

ICFP M2 - STATISTICAL PHYSICS 2 – Lecture n° 9

Localization

Grégory Schehr, Guilhem Semerjian

March 2020

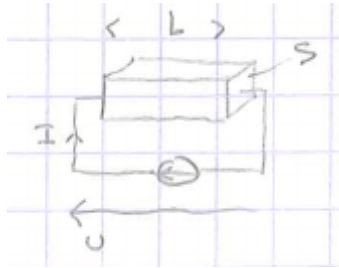
A recurrent pattern in this course has been to ask the question: “what happens if one adds (strong) disorder to something ? ” Following this pattern we shall consider in this lecture the effects of disorder on the transport of some quantities, or the propagation of some waves. This is a vast subject, which was the sole topic of the series of M2 lectures by N. Cherroret and C. Texier, this lecture will not attempt to cover it in full details, but rather to introduce the aspects of this field that are most connected to the rest of the course.

1 Phenomenology of transport in disordered media

Let us start by reviewing some basic aspects of transport, focusing first on condensed matter systems.

1.1 Conductivity

Consider a piece of material submitted to some voltage, as depicted in the figure:

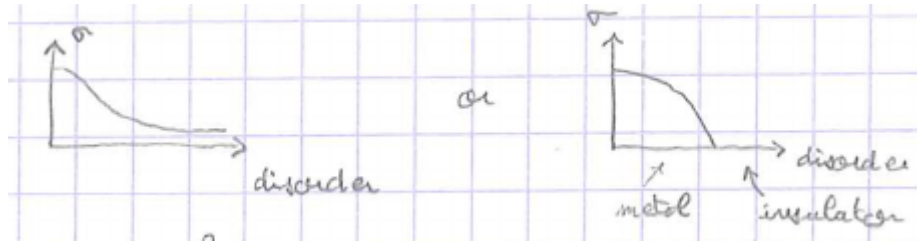


The voltage and the intensity of the current are linked by Ohm’s law, $U = RI$, with in this geometry $R = \frac{1}{\sigma} \frac{L}{S}$, with L the length of the sample, S its area, and σ the conductivity of the material. At the microscopic level σ links the charge current density and the electrical field according to $\vec{j} = \sigma \vec{E}$.

Note that the computation of σ from first principles, starting from a microscopic model of the material, is an extremely difficult task, because it involves a very large number of particles in interaction, that should be treated in a quantum mechanical way, and most importantly because it is an intrinsically out-of-equilibrium phenomenon. Indeed the flow of a current in a system cannot be explained by the laws of equilibrium statistical mechanics, and out-of-equilibrium physics lacks the general principles (canonical ensembles in particular) of its equilibrium counterpart.

In the phenomenological Drude modelling of this problem one introduces a “friction” force that compensates the effect of the electric field on the charge carriers (electrons for instance) to reach a steady state, and expresses the conductivity as $\sigma = \frac{ne^2\tau}{m}$, with n , e and m the density, charge and mass of the carriers. The parameter τ is a characteristic time scale of the “friction” exerted on the moving charges, that comes from the “collisions” (more precisely scattering events) between the charges and the imperfections of the crystal lattice (remember that by Bloch’s theorem a perfectly periodic lattice admits propagating waves). These imperfections come either from the thermal vibrations of the lattice (phonons) or from the defects of the lattice (dislocations, replacement of an atom by an impurity).

When the concentration of impurities in a material increases, it is intuitively clear that the conductivity σ should decrease, the “collisions” of the charge carrier against these impurities being more frequent. This however leaves open two possibilities, represented schematically on the figure:



- either σ decreases smoothly when the disorder is increased, the material remains always metallic (conductor) even if more and more poorly (left panel)
- or there is a phase transition towards an insulating phase when the disorder exceeds some critical value (right panel)

Which of these two scenarii is realized is a rather fundamental question which was first raised by Anderson in a seminal paper in 1958, this field is now mainly known under the name of Anderson localization.

1.2 Diffusion

In the previous paragraph we considered the response of a system submitted to a non-zero electric field \vec{E} , and characterized it through the conductivity σ , the proportionality factor between the excitation \vec{E} and the response \vec{j} .

