# Anderson Localization

Jonathan San Miguel
*Department of Physics, Stanford University, Stanford, CA 94305*
(Dated: June 24, 2020)

**Submitted as coursework for PH470, Stanford University, Spring 2020**

We review Anderson's seminal paper "Absence of Localization in Certain Random Lattices", which presents the theory of Anderson localization. On a lattice of non-interacting atoms in three dimensions with disorder scale $W$ and hopping scale $V$, Anderson finds a critical hopping $V_c$ below which states become completely localized. This occurs as long as the couplings fall off spatially faster than $V(r) \sim r^{-3}$. We also briefly review some of the progress and further work since Anderson's paper, mainly following Elihu Abrahams' book *50 Years of Anderson Localization*.



## I. INTRODUCTION

Consider a system of non-interacting atoms, living on a lattice, with a single-particle Hamiltonian given by

$$H_{ij} = \sum_k E_k \delta_{jk} \delta_{ij} + V_{ij}. \tag{1}$$

At each lattice site $i$ is an onsite an energy $E_i$, and between each pair of sites is a hopping term $V_{ij}$. Now say we start with a wavefunction localized near some lattice site. Will it remain there at late times, or will it tend to spread out to other sites?

The latter process is called diffusion, and it has many important implications. A lack of diffusion implies a lack of transport in the system. Such a lack of transport leads to many phenomena, such as metal-insulator transitions and even quantum Hall plateaus[1]. Diffusion is also important for thermalization. If an atom remains localized for all times, the system cannot explore all of phase space, and the ergodic hypothesis cannot be satisfied.

So far, we have not discussed the form of the onsite energies and hopping terms in (1). It is easy to see, however, that if we require these terms to be translationally invariant, then the system will diffuse with even the smallest $V_{ij}$. This is because localized states are extremely spread out in the momentum basis.

In order for diffusion not to occur, translational invariance must be broken. Philip Anderson's 1958 paper, "Absence of Diffusion in Certain Random Lattices," discusses one of the most important examples of this: quenched randomness[1,2]. Anderson looks at Hamiltonian (1) on a 3-dimensional lattice, with random onsite energies falling inside a range $[-W, W]$. The central question of the work is this: how does tuning the ratio of disorder strength to coupling, $W/V_{ij}$, affect diffusion?

To answer this question, Anderson studies several different kinds of coupling $V_{ij}$. He finds that if $V$ falls off with distance $r$ more slowly than $1/r^3$, then diffusion occurs no matter the strength of the disorder. However, if the coupling falls off more quickly, then there is a sharp phase transition: if $V < V_0$ (for a given $W$), no diffusion occurs at all. This $V_0$ cannot be calculated exactly; however, Anderson and others have estimated it numerically for various forms of the coupling.

In this review, we will work through some of the main features of Anderson's paper. We will not present the full derivation, which is quite involved. Instead, we will highlight some of the important insights made by Anderson that first allowed for a detailed understanding of localization. We will also not burden ourselves considerably with some of the mathematical difficulties noted in Anderson's original work. Rather, we hope the reader will be satisfied by the extensive numerical and experimental evidence for the theory. Finally, we will also discuss a fraction of some of the work done since 1958, which has built greatly upon Anderson's original contributions. This last review is mainly based off of Elihu Abrahams' book *50 Years of Anderson Localization*[1], as well as a review of many body localization due to Abanin et al.[3]

## II. METHODOLOGY

In order to understand diffusion, we must first precisely quantify it. Let us label the wavefunction on each site $i$ as $a_i$ Now consider a particle starting with $|a_0| = 1$, $a_i = 0$, for all $i \neq 0$. Then we say that such a state diffuses if, in the infinite-size limit, as $t \to \infty$, $|a_0| \to 0$.

