e^{it\hat{D}}Oe^{-it\hat{D}}\| \leq K(O)e^{-rn_*}(t + K'(O))^{d+1}, \quad \text{for } t \in T\mathbb{N}.$$

This theorem is for us the most clear expression of the fact that $\hat{D}$ really describes the dynamics for very long times. We do not see how to improve the dependence of the bound on t, unless one would manage to replace the Lieb-Robinson bound by a diffusive bound. Phrased in a different way, this result says that the dynamics generated by $\hat{D}$ is close to the actual dynamics up to a time that grows quasi-exponentially in ν. Theorem 2.3 is actually a particular case of a more general statement valid for any $t \in \mathbb{R}$, provided that $e^{it\hat{D}}Oe^{-it\hat{D}}$ is replaced by $e^{it\hat{D}}Y(t)OY^*(t)e^{-it\hat{D}}$.

3. Results for Time-Independent Systems

3.1. Setup. We turn to a time-independent setup.

Let us have a family of 'number operators' N_x acting on site x. By number operators, we simply mean that $\sigma(N_x) \subset \mathbb{Z}$, for every x, with $\sigma(\cdot)$ the spectrum. Set

$$N := \sum_x N_x$$

The idea is that a multiple of this operator dominates the Hamiltonian of our system; the total Hamiltonian is

$$G := \nu N + H$$

with ν large compared to the local energy scales of H. We decompose $H = D + V$ with

$$D := \langle H \rangle = \sum_{n \in \mathbb{Z}} \chi(N = n)H\chi(N = n), \quad V = H - \langle H \rangle$$

such that $[N, D] = 0$. The choice of not simply calling the total Hamiltonian H is to make the analogy with the time-dependent case maximal, as will be clear later. Relevant examples of this setup will be discussed below.

3.2. General result. To state the result, we now exploit the notational similarity with the time-dependent case, which allows for most definitions and formulas to be identical in both cases. The norms $\|\cdot\|_\kappa$ are defined as before, except that the supremum over time is now omitted (there is no time-dependence).

We assume that the parameter ν is large compared to some local energy scales, namely that there is a decay rate $\kappa_0 > 0$ such that

$$\nu \geq \frac{9\pi \|V\|_{\kappa_0}}{\kappa_0}, \qquad n_* \geq 1, \tag{3.1}$$

where

$$n_* := \left\lfloor \frac{\nu/\nu_0}{(1 + \ln \nu/\nu_0)^3} \right\rfloor - 2, \qquad \text{with } \nu_0 := \frac{54\pi}{\kappa_0^2} \left(\|D\|_{\kappa_0} + 2\|V\|_{\kappa_0} \right).$$

In the theorem below (and further in the text), C refer to numerical constants that can be chosen independent of all model parameters. Most importantly, they are independent of the volume Λ.

Theorem 3.1. *Assume that* (3.1) *holds, then there are Hermitian operators (potentials)* $\widehat{H}, \widehat{D}, \widehat{V}$ *and a unitary* Y, *such that*

$$Y(\nu N + H)Y^* = \nu N + \widehat{H} = \nu N + \widehat{D} + \widehat{V}$$

with

1. $\widehat{D} = \langle \widehat{H} \rangle$, *i.e.* $[\widehat{D}, N] = 0$.
2. Set the decay rate $\kappa_{n_*} := \kappa_0 (1 + \log(n_* + 1))^{-1}$, *then*

$$\|\widehat{D} - D\|_{\kappa_{n_*}} \leq C(\nu_0/\nu), \qquad \|\widehat{V}\|_{\kappa_{n_*}} \leq (2/3)^{n_*} \|V\|_{\kappa_0}. \tag{3.2}$$

3. The unitaries Y *are close to the identity and quasilocal in the sense that, for any operator (potential)* Z,

$$\|YZY^* - Z\|_{\kappa_{n_*}} \leq C(\nu_0/\nu)\|Z\|_{\kappa_0}.$$

3.3. Physical consequences. Thanks to the above theorem, we identify two extensive quantities that stay almost conserved for quasi-exponentially long times in ν. Let

$$\mathcal{D} = Y^*\widehat{D}Y, \qquad \mathcal{N} = Y^*NY.$$

The following result derives from Theorem 3.1 in the same way Theorem 2.2 follows from Theorem 2.1.

Theorem 3.2. *For any* $t \geq 0$,

$$\frac{1}{|\Lambda|} \|U^*(t)\mathcal{D}U(t) - \mathcal{D}\| \leq t K_0 (2/3)^{n_*}, \tag{3.3}$$

$$\frac{1}{|\Lambda|} \|U^*(t)\mathcal{N}U(t) - \mathcal{N}\| \leq t K_0'(2/3)^{n_*}, \tag{3.4}$$

with $K_0 = C\|D\|_{\kappa_0}\|V\|_{\kappa_0}$, $K_0' = C(\sup_x \|N_x\|)\|V\|_{\kappa_0}$, *and* n_* *as in Theorem 3.1.*

Let us notice that

$$N = \frac{1}{\nu}G - \frac{1}{\nu}H \tag{3.5}$$

so N simply equals a conserved quantity plus a term of norm proportional to $|\Lambda|/\nu$; hence N is always conserved up to an error of order $|\Lambda|/\nu$. Our second theorem states that the evolution of local observables is well-described by the effective Hamiltonian $\nu N + \widehat{D}$, modulo an error of order $1/\nu$ and up to quasi-exponentially large times.

