DISSERTATION

Predicting quantum many-body dynamics out of equilibrium

Lennart Dabelow

30 September 2020

UNIVERSITÄT
BIELEFELD
Fakultät für Physik

Universität Bielefeld
Fakultät für Physik · Theorie der kondensierten Materie

Predicting quantum many-body dynamics out of equilibrium

Understanding how the macroscopically observable behavior of systems with many degrees of freedom emerges from the laws governing their microscopic constituents is an intriguing fundamental problem. We approach this issue by investigating how the dynamics of many-body quantum systems is affected by weak-to-moderate perturbations in three different nonequilibrium setups. First, we study the relaxation towards equilibrium under the influence of time-independent perturbations. Second, we consider so-called echo protocols, where the system relaxes for a certain time followed by an effective time reversal during another period of equal duration, spoiled by small inaccuracies in the state at the point of reversal or in the dynamical laws. Third, we analyze the response to external driving in the form of a time-dependent coupling strength for perturbations of a similar kind as in the first setup. These settings cover a large variety of different phenomena and applications such as the relaxation of system-bath compounds, prethermalization, magnetic resonance imaging, quantum quenches, and periodically modulated external fields. Adopting typicality arguments, we derive analytical predictions for the observable dynamics in all three scenarios. Furthermore, we analyze these predictions and verify them by comparison with numerical and experimental data for several different models and observables.

Wie sich das makroskopisch beobachtbare Verhalten von Systemen mit vielen Freiheitsgraden aus den Gesetzmäßigkeiten ergibt, die deren mikroskopische Bestandteile beschreiben, ist eine faszinierende grundlegende Fragestellung. Dieser nähern wir uns hier, indem wir in drei verschiedenen Nichtgleichgewichts-Szenarien ergründen, wie sich die Dynamik von Quanten-Vielteilchen-Systemen unter dem Einfluss schwacher Störungen ändert. Erstens untersuchen wir den Effekt von zeitunabhängigen Störungen auf die Relaxation ins Gleichgewicht. Zweitens betrachten wir sogenannte Echo-Protokolle, bei denen ein System für eine gewisse Zeit relaxiert, worauf eine Phase gleicher Länge folgt, während der es sich effektiv in umgekehrter Zeitrichtung entwickelt, wobei jedoch kleine Ungenauigkeiten im Zustand am Umkehrzeitpunkt oder in den dynamischen Eigenschaften auftreten können. Drittens analysieren wir, wie ein gegebenes System auf Störungen ähnlich wie im ersten Fall, jedoch mit zeitlich variabler Intensität reagiert. Diese Szenarien decken eine Vielzahl verschiedener Phänomene und Anwendungen ab, zum Beispiel die Relaxation von zusammengesetzten Systemen mit Wärmebad, Präthermalisierung, Magnetresonanztomografie, schnelle Parameteränderungen (engl. „quenches“) oder periodisch modulierte externe Felder. Unter Ausnutzung von Typikalitätsargumenten leiten wir in allen drei Fällen analytische Vorhersagen für die beobachtbare Dynamik her. Darüber hinaus analysieren wir diese Vorhersagen und vergleichen sie mit numerischen und experimentellen Daten für verschiedene Modellsysteme und Observablen.

Dissertation

zur Erlangung des akademischen Grades eines
Dr. rer. nat.

eingereicht am 30.09.2020

von
M. Sc. Lennart Justin Dabelow
ldabelow@physik.uni-bielefeld.de

Gutachter:

1. Prof. Dr. Peter Reimann, Universität Bielefeld (Promotionsbetreuer)
2. Prof. Dr. Jürgen Schnack, Universität Bielefeld
3. Prof. Dr. Jens Eisert, Freie Universität Berlin

Weitere Mitglieder der Prüfungskommission:

4. Prof. Dr. Thomas Dahm, Universität Bielefeld
5. Prof. Dr. Walter Pfeiffer, Universität Bielefeld

Datum der Disputation: 29.01.2021

Gedruckt auf alterungsbeständigem Papier °° ISO 9706

Ich versichere, dass ich vorliegende Dissertation selbständig und nur unter Verwendung der angegebenen Hilfsmittel und Quellen verfasst sowie Zitate und gedankliche Übernahmen kenntlich gemacht habe. Diese Arbeit ist weder vollständig noch in Teilen anderweitig als Prüfungsleistung verwendet und nicht in englischer oder einer anderen Sprache veröffentlicht worden.

Ferner erkläre ich, dass mir die Promotionsordnung der Fakultät für Physik der Universität Bielefeld bekannt ist und keine unmittelbar oder mittelbar geldwerten Leistungen für Arbeiten, die im Zusammenhang mit dem Zustandekommen oder dem Inhalt der Dissertation stehen, an Dritte ergangen sind.

Bielefeld, 30.09.2020

Publications and contributions

Following common practice, parts of the results presented in this thesis have already been published as articles in scientific journals. This concerns, in particular, the results from Chapters 3 and 4, which are partly included in the following publications (in chronological order):

- P. Reimann and L. Dabelow, "Typicality of prethermalization," Phys. Rev. Lett. **122**, 080603 (2019), cf. Chapter 3, chiefly Sec. 3.5;
- L. Dabelow and P. Reimann, "Relaxation theory for perturbed many-body quantum systems versus numerics and experiment," Phys. Rev. Lett. **124**, 120602 (2020), cf. Chapter 3, chiefly Secs. 3.2 through 3.7;
- L. Dabelow and P. Reimann, "Predicting imperfect echo dynamics in many-body quantum systems," Z. Naturforsch. A **75**, 403 (2020), cf. Chapter 4, chiefly Sec. 4.4;
- L. Dabelow and P. Reimann, "Persistent many-body quantum echoes," Phys. Rev. Research **2**, 023216 (2020), cf. Chapter 4, chiefly Secs. 4.3 and 4.5;
- L. Dabelow, P. Vorndamme, and P. Reimann, "Modification of quantum many-body relaxation by perturbations exhibiting a banded matrix structure," Phys. Rev. Research **2**, 033210 (2020), cf. Chapter 3, chiefly Secs. 3.2, 3.4.2, 3.6.2, and 3.7.

As indicated by the publications' authorship, I obtained these and other results in close collaboration with my supervisor Peter Reimann.

The research projects of Chapters 3 and 5 were originally conceived by Peter Reimann and subsequently extended to some degree by me. The project underlying Chapter 4 was mostly conceived by me.

Regarding the derivations, as a rule of thumb, I will report my own contributions in more detail and tend to refer to the published works otherwise. For coherence of the presentation, however, I will not strictly adhere to this principle and will include concepts and calculations contributed by Peter Reimann whenever deemed appropriate. Numerical simulation results are due to me unless they were extracted from third-party publications as indicated.