This can be conveniently studied with the Laplace transform, given by

$$f_j(s) = \int_0^\infty dt e^{-st} a_j(t). \tag{2}$$

The Hamiltonian (1) will have several exact energies

$\tilde{E}_k$. Then the amplitude of the eigenstate with energy $\tilde{E}_m$ on site $i$ is given by

$$|a_i|_{\tilde{E}_m} = \lim_{\sigma\to 0} \sigma f_i(\sigma - i\tilde{E}_m). \tag{3}$$

Now say the state starting at 0 is localized. Since these amplitudes are conserved in time, there should be some energy $\tilde{E}$ such that $|a_0|_{\tilde{E}}$ is nonvanishing in the infinite system size limit. If the state diffuses, on the other hand, none of the eigenstates can have finite amplitude on 0.

To understand whether (3) vanishes, we will rely on perturbation theory in the coupling $V_{ij}$. Just as eigenvalues may be expanded perturbatively, the Laplace transform may be written

$$\begin{aligned} f_0(s) =& \frac{i}{is - E_0} + \\ & \frac{1}{is - E_0} \sum_k V_{0k} \times \\ & \left( \frac{V_{k_0}}{is - E_k} + \sum_l \frac{1}{is - E_k} V_{kl} \frac{1}{is - E_l} V_{l0} + ... \right) f_0(s) \\ \equiv & \frac{i}{is - E_0} + \frac{1}{is - E_0} V_c(0) f_0(s). \end{aligned} \tag{4}$$

$V_c(0)$ is the central quantity that Anderson studies. To see its importance, we will substitute (4) into (3), giving

$$|a_0|_{\tilde{E}} = \lim_{\sigma\to 0} \frac{\sigma}{\sigma + iE_0 - i\tilde{E} + iV_c(0)}. \tag{5}$$

Now, say we are probing one of the exact eigenenergies $\tilde{E}$. Then since the amplitude is real, the imaginary parts of the denominator of (5) must cancel, leaving

$$|a_0|_{\tilde{E}} = \lim_{\sigma\to 0} \frac{\sigma}{\sigma - \Im(V_c(0))}. \tag{6}$$

Then the condition for Anderson localization to occur is that $\Im(V_c(0)) \to 0$ as $\sigma \to 0$.

## III. RESULTS

We next to turn to the key results of Anderson's paper, but now using the language of perturbation theory. Anderson splits the perturbation series for $V_c(0)$ (4) into two parts: the lowest order term and all higher terms. While all terms must be examined to assess convergence, the convergence of the first term alone can impose a necessary, though not sufficient, condition. The advantage of this treatment is that the derivation of this condition is much simpler without an infinite series. It will also illuminate some of the critical behavior of the model.

In both parts, Anderson studies $V_c(0)$ as a random variable over ensembles of $E_k$. The goal will be to find results that are true in all but a vanishing set of ensembles, which become increasingly unlikely in the infinite system size limit. In order to do this, we will need the probability distribution of $V_c(0)$.

### A. Lowest Order Term

The lowest order term of $\Im(V_c(0))$ is given by

$$\Im(V_c(0)) = -s \sum_k \frac{|V_{0k}|^2}{s^2 + E_k^2} + ... \equiv -sX(s). \tag{7}$$

The quantity $X(s)$ can be easily shown in the perturbation theory (4) to be

$$X(s) = \sum_{j\neq 0} \frac{|f_j(s)|^2}{|f_0(s)|^2}, \tag{8}$$

to lowest order. From (3), this implies that if $X(s)$ is finite, there is no transport.

Anderson estimates the probability distribution $P(x)$ for coupling that scales as $V_{0k} \sim 1/r^q$, where $r$ is the distance of site $k$ to site 0. The sites are taken to be randomly distributed in space. The on-site energies $E_k$ are assumed to lie in a uniform distribution between $-W$ and $W$.