Theorem 3.3. *For any $0 < r_1 < \frac{1}{d+1}\ln(3/2)$, and local operator O, there is a $K_3(O)$ such that*

$$\|U^*(t)OU(t) - e^{it(\nu N + \widehat{D})}Oe^{-it(\nu N + \widehat{D})}\| \leq \frac{1}{\nu}K_3(O), \quad \text{for } t \leq e^{r_1 n_*}.$$

To prove this theorem, we start from

$$YU^*(t)Y^*OYU(t)Y^* = e^{it(\nu N + \widehat{H})}Oe^{-it(\nu N + \widehat{H})}$$

en then use Lieb-Robinson bounds to replace the right hand side by $e^{it(\nu N + \widehat{D})}Oe^{-it(\nu N + \widehat{D})}$, up to an error that remains small for $t \leq e^{r_1 n_*}$. This is analogous to the proof of Theorem 2.3. Then, one gets rid of the Y, Y^* by using that these unitaries are locally close to identity, in particular, $\|Y^*OY - O\| \leq \frac{1}{\nu}K(O)$. Such statements are similar to Statement 3) of Theorem 3.1 and their proof is an obvious variation.

3.4. Examples.

3.4.1. Fermi–Hubbard chain.
We consider the Fermi–Hubbard Hamiltonian, for $J, U \in \mathbb{R}$,

$$G = J\sum_{x \sim y, \sigma} c_{x,\sigma}c_{y,\sigma}^* + U\sum_x n_{x,\downarrow}n_{x,\uparrow}$$

where $\sigma \in \{\uparrow, \downarrow\}$, $c_{x,\sigma}, c_{x,\sigma}^*$ are fermionic field operators, i.e.

$$\{c_{x,\sigma}, c_{x',\sigma'}^*\} = \delta_{x,x'}\delta_{\sigma,\sigma'}, \quad \{c_{x,\sigma}, c_{x',\sigma'}\} = 0, \quad \{c_{x,\sigma}^*, c_{x',\sigma'}^*\} = 0$$

and $n_{x,\sigma} = c_{x,\sigma}^*c_{x,\sigma}$ and $n_x = \sum_\sigma n_{x,\sigma}$. Apart from the energy, there are two conserved quantities, namely $n_\sigma = \sum_x n_{x,\sigma}$. We now assume that $U \gg J$. We set $\nu \equiv U$ and

$$N_x := n_{x,\downarrow}n_{x,\uparrow} = \chi(n_x = 2), \quad N = \sum_x N_x,$$

i.e. N is the number of doublons in the system. Now we find D as the part of H that commutes with N:

$$D = T_s + T_d := J\sum_{x \sim y, \sigma} c_{x,\sigma}c_{y,\sigma}^* \chi(n_x + n_y = 1) + J\sum_{x \sim y, \sigma} c_{x,\sigma}c_{y,\sigma}^* \chi(n_x + n_y = 3)$$

where T_s, T_d stand for the 'singlon' and 'doublon' kinetic energies, respectively. In fact, both T_s, T_d commute with N. The total Hamiltonian is hence $\nu N + T_s + T_d + V$, in accordance with the general abstract setup.

As observed in (3.5), the density of doublons, $\frac{N}{|\Lambda|}$, is conserved up to a quantity of order $\mathcal{O}(1/\nu)$ by energy conservation. According to Theorem 3.2, we conclude that $\mathcal{N}$, a dressed version of the number of doublons N, as well as $D = T_s + T_d$ are extensive quasi-conserved quantities, up to an (almost) exponentially small quantity for (almost) exponentially long times in ν. This remarkable feature shows thus the appearance of a long pre-thermal regime in the Fermi–Hubbard chain in the regime where $J/U \ll 1$.

3.4.2. XYZ chain with large magnetic field. We consider the spin s-chain with Hamiltonian

$$G = \sum_x J_1 S_x^1 S_{x+1}^1 + J_2 S_x^2 S_{x+1}^2 + J_3 S_x^3 S_{x+1}^3 + h \sum_x S_x^3$$

with S^α the spin-s representation of $SU(2)$ acting on $\mathbb{C}^{2s+1}$ and S_x^α copies thereof on site x. We choose $h \gg J_\alpha$ and we set $\nu \equiv h$ and

$$N_x := S_x^3, \qquad \text{i.e. the magnetization}$$

The operator D is then given by

$$D = \sum_x J(S_x^1 S_{x+1}^1 + S_x^2 S_{x+1}^2) + J_3 S_x^3 S_{x+1}^3, \qquad 2J = J_1 + J_2$$

which indeed commutes with the magnetization N. So we see here that there is an emergent $U(1)$ symmetry in D, corresponding to the conservation of N.