Contents

1 Motivation **1**

2 Context and concepts **5**
2.1 Quantum and statistical mechanics . . . 5
2.2 Equilibration and thermalization . . . 12
2.2.1 Equilibration . . . 12
2.2.2 Thermalization . . . 16
2.2.3 Absence of thermalization and integrability . . . 19
2.3 Typicality and concentration of measure . . . 23
2.3.1 General considerations . . . 24
2.3.2 Applications . . . 25

3 Perturbed relaxation **32**
3.1 Aims and setup . . . 32
3.2 Prerequisites . . . 35
3.3 Perturbation ensembles . . . 41
3.4 Eigenvector overlap moments . . . 43
3.4.1 Resolvent approach and supersymmetry methods . . . 45
3.4.2 Second moment and overlap distribution . . . 48
3.4.3 Fourth moment . . . 57
3.4.4 Alternative approach and higher-order moments . . . 66
3.5 Prethermalization . . . 68
3.6 Typical time evolution . . . 71
3.6.1 Expectation-value dynamics . . . 72
3.6.2 Response profile . . . 76
3.6.3 Discussion . . . 81
3.7 Examples . . . 85
3.8 Connections, flaws, and prospects . . . 92

4 Echo dynamics **96**
4.1 Echo protocols and irreversibility . . . 96
4.2 Classical examples . . . 99
4.3 Imperfect preparation . . . 102
4.3.1 Typical echo signal . . . 103
4.3.2 Examples . . . 106
4.4 Imperfect reversal . . . 108
4.4.1 Typical echo signal . . . 109
4.4.2 Examples . . . 110
4.5 Combined effect and discussion . . . 111

5 Driven systems **115**
5.1 Setup and methods . . . 115
5.2 Prerequisites and expected applicability . . . 118
5.3 Typical response . . . 122
5.3.1 Prediction for driven time evolution . . . 122
5.3.2 Response profile . . . 123
5.3.3 Validity analysis . . . 125
5.4 Examples . . . 130
5.5 Discussion . . . 133

6 Conclusions **137**

Acknowledgments **141**

A Abbreviations and conventions **142**
A.1 Abbreviations . . . 142

A.2 Concepts . . . 142
A.3 Symbols . . . 144

B Supersymmetry methods **148**
B.1 Anticommuting numbers and graded algebra . . . 148
B.2 Linear algebra in superspaces . . . 149
B.3 Differentiation and integration . . . 150

C Gaussian integrals **153**
C.1 Commuting variables . . . 153
C.2 Anticommuting variables . . . 154
C.3 Supersymmetric variables . . . 155
C.4 Hubbard-Stratonovich transformation . . . 156

D Saddle-point approximation **158**
D.1 Laplace's method . . . 158
D.2 Saddle-point method for complex integrals . . . 159
D.3 Supersymmetric extension . . . 161

E Details on derivations **163**
E.1 Ensemble variance of the dynamics under diagonal perturbations (Sec. 3.2) . . . 163
E.2 Saddle-point integral for the fourth overlap moment (Sec. 3.4.3) . . . 163
E.3 Example for the alternative overlap-moment approximation (Sec. 3.4.4) . . . 167
E.4 Exploiting ensemble properties for the prethermalization bound (Sec. 3.5) . . . 168
E.5 Bound for the remnant term in the ensemble-averaged dynamics (Sec. 3.6.3) . . . 170
E.6 Ensemble variance of the echo signal under imperfect preparation (Sec. 4.3.1) . . . 172

References **175**

1 Motivation

It is a fascinating and ubiquitous phenomenon that the complex interplay of many essentially independent agents can result in astonishingly stable and regular behavior when viewed on a larger scale. One may think of, for example, an organ made up of cells, a colony of ants, or the global economy emerging from individual customers and enterprises. It is generally hopeless to follow the behavior of every single constituent, and nevertheless the functioning of their assembly as a whole can be characterized by relatively simple laws involving a manageable number of variables. Understanding how these effective laws arise from the complicated interactions of the basic entities is an intriguing, yet challenging endeavor.

While the above examples all involved living organisms, the same principles in fact apply to literally everything around us: All objects of our everyday experience, including ourselves, are composed of atoms and molecules. Their individual behavior and interactions are exceedingly complicated and seemingly chaotic, but somehow they conspire to form macroscopic objects and, what is more, we can often predict the behavior of these objects amazingly well without knowing what all the atoms do precisely. Again, it is a fascinating question how this macroscopic regularity emerges from microscopic complexity. The present thesis is supposed to contribute a very tiny piece of this puzzle.

Fortunately, we are actually quite well off as far as the starting conditions for this undertaking are concerned: Unlike in the previous biological and economic examples, we know the microscopic laws that govern the behavior of the basic constituents, atoms and molecules, in remarkable detail. That is to say, there is a full-fledged theory called *quantum mechanics* [1, 2] which, based on a few fundamental principles, allows us to predict the outcome of processes involving just a few atoms with spectacular accuracy. However, basic perceptions of our everyday experience like the elapsing of time or the tendency for imbalances to equalize are not built into those fundamental laws. The aim to understand the origins of the latter two macroscopic phenomena in particular forms the broader context for the subsequently presented research.

Time goes by. Microscopic processes on the level of individual atoms and molecules are *reversible*: If we were presented with a movie showing the interaction of a few atoms, we could not possibly decide whether the film is being played forwards or backwards because processes at this level occur equally likely or frequently in either direction. In the situation depicted in Fig. 1.1a, for example, an atom absorbs a photon and goes into an excited state (left to right). However, the inverse process, whereby an atom in an excited state relaxes and emits a photon (right to left), is equally valid and observable. Given just the sequence of events, neither order is somehow "preferred" by nature. (We remark that the setting is deliberately kept vague here.)

Consider now the process in Fig. 1.1b. A balloon "absorbs" a dart and goes into a rather unexcited state ("pop!"). Confronted with a movie of this event, we would immediately declare this to be the "right" direction, whereas it would be absurd to actually observe the reversed sequence in reality. Yet the entire process could in principle be decomposed into microscopic, reversible steps

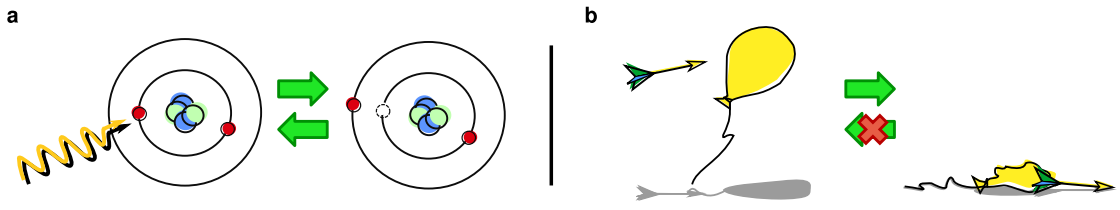


Figure 1.1: Microscopic reversibility and macroscopic irreversibility. **a.** According to the microscopic laws governing the behavior of individual atoms and molecules, fundamental processes are reversible: It is equally possible for an atom to absorb a photon and send one of its electrons to an excited state as it is for an electron in an excited state to relax to the ground state, causing the atom to emit a photon. **b.** Macroscopic processes, which are ultimately composed of many microscopic ones, appear irreversible: A balloon hit by a dart pops and falls to the ground, but we never observe a floppy balloon inflating spontaneously and casting off a dart.

involving, say, only a few atoms each. Hence, again, both directions are equally valid, but one of them is apparently much more common.

The vast majority of processes we experience in our everyday life are of this kind and thus practically *irreversible*: A sequence of events or observations on a system naturally occurs in one direction (i.e., "forward in time"), but it is extremely difficult if not impossible in practice to set things up such that the same events happen in reverse order ("backward in time"). One may also think of a glass being tipped over the edge of a table and shattering on the floor, the blending of two liquids, or a bowl of melting ice cream in a warm room. Particularly the last example illustrates that this preferential direction or "arrow of time" exists regardless of the action or intervention of an external agent such as a human being.

As observed above, however, this arrow of time is not part of the microscopic laws governing the behavior of atoms, molecules, and their constituents. To be more precise, as far as our current understanding of the fundamental laws is concerned, there do exist microscopic time asymmetries in the so-called electroweak interaction, but their effect is far too small to explain the macroscopically observed irreversibility. More importantly, the related so-called CPT symmetry (charge, parity, time reversal) is still bearing up with essentially the same philosophical consequences. Moreover, as we will see, already the frameworks of nonrelativistic quantum mechanics or even classical mechanics (Newton's laws) entail a preferred direction macroscopically despite being perfectly symmetric on the fundamental level. Hence irreversibility apparently emerges as a result of the complex interactions between large numbers of microscopic degrees of freedom.