Two regimes are relevant. In the first, $V$ falls off faster than $r^{-3}$. In this case, the $s \to 0$ limit of $X(s)$ may be taken, and the result is that for large $X$,

$$P(X) \sim \frac{1}{X^{3/2}}. \tag{9}$$

The reader might notice that the average of this function diverges. This does not change the fact, however, that we can make the cumulative probability distribution arbitrarily close to 1 for finite values of $X$. This means if we draw a random $X$ from the distribution, the probability of it diverging is 0. To first order, at least, this means that the state is localized at all times, though perhaps spread out over some finite volume.

On the other hand, if $V$ falls of slower than $r^{-3}$, $P(X)$ itself diverges as $s \to 0$. In this case, regardless of the disorder strength $W$, diffusion occurs.

At the interface between these two regimes is $V(r) = Ar^{-3}$, for some constant $A$. Such functional form is experimentally relevant, corresponding to dipole interactions. Here, Anderson finds that the most probable value of $X$ scales as

$$X_{MP} \sim \left[ \sinh^{-1}\left(\frac{W}{2s}\right) \right]^2. \tag{10}$$

As $s \to 0$, this scales as $\log(s)^2$. This means that in the critical regime, there is still a tendency to diffuse, albeit slowly.

To briefly summarize, the first term of $\Im(V_c(0))$ always converges, as long as $V(r)$ falls off faster than $r^{-3}$. At the moment, this is only a restriction on the functional form of the couplings, and not the strength of the disorder $W$. In order to understand the latter, we will need to move to higher order terms.

### B. Higher Order Terms

The rest of the series is given by

$$V_c(0) = \sum_{k,l} \frac{1}{is - E_k} V_{kl} \frac{1}{is - E_l} V_{l0} + \sum_{k,l,m} \frac{1}{is - E_k} V_{kl} \frac{1}{is - E_l} V_{lm} \frac{1}{is - E_m} V_{m0} + ... \tag{11}$$

As is standard in perturbation theory, we can view this graphically as a series of hops between sites (Figure 1). The series we are interested in now is a sum over all paths of two or more hops.

Unfortunately, the terms in (11) do not all get smaller as we go to more and more hoppings. The reason is that, however unlikely, there may be some specific looping path with a very large amplitude. This path may be repeated endlessly, and as a result, we need to account for its contribution with higher and higher order terms. Such a path is depicted (path A) in figure 1.

In order to resolve this, Anderson resorts to a common technique in perturbation theory: resummation. The goal is to rewrite a series of a similar form to (11), but without repeating any sites. For example, a term in this sum may look like

$$\sum_{\substack{k\neq 0 \\ l\neq 0,k \\ m\neq 0,k,l...}} \frac{1}{e_k} V_{kl} \frac{1}{e_l} V_{lm} \frac{1}{e_m} V_{m0} ... \tag{12}$$

Graphically, such a term would resemble path B in figure 1. It turns out that such a resummation may be done, but with a modification of the denominators $e_k$. If we think of the denominator as a propagator, this is analogous to a mass type renormalization, where the loops correspond to one-particle irreducible diagrams. This renormalization takes the form,

$$e_k = is - E_k - V_c^{0,l,m,\cdots}(j), \tag{13}$$

Where $V_c^{0,l,m,\cdots}(j)$ is defined as before, except excluding any hopping to sites $0, l, m, ...$, the sites to the right of this term in (12). This renormalization is the sum of all processes that start and end on site $j$, except for those including sites that have already occurred in previous terms in the product.

