Many-Body Localization and its Discontents

John Z. Imbrie

University of Virginia

Stanford Institute for Theoretical Physics
March 7, 2022

Overview

A quantum system is said to be many-body localized (MBL) if it remains close to its initial state, i.e., it fails to thermalize. In 2016, I published a proof that certain one-dimensional spin chains have an MBL phase (the proof depended on a certain assumption on level statistics). Some recent numerical studies have raised questions about whether there is a true MBL phase. I will attempt to summarize the issues raised, but the fact remains that the mechanisms for the breakdown of MBL phase are well understood theoretically. In recent work with Morningstar and Huse (PRB, 2020), we develop specific RG flow equations. These are similar to the Kosterlitz-Thouless (KT) flow as previously shown, but there are important differences that place the MBL transition in a new universality class.

Outline

Phenomenology of MBL

For a many-body quantum system with disorder, we may observe the following, which may be thought of as essential features of many-body localization (MBL):

1. Absence of transport
2. Anderson localization in configuration space (as in, e.g. IPR measures)
3. Area law entanglement
4. Violation of ETH (eigenstate thermalization hypothesis)
5. Absence of level repulsion
6. Logarithmic growth of entanglement for an initial product state

Typical example: disordered spin chain

Spin chain with random interactions and a weak transverse field on $\Lambda = [-K, K] \cap \mathbb{Z}$:

$$H = \sum_{i=-K}^{K} h_i S_i^z + \sum_{i=-K}^{K} \gamma_i S_i^x + \sum_{i=-K-1}^{K} J_i S_i^z S_{i+1}^z.$$

This operates on the Hilbert space $\mathcal{H} = \bigotimes_{i \in \Lambda} \mathbb{C}^2$, with

$$S_i^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, S_i^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

operating on the i^{th} variable.

Assume $\gamma_i = \gamma \Gamma_i$ with γ small. Random variables h_i, Γ_i, J_i are independent and bounded, with bounded probability densities.

Ergodicity breaking and the emergence of an extensive set of local integrals of motion (LIOMs)

A fully MBL system has a complete set of of conserved quantities (quasilocal in nature) – a complete failure of ergodicity.

How do we know if a system has a complete set of quasilocal LIOMs? Can we construct them?

We seek a quasilocal unitary that diagonalizes H. That is, $D = U^\star HU$ is diagonal, and quasilocality means that the effect of U on a set of spins that span a distance L in the lattice should be (identity) + (exponentially small in L). There may be rare, nonpercolating regions where this property fails (resonant regions).

Then we may define LIOMs $\tau_i = US_i^z U^\star$.

It is clear that $[H, \tau_i] = [D, S_i^z] = 0$.

Likewise $[\tau_i, \tau_j] = 0$.

Properties 1-6 listed above for MBL should follow if one can find a complete set of LIOMs[1]

[1] Huse, Nandkishore, Oganesyan, PRB '14; Serbyn, Papic, Abanin, PRL '13

One spin

For guidance, consider what happens for a single spin. Then

$$H = \begin{pmatrix} h & \gamma \\ \gamma & -h \end{pmatrix}$$

and for $\gamma \ll h$ the eigenfunctions are close to $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$. The eigenfunctions resemble the basis vectors. This means the basis vectors can be used to label the eigenfunctions.

At the other extreme, if $\gamma \gg h$ the eigenfunctions are close to $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$. With complete hybridization, there is no meaningful way to associate eigenfunctions with basis vectors.

Perturbative and non-perturbative approaches

One may construct LIOMs perturbatively[2].

But rare regions where perturbation theory breaks down have the potential to spoil MBL. I gave a nonperturbative proof of MBL in a 1d spin chain[3] (which, however, depends on a physically reasonable assumption on eigenvalue statistics – essentially that the level spacings in a system of n spins are no smaller than some exponential in n.)

It is especially important to have a nonperturbative proof of an MBL phase, as some are questioning the numerical evidence for MBL[4].

---

[2] Integrals of motion in the many-body localized phase, Ros, Müller, Scardicchio NP '15

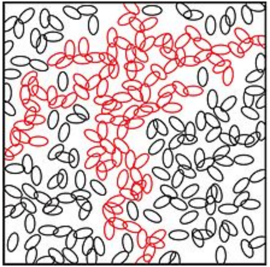
[3] Imbrie, *On many-body localization for quantum spin chains*, JSP '16

[4] *Quantum chaos challenges many-body localization*, Šuntajs, Bonča, Prosen, Vidmar arXiv:1905

Percolation picture validated for large disorder or weak interactions in 1d

Proof controls the probability of resonance for processes, and shows that the graph of resonances is non-percolating.