Striving for equality. A phenomenological characteristic of macroscopic irreversibility is that imbalances in a given setting tend to disappear as time progresses if the system is left on its own. Moreover, once such imbalances have disappeared, they will not re-emerge spontaneously, and instead the dynamics comes to rest. Returning to the ice-cream example, there is a temperature difference between the contents of the bowl and its surroundings initially, but this difference gradually diminishes over time and eventually vanishes, commonly accompanied by the ice cream's melting. As soon as the temperatures are equalized, in turn, the (melted) ice cream will essentially stay "as is" and will not move visibly in the bowl. This is not quite true, of course, because if we wait longer, it will eventually evaporate, meaning that all its molecules will spread somewhat uniformly across the surrounding room (if the door was closed). Put differently, the tendency to equalize continues on larger scales until we have reached an even more balanced state, and the process does not necessarily stop there either. The question of when things are at rest thus depends on the perspective and the properties one is interested in, too, i.e., different properties may balance on different time scales.

The theoretical framework to classify the long-term properties of macroscopic systems is provided by *statistical mechanics* and *thermodynamics* [3–5]. Technically speaking, the system approaches a state of *equilibrium*, characterized by an essential homogeneity and stationarity of its macroscopically perceived properties. Yet the microscopic constituents of the system actually do not ever come to rest at all. The mere fact that a system *has* a temperature, even though it may be the same everywhere, already indicates that its atoms and molecules are constantly jiggling around since temperature is just a measure of the intensity with which they do so. The puzzle is thus, once again, how our macroscopic perception of equilibrium and constancy can be reconciled with the perpetual and seemingly undirected, time-symmetric motion at the microscopic level.

How to relax. Characterizing the state of equilibrium and comprehending how it can appear stationary to us even though things will never calm down microscopically is one aspect of the problem. Indeed, recent years have witnessed remarkable progress regarding our understanding of these equilibrium states, some of which will be reviewed in Chapter 2. Such a characterization of equilibrium, however, is not at the focus of this thesis. Instead, we will investigate how equilibrium is approached in the first place. In terms of the ice-cream example, we are interested in describing the process of melting rather than the final product of ice-cream soup (even though, admittedly, the eventually developed theory will not apply to this particular example involving a phase transition). In more technical terms, we intend to track down how a system that is in a nonequilibrium state at some point in time eventually relaxes to some macroscopically stationary configuration. Specifically, we will start from some macroscopic system whose relaxation behavior in a given setup

is known, and explore how this process is modified if we change the setting slightly, meaning that we perturb or distort the system to some extent.

The first principle goal of this thesis is to explain how equilibrium is reached in such a modified setting based on the known behavior of the original system. In other words, we look for a prediction of the relaxation dynamics of perturbed systems with many degrees of freedom.

At first sight, this may seem a daunting, if not hopeless intent. The unimaginably large number of atoms and molecules in any setting of our macroscopic everyday experience entails that it is plainly impossible to keep track of every individual degree of freedom. Then again, it is a phenomenological fact that the precise microscopic details are seemingly irrelevant with regard to the macroscopic behavior: No repetition of a macroscopic experiment will start out from the exact same microscopic configuration, and yet our ice cream keeps melting again and again. Furthermore, somewhat different substances typically behave similarly, too—as far as the melting is concerned, we could not care less whether our ice cream is chocolate or strawberry. On the other hand, a billiard ball will certainly behave differently from a ball of ice cream under otherwise identical circumstances, even though both are just lumps of atoms, if you will. Hence some of the microscopic details actually do matter, and the crucial point is to separate the relevant from the irrelevant ones.

The general strategy by which we will arrive at predictions of the relaxation process exploits this indifference to certain microscopic details. The idea is to consider classes of systems which share all the relevant characteristics for the question under study, but are otherwise as general as possible. In practice, unfortunately, we cannot follow this ideal path exactly, but we will nonetheless obtain descriptions for the overwhelming majority of systems from suitable classes, in the sense that the fraction of exceptions becomes ever smaller the larger the size of the system is.

Being driven. Besides the relaxation process, we will investigate how changing certain properties over time affects the dynamics in a given setup. In this case, the system is not left on its own, but rather is exposed to time-dependent variations of some of its parameters, for instance due to externally applied forces.

Our second principal goal is to characterize the response of many-body quantum systems to such time-dependent driving. Adopting a similar approach as for the relaxation process under time-independent perturbations, we will consider large classes with similar driving characteristics simultaneously in order to establish a prediction for the observable dynamics.

Going back in time. In a third part, we will address the issue of macroscopic irreversibility more directly. The general idea is to compare the usual observable relaxation of a large system with a (perhaps hypothetical) time-reversed evolution. As mentioned before, the time-reversed process is equally valid and follows the same microscopic laws. To understand why this "backward process" is still not observed macroscopically, we will examine how small inaccuracies in the reversed setup spoil the ability to rewind the original (forward) dynamics.

The third principal goal is thus to assess the stability of microscopic reversibility in many-body quantum systems. The key results will again be characterizations of the typical modifications the time-reversed dynamics exhibits due to different types of inaccuracies.

Fact checking. In all three settings, we thus obtain theoretical predictions for the dynamics of systems with many degrees of freedom. Like any analytical theory, these predictions can ultimately demonstrate their value only by standing up to thorough testing. Common strategies to do so are experiments or numerical simulations. Experiments certainly provide the more direct connection to the real world. Yet, given that the laws describing the individual constituents of a system are well established experimentally, simulations of these laws become numerical experiments and can be an equally valid and sometimes more viable way to test the theory for larger system sizes. However, for both experiments and simulations, the large number of degrees of freedom poses serious challenges. Experimentally, it becomes exceedingly difficult to prepare and maintain a controlled environment that confines a well-defined number of atoms and isolates them from any external perturbations. Numerically, calculations quickly exceed memory capacities because the fundamental laws dictate that the number of variables needed to fully describe a system grows exponentially with the degrees of freedom. Verifying the predictions in a controlled way for macroscopic systems consisting of roughly 10^{23} individual atoms is thus practically impossible.

Fortunately, whereas large numbers of microscopic constituents are assumed in the derivations, it turns out that a few tens of degrees of freedom can be sufficient to put the theory to test. Thanks to impressive experimental advances in recent years as well as ever increasing computational capabilities and sophisticated simulation techniques, we thus can and will validate our predictions by comparison with both concrete real-world and numerical experiments.

Coming up next. The present introduction was meant to give a rough and generally accessible overview regarding the motivation and background of the questions studied in this thesis. We will proceed in Chapter 2 to put these considerations on firmer theoretical grounds. Notably, we will properly define various concepts introduced in layman's terms above, e.g., the microscopic theory of atoms and molecules, the notion of equilibrium, or the simultaneous description of classes of similar systems. Moreover, since this is not the first study of the relaxation behavior in large systems, Chapter 2 will provide a brief overview of the general context and pertinent previous developments as well.

The principal results of this thesis will be presented in Chapters 3 through 5. As mentioned above, these results constitute predictions for the dynamical behavior of large systems which are somehow modified or perturbed from a certain known reference scenario. In Chapter 3, we will target the relaxation process in such a setting and describe how systems that are left on their own approach equilibrium if they exhibit some sort of imbalances initially. In Chapter 4, we will address more specifically the origins of the macroscopically observed irreversibility by studying how small imperfections in the preparation of a system or in the dynamical laws impede a reversal of time even though it is principally allowed on the microscopic level. In Chapter 5, we will investigate the response of large systems to external driving, i.e., a forced, time-dependent variation of some ambient or intrinsic properties.

While each individual chapter will entail conclusions about the corresponding results, the findings will be summarized and discussed in a broader context in the closing Chapter 6.