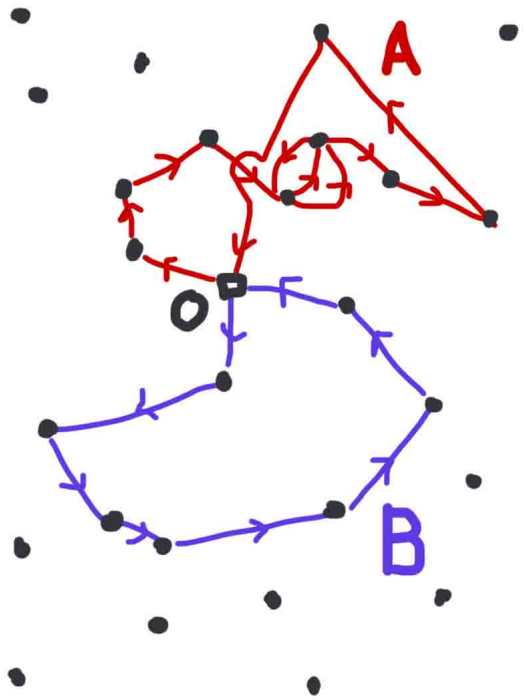


FIG. 1: Possible paths. In the path A, the same loop is repeated multiple times. In path B, there are no repeated lattice sites.[2]

This is still not obviously convergent, although it is clear that we can no longer sum over paths like path A. Indeed, (12) is still a random variable, and in general has a possibility of being quite large. However, Anderson shows that as we increase the number of hops in series of this form, the probability of very large contributions decreases exponentially.

This can be shown in two coupling regimes. The first is a "constant connectivity" system where each site is coupled only to its $K$-nearest neighbors, with a constant coupling $V$. Here, the probability distribution of terms such as (12) is reduced to only studying the distribution of the $e_k$'s, as well as the average number of paths of a given length, which may be estimated with graph theory.

The second regime is more physical: $V(r) \sim r^{-3N}$, with $N > 1$. This case, however, must be estimated with further approximations. In particular, Anderson assumes $V(r)$ to be randomly distributed, even though the restriction on nonrepeating paths from (12) introduces correlations between $r$'s. However, it can be shown that this approximation is a strict overestimate. The second approximation is to introduce a minimum distance $a$ between atoms. While this is an underestimate, it is reasonable to expect this feature in physical systems as well.

Both regimes are studied with the number distribution $n(T)$, where $n(T)dT$ represents the number of terms with $L$ hops with magnitude between $T$ and $T + dT$. In each

case, Anderson finds

$$n(T)dT = F(W/V)^L \frac{dT}{T^2} L(T), \tag{14}$$

Where $F$ is some function of the Hamiltonian parameters, and $L(T)$ varies slowly with $T$ (at large $T$). Additionally, $F$ decreases with large $W/V$. What does (14) mean? Say there is a critical value $V_0$, where

$$F(W/V_c)^L L(1) = 1. \tag{15}$$

Now, say we decrease $V_0$ just slightly below this critical value. Then at large $T$, the number density at values greater than $T = 1$ is now decreasing as $e^{-L}$. In fact, we can do better: the number density above some threshold, $n(T > (Ce)^{-L})$ falls off exponentially. This means that as $L$ gets larger, the terms are, except with vanishing probability, constrained within a smaller and smaller window. The interpretation, of course, is that as we average over more and more paths, small variations in the disorder become more and more insignificant.

At this point, we have only considered the number density for single terms with $L$ hops. However, in our actual series for $V_c(0)$, we will have many terms at each path length (infinitely many, in the case of power law $V(r)$. Let $\Sigma(L)$ be the sum of all terms with $L$ hops. A result from the theory of the Holtsmark distribution[4] is that the probability distribution of the sum of many terms of different signs (with some restrictions on smoothness) is the same as the probability distribution on the largest term. Now, since the number of terms above $e^{-L}$ are decreasing as $e^{-L}$, the probability of $\Sigma(L)$ being order larger than $e^{-L}$ is of order $e^{-L}$.

This means that, with all but vanishing probability, the series (12) will converge in both the case of power law (falling faster than $r^{-3}$) and nearest neighbor couplings, as long as $V$ is smaller than some critical value $V_0$.

As we saw in section A, this critical $V$ has some of the character of a phase transition. In the infinite size limit, below it, almost every state is localized, but above it, almost every state is delocalized. Estimating this parameter, however, is difficult, because the functions $F$ and $L$ are not known exactly. Anderson's analytical results should be seen mainly as a way to show the existence of a localized regime, not as an exact prediction.