Consider now the somehow dual question: what happens to a particle, in the absence of any field \vec{E} ? Without loss of generality we assume that the particle is in $\vec{r} = \vec{0}$ at $t = 0$. Quite generically we expect the particle to have a diffusive behavior at large times, its average mean square displacement behaving as $\langle \vec{r}(t)^2 \rangle \sim 2Dt$ as $t \rightarrow \infty$, which defines the diffusion constant D .

It is important to keep in mind that σ and D are not independent. The two a priori different situations they describe (response to a field vs spontaneous fluctuations without field) are indeed tied together by the Fluctuation Dissipation Theorem (FDT, associated to Einstein and Kubo), which is a fundamental property of statistical mechanics and exist under different forms in various contexts. Here one can quote it as $D = \frac{kT}{ne^2} \sigma$.

In particular in an insulator, defined by $\sigma = 0$, one has necessarily $D = 0$, particles do not diffuse at large times, they thus remain “localized” around their initial position, hence a partial justification of the name of this topic. On the contrary, in a conductor, both σ and D are strictly positive, particles can diffuse away from their initial position in the absence of an external drive.

1.3 Dimension matters

Before describing in more details the models that can be investigated to answer Anderson’s question, we can already anticipate that the dimensionality will play a crucial role. One can think of the metallic phase as an ordered state that will be perturbed by the addition of disorder. But as usual fluctuations (thermal or due to disorder) destroy ordered phase more efficiently in lower dimensions (recall that the pure Ising model in $d = 1$ has no ordered phase at any positive temperature, whereas for $d \geq 2$ it has a low temperature magnetized phase). This is also the case for the localization phenomenon:

- for $d = 1$ and $d = 2$ an infinitesimal amount of disorder is enough to induce localization (absence of conductivity),
- for $d \geq 3$ there is a phase transition between metallic and insulating phases as a function of the disorder strength.

Note that considering dimensions lower than 3 is not an academic question, as effectively 2d (graphene) and 1d (carbone nanotubes) condensed matter systems do exist.

1.4 Other systems

The presentation above has been focused on electrons in condensed matter systems, but actually Anderson localization occurs and has been studied both theoretically and experimentally in many other contexts, in particular for the propagation of classical waves (light and sound) in disordered media, and for cold atoms in disordered potentials (created by a speckle phenomenon).

2 Microscopic models

Let us now turn to a more precise description of simple models that arise naturally in this context.

2.1 In the continuum

Maybe the most natural approach consists in studying the Schrödinger equation for a particle of mass m in a disordered potential $V(\vec{r})$,

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = \left(-\frac{\hbar^2}{2m} \Delta + V(\vec{r}) \right) \psi(\vec{r}, t) \quad (1)$$

where the wave function ψ is a square-normalizable function ($\psi \in L^2(\mathbb{R}^d)$). Various models of the disordered potential can be used, for instance one can assume that point-like scatterers are placed randomly in space, yielding

$$V(\vec{r}) = \sum_n a_n \delta(\vec{r} - \vec{x}_n) , \quad (2)$$

with a_n and \vec{r}_n the amplitude and the location of the n -th scatterer, both can be taken random, the \vec{r}_n can be for instance drawn as a Poisson Point Process (there is a scatterer in any infinitesimal volume $d\vec{x}$ with probability proportional to $\text{vol}(d\vec{x})$, independently for each infinitesimal volume).

In this context the question of the existence of transport can be reformulated as follows : imagine sending a wave packet onto such a disordered potential. There will be multiple scattering, with many different paths between the various scattering centers. The question is whether these multiple paths will lead to constructive interference, in which case the wavepacket will be transported through the medium, or whether the interferences will be destructive. If the potential is periodic Bloch's theorem asserts that the interferences are constructive, but in the disordered case the answer is not obvious and can depend on the strength of the disorder.