4. Renormalization of Hamiltonians

4.1. Recursion formulae. We now describe our main scheme to transform the Hamiltonian. It is inspired by [3,5,21]. Let us rename the operators $H(t)$, D, $V(t)$ as $H_0(t)$, D_0, $V_0(t)$ and we will now construct $H_n(t)$, D_n, $V_n(t)$, with n up to n_*. We mostly drop the t in the notation and it is simply understood that D_n are time-independent, whereas other operators are T-periodic. At each scale we define D_n, V_n from H_n by setting

$$D_n := \langle H_n \rangle := (1/T) \int_0^T H_n(t), \qquad V_n = H_n - D_n.$$

H_{n+1} is constructed out of H_n by

$$H_{n+1} := e^{-A_n} H_n e^{A_n} - i e^{-A_n} \partial_t e^{A_n} \tag{4.1}$$

where A_n is determined by

$$V_n - i \partial_t A_n = 0, \qquad A_n(t = 0) = 0 \tag{4.2}$$

and we note that A_n is indeed T-periodic because $\langle V_n \rangle = 0$. Note that though we have not demanded $t \mapsto V(t)$ to be smooth, but simply bounded and measurable, the use of derivatives above is justified. Indeed, the integral of a bounded, measurable function f is differentiable almost surely (because it has bounded variation on intervals) and the derivative equals f almost surely.

We have now defined all operators H_n. To appreciate why such a procedure is useful, i.e. why the V_n decrease with n, we unwrap the recursion relation a bit. We define the transformations (O is an arbitrary operator)

$$\gamma_n(O) := e^{-A_n} O e^{A_n} = e^{-\operatorname{ad}_{A_n}} O.$$

and

$$\alpha_n(O) := \int_0^1 ds\, e^{-sA_n} O e^{sA_n} = \int_0^1 ds\, e^{-s\operatorname{ad}_{A_n}} O.$$

The latter involves a dummy time s that has nothing to do with the cycle time t; the transformation α_k is defined pointwisely for any t in the cycle. The use of α_n is that

$$e^{-A_n} \partial_t e^{A_n} = \alpha_n(\partial_t A_n)$$

as one easily checks by an explicit calculation. If $A_n(t)$ for different t would commute among themselves, then we would simply find back the familiar expression

$$e^{-A_n} \partial_t e^{A_n} = \partial_t A_n, \qquad \text{(wrong in general)}$$

Recasting (4.1) with the help of the above notation, we get

$$H_{n+1} = \gamma_n(H_n) - i\alpha_n(\partial_t A_n) \tag{4.3}$$

$$= \gamma_n(D_n) + (\gamma_n(V_n) - V_n) + (V_n - i\partial_t A_n) - i(\alpha_n(\partial_t A_n) - \partial_t A_n) \tag{4.4}$$

$$= \gamma_n(D_n) + (\gamma_n(V_n) - V_n) - (\alpha_n(V_n) - V_n). \tag{4.5}$$

For later convenience, we remark that, upon splitting $H_{n+1} = D_{n+1} + V_{n+1}$,

$$D_{n+1} = D_n + \langle W_n \rangle, \qquad V_{n+1} = W_n - \langle W_n \rangle, \tag{4.6}$$

with

$$W_n := (\gamma_n(D_n) - D_n) + (\gamma_n(V_n) - V_n) - (\alpha_n(V_n) - V_n). \tag{4.7}$$

This defines the iteration scheme and we now move to bounds.

4.2. Iterative bounds. The relation (4.7) makes it particularly intuitive that W_n and hence V_{n+1} are of higher order in $(\frac{1}{\nu})$ than V_n. Indeed, from (4.2) and periodicity of $V_n(t)$,

$$\|A_n\|_\kappa \le (T/2)\|V_n\|_\kappa, \qquad \text{for any } \kappa > 0$$

and (4.7) contains three exponentials of A_k with zero'th order terms removed, so that keeping only the first order suggests a bound like

$$\|V_{n+1}\|_\kappa \approx (T/2)\|V_n\|_\kappa (\|D_n\|_\kappa + 2\|V_n\|_\kappa). \tag{4.8}$$

To make this precise, we will allow the decay rate κ on the left-hand side to be slightly smaller than that on the right hand side. We consider a family of norms by fixing a strictly decreasing sequence of decay rates $\kappa_n > 0, n \ge 1$. Eventually we will choose $\kappa_n = (1 + \log(n+1))^{-1}\kappa_0$ with κ_0 chosen appropriately for the initial Hamiltonians (cfr. statement of Theorem 2.1), but for the time being it is convenient to keep the sequence κ_n general. We will abbreviate $\|\cdot\|_{\kappa(n)}$ by $\|\cdot\|_n$. The following lemma is our prime (and only) tool:

Lemma 4.1. *Let Z, Q be potentials and assume that $3\|Q\|_n \le \delta\kappa_n := \kappa_n - \kappa_{n+1}$. Then*

$$\|e^Q Z e^{-Q} - Z\|_{n+1} \le \frac{18}{\delta\kappa_n\kappa_{n+1}}\|Z\|_n\|Q\|_n.$$

Since $\|Z\|_{n+1} \le \|Z\|_n$, we also get

$$\|e^Q Z e^{-Q}\|_{n+1} \le \left(1 + \frac{18}{\delta\kappa_n\kappa_{n+1}}\|Q\|_n\right)\|Z\|_n.$$

We postpone the combinatorial proof, relying on cluster expansions, to Sect. 5.1.