Finally, there are also various appendices, most of which supply technical background and details of the calculations. Appendix A, however, may turn out useful as a general reference of abbreviations and conventions used throughout this thesis, comprising, in particular, a list of repeatedly used symbols.

2 Context and concepts

Before the game can start, we need to agree on a playing field, select our players, and devise a line-up and strategy. The present chapter is supposed to do precisely this: We will specify the physical setting we intend to describe and introduce its mathematical modeling (Sec. 2.1). The appropriate theoretical frameworks in this context are *quantum mechanics* as well as *thermodynamics* and *statistical mechanics*. Thereafter, we will review important concepts and recurring assumptions for the analysis in the subsequent chapters (Sec. 2.2), providing context to the questions studied in this thesis. Of particular relevance are the notions of *equilibration* and *thermalization* since they arguably describe the standard relaxation paths in macroscopic systems. Finally, we will sketch the principal methods employed in the ensuing derivations, notably the *typicality* approach (Sec. 2.3). Upon collection of these pertinent conceptual ideas, we will include a brief and undoubtedly subjectively biased account of their history, focusing on aspects most relevant with respect to the intended investigation. A broader and somewhat more detailed overview of these and related concepts may be found, for instance, in the reviews [6–10].

2.1 Quantum and statistical mechanics

Microscopic degrees of freedom. As announced in the introduction, the research goal of this thesis is to achieve a better understanding of how the macroscopically observed behavior of many-body systems emerges from the laws governing their microscopic constituents. In principle, it may already be disputable what these microscopic constituents are. The Standard Model of elementary particle physics [11–13] postulates that all of matter, at least as we presently know it, is composed of *quarks* and *leptons*, with masses provided via the *Higgs boson* and interactions mediated by *gauge bosons*. While this Standard Model offers an utterly precise description of many fundamental processes, it is well-known to be incomplete, still, prominently missing, in particular, a reconciliation with general relativity and a description of ominous dark matter and dark energy. Moreover, describing processes of our macroscopic experience in terms of those elementary particles is a plainly hopeless endeavor.

At the length and energy scales of our everyday experience, quarks practically exist exclusively in bound states of *protons* and *neutrons*, and the only relevant lepton is the *electron*. The theoretical framework for their description is *nonrelativistic quantum mechanics* [1, 2], which has been validated to successfully model the behavior of atoms and molecules in innumerable experiments since its conception in the beginning of the last century. Hence we will take this quantum mechanical formalism as our "fundamental" laws, i.e., as the starting point for the analysis. Of course, one would perhaps want to derive these laws from an even more fundamental relativistic quantum field theory, but that is a different story [14].

State, observable, and Hamiltonian. The mathematical backbone of any quantum mechanical description is a *Hilbert space* $\mathcal{H}$, i.e., a complex vector space with an inner product $\langle \cdot | \cdot \rangle$. Its elements are denoted by $|\psi\rangle \in \mathcal{H}$ and encode the *pure states* that the system under study can assume. For this purpose, the $|\psi\rangle$ are taken to be normalized such that $\||\psi\rangle\|^2 := \langle\psi|\psi\rangle = 1$. More generally, we will also allow so-called *mixed states* $\rho = \sum_n p_n |\psi_n\rangle\langle\psi_n|$, where the $|\psi_n\rangle$ are arbitrary (normalized) state vectors, the $\langle\psi_n|$ are the associated dual vectors, and $p_n \in [0,1]$ such that $\sum_n p_n = 1$. In the following, we will almost always use such density operators ρ to denote the state of the system, which may generally be either pure (i.e., $\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle \in \mathcal{H}$) or mixed.

The abstract Hilbert space is connected to the physical world by means of linear operators on that space, notably self-adjoint operators which model (in principle) measurable properties of the system such as particle positions, momenta, energy, magnetization, etc. These are called *observables*. Being self-adjoint, the spectrum of any such observable $A : \mathcal{H} \to \mathcal{H}$ is real, i.e., all eigenvalues are real numbers. Furthermore, the corresponding eigenvectors form a complete set of states from which an orthonormal *basis* of the Hilbert space $\mathcal{H}$, the so-called *eigenbasis* of A,

can be chosen, meaning that all $|\psi\rangle \in \mathcal{H}$ can be decomposed as a linear combination of the basis vectors. When performing a measurement of the observable A on a system in the state ρ, the outcome is one of the eigenvalues a_n of A, realized with probability $\mathrm{tr}[\rho\Pi_n]$, where Π_n is the projection operator onto the eigenspace of A corresponding to the eigenvalue a_n, i.e., $\Pi_n^2 = \Pi_n$ and $A\Pi_n|\psi\rangle = a_n\Pi_n|\psi\rangle$ for all $|\psi\rangle \in \mathcal{H}$. Due to this intrinsic probabilistic nature of quantum mechanics, it is generally not possible to predict the result of a measurement on a system in a state ρ with certainty. Of particular interest is therefore the average outcome of such a measurement, the *expectation value*

$$\langle A\rangle_\rho := \mathrm{tr}[\rho A] = \sum_n a_n \,\mathrm{tr}[\rho\Pi_n]\,, \tag{2.1}$$

where the sum on the right-hand side is over all distinct eigenspaces of A. The observable expectation value (2.1) is the first of three specific types of "averages" we will encounter in the following. We remark that, in principle, the spectrum of A may have continuous parts, so the sum in (2.1) may be supplemented by an integral contribution, but such technicalities will be of no importance for our purposes.

A particularly important observable is the energy operator or *Hamiltonian* $H : \mathcal{H} \to \mathcal{H}$. It encodes the system's degrees of freedom and their interactions, usually in the form of kinetic and potential energy contributions. The systems we are interested in are in general large, but finite (finite number of particles, finite volume, ...). In this case the spectrum of H is discrete, i.e., there exists an at most countably infinite set $\{|n\rangle\}$ of eigenvectors with eigenvalues E_n, the possible energy values of the system. Moreover, this spectrum is bounded from below, meaning that there exists a state $|0\rangle$ with energy E_0, called the *ground state*, such that $E_n \geq E_0$ for all n. Lastly, we will usually assume that the spectrum of H is *nondegenerate*, so the eigenvalues of all $|n\rangle$ are pairwise distinct, which is generically the case for interacting many-body systems. As explained above Eq. (2.1), the probability to measure the energy E_n on a system in the state ρ is thus given by $\mathrm{tr}[\rho\Pi_n] = \langle n|\rho|n\rangle =: \rho_{nn}$, and ρ_{nn} is called the *population* or *occupation* of the energy level E_n. Note that we will use the notation $A_{mn} := \langle m|A|n\rangle$ to denote the matrix elements of an arbitrary observable A in the eigenbasis of the Hamiltonian H in this section.

The mathematical formalities such as the precise structure of the relevant Hilbert space, normalization and positive definiteness of density operators, self-adjointness of observables, etc. will be tacitly taken for granted in the following, and the same holds for many additional formal aspects not even touched upon here, e.g., the existence of orthonormal basis vectors, the spectral decomposition of unbounded operators, and so on. In particular, a physical system is thus usually defined via its Hamiltonian, and the underlying Hilbert space is understood implicitly.

Local and few-body observables. A basic feature of our fundamental understanding of nature is that interactions are *local* and of *few-body* type. *Locality* roughly means that the strength of the interaction between two elementary degrees of freedom ("particles") decreases with their spatial separation. This is often modeled by considering space as discretized by introducing a lattice of accessible sites for the particles together with a distance measure on that lattice. Local interactions can then typically be described by operators supported on a finite, nonextensive (independent of the system size) number of sites. Sometimes so-called "quasilocal" operators are employed instead, for which the interaction strength decays exponentially or faster with the distance between the involved sites. The *few-body character* is expressed by the fact that interactions can usually be broken down to contributions involving only a small number of degrees of freedom (often just two), meaning that a corresponding few-body operator only probes or modifies a few degrees of freedom. Typical Hamiltonians are therefore sums of local and few-body operators, and the same holds for commonly measurable quantities. We thus refer to such operators as *physical observables*.