Still, Anderson does make numerical calculations to find the value of $V_0$. These numerics are done by iteratively solving (15) for the case of nearest neighbor interactions. The results are in figure 2. Anderson finds that as connectivity $K$ increases, the critical ratio of $W/V$ increases as $K\log(K)$. This general increase can be expected, as increasing the number of nearest neighbors increases the interactions felt by any given site.

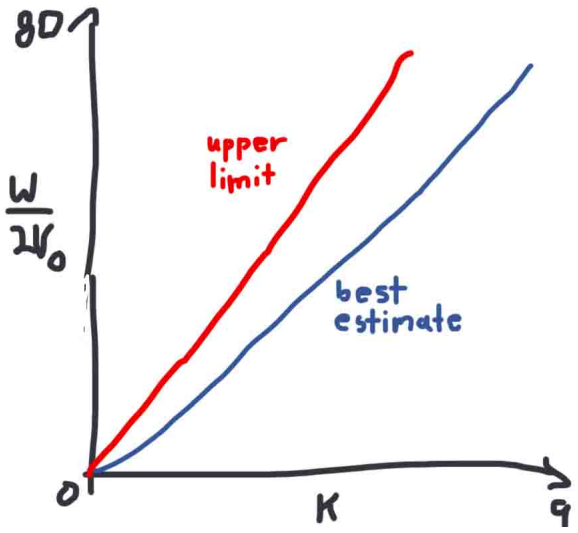


FIG. 2: Numerical estimates of the critical ratio $V/W$ versus connectivity $K$, for nearest-neighbor type couplings.[2]

## IV. CONCLUSIONS AND FURTHER WORK

In "Absence of Diffusion", Anderson shows that the quantity $\Im(V_c(0))$ (4) converges, which in turn implies localization. This is true probabilistically, meaning that while an average state is localized, it may still be possible that some vanishing fraction of states are extended.

There are two requirements for this convergence to occur. First, the coupling $V(r)$ must fall off faster than $r^{-3}$, and second, $V/W$ must be less than some critical value. At this critical value, there is a sharp transition between localization and diffusion.

These are already strong, testable predictions, but since 1958, when Anderson published his seminal work, our understanding of this problem has increased significantly. First, Anderson himself, along with Abrahams, Licciardello, and Ramakrishnan, studied the problem in dimensions other than three, in 1979[5]. They found that in one and two dimensions, any amount of disorder causes localization, and there is no critical value $V_0$.

Second, in the intervening decades, numerics have improved dramatically. Anderson localization has been extensively confirmed in simulations, for example by Yoshino and Ozaki for the two-dimensional case[6], or MacKinnon and Kramer in two and three dimension[7]. Such simulations have also allowed for increasingly precise estimates of the critical exponents near the localization transition, as well as the critical coupling.

Finally, Anderson localization has been experimentally demonstrated. Indeed, one of the original motivations of Anderson's paper was an experiment by G. Feher finding a lack of spin diffusion in silicon[8]. The Hamiltonian (1) is rather general, however, and Anderson localization has been found in many systems, from Bose-Einstein condensates[9] to photonic crystals[10].

When Anderson first proposed his theory, it also led

to many new questions about statistical mechanics. In particular, does Anderson localization violate the ergodic hypothesis? If states fail to diffuse, then how can such a system thermalize?

Answering this point requires the introduction of interactions, which creates a fundamentally different problem[3]. Since Hamiltonian (1) is noninteracting, it is integrable: there are individual conserved quantities for every particle. We cannot meaningfully speak of thermalization or ergodicity in such a setting, as our system will be forever restricted to a small manifold in phase space.