2.2 On a lattice: the Anderson model

As usual it is technically simpler to work in a discrete setting than in a continuous one. We shall thus modify the previous model and assume that the particle can only be at some discrete positions (as in the tight-binding approximation of condensed matter physics), modelled as a graph Λ , for instance \mathbb{Z}^d or a finite portion of it. The wavefunction ψ is defined by its values ψ_i at all vertices of the graph, $\psi \in \ell^2(\Lambda)$ and we introduce an operator H that acts on ψ according to:

$$(H\psi)_i = J \sum_{j \in \partial i} \psi_j + V_i \psi_i , \quad (3)$$

where ∂i denotes the set of nearest neighbors of i along the edges of the graph, and V_i is a random potential on site i . The first term is proportional to a discrete version of the Laplacian (one should subtract $2d\psi_i$ to have exactly the discrete Laplacian, but this would only lead to a global shift of the energy scale). The constant J is an energy scale that sets the strength of the hopping between neighboring sites (kinetic energy). This Hamiltonian is thus a discrete version of (1), and is known as the Anderson model. Usually the random variables V_i , that represents a pinning energy that attracts or repel the particle on site i , are taken as independent identically distributed. A standard choice for their law is to take them uniformly distributed on the interval $[-W/2, W/2]$, the parameter W controlling in this way the strength of the disorder.

When Λ is finite, for instance a portion of linear size L of the lattice \mathbb{Z}^d , one can view H as a square matrix of size $N = L^d$, acting on vectors ψ of size N by the usual multiplication rule. An equivalent definition of H is then given via its matrix elements,

$$H_{ij} = \begin{cases} J & \text{if } i \text{ and } j \text{ are nearest neighbors} \\ V_i & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

As the diagonal elements V_i are random variable H is a random matrix, as the ones studied in the previous lecture. It is however important to underline two crucial differences with the random matrix ensembles introduced previously (i.e. GOE, GUE, invariant ensembles, Wigner matrices):

- the non-zero off-diagonal elements of H have a finite-dimensional Euclidean structure, whereas the usual random matrices do not (they are mean-field like)
- the Anderson model has two energy scales, J and W , whereas usual random matrices have only one (that can be absorbed in a global multiplicative coefficient); this second remark will play an important role later on.

As anticipated the Anderson model can exhibit phase transitions; these can only occur in the thermodynamic limit. In this context there are two ways of reaching it: either one starts with a finite size L , deal with H as a matrix, and then send L to ∞ , or one works directly with an infinite lattice as \mathbb{Z}^d , considering H as a linear operator on the infinite-dimensional vector space $\ell^2(\mathbb{Z}^d)$. The second approach is mathematically more demanding but sometimes necessary to obtain precise results.

Let us also mention that the Anderson model can be defined on any graph, it has been in particular studied also in random graphs of the type introduced previously in this course, and on the related infinite trees. These constructions are often called Bethe lattices in this context, and provide a mean-field version of the model.

In the following we will only consider the lattice Anderson model, and put $\hbar = 1$ for simplicity.

2.3 Observables

Imagine you are given an instance of the Anderson model, i.e. a graph Λ and the energies V_i on the vertices, and asked to determine whether it represents a metal or an insulator; what should you compute? This is what we are going to see now. For simplicity we do not want to put a current in the system, we shall thus take the “diffusion” point of view on this question. We will also assume that the graph is finite, with N vertices.

- The state of the system is represented by the wavefunction $|\psi(t)\rangle$, here a vector of size N , that evolves according to Schrödinger’s equation $i\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle$; this equation being linear all the relevant information on H is given by its eigendecomposition. As H is Hermitian we have a basis of eigenvectors $|v_\alpha\rangle$, for $\alpha = 1, \dots, N$, such that $H|v_\alpha\rangle = \omega_\alpha|v_\alpha\rangle$, with ω_α the real associated eigenvalues. We assume the eigenvectors to be orthonormal, $\langle v_\alpha|v_\beta\rangle = \delta_{\alpha,\beta}$. We will also denote $\{|i\rangle\}$ the canonical basis, i.e. $|i\rangle$ is the normalized vector for a particle present only on the vertex i of the graph.
- Let us compute in particular $p_{i \rightarrow j}(t)$ defined as the probability that a particle which starts at $t = 0$ on vertex i is present on vertex j at a later time t :

$$p_{i \rightarrow j}(t) = |\langle j|e^{-iHt}|i\rangle|^2 \quad (5)$$

$$= \left| \sum_{\alpha=1}^N \langle j|v_\alpha\rangle \langle v_\alpha|i\rangle e^{-i\omega_\alpha t} \right|^2 \quad (6)$$

$$= \sum_{\alpha=1}^N |\langle j|v_\alpha\rangle|^2 |\langle v_\alpha|i\rangle|^2 + \sum_{\alpha \neq \beta} e^{i(\omega_\beta - \omega_\alpha)t} \langle j|v_\alpha\rangle \langle v_\alpha|i\rangle \langle i|v_\beta\rangle \langle v_\beta|j\rangle \quad (7)$$