We emphasize, however, that this labeling does not imply that other types of observables are physically meaningless. For instance, many symmetry properties entail conserved quantities that are given by nonlocal or many-body operators, and theoretical concepts may sometimes be best characterized by nonlocal or many-body observables. Hence we merely wish to express the fact that such "unphysical" observables usually defy direct measurement. Furthermore, we remark that the naming conventions are not consistent in the literature, e.g., "local" can sometimes mean "few-body" in our sense as well.

Time evolution. Besides characterizing the energy, the Hamiltonian also mediates the time evolution of the system (in the absence of measurements): The density operator $\rho(t)$ describing the system's state at time t satisfies the *Liouville-von Neumann equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) = \frac{\mathrm{i}}{\hbar}\,[\rho(t), H]\,, \tag{2.2}$$

where $\hbar$ is the reduced Planck constant, i is the imaginary unit, and H may in general depend on t, too. Moreover, $[A, B] := AB - BA$ denotes the *commutator* of the operators A and B. In the following, we will exclusively employ units with $\hbar = 1$ (see also Appendix A for an overview of the abbreviations and conventions used throughout this thesis). Given the state $\rho(t_0)$ at an arbitrary point in time t_0, the state $\rho(t)$ at any other time t can thus be obtained by integrating the Liouville-von Neumann equation (2.2), again as long as there are no measurements performed on the system. The relationship between $\rho(t)$ and $\rho(t_0)$ can be formally expressed as $\rho(t) = \mathcal{U}(t, t_0)\,\rho(t_0)\,\mathcal{U}(t, t_0)^\dagger$ by introducing the unitary *time evolution operator* or *propagator* $\mathcal{U}(t, t_0)$ satisfying

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{U}(t, t_0) = -\mathrm{i}\,H\,\mathcal{U}(t, t_0)\,, \quad \mathcal{U}(t_0, t_0) = \mathbb{1} \tag{2.3}$$

as well as $\mathcal{U}(t, t_0) = \mathcal{U}(t, t_1)\mathcal{U}(t_1, t_0)$ and $\mathcal{U}(t, t_0)^{-1} = \mathcal{U}(t, t_0)^\dagger = \mathcal{U}(t_0, t)$. Here $\mathbb{1}$ is the identity operator on $\mathcal{H}$, $\mathcal{U}^{-1}$ and $\mathcal{U}^\dagger$ denote the inverse and adjoint operators of $\mathcal{U}$, respectively, and $t_1 \in \mathbb{R}$ is an arbitrary third time point. We will commonly take the reference time $t_0 = 0$ and also write $\mathcal{U}(t) := \mathcal{U}(t, 0)$.

As indicated above, the unitary time evolution mediated by $\mathcal{U}(t, t_0)$ is interrupted by measurements of the system. Notably, these measurements render the dynamics irreversible because the state of the system *after* a measurement of the observable A that yielded the eigenvalue a_n is obtained by projecting onto the corresponding eigenspace, $\rho(t) \mapsto \Pi_n \rho(t) \Pi_n$, meaning that it is generally impossible to restore the information about the pre-measurement state. The subtleties of this measurement process (let alone its interpretation and philosophical implications [15]) will be of no concern in this thesis. Instead, we will exclusively deal with time-dependent expectation values $\langle A\rangle_{\rho(t)}$ (cf. Eq. (2.1)) of the unitarily evolved state $\rho(t)$, thereby addressing the question what a measurement of the observable A would yield on average if we were to perform it at time t. In other words, we consider the system to be left on its own, completely isolated from external influences, and ask how properties which we could observe in principle change with time.

A particularly important case with regard to the isolated many-body quantum systems we intend to study are Hamiltonians H which are *time independent*, meaning that the degrees of freedom and their interactions are constant in time. In this case, the propagator from (2.3) is obtained straightforwardly as $\mathcal{U}(t, t_0) = \mathrm{e}^{-\mathrm{i}H(t-t_0)}$, and the states $\rho(t)$ and $\rho(t_0)$ are related by $\rho(t) = \mathrm{e}^{-\mathrm{i}H(t-t_0)}\rho(t_0)\,\mathrm{e}^{\mathrm{i}H(t-t_0)}$, i.e., we immediately have a formal solution of the Liouville-von Neumann equation (2.2). Evaluating the trace (2.1) with $\rho = \rho(t)$ in the eigenbasis of H (see above) and recalling the notation $A_{mn} = \langle m|A|n\rangle$ and $\rho_{mn}(t) = \langle m|\rho(t)|n\rangle$, the time-dependent expectation value can thus be written as

$$\langle A\rangle_{\rho(t)} = \sum_{m,n} \mathrm{e}^{\mathrm{i}(E_n - E_m)(t-t_0)}\,\rho_{mn}(t_0) A_{nm}\,. \tag{2.4}$$

A special class of observables are self-adjoint operators Q that *commute* with the Hamiltonian H, meaning that $[H, Q] = 0$. In this case, there exists a common eigenbasis for H and Q so that $Q_{mn} = \langle m|Q|n\rangle = \delta_{mn} q_n$ with the Kronecker delta δ_{mn} and the eigenvalues q_n of Q. From Eq. (2.4), we understand that $\langle Q\rangle_{\rho(t)} = \langle Q\rangle_{\rho(t_0)} = \text{const}$, i.e., Q is a *conserved quantity* or, by analogy with classical mechanics, an *integral of motion*. Obviously, the (time-independent) Hamiltonian H is itself a conserved quantity, reflecting *conservation of energy*.

The recurrent key goal of the present thesis is to make the abstract general solution (2.4) for the time-dependent expectation values $\langle A\rangle_{\rho(t)}$ of the (nonconserved) observable A more concrete in specific setups by exploiting generic properties of isolated many-body systems, notably their large number of degrees of freedom and well-defined macroscopic energy. In particular, Chapter 3 and essentially also Chapter 4 will deal with time-independent Hamiltonians and thus start from Eq. (2.4) to devise predictions for the system dynamics. Nevertheless, we will also consider so-called driven systems with explicitly time-dependent Hamiltonians in Chapter 5.

Equilibrium thermodynamics and statistical mechanics. Aiming at a description of macroscopic systems, we may ultimately have in mind degrees of freedom f on the order of 10^{23} or larger. However, the results obtained in the following usually hold (in good approximation) in considerably smaller systems, too. Notably, the specific numerical and experimental examples with which we will compare our theoretical predictions are usually much smaller, on the order of at most a hundred degrees of freedom. The term *many-body system* is thus defined rather loosely to span a wide range of system sizes.

Basic properties of macroscopic systems are described by the theoretical framework of *thermodynamics* and *statistical mechanics* [3–5]. It is a well-established phenomenological observation that many-body systems commonly tend to equilibrate over time if left on their own, meaning that, even if they exhibit a possibly complicated time dependence initially, the values of macroscopically measurable parameters (e.g., energy, pressure, magnetization, ...) eventually become stationary. If all these macroscopic properties remain constant, the system is said to be *in equilibrium*. Understanding how this one-way dynamics towards equilibrium emerges from the microscopically reversible unitary time evolution is the grand motivation for all the more specific questions investigated in Chapters 3 through 5. Likewise, the concepts to be introduced in the subsequent Secs. 2.2 and 2.3 address aspects of that puzzle.