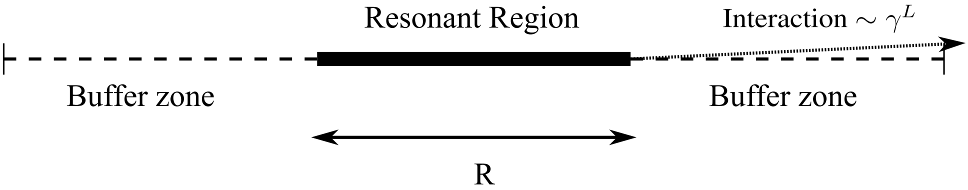
Then is possible to define quasilocal similarity transformations on H that diagonalize it, deforming the tensor product basis vectors into the exact eigenfunctions.



Resonant regions (= Griffiths regions) need buffer zones

These are regions where we have failure of the bounds needed to control the rotations.

Buffer zones are needed so that the smallness $\sim \gamma^L$ of a graph crossing the buffer is much smaller than the typical $\Delta E = 2^{-R}$ in the resonant region.



The buffer zone is expected to be thermalized by the resonant region.

In 1-d the buffer zone has volume comparable to that of the resonant block, so we can diagonalize H in the combined region, eliminating internal interactions while keeping the level-spacing larger than the interactions with spins outside.

Renormalization group picture

In RG terms, the rotations removing terms in the Hamiltonian up to order γ^L is analogous to "integrating out" short distance degrees of freedom in traditional RG.

At the same time, resonant regions up to some size R are "eliminated" once L is large enough so that the remaining interaction terms are smaller than the level spacing in the region (with its buffer zone, total size $R + 2L$). At that point, the region hosts a "metaspin" which takes 2^{R+2L} values, but the interactions are so small that there is little hybridization with spins elsewhere.

Deep in the localized region, this RG has the property that the density of remaining resonant regions (including their buffer zones with width given by the running RG length L) goes to zero with L.

Note two effects are in play:
(1) Elimination of smaller resonant regions reduces the density.
(2) Fattening of the buffer zones on the remaining regions increases the density.
My MBL proof shows that (1) dominates (2) deep in the weak coupling/strong disorder region, and the density goes to zero as $L \to \infty$.

Moving toward the transition: the avalanche effect

For weaker disorder/stronger interactions, the decay rate can be reduced to the point where no buffer size can insulate the resonant region from the rest of the chain: the avalanche instability[5].

Flip rates for off-diagonal matrix elements connecting the resonant region with spins outside the buffer zone should behave as $2^{-2L/\zeta}$ for some decay length ζ. For this to be small compared with the level spacing $\sim 2^{-(R+2L)}$, we need $\zeta^{-1} > 1$. Let $x = \zeta^{-1} - 1$ be the excess decay rate. The buffer size must satisfy

$$L \geq \frac{R}{2x}.$$

As γ increases, the excess decay rate $x \to 0$ and then the buffer size L will diverge.

At some point, then, increasing γ causes (2) to dominate (1); *i.e.* the fattening effect dominates the eliminations, and the density of resonant regions grows with L.

[5] *Many-body delocalization as a quantum avalanche*. Thiery, Huveneers, Müller, De Roeck, PRL '18

Many-body localization and its discontents, I

Šuntajs *et al*, PRE 2020[6]., argued that the numerics show a scaling collapse when $W \sim L$, meaning that any fixed disorder strength would thermalize for sufficiently large L.

But Abanin *et al*, Ann. Phys 2021[7] disputed this conclusion, pointing out that some drift of W with L is expected and in fact occurs in systems known to have a transition.

Furthermore, if $W \sim L$ then the system would be in the regime covered by my proof when L exceeds some threshold. The fact that there is drift for the small L's accessible to numerics does not imply continued drift for all L.

---

[6] *Quantum chaos challenges many-body localization*, Šuntajs, Bonča, Prosen, Vidmar

[7] Distinguishing localization from chaos: challenges in finite-size systems, Abanin, Bardarson, De Tomasi, Gopalakrishnan, Khemani, Parameswaran, Pollmann, Potter, Serbyn, Vasseur

Many-body localization and its discontents, II

More recently, Sels and Polkovnikov, PRE 2021[8], found further evidence in numerics (using novel measures) to support the view that there is no transition.