4.3. Bounds on transformed potentials. We set

$$v(n) := \|V_n\|_n, \qquad d(n) := \|D_n\|_n, \qquad \delta d(n) := \|D_{n+1} - D_n\|_{n+1}.$$

We do not need to introduce any shorthand for $\|A_n\|_n$ since $\|A_n\|_n \le (T/2)v(n)$, as follows from (4.2). From (4.7) to Lemma 4.1, we then get

$$\|W_n\|_{n+1} \le \frac{T}{2}m(n)v(n)\left[d(n) + 2v(n)\right], \qquad m(n) := \frac{18}{\delta\kappa_n\kappa_{n+1}}$$

provided that $(3T/2)v(n) \le \delta\kappa_n$. In that case, (4.6) yields

$$2\delta d(n), v(n+1) \le Tm(n)v(n)\left[d(n) + 2v(n)\right]$$

To get recursive bounds, it is handy to demand that

$$Tm(n)\left[d(n) + 2v(n)\right] < 2/3, \tag{4.9}$$

because then

$$v(n+1), \delta d(n) \le (2/3)v(n) \tag{4.10}$$

and hence

$$d(n+1) + 2v(n+1) \le d(n) + 2v(n) \tag{4.11}$$

which makes it easy to check the validity of (4.9) at the next order. Indeed, we see that if

$$(3T/2)m(j)\left[d(0) + 2v(0)\right] \le 1, \qquad (3T/2)v(j) \le \delta\kappa_j, \qquad \text{for any } 0 \le j \le n \tag{4.12}$$

then

$$\delta d(n), v(n+1) \le v(0)(2/3)^{n+1}.$$

Let us now use the specific choice $\kappa_n := \frac{\kappa_0}{1+\ln(n+1)}$, then the second condition of (4.12) is always satisfied provided that

$$(3T/2)C_0 v(0) \le \kappa_0, \qquad \text{with } C_0 = \sup_{n\ge 0}(2/3)^n(n+2)\ln(2+n) \le 3$$

Thus we find that the conditions

$$9\pi v(0)\kappa_0^{-1} \le \nu, \qquad \nu_0 \le \nu, \qquad \text{with } \nu_0 := \frac{54\pi}{\kappa_0^2}\left[d(0) + 2v(0)\right]$$

allow us to start the procedure and we can continue, i.e. (4.12) remains satisfied, i.e. at least as long as $(1 + \ln(2+n))^3(n+2) \le \nu/\nu_0$, hence at least up to $n = n_*$ with

$$n_* = \left\lfloor \frac{\nu/\nu_0}{(1 + \ln\nu/\nu_0)^3} \right\rfloor - 2.$$

4.4. Proof of Theorem 2.1. The above calculation completes an important part of the proof of Theorem 2.1, namely the bound on $\|V_n\|_n$ at $n = n_*$. The bound on $\widehat{D}$ follows by summing the $d(n)$. Let us now derive the stated bounds on Y. By repeated application of Lemma 4.1 and the bounds derived above, we have the first inequality in (for some $c > 0$)

$$\|e^{\mathrm{ad}_{A_n}} \ldots e^{\mathrm{ad}_{A_0}} Z\|_{n+1} \leq \|Z\|_0 \prod_{j=0}^{n} (1 + C(\nu_0/\nu)e^{-cj}) \leq C\|Z\|_0 \qquad (4.13)$$

The second inequality follows because $(\nu_0/\nu) \leq 1$ (and redefining C). Let

$$E_{n+1} := e^{\mathrm{ad}_{A_n}} \ldots e^{\mathrm{ad}_{A_0}} Z - Z,$$

so that

$$E_{n+1} - E_n = e^{\mathrm{ad}_{A_n}} (e^{\mathrm{ad}_{A_{n-1}}} \ldots e^{\mathrm{ad}_{A_0}} Z) - (e^{\mathrm{ad}_{A_{n-1}}} \ldots e^{\mathrm{ad}_{A_0}} Z)$$

Then we use again Lemma 4.1 and (4.13) (together with the bounds leading to (4.13)) to get

$$\|E_{n+1} - E_n\|_{n+1} \leq Ce^{-cn}(\nu_0/\nu)\|Z\|_0$$

Item 3) of Theorem 2.1 follows now by

$$\|E_{n_*+1}\|_{n_*+1} \leq \sum_{j=0}^{n_*} \|E_{j+1} - E_j\|_{j+1} \leq C(\nu_0/\nu)\|Z\|_0.$$

5. Proofs

5.1. Proof of Lemma 4.1. As in the previous section, we drop the dependence on t from the notation. Starting from (5.1), one checks that we get indeed the correct bounds on potentials when reinstating (the supremum over) t.