We will denote $\tilde{p}_{i \rightarrow j}$ the first term in the last line, and think of it as the large time limit of $p_{i \rightarrow j}(t)$. This is slightly abusing, as the second term is, for all finite systems, a periodic function of t . Nevertheless we have in mind the large N limit, in the presence of disorder, in which case the oscillations of the second term become very strong and it is reasonable to assume that they average out to zero. Let us make two remarks on the expression of \tilde{p} :

- in order to have $\tilde{p}_{i \rightarrow j} > 0$ one needs the existence of some eigenvector $|v_\alpha\rangle$ of H which overlap both with i and j , i.e. such that $\langle i|v_\alpha\rangle$ and $\langle j|v_\alpha\rangle$ are different from zero. In other words the diffusion at long times between distant sites i and j is only possible if some eigenvectors connect them, i.e. are “delocalized” in a sense that will be made more precise later on.
 - reciprocally for $j = i$ the probability that the particle remains at its starting point is $\tilde{p}_{i \rightarrow i} = \sum_{\alpha=1}^N |\langle i|v_\alpha\rangle|^4$.
- Let us introduce some quantities which are commonly used to summarize in a compact way the relevant information about the eigendecomposition of H :

- The Density Of States (DOS) is defined as

$$\rho(E) = \frac{1}{N} \sum_{\alpha=1}^N \delta(E - \omega_\alpha) . \quad (8)$$

It is the empirical distribution of the eigenvalues, it washes out all information about the eigenvectors, but tells what are the possible energy levels of the system.

- The Local Density Of States (DOS) is defined, for a site i of the system, as

$$\rho_i(E) = \sum_{\alpha=1}^N |\langle i|v_\alpha\rangle|^2 \delta(E - \omega_\alpha) . \quad (9)$$

This contains now some information on the eigenvectors, and tells how much the states of energy E overlap with site i . One can show that the spatial average of the LDOS is equal to the DOS, $\frac{1}{N} \sum_{i=1}^N \rho_i(E) = \rho(E)$, because $\sum_i |\langle i|v_\alpha\rangle|^2 = \langle v_\alpha | \sum_i |i\rangle \langle i| v_\alpha \rangle = \langle v_\alpha | v_\alpha \rangle = 1$.

- The Inverse Participation Ratio of the α -th eigenvector is defined as

$$\text{IPR}_\alpha = \sum_{i=1}^N |\langle i|v_\alpha\rangle|^4 . \quad (10)$$

Recall that we have already encountered such a quantity in TD2, and explained its interpretation as follows. Suppose that $|v_\alpha\rangle$ is non-zero, and constant, on a number N_α of sites. Because of the normalization condition one has

$$\langle i|v_\alpha\rangle = \begin{cases} \frac{1}{\sqrt{N_\alpha}} & \text{on } N_\alpha \text{ sites } i \\ 0 & \text{on the other sites} \end{cases} , \text{ hence } \quad \text{IPR}_\alpha = N_\alpha \frac{1}{N_\alpha^2} = \frac{1}{N_\alpha} . \quad (11)$$

In general the inverse of the IPR is the order of magnitude of the number of sites on which the eigenvector has a sizable value, or that “participates” to the vector, which explains the name of the quantity. In the thermodynamic limit $N \rightarrow \infty$ these IPR can have dramatically different behaviors: if $\text{IPR}_\alpha \rightarrow 0$ one says that the vector $|v_\alpha\rangle$ is delocalized, or extended, because it overlaps with a diverging number of sites. On the contrary if the IPR remains strictly positive the vector is said to be localized, as it has a sizable value only on a finite number of sites. The previous computations should help to see that these (de)localization properties of the eigenvectors translate directly into the diffusion or localization of the particles evolving according to Schrödinger’s dynamics.