However, these alternative views do not actually take into account that the system sizes studied are insufficient to see the predicted phenomena (avalanches) that we believe actually drive the transition. In the avalanche picture there are rare events that can have a profound effect, thermalizing the entire system. If they are rare, then they will be hard to see in small systems. To get past the infinite volume transition, one would need to crank up the disorder (or reduce interactions) well past what one would think based on small system observations.

[8] Dynamical obstruction to localization in a disordered spin chain

Many-body localization and its discontents, III

It seems that at modest system sizes one is only probing crossover phenomena, and that the nature of the transition can only be probed by going to much larger system sizes – too large to be done on the computer. In order to at least partially get around the problem of the need for extremely large L, Morningstar, *et al* arxiv:2107[9] probe a measure of avalanche instability, *i.e.* system is avalanche-ready even if there is no trigger to set it off. (Also other intermediate measures that go beyond the traditional RM to Poisson level statistics measures.) The conclusion is that the $L = \infty$ transition is quite far from the regimes studied in numerics. One should not assume that the drift of the apparent transition continues forever – eventually you would get to the regime covered by my proof.

[9] Avalanches and many-body resonances in many-body localized systems, Morningstar, Colmenarez, Khemani, Luitz, Huse

Simplified strong-disorder RG picture

At a given cutoff Λ, the line consists of alternating localized intervals (L-blocks) and thermalized intervals (T-blocks). Assumes bimodality is strongly attractive near transition. Assume the decay rate deficit x is constant in space.[10]

- L-blocks represent intervals where quasilocal basis changes have been defined.
- T-blocks have minimum length Λ; they represent intervals where the basis change cannot be defined due to too-strong interactions with the environment.
- As $\Lambda \to \Lambda + d\Lambda$, T-blocks of length $\in [\Lambda, \Lambda + d\Lambda]$ are *erased* (absorbed into neighboring L-blocks) if they are *isolated*, that is, separated by more than the buffer size Λ/x from other T-blocks.
- If a T-block is *not isolated*, then it pairs with a neighboring T-block that lies within the distance Λ/x to form a larger T-block (eliminating the intervening L-block). Such blocks do not have enough room to localize separately.
- The avalanche parameter x flows downward with the RG because erased T-blocks interrupt the decay of interactions.

[10] This approximation can be justified near the transition using Chayes-Harris arguments, once we have solved for the length divergence.

Functional RG

Due to the quenched (iid) randomness, we can assume that T-blocks appear "at random" with an exponential distribution in space for each subsequent T-block (outside of the minimum distance Λ/x as determined by the RG rules). Letting R_Λ denote the rate for this exponential distribution, we have that $R_\Lambda \exp(-R_\Lambda w) dw$ is the probability that length of an L-block lies in $[\Lambda/x + w, \Lambda/x + w + dw]$.

This rate can be broken down according to the length ℓ of the T-block that appears after the L-block: $R_\Lambda = \int_\Lambda^\infty r_\Lambda(\ell) d\ell$.

The full functional RG describes the flow with Λ of the function $r_\Lambda(\ell)$ and x

$$\frac{dx}{d\Lambda} = -\frac{\Lambda r_\Lambda(\Lambda)(1+x)}{1+\Lambda R_\Lambda/x}$$

$$\frac{dr_\Lambda(L)}{d\Lambda} = \frac{1}{x}\left(\frac{dx}{d\Lambda} - R_\Lambda\right) r_\Lambda(L) + \frac{1}{x}\Theta(L - [2+x^{-1}]\Lambda) \int_\Lambda^{L-(1+x^{-1})\Lambda} d\ell r_\Lambda(\ell) r_\Lambda(L-\ell-\Lambda/x).$$

Reduction to two parameters

The rate $r_\Lambda(\Lambda)$ has dimensions $1/(\text{length})^2$, so let us define a dimensionless rate

$$y = y_\Lambda = \Lambda^2 r_\Lambda(\Lambda).$$

We anticipate that $y = 0$, $x \geq 0$ will be the MBL fixed line, due to the vanishing density of T-blocks. The phase transition will be governed by the point $x = y = 0$, where the fixed line becomes unstable because the interaction decay rate reaches the critical value for avalanches.