We assume that $Q \neq 0$, $Z \neq 0$, else the claim is trivial. All sets $S \subset \Lambda$ that appear below are assumed to be connected. We expand the exponential

$$e^{\mathrm{ad}_Q}(Z_{S_0}) = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{S_1,\ldots,S_m} \mathrm{ad}_{Q_{S_m}} \ldots \mathrm{ad}_{Q_{S_2}} \mathrm{ad}_{Q_{S_1}} Z_{S_0}.$$

The integrand vanishes unless S_j has nonempty overlap with $\cup_{i=0}^{j-1} S_i$. Localising and taking norms, we get

$$\|(e^{\mathrm{ad}_Q} Z_{S_0} - Z_{S_0})_P\| \leq \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{S_1,\ldots,S_m}^{c,P} \|Z_{S_0}\| \prod_j (2\|Q_{S_j}\|), \qquad (5.1)$$

where $\sum_{\ldots}^{c,P}$ indicates that the family of sets $S_0, \ldots, S_m$ is connected (they cannot be split into two nonempty mutually disjoint collections) and the union is $P = \cup_{j=0,\ldots,m} S_j$. Multiplying with $e^{\kappa_{n+1}|P|}$ and summing over S_0, we get

$$e^{\kappa_{n+1}|P|} \|(e^{\mathrm{ad}_Q} Z - Z)_P\| \leq \sum_{m=1}^{\infty} \frac{e^{-\kappa(n+1)m}}{m!} (\frac{3}{\delta\kappa_n})^{m+1} \|Z\|_n \|Q\|_n^m) \sum_{S_0,\ldots,S_m}^{c,P} \prod_j \nu(S_j)$$

$$(5.2)$$

where we introduced the shorthand

$$v(S) := (\frac{\delta\kappa_n}{3}) e^{\kappa(n+1)|S|} \left(\frac{\|Z_S\|}{\|Z\|_n} + \frac{2\|Q_S\|}{\|Q\|_n} \right)$$

and we exploited the fact that

$$|P| \leq -m + \sum_{j=0}^{m} |S_j|$$

Next, we use the assumption that $3\|Q\|_n \leq \delta\kappa_n$ together with $e^{-\kappa(n+1)m} \leq \frac{2}{\kappa(n+1)(m+1)}$ to get

$$(5.2) \leq \underbrace{2(3/\delta\kappa_n)^2(\kappa(n+1))^{-1}\|Z\|_n\|Q\|_n)}_{=:L_n} \sum_{m=1}^{\infty} \frac{1}{(m+1)!} \sum_{S_0,...S_m}^{c,P} \prod_{j} v(S_j).$$

Taking the sum over P with $P \ni x$, we get then

$$\|e^{\mathrm{ad}_Q} Z - Z\|_{n+1} \leq L_n \sup_x \sum_{m=1}^{\infty} \frac{1}{(m+1)!} \sum_{S_0,...S_m}^{c} \left(\sum_{j=0}^{m} \chi(x \in S_j) \right) \prod_{j=0}^{m} v(S_j). \quad (5.3)$$

where 'c' in the sum $\sum_{S_0,...,S_m}^{c}$ indicates, as above, that the collection of sets is connected. To perform the sum over connected graphs rooted in S_j, we use some standard combinatorial tools: polymer expansions. Note that by the definition of the $\|\cdot\|_n$, we have

$$\sum_{S:S\cap S'\neq\emptyset} v(S) \exp(\delta\kappa_n|S|) \leq \delta\kappa_n|S'| \qquad \text{for any } S', \quad (5.4)$$

and this assures that one can inductively sum the graphs. A convenient reference is [38]. Theorem 1 in [38] (more precisely, eq. (5) in the proof of said theorem) states that

$$\sum_{n=1}^{\infty} \frac{1}{(m+1)!} \sum_{S_0,...S_m}^{c} \left(\sum_{j=0}^{m} \chi(x \in S_j) \right) \prod_{j=0}^{m} v(S_j) \leq \delta\kappa_n.$$

To translate the relevant result there to our work, we identify $\mathbb{A}$ with the set of connected subsets of $\Lambda \subset \mathbb{Z}^d$, the measure μ on $\mathbb{A}$ with the discrete measure with weights $v(S)$, the function $-\zeta(\cdot,\cdot)$ with the indicator that two sets are non-disjoint and the function $a(S) = \delta\kappa_n|S|$. The criterion (3) in [38] is then precisely (5.4) and we use it for singletons $S' = \{x\}$.