Even if the system is in equilibrium, the microscopic constituents will usually still undergo complicated dynamics, i.e., the microscopic configuration or *microstate* will not be stationary at all, although the macroscopically perceived status or *macrostate* does not change. Consequently, a given macrostate can have many possible microscopic realizations. Generally speaking, a macrostate thus formulates constraints on the set of all microstates, typically in the form of fixed values for certain *state variables*, i.e., certain system properties or control parameters such as energy, temperature, volume, pressure, etc. In the standard formalism of equilibrium statistical mechanics, this leads to the notion of so-called *thermodynamic ensembles*, which constitute classes of microstates supplemented with a probability distribution depending on the state variables. The important example of an *isolated system*, for instance, is characterized by constant total energy $\mathcal{E}$, particle number $\mathcal{N}$, and volume $\mathcal{V}$. (In principle, the macrostate may be characterized by additional variables/constraints, but we tacitly restrict ourselves to the traditional ones here.) Within the quantum mechanical description, the compatible microstates should thus be eigenstates of the associated operators with eigenvalues $\mathcal{E}$, $\mathcal{N}$, and $\mathcal{V}$. More generally, one might consider macroscopically small windows $[\mathcal{E},\mathcal{E}+\Delta_\mathcal{E}]$, $[\mathcal{N},\mathcal{N}+\Delta_\mathcal{N}]$, $[\mathcal{V},\mathcal{V}+\Delta_\mathcal{V}]$ within which the admissible microstate should lie, reflecting the finite precision of any macroscopic measurement.

A basic assertion of statistical mechanics then is that the equilibrium properties of the system can be determined by averaging over all microstates in the ensemble according to the corresponding probability distribution. For isolated systems, for example, this means taking the expectation value with respect to the so-called *microcanonical density operator*

$$\rho_\mathrm{mc} := \Pi_{\mathcal{E},\mathcal{N},\mathcal{V}}/\Omega(\mathcal{E},\mathcal{N},\mathcal{V})\,. \tag{2.5}$$

Here $\Pi_{\mathcal{E},\mathcal{N},\mathcal{V}}$ is the projector onto the space of all compatible microstates, and $\Omega(\mathcal{E},\mathcal{N},\mathcal{V})$ is the dimension of this subspace. Hence the microcanonical ensemble treats all microstates conforming with the macroscopic constraints on equal footing, which is known as the "principle of equal *a priori* probabilities" and seems reasonable in the absence of any additional knowledge about the equilibrium state. The same idea can be formulated equivalently as a "maximum entropy principle:" The *von Neumann entropy*

$$\mathcal{S}(\rho) := -k_\mathrm{B}\,\mathrm{tr}[\rho\,\ln\rho] \tag{2.6}$$

with Boltzmann's constant k_B can be understood as a quantifier of uncertainty in a (macroscopic) density operator ρ about the comprised microstates: A larger value of $\mathcal{S}(\rho)$ corresponds to a less detailed specification and thus greater uncertainty about the actual microstate. These notions have been formalized and generalized in the context of *information theory* [16]. The microcanonical density operator ρ_mc maximizes the von Neumann entropy in the sense that $\mathcal{S}(\rho_\mathrm{mc}) \geq S(\rho)$ for all density operators ρ that can be constructed from the admissible microstates. Hence ρ_mc is the macroscopic state that assumes the least about the compatible microstates, resonating with Occam's razor. Observing that

$$\mathcal{S}(\rho_\mathrm{mc}) = k_\mathrm{B}\,\ln\Omega(\mathcal{E},\mathcal{N},\mathcal{V}) =: \mathcal{S}(\mathcal{E},\mathcal{N},\mathcal{V})\,, \tag{2.7}$$

one readily recovers the *Boltzmann entropy*,[1] for which we use the same symbol $\mathcal{S}$, but a change of perspective is entailed in considering it to be a function of the thermodynamic state variables characterizing the macrostate. Encoding the equilibrium state as its maximum, the entropy is an instance of a *thermodynamic potential*, pertaining to isolated systems with fixed $\mathcal{E}$, $\mathcal{N}$, $\mathcal{V}$.

The state variables $\mathcal{E}$, $\mathcal{N}$, and $\mathcal{V}$ characterizing isolated systems are all *extensive*, meaning that their magnitude is roughly proportional to the degrees of freedom f. It is sometimes convenient to work with *intensive* state variables instead, whose values are independent of the system size. For every extensive variable, there exists a conjugated intensive variable describing the physical quantity that will be balanced upon equilibration if two (or more) subsystems can exchange the corresponding extensive quantity. For example, the conjugate variable of the energy $\mathcal{E}$ is the temperature T (more precisely, the inverse temperature $\beta = 1/k_{\mathrm{B}}T$): The equilibrium configuration of a composite system consisting of two subsystems which can exchange energy is given by a state with equal temperatures in the two subsystems. Similarly, the chemical potential μ and the pressure p are conjugated to the particle number $\mathcal{N}$ and the volume $\mathcal{V}$, respectively. These changes of state variables call for changes of the pertinent thermodynamic potential, which are mediated mathematically by Legendre transformations. Instead of an isolated system we may, for example, consider a *closed system*, which still has fixed particle number $\mathcal{N}$ and volume $\mathcal{V}$, but can exchange energy with its environment (commonly called a reservoir or *heat bath*). The Legendre transform of the (dimensionless) entropy $\tilde{\mathcal{S}} := \mathcal{S}/k_{\mathrm{B}}$ with respect to the energy $\mathcal{E}$ is the (dimensionless) free energy $\tilde{\mathcal{F}}(\beta, \mathcal{N}, \mathcal{V}) := \beta\mathcal{E} - \tilde{\mathcal{S}}(\mathcal{E}, \mathcal{N}, \mathcal{V})$, where the inverse temperature is obtained via $\beta := \partial\tilde{\mathcal{S}}(\mathcal{E}, \mathcal{N}, \mathcal{V})/\partial\mathcal{E}$, and this relation is to be used as well to express $\mathcal{E}$ as a function of β, $\mathcal{N}$, and $\mathcal{V}$ in the definition of $\tilde{\mathcal{F}}(\beta, \mathcal{N}, \mathcal{V})$. Note that it is historically more common to work with the free energy $\mathcal{F} := \tilde{\mathcal{F}}/\beta$ instead, which satisfies $\mathcal{F} = \mathcal{E} - T\mathcal{S}$, where $T := k_{\mathrm{B}}/\beta$ is the temperature. The equilibrium state of the closed system is then given by the state that minimizes the free energy, leading to the *canonical density operator* or *Gibbs ensemble*

$$\rho_{\mathrm{can}} := \mathrm{e}^{-\beta H}/Z \tag{2.8}$$

with the *(canonical) partition function* $Z := \mathrm{tr}(\mathrm{e}^{-\beta H})$. Alternatively, the canonical ensemble (2.8) arises again by maximizing the von Neumann entropy (2.6), but instead of restricting to microstates with a given energy value (within a macroscopically small window) as in the microcanonical setup, we require that the *average* energy should be fixed, $\langle H \rangle_\rho = \mathcal{E}$. The dimensionless free energy is then simply the (negative) Lagrange function of the maximization problem for $\tilde{\mathcal{S}}(\rho)$ under this constraint and β is the corresponding Lagrange multiplier, i.e., $\tilde{\mathcal{F}}(\rho) = -\tilde{\mathcal{S}}(\rho) + \beta \langle H \rangle_\rho$. The properly normalized density operator solving this optimization problem is indeed ρ_{can} from (2.8).