The study of localization in interacting, disordered systems is called many-body localization. This problem has been studied for many decades, going back to the work of Fleishman and Anderson[11]. However, the modern treatment of many-body localization was first given by Basko, Aleiner, and Altshuler in 2006[12,13]. Their work shows that disorder can induce localization in interacting, and, in particular, non-integrable systems. Since then, both experiments[14,15] and numerics[16,17] have lent further credence to the theory of many-body localization, and it remains an important area of study in modern physics.

The theory of Anderson localization can perhaps be characterized by both its great impact and its theoretical difficulty. It was one of the main works cited in Anderson's 1977 Nobel Prize, yet even in his Nobel lecture, Anderson noted that "it has yet to receive adequate mathematical treatment"[18]. Even today, over 40 years after Anderson's Nobel Prize and and 60 years after he first proposed the theory, Anderson localization is still an active research topic, and there remain many unsolved questions. It has also been the progenitor of many important problems in physics, such as many-body localization. Finally, as both numerics and quantum simulation improve, we can expect to continue to learn about Anderson localization well into the future.


[1] E. Abrahams, *50 Years of Anderson Localization* (2010).

[2] P. W. Anderson, Phys. Rev. **109**, 1492 (1958), URL https://link.aps.org/doi/10.1103/PhysRev.109.1492.

[3] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Rev. Mod. Phys. **91**, 021001 (2019), URL https://link.aps.org/doi/10.1103/RevModPhys.91.021001.

[4] R. Klages, G. Radons, and I. M. Sokolov, *Anomalous Transport: Foundations and Applications* (2008), ISBN 978-3-527-40722-4.

[5] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979), URL https://link.aps.org/doi/10.1103/PhysRevLett.42.673.

[6] S. Yoshino and M. Okazaki, Journal of the Physical Society of Japan **43**, 415 (1977), URL https://doi.org/10.1143/JPSJ.43.415.

[7] A. MacKinnon and B. Kramer, Phys. Rev. Lett. **47**, 1546 (1981), URL https://link.aps.org/doi/10.1103/PhysRevLett.47.1546.

[8] G. Feher and E. A. Gere, Phys. Rev. **114**, 1245 (1959), URL https://link.aps.org/doi/10.1103/PhysRev.114.1245.

[9] J. Billy, V. Josse, Z. Zuo, A. Bernard, B. Hambrecht, P. Lugan, D. Clément, L. Sanchez-Palencia, P. Bouyer, and A. Aspect, Nature **453**, 891 (2008), ISSN 1476-4687, URL https://doi.org/10.1038/nature07000.

[10] T. Schwartz, G. Bartal, S. Fishman, and M. Segev, Nature **446**, 52 (2007), ISSN 1476-4687, URL https://doi.org/10.1038/nature05623.

[11] L. Fleishman and P. W. Anderson, Phys. Rev. B **21**, 2366 (1980), URL https://link.aps.org/doi/10.1103/PhysRevB.21.2366.

[12] D. Basko, I. Aleiner, and B. Altshuler, Annals of Physics **321**, 1126 (2006), ISSN 0003-4916, URL http://www.sciencedirect.com/science/article/pii/S0003491605002630.

[13] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, *On the problem of many-body localization* (2006), cond-mat/0602510.

[14] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, Science **349**, 842 (2015), ISSN 0036-8075, URL https://science.sciencemag.org/content/349/6250/842.

[15] A. Lukin, M. Rispoli, R. Schittko, M. E. Tai, A. M. Kaufman, S. Choi, V. Khemani, J. Léonard, and M. Greiner, Science **364**, 256 (2019), ISSN 0036-8075, URL https://science.sciencemag.org/content/364/6437/256.

[16] D. M. Kennes and C. Karrasch, Phys. Rev. B **93**, 245129 (2016), URL https://link.aps.org/doi/10.1103/PhysRevB.93.245129.

[17] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, Phys. Rev. Lett. **113**, 107204 (2014), URL https://link.aps.org/doi/10.1103/PhysRevLett.113.107204.

[18] P. W. Anderson, *Local moments and localized states* (1977).