5.2. Proofs of Corollary 2.2. By Theorem 2.1,

$$U^*(t)Y^*(t)\widehat{D}Y(t)U(t) - \widehat{D} = \mathrm{i} \int_0^t \mathrm{d}s\, \widehat{U}^*(s) [\widehat{V}(s), \widehat{D}] \widehat{U}(s),$$

and hence the operator norm of the above expression is bounded by

$$\int_0^t \mathrm{d}s \, \|[\widehat{V}(s), \widehat{D}]\| \le 2t|\Lambda| \|\widehat{D}\|_0 \|\widehat{V}\|_0 \tag{5.5}$$

$$\le Ct|\Lambda| \|D\|_{\kappa_0} \|V\|_{\kappa_0} (2/3)^{n_*}, \tag{5.6}$$

where the first inequality is a straightforward calculation and the second is item 2) of Theorem 2.1. For stroboscopic times $t \in T\mathbb{N}$, for which $Y(t) = 1$, this gives (2.3). To get (2.4), we use (5.6) together with a bound on $\|D - \widehat{D}\|$ following from (2.2), and a bound on $\|D - Y(t)DY^*(t)\|$ following from item (3) of Theorem 2.1.

5.3. Proof of Theorem 2.3. We use the Duhamel formula to write

$$\widehat{U}^*(t)O\widehat{U}(t) - \mathrm{e}^{\mathrm{i}t\widehat{D}}O\mathrm{e}^{-\mathrm{i}t\widehat{D}} = \int_0^t \mathrm{d}s \, W^*(s)[\widehat{V}(s), \mathrm{e}^{\mathrm{i}s\widehat{D}}O\mathrm{e}^{-\mathrm{i}s\widehat{D}}]W(s), \tag{5.7}$$

with $W(s) = \widehat{U}(s)^{-1}\widehat{U}(t)$. Hence, to get a bound on the left-hand side, it suffices to estimate the norm of the commutator. The latter is small by the smallness of $\widehat{V}(s)$, but we also need a Lieb-Robinson bound [27] to avoid getting terms proportional to volume:

Lemma 5.1 (Lieb-Robinson bound). *Let $Z = \sum_S Z_S$ be a Hermitian operator (potential) with $\|Z\|_{2\kappa} < \infty$ for some $\kappa > 0$. Let the operators A, B act within $X, Y \subset \Lambda$, respectively. Then,*

$$\|[A, \mathrm{e}^{\mathrm{i}tZ}B\mathrm{e}^{-\mathrm{i}tZ}]\| \le \|A\|\|B\|\mathrm{e}^{-\kappa(d(X,Y)-vt)} \min(|X|, |Y|)$$

with Lieb-Robinson speed $v = v(Z, \kappa) := C(d)(\kappa^{-(d+2)}\mathrm{e}^{\kappa})\|Z\|_{2\kappa}$ and $C(d)$ only depending on the spatial dimension d.

This is a direct consequence of Theorem 2.1 (in particular, eq. 2.17 following it) in [31], with obvious adaptations to the notation and setup of the present paper (for example; the function F is chosen as $F(r) = \mathrm{e}^{-\kappa r}(1 + r)^{-(d+1)}$).

We start from (5.7) and we choose O supported in a set S_0. By unitarity of $W(s)$, we bound the right-hand side as

$$\sum_S \|[\widehat{V}_S(s), \mathrm{e}^{\mathrm{i}s\widehat{D}^{(r)}}O\mathrm{e}^{-\mathrm{i}s\widehat{D}^{(r)}}]\| \tag{5.8}$$

and we can bound the summands by either the above Lieb-Robinson bound, or a trivial norm bound, yielding

$$|S|\mathrm{e}^{-\kappa(d(S,S_0)-vs)}\|\widehat{V}_S(s)\|\|O\|, \quad \text{or} \quad 2\|\widehat{V}_S(s)\|\|O\|. \tag{5.9}$$

with the LR velocity $v = v(\widehat{D}, \kappa)$, see theorem above. We estimate (5.8) by

$$C(1 + \tfrac{1}{\kappa})|S_0|\|\widehat{V}\|_{\kappa}\|O\| \int_0^t \mathrm{d}s \, \xi^d, \quad \xi := vs + |S_0| \tag{5.10}$$

using the first and second bound of (5.9) for S such that $d(S, S_0) - vs \le 0$ and $d(S, S_0) - vs > 0$, respectively. Now we choose $2\kappa = \kappa_{n_*}$, and the bound on $\|\widehat{V}\|_{\kappa_{n_*}}$ from Theorem 2.1, to bound (5.10) by $K(O)\mathrm{e}^{-kn_*}(t + K'(O))^{d+1}$, with $K(O), K'(O)$ depending on our model parameters and O, but not on v and t.

Taking then $t \in T\mathbb{N}$ so that $\widehat{U}(t) = U(t)$, we get the theorem from (5.7).