Similar transformations may be employed for the other extensive state variables such as $\mathcal{N}$ and $\mathcal{V}$, and the associated intensive variables again arise either as conjugated variables of the Legendre transform or Lagrange multipliers for the related constrained optimization problem. Moreover, at least for sufficiently large systems, all the thermodynamic ensembles obtained this way are essentially equivalent: As explained above, the extensive variables either assume similar values for all individual microstates or they are fixed on average. But since the corresponding probability distributions become extremely narrow if many degrees of freedom are involved [4], fixing the average is practically the same as restricting to a small window of admissible eigenvalues. We will re-encounter this *concentration of measure* property in Sec. 2.3 in the form of the "typicality method," which in turn will be the basis for all the main results obtained in this thesis. Although we introduced the microcanonical formalism as a description of isolated systems, the canonical formalism for closed systems, etc., it is thus in fact a matter of convenience which one to choose, provided that the system is sufficiently large, which is essentially a prerequisite for the statistical treatment to be reasonable in the first place.

Standard thermodynamics thus characterizes the equilibrium properties of large systems. In this thesis, however, we are not so much interested in these equilibrium properties, but rather in the relaxation process which takes a system initially out of equilibrium to a stationary state that is potentially described by a thermodynamic ensemble. In a first step, we should therefore explore

[1] Note that the correspondence between von Neumann's and Boltzmann's entropies only holds in equilibrium. An individual (pure) microstate $\rho = |\psi\rangle\langle\psi|$, for instance, has vanishing von Neumann entropy, whereas the Boltzmann entropy is that of the corresponding macrostate.

under which circumstances many-body quantum systems show such a relaxation behavior at all. This will be the focus of the ensuing Sec. 2.2. Before, however, we will briefly collect the essential properties of the *isolated many-body quantum systems* we intend to study in the following and comment on experimental platforms which allow to probe quantum many-body phenomena in a controlled environment.

Isolated many-body quantum systems. The degrees of freedom of a quantum system and their interactions are encoded in the Hamiltonian H. As explained above, a key property of isolated many-body systems is that they exhibit a well-defined macroscopic energy, which implies that there exists a macroscopically small *energy window*

$$I_{\mathcal{E}} := [\mathcal{E}, \mathcal{E} + \Delta_{\mathcal{E}}] \tag{2.9}$$

comprising all energy levels E_n of H that are significantly populated by the state $\rho(t)$ at any time. Note that for a time-independent Hamiltonian, the occupations $\rho_{nn}(t)$ (see below Eq. (2.1)) are independent of t, too. In view of (2.1) and observing that $|\rho_{mn}(t)|^2 \leq \rho_{mm}(t)\,\rho_{nn}(t)$ due to the Cauchy-Schwarz inequality, occupations outside of $I_{\mathcal{E}}$ are therefore negligible with regard to the dynamics of expectation values. For the questions of interest in this thesis, we can thus usually restrict ourselves to the N-dimensional Hilbert space $\mathcal{H}_{\mathcal{E}} := \mathrm{span}\{|n\rangle : E_n \in I_{\mathcal{E}}\}$ called the *energy shell*. In general, there may exist further conserved quantities besides H, i.e., physical observables Q such that $[H, Q] = 0$, e.g. the number of particles, the magnetization, or other state variables of traditional thermodynamics. In the spirit of the above introduced thermodynamic ensembles, such a conserved quantity Q should then be accounted for in the choice of the energy shell $\mathcal{H}_{\mathcal{E}} \equiv \mathcal{H}_{\mathcal{E},\mathcal{Q}}$, meaning that $\langle n|Q|n\rangle \in I_{\mathcal{Q}}$ for some macroscopically small window $I_{\mathcal{Q}} = [\mathcal{Q}, \mathcal{Q} + \Delta_{\mathcal{Q}}]$; see also Sec. 2.2.3 below for additional comments on the role of conserved quantities.

As a consequence of the system's many-body character, the energy spectrum is extremely dense. Therefore, as long as we are not extremely close to the ground state or a possible upper end of the spectrum, the energy window $I_{\mathcal{E}}$ is still microscopically large in the sense that the number of levels N with $E_n \in I_{\mathcal{E}}$ (i.e., the dimension of $\mathcal{H}_{\mathcal{E}}$) is exponentially large in the system's degrees of freedom f [4],

$$N = 10^{\mathcal{O}(f)} \gg 1\,. \tag{2.10}$$

The distribution of the energy levels within the window $I_{\mathcal{E}}$ is encoded in the *density of states* (DOS)

$$D(E) := \sum_{n:E_n \in I_{\mathcal{E}}} \delta(E - E_n)\,. \tag{2.11}$$

Given an arbitrary function $h(E)$, summations over energy levels within $I_{\mathcal{E}}$ may thus be expressed as integrals over the DOS, i.e.,

$$\sum_{n:E_n \in I_{\mathcal{E}}} h(E_n) = \int \mathrm{d}E\, D(E)\, h(E)\,. \tag{2.12}$$

Due to the extremely dense spectrum, $D(E)$ can usually be approximated excellently by a smooth function, notably if $h(E)$ in (2.12) is sufficiently slowly varying in E. In this case, which we will practically always take for granted in the subsequent investigations, $\int \mathrm{d}E\, D(E)\, h(E)$ becomes a proper integral. In fact, we will frequently (but not always) assume that the energy window is sufficiently small so that $D(E)$ can be well approximated by a constant, i.e.,

$$D(E) \approx \varepsilon^{-1} \tag{2.13}$$

with the *mean level spacing* ε. The assumption that sums over E_n can be approximated by an integral as in (2.12) with $D(E) = \varepsilon^{-1}$ then holds as long as $h(E)$ is a slowly varying function of E compared to the mean level spacing ε.

Finally, the exceedingly high level density usually implies that it is virtually impossible to populate only a few energy levels significantly, i.e., even the utmost careful experimental preparation will have a finite precision that is still much larger than the mean level spacing ε [17, 18]. Consequently, the occupations $\rho_{nn}(0)$ will generally be distributed across a large number of energy levels, typically on the order of the dimension N of the appropriate energy window from (2.10)

and thus exponentially large in the degrees of freedom f [17–21]. Since $\sum_n \rho_{nn}(0) = 1$, the largest individual population $p_{\max}$ will thus be much smaller than unity,

$$p_{\max} := \max_n \rho_{nn}(0) = 10^{-\mathcal{O}(f)} \ll 1\,. \tag{2.14}$$

A noteworthy exception occurs if there is a gap between the ground state and the first excited state of the many-body spectrum [20, 22], but as insinuated above Eq. (2.10), states too close to the edges of the spectrum will usually be excluded from our discussion anyway. A related concept to quantify the extent to which the state $\rho(0)$ spreads across the energy levels of the Hamiltonian H is the *effective dimension* [23]

$$N_{\text{eff}} := \left(\sum\nolimits_n \rho_{nn}(0)^2\right)^{-1}\,, \tag{2.15}$$

which estimates the number of energy levels which appreciably contribute to the state $\rho(0)$. In particular, the condition $p_{\max} \ll 1$ from (2.14) is equivalent to $N_{\text{eff}} \gg 1$ [20].

To conclude, we remark that the basic reason for focusing on isolated systems in the following is again our intention to understand relaxation from a fundamental point of view within the validity of nonrelativistic quantum mechanics. It is undoubted that physical systems are almost never perfectly isolated in practical applications and may thus call for a modified theoretical modeling, e.g., as an open system [24, 25]. However, such theoretical approaches are usually somehow phenomenologically motivated and not "purely quantum." From a foundational point of view, it is thus desirable to concentrate on isolated systems, which can be described in a self-contained way within the fundamental framework of quantum mechanics. Moreover, the important example of a system of interest coupled to a larger environment ("bath") can be embedded naturally into the adopted formalism by considering the joint system-plus-bath compound as an isolated "supersystem" and investigating the relaxation behavior of observables supported on the smaller system of interest. Not least, impressive experimental improvements in recent years have managed to isolate systems of reasonably many degrees of freedom reasonably well for reasonably long times [26–33], so the predictions developed here are indeed experimentally testable.