2.4 Phase transition in the Anderson model

Let us recall the definition (3) of the Anderson model, that we rewrite with bra-ket notations as

$$H = J \sum_{i,j \text{ p.v.}} (|i\rangle\langle j| + |j\rangle\langle i|) + \sum_i V_i |i\rangle\langle i|, \quad (12)$$

where the first sum runs over all pairs of neighboring vertices, and the V_i are i.i.d. random variables, uniform on $[-W/2, W/2]$. For concreteness we assume the graph to be a portion of \mathbb{Z}^d , of linear size L , containing $N = L^d$ sites, with periodic boundary conditions.

The two terms in H have very different properties, it is instructive to study first the two limits where only one of the two terms is present:

- If $J = 0$ the Hamiltonian is diagonal in the canonical (site) basis, the energies ω_α are equal to the local potentials V_i , all states are maximally localized, i.e. $\text{IPR}_\alpha = 1$ for all α . The density of states $\rho(E)$ converges in the thermodynamic limit to the distribution of the V_i 's, here the uniform law on $[-W/2, W/2]$. On the contrary the LDOS is very singular, it contains only a delta peak at the potential energy of the site, $\rho_i(E) = \delta(E - V_i)$. One has $p_{i \rightarrow j}(t) = \delta_{i,j}$ at all times, as no eigenvector connect different sites.
- On the contrary when $W = 0$ there is only the hopping term, the Hamiltonian is a pure tight-binding model which, thanks to the invariance by translation, can be diagonalized in the Fourier basis. The eigenvectors can indeed be chosen as

$$\langle i | v_\alpha \rangle = \frac{1}{\sqrt{N}} e^{i\vec{k}_\alpha \cdot \vec{r}_i}, \quad (13)$$

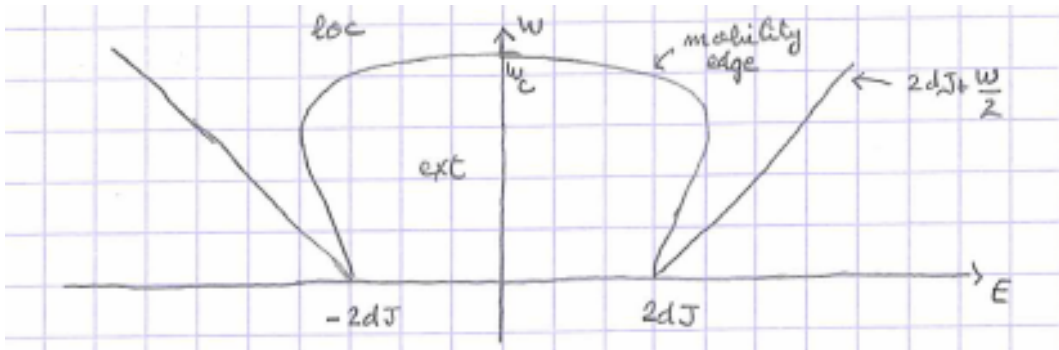
where \vec{r}_i is the real-space position of the site i , and \vec{k}_α is one of the vectors of the form $(\frac{2\pi}{L}n_1, \dots, \frac{2\pi}{L}n_d)$, where the n 's are integers, and the \vec{k} can be restricted to the first Brillouin zone $[-\pi, \pi]^d$. The energy of such an eigenvector is then found to be

$$\omega(\vec{k}) = J \sum_{\mu=1}^d 2 \cos(k_\mu). \quad (14)$$

The density of states $\rho(E)$ can be expressed, in the thermodynamic limit, as an integral over the first Brillouin zone, with a change of variables between \vec{k} and E given by this dispersion relation. The precise form of $\rho(E)$ is not important, what we need for the following is that its support is the interval $[-2dJ, 2dJ]$, as is clear from the dispersion relation.

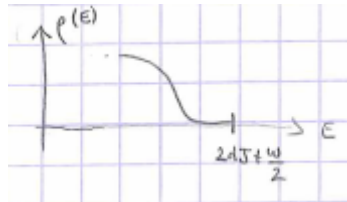
One can check from the plane wave form (13) that these eigenvectors are maximally delocalized, $\text{IPR}_\alpha = 1/N$ for all α , and that the LDOS and the DOS coincide, i.e. $\rho_i(E) = \rho(E)$ for all sites i . Finally $\tilde{p}_{i \rightarrow j} = 1/N$ for all i and j , at long times a particle is completely delocalized on all sites.