The dominant mode of production of T-blocks of size Λ/x should be the combination of component T-blocks of size close to Λ. This leads to a recursion relation

$$r_\Lambda(\Lambda/x) = R_\Lambda^2. \tag{1}$$

For similar reasons, $r_\Lambda(\ell)$ should depend weakly on Λ between $x\ell$ and ℓ. This means that $r_\Lambda(\ell) \approx y_\ell/\ell^2$ for $\Lambda \leq \ell \leq \Lambda/x$ and $R_\Lambda \approx \Lambda r_\Lambda(\Lambda)$. Combining these facts with the recursion (1), we obtain a recursion for y:

$$y_{\Lambda/x} = \left(\frac{y_\Lambda}{x_\Lambda}\right)^2. \tag{2}$$

Behavior of the recursion/flow

As is customary, we use $t = \log \Lambda$ to parametrize the RG.
The recursion/flow can then be written as:

$$\frac{dx}{dt} = -y, \qquad y_{\Lambda/x} = \left(\frac{y_\Lambda}{x_\Lambda}\right)^2, \tag{3}$$

with the equation for x representing the decrease in decay rate due to the erasure of T-blocks at the cutoff Λ.

If we start on the curve $y = x^{2+\delta}$, then the image under the recursion is close to the curve $y = x^{2+2\delta}$. Hence the separatrix is asymptotic to the curve $y = x^2$.

The flow along the separatrix is then determined, with $x \sim t^{-1}$, $y \sim t^{-2}$.

Diverging length

A diverging length may be defined as the point where an orbit departs the vicinity of the separatrix, from an initial small displacement δ_0. We find that this length is

$$\Lambda = e^t = \delta_0^{-\log_2 \log_2 \delta_0^{-1}}.$$

This evidently diverges faster than any power of δ_0, so we have in effect $\nu = \infty$.

This may be distinguished from the KT form: $\Lambda = \exp(\text{const} \cdot \delta_0^{-1/2})$.

Like the KT flow, there is logarithmic slowdown along the separatrix and $\nu = \infty$. However in that case progress is slow both along the separatrix and orthogonal to it.

Here we have exponential divergence from the separatrix, albeit proceeding through the logarithmically-slowed RG time that is dictated by the separatrix flow.

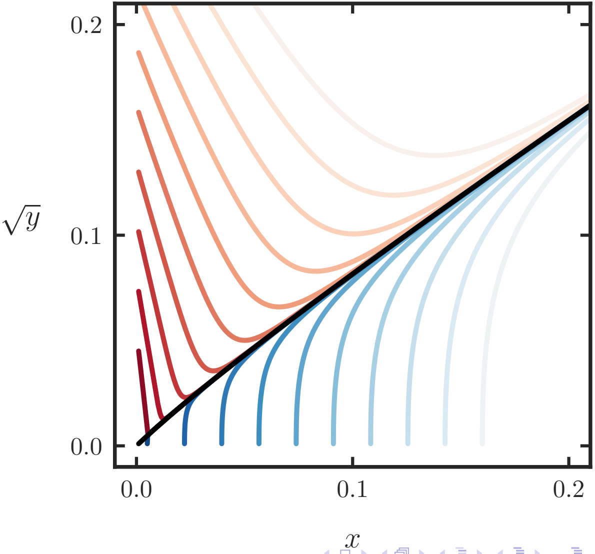
Equivalent flow equations

The following flow equation for y leads to the same critical behavior as the recursion:

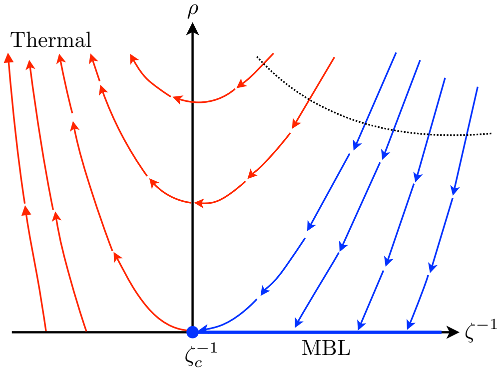
$$\begin{aligned}\frac{dy}{dt} &= -(\log 2) y \delta \\ &= -(\log 2) y \left(\frac{\log y}{\log x} - 2 \right).\end{aligned}$$

The flow equation for x remains as before:

$$\frac{dx}{dt} = -y$$



Parallels with the KT transition



Like the vortices, T-blocks represent nonperturbative effects, and the tendency of these effects to grow or shrink with the flow determines the phase reached from any starting point in the diagram. Vortex binding is analogous to T-block erasure as discussed above.

When bound, vortices renormalize the stiffness (screening). Likewise, when eliminated, T-blocks renormalize the decay rate (anti-screening).

Note: Earlier works (Dumitrescu et al, Goremykina et al., 2019) suggested a KT picture for the transition but assumed analytic flow equations. Our flow involves a factor $(\log y)/(\log x) - 2$, which puts this problem in a different universality class from KT.