5.4. Proofs for the time-independent setup. Let us see that the algebra in the time-independent setup is identical to the one in the time-dependent problem, so that the proof of the theorems can be taken over in a straightforward way.

$$G = \nu N + H = \nu N + D + V, \quad \text{where} \quad [D, N] = 0$$

with ν large compared to the local energy scales of H. The goal is to transform this Hamiltonian into

$$\nu N + \widehat{H} = \nu N + \widehat{D} + \widehat{V}$$

with $\widehat{V}$ quasi-exponentially small in ν and $\widehat{D}$ commuting with N. As in the time-dependent case, let us rename H, D, V as H_0, D_0, V_0 and we now construct inductively the sequence of operators H_n, D_n, V_n. At each step, the splitting $H_n = D_n + V_n$ is defined by

$$D_n = \langle H_n \rangle, \qquad V_n = H_n - D_n,$$

with

$$\langle O \rangle = (1/T) \int_0^T dt\, e^{it\nu N} O e^{-it\nu N}, \qquad T := 2\pi/\nu.$$

Indeed, this operation eliminates any off-diagonal elements in N-basis, hence we have $[\langle O \rangle, N] = 0$. The renormalization step defining H_{n+1} is

$$e^{-A_n}(\nu N + H_n)e^{A_n} = \nu N + H_{n+1}$$

with A_n determined so that it satisfies the equation

$$[\nu N, A_n] + V_n = 0.$$

We choose the solution given by

$$A_n := -\frac{i}{T} \int_0^T dt \int_0^t ds\, e^{is\nu N} V_n e^{-is\nu N}$$

Indeed, we find that

$$[\nu N, A_n] = -\frac{i}{T} \int_0^T dt \int_0^t ds\, \frac{d}{ds}(e^{is\nu N} V_n e^{-is\nu N}) = -\langle V_n \rangle + V_n = V_n$$

where the last equality $\langle V_n \rangle = 0$ is by the definition of V_n. A straightforward bound gives

$$\|A_n\| \leq \frac{T}{2}\|V_n\|$$

Defining now γ_n, α_n just as before, i.e.

$$\gamma_n(O) := e^{-A_n} O e^{A_n} = e^{-\mathrm{ad}_{A_n}} O,$$

and

$$\alpha_n(O) := \int_0^1 ds\, e^{-s A_n} O e^{s A_n} = \int_0^1 ds\, e^{-s\,\mathrm{ad}_{A_n}} O,$$

we get indeed

$$H_{n+1} = \gamma_n(H_n) + \gamma_n(\nu N) - \nu N \tag{5.11}$$

$$= \gamma_n(H_n) - \alpha_n([A_n, \nu N]) \tag{5.12}$$

$$= \gamma_n(D_n) + (\gamma_n(V_n) - V_n) + (V_n - [A_n, \nu N]) - (\alpha_n([A_n, \nu N]) - [A_n, \nu N]) \tag{5.13}$$

$$= \gamma_n(D_n) + (\gamma_n(V_n) - V_n) - (\alpha_n(V_n) - V_n). \tag{5.14}$$

This equation is identical to the one for the time-dependent setup and all the bounds can be copied.

References

1. Abanin, D.A., De Roeck, W., Ho, W.W., Huveneers, F.: Effective hamiltonians, prethermalization, and slow energy absorption in periodically driven many-body systems. Phys. Rev. B **95**, 014112 (2017)
2. Abanin, D.A., De Roeck, W., Huveneers, F.: Exponentially slow heating in periodically driven many-body systems. Phys. Rev. Lett. **115**(25), 256803 (2015)
3. Abanin, D.A., De Roeck, W., Huveneers, F.: Theory of many-body localization in periodically driven systems. Ann. Phys. **372**, 1–11 (2016)
4. Aidelsburger, M., Lohse, M., Schweizer, C., Atala, M., Barreiro, Julio T., Nascimbène, S., Cooper, N.R., Bloch, I., Goldman, N.: Measuring the chern number of hofstadter bands with ultracold bosonic atoms. Nat. Phys. **11**(2), 162–166 (2015)
5. Bambusi, D., Graffi, S.: Time quasi-periodic unbounded perturbations of schrödinger operators and kam methods. Commun. Math. Phys. **219**(2), 465–480 (2001)
6. Bourgain, J., Wang, W.-M.: Anderson localization for time quasi-periodic random schrödinger and wave equations. Commun. Math. Phys. **248**(3), 429–466 (2004)
7. Bru, J.-B., Pedra, W.: Lieb-robinson bounds for multi-commutators and applications to response theory. Springer Briefs in Math. Phys. **13**, 1–110 (2017)
8. Bukov, M., D'Alessio, L., Polkovnikov, A.: Universal high-frequency behavior of periodically driven systems: from dynamical stabilization to floquet engineering. Adv. Phys. **64**(2), 139–226 (2015)
9. Bukov, M., Gopalakrishnan, S., Knap, M., Demler, E.: Prethermal floquet steady states and instabilities in the periodically driven, weakly interacting bose-hubbard model. Phys. Rev. Lett. **115**(20), 205301 (2015)
10. Carati, A., Galgani, L., Maiocchi, A., Gangemi, F., Gangemi, R.: Persistence of regular motions for nearly integrable hamiltonian systems in the thermodynamic limit. Regul. Chaot. Dyn. **21**(6), 660–664 (2016)
11. Chudnovskiy, A.L., Gangardt, D.M., Kamenev, A.: Doublon relaxation in the bose-hubbard model. Phys. Rev. Lett. **108**(8), 085302 (2012)
12. Combescure, M.: Spectral properties of a periodically kicked quantum hamiltonian. J. Stat. Phys. **59**(3-4), 679–690 (1990)
13. D'Alessio, Luca, Rigol, Marcos: Long-time behavior of periodically driven isolated interacting lattice systems. Phys. Rev. X **4**, 041048 (2014)
14. De Roeck, W., Huveneers, F.: Asymptotic localization of energy in nondisordered oscillator chains. Commun. Pure Appl. Math. **68**(9), 1532–1568 (2015)
15. Ducatez, R., Huveneers, F.: Anderson localisation for periodically driven systems. (2016). arXiv preprint arXiv:1607.07353
16. Eckardt, A., Anisimovas, E.: High-frequency approximation for periodically driven quantum systems from a floquet-space perspective. New J. Phys. **17**(9), 093039 (2015)
17. Eliasson, H.L., Kuksin, S.B.: On reducibility of schrödinger equations with quasiperiodic in time potentials. Commun. Math. Phys. **286**(1), 125–135 (2009)
18. Fishman, S., Grempel, D.R., Prange, R.E.: Chaos, quantum recurrences, and anderson localization. Phys. Rev. Lett. **49**(8), 509 (1982)