It should be clear now that the two terms of the Hamiltonian have eigenstates of opposite physical characteristics. When both are present at the same time there will be a competition between them, and possibly phase transitions separating phases where one dominates the other. As a matter of fact the phase diagram of the Anderson model for $d \geq 3$, which is the main message to remember from this lecture, is represented on the following figure:



Let us emphasize the main points to take from this phase diagram, where the vertical axis is W , the strength of the disorder, and the horizontal axis is E , the energy of the eigenstates:

- the density of states $\rho(E)$ is non-zero on the interval $[-2dJ - W/2, 2dJ + W/2]$. Indeed when $W = 0$ the band of plane waves covers the energies $[-2dJ, 2dJ]$, and the existence of arbitrary large cubes where all the local potentials V_i are close to their maximal value $W/2$ (or minimal values $-W/2$) implies that states exist for energies arbitrary close to the edges $\pm(2dJ + W/2)$.
- depending on the values of W and E in this authorized interval the nature of the corresponding eigenstate changes drastically. Some correspond to delocalized eigenvectors, in the region denoted “ext” (for extended) on the figure, while the eigenstates in the “loc” phase are localized. The line separating these two phases is called the mobility edge, a phase transition occurs when it is crossed. When $W = 0$ of course the band $[-2dJ, 2dJ]$ corresponds to delocalized states. When W is increased but remains small compared to J there are both extended states in the middle of the band, and localized states close to the energy edges. When W is further increased above a critical value W_c all extended states disappear, only localized states survive.
- the transition between the localized and delocalized phases cannot be seen on the DOS $\rho(E)$. If one however zooms on the distribution of eigenvalues before taking the thermodynamic limit, and look at the correlation between energy levels on the mean spacing scale, one sees a drastic difference between the two phases. Indeed the delocalized regime exhibits level repulsion between eigenvalues, exactly as in the usual random matrix case, while the localized eigenvalues do not repel each other and exhibits Poisson statistics, as if they were independent. This is reasonable, localized states have non-zero values on sites that are typically far apart in real space, hence there is a negligible overlap between two distinct localized eigenvectors.
- As said above $\rho(E)$ is non-zero up to the edges $\pm(2dJ + W/2)$ but it is very small close to them, as schematized on this plot :



In fact the behavior close to the edge is called a Lifshitz tail,

$$\rho(2dJ + W/2 - \delta) \sim A \exp \left[-\frac{B}{\delta^{d/2}} \right] \quad \text{as } \delta \rightarrow 0, \quad (15)$$

where A and B are two positive constants. This is an essential singularity with all the derivatives of ρ being equal to 0. Physically it comes from the fact that all sites in a large cube must have a local energy V_i close to its maximal value $W/2$ to obtain an eigenvalue close to the edge, an event which has a probability exponentially small in the volume of the cube.

3 Conclusions

This localization transition has been the subject of intense research efforts, both experimentally and theoretically, in physics and in mathematics. It is still an active field in various directions, let us mention two of them.

At the level of mathematical rigor it is now well-understood how to prove the existence of localized regimes, at high energies or for large disorder. However the proof of the existence of delocalized eigenvectors in the presence of disorder is restricted to some very special cases.

An important line of research is devoted, in physics and in mathematics, to the so-called Many-Body Localization (MBL) problem. To explain its origin let us first emphasize that we studied in

this lecture the localization of a single quantum particle, isolated from any environment: we used the Schrödinger equation, at zero temperature. If there are many quantum particles, still in a closed system, with no interactions between them, the situation does not change much: the many-body eigenstates are just obtained by symmetrizing or anti-symmetrizing (depending on whether the particles are bosons or fermions) the tensor products of single particle states.

But suppose now that such a disordered system is coupled to an heat bath at temperature $T > 0$. Then no strict localization can survive: arbitrary localization barriers ΔE can be overcome, even with a small probability $e^{-\Delta E/T}$, and hence leads to delocalization (a mechanism known as “variable range hopping”).

Consider finally a very large, isolated system. The basic hypotheses of statistical mechanics (ergodicity, thermalization, . . .) postulate that such a large system acts as an heat bath for its subsystems (provided there are interactions in the system, and in absence of integrability), and hence should be delocalized. The point is that there seems to exist non-integrable interacting systems that have localized states, called MBL systems, hence violating the basic hypotheses of statistical mechanics. Their understanding thus raises fundamental questions in this field.