19. Goldman, N., Dalibard, J.: Periodically driven quantum systems: effective hamiltonians and engineered gauge fields. Phys. Rev. X **4**(3), 031027 (2014)
20. Howland, J.S.: Stationary scattering theory for time-dependent hamiltonians. Math. Ann. **207**(4), 315–335 (1974)
21. Imbrie, J.Z.: On many-body localization for quantum spin chains. J. Stat. Phys. **163**(5), 998–1048 (2016)
22. Kapitza, P.L.: Dynamic stability of a pendulum with an oscillating point of suspension. J. Exp. Theor. Phys. **21**(5), 588–597 (1951)
23. Kitagawa, T., Berg, E., Rudner, M., Demler, E.: Topological characterization of periodically driven quantum systems. Phys. Rev. B **82**(23), 235114 (2010)
24. Kuwahara, T., Mori, T., Saito, K.: Floquet–Magnus theory and generic transient dynamics in periodically driven many-body quantum systems. Ann. Phys. **367**, 96–124 (2016)
25. Lazarides, A., Das, A., Moessner, R.: Equilibrium states of generic quantum systems subject to periodic driving. Phys. Rev. E **90**(1), 012110 (2014)
26. Lazarides, A., Das, A., Moessner, R.: Fate of many-body localization under periodic driving. Phys. Rev. Lett. **115**(3), 030402 (2015)
27. Lieb, E.H., Robinson, D.W.: The finite group velocity of quantum spin systems. Commun. Math. Phys. **28**(3), 251–257 (1972). doi:10.1007/BF01645779
28. Lindner, N.H., Refael, G., Galitski, V.: Floquet topological insulator in semiconductor quantum wells. Nat. Phys. **7**(6), 490–495 (2011)
29. Magnus, W.: On the exponential solution of differential equations for a linear operator. Commun. Pure Appl. Math. **7**(4), 649–673 (1954)
30. Mori, T., Kuwahara, T., Saito, K.: Rigorous bound on energy absorption and generic relaxation in periodically driven quantum systems. Phys. Rev. Lett. **116**(12), 120401 (2016)
31. Nachtergaele, B., Ogata, Y., Sims, R.: Propagation of correlations in quantum lattice systems. J. Stat. Phys. **124**(1), 1–13 (2006)
32. Ponte, P., Chandran, A., Papić, Z., Abanin, D.A.: Periodically driven ergodic and many-body localized quantum systems. Ann. Phys. **353**, 196–204 (2015)
33. Ponte, P., Papić, Z., Huveneers, F., Abanin, D.A.: Many-body localization in periodically driven systems. Phys. Rev. Lett. **114**(14), 140401 (2015)
34. Rahav, S., Gilary, I., Fishman, Shmuel: Effective hamiltonians for periodically driven systems. Phys. Rev. A **68**(1), 013820 (2003)
35. Sensarma, R., Pekker, D., Altman, E., Demler, E., Strohmaier, N., Greif, D., Jördens, R., Tarruell, L., Moritz, H., Esslinger, T.: Lifetime of double occupancies in the Fermi–Hubbard model. Phys. Rev. B **82**(22), 224302 (2010)
36. Simon, B.: The Statistical Mechanics of Lattice Gases, vol. 1. Princeton University Press, Princeton (2014)
37. Soffer, A., Wang, W.-M.: Anderson localization for time periodic random Schrödinger operators. Commun. Partial. Differ. Equ. **28**(1–2), 333–347 (2003). doi:10.1081/PDE-120019385
38. Ueltschi, D.: Cluster expansions and correlation functions. Mosc. Math. J. **4**(2), 511–522 (2004)
39. Verdeny, A., Mielke, A., Mintert, F.: Accurate effective hamiltonians via unitary flow in floquet space. Phys. Rev. Lett. **111**(17), 175301 (2013)

Communicated by H. Spohn