Experimental and numerical explorations. Maintaining a controlled and truly isolated testbed of sufficiently many microscopic degrees of freedom to probe quantum statistical properties and especially many-body dynamics is a highly nontrivial task. A very fruitful and versatile approach utilizes *ultracold atoms* confined by *magnetic* and/or *optical traps* [26–28, 30, 34]. In these setups, magnetic field gradients or counterpropagating laser beams are employed to generate spatially varying magnetic or electric dipole potentials for dilute gases of neutral atoms in a vacuum chamber at millikelvin temperatures or below. Particle densities and momentum distributions can then be assessed by probing fluorescence or absorption of photons directly in the trap or in time-of-flight expansions, and further observables are accessible in many situations by means of sophisticated imaging techniques [30, 34]. Being the most matured technology to control many-body quantum systems with a good degree of flexibility, such cold-atom experiments have so far been the principal tool to validate the theoretical concepts that will be presented below in real-world systems. In particular, these techniques allow to set up one-, two-, or three-dimensional lattice systems of bosons or fermions with tunable interaction strengths and can thereby serve as *quantum simulators* [35, 36] for a variety of popular condensed-matter models.

Nuclear or electronic spins of certain crystals or molecules can also provide a way to study quantum many-body dynamics, albeit with less detailed control over the individual degrees of freedom. Notably, it is still possible to adjust the effective interaction strength in such setups by applying elaborate pulse sequences of external magnetic fields, prominently employed in so-called *magic-* or *polarization-echo experiments* [37–44] as well as, more recently, to monitor *out-of-time-ordered correlators* [45–47]. We will come back to those echo experiments, in particular, in Chapter 4.

Another potentially versatile type of quantum simulators is offered by universal *quantum computers* [35, 36, 48, 49], which consist of a set of two-state systems ("qubits") with the possibility to apply essentially arbitrary unitary transformations ("gates") on their joint Hilbert space and to initialize and measure their state at the beginning and end of the transformation, respectively. Promising setups to realize such highly manipulable qubits include Josephson junctions [50–52] and

trapped ions [53, 54] (see also Ref. [36] for an overview of flexible quantum-simulator approaches in particular). While the presently available devices still suffer from relatively short coherence times due to insufficient isolation, they have already been employed in proof-of-principle studies to simulate many-body dynamics, in particular, albeit with low accuracy (see, for instance, Refs. [54–57]). Thus, given the impressive progress in recent years, universal quantum computers may become a powerful way to explore largely arbitrary and truly many-body model systems in the future.

Finally, numerical simulations can help exploring explicit models and verifying theoretical predictions. Naturally, such simulations provide less direct evidence than an actual experiment. Nevertheless, the fact that quantum mechanics as the general framework for describing the behavior of individual atoms and molecules is extremely well secured experimentally ensures that simulations building on that framework can commonly be taken as a reliable complement to experiments. Not least, agreement between the two is usually pretty good in situations where both are feasible (see, for example, Refs. [32, 58–62]). Yet modeling the behavior of many-body systems on a computer comes with its own challenges, most notably the exponential growth of the Hilbert space with the degrees of freedom. Thanks to sophisticated algorithms, it is still possible to reach into regimes in which the many-body character becomes decisive. Suitable methods for the simulation of many-body dynamics include exact diagonalization, time-evolving block decimation [63–65] and time-dependent density-matrix renormalization-group calculations [66–68], numerical linked-cluster expansions [69–71], dynamical mean-field theory [72], or dynamical typicality [71, 73–75] (see also Sec. 2.3.2 below).

2.2 Equilibration and thermalization

While traditional thermodynamics describes the equilibrium properties of large generic systems, it does not satisfactorily answer the equally important questions of why and how these systems reach thermal equilibrium in the first place. In this section, we review key results that explain aspects of these issues based on a quantum mechanical modeling of many-body systems. As is common practice, we split the problem into two parts: *Equilibration* (Sec. 2.2.1) refers to the questions in which sense and under which circumstances an initially out-of-equilibrium system reaches a stationary state as time progresses. *Thermalization* (Sec. 2.2.2) deals with the question of whether or not this stationary state is in agreement with the predictions about thermal equilibrium from traditional thermodynamics. Situations where the answer to the latter question is negative lead to interesting and "unexpected" types of nonthermalizing systems (Sec. 2.2.3). An important class of results from this context obtained by means of so-called typicality methods will be omitted for the most part here and deferred to the subsequent Sec. 2.3.

2.2.1 Equilibration

Definition. Generally speaking, a system is said to *equilibrate* if its dynamics becomes stationary at long times. The quasiperiodic nature of time evolution in quantum mechanics (see, e.g., Eq. (2.4) and Refs. [76, 77]) implies that such equilibration cannot occur in the strict sense that the state $\rho(t)$ asymptotically approaches a well-defined limit as $t \to \infty$. In fact, this is not a peculiarity of quantum mechanics and already arises in classical (Hamiltonian) dynamics, prominently quantified, for example, by the *Poincaré recurrence theorem* [3]. Yet the time scales on which such revivals or nonequilibrium fluctuations occur are usually unimaginably large in generic many-body systems, and observable system properties often become essentially stationary after an initial relaxation phase.

In the literature, a few different notions of *equilibration* can be found to formalize this "quasistationarity," including, among others, equilibration of subsystems [19, 21, 78], during intervals [7], or with respect to distinguishability measures [23, 79, 80], local relaxation [81–83], or proximity to macroscopic equilibrium subspaces [84–87]. For our purposes, a suitable definition is that a system with Hamiltonian H, prepared in the state $\rho(0)$ at $t = 0$, is said to equilibrate if there exists a state $\bar{\rho}$ such that the time-dependent expectation values $\langle A \rangle_{\rho(t)}$ of *experimentally realistic*

observables A (see below) are practically indistinguishable from or at least very close to $\langle A\rangle_{\bar\rho}$ for nearly all later times t [17, 20, 23]. In particular, fluctuations are supposed to decrease as the system size is increased. Nonequilibrium expectation values should therefore become exceedingly small or exceedingly rare at sufficiently late times. The state $\bar\rho$ then coincides, by construction, with the *time-averaged* (or *dephased*) state

$$\overline{\rho(t)} := \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau \mathrm{d}t\; \rho(t)\,, \tag{2.16}$$

provided that this average is well-defined. Here we implicitly introduced time averages as the second important type of averaging procedure occurring in this thesis, for which we generally reserve the overbar notation.

If the Hamiltonian H is nondegenerate, we can exploit its eigenbasis $\{|n\rangle\}$ to write $\bar\rho := \overline{\rho(t)}$ in the form (see also Eq. (2.4) and the discussion above it)

$$\bar\rho = \sum_n \rho_{nn}(0)\, |n\rangle\langle n|\,, \tag{2.17}$$

hence $\bar\rho$ is just the time-independent component of $\rho(t)$. Note that in case of a degenerate H, we can still choose the basis states $|n\rangle$ such that the initial state $\rho(0)$ (or, more precisely, its matrix representation $\rho_{mn}(0) = \langle m|\rho(0)|n\rangle$) is diagonal within every degenerate subspace of H, so that the time-averaged expectation value $\overline{\langle A\rangle_{\rho(t)}}$ still takes the form $\langle A\rangle_{\bar\rho}$ with $\bar\rho$ from (2.17). In view of Eq. (2.17), the time-averaged state $\bar\rho$ is also sometimes called the *diagonal ensemble*.

Experimentally realistic observables. The above definition of equilibration referred to "experimentally realistic" observables whose expectation values should become quasistationary. Loosely speaking, this means that the observable A should model a (macroscopic) measurement apparatus as it could be used to probe a certain property of the many-body system under study [17, 18, 88]. Indeed, without such a restriction, it is always possible to construct, for any given H and $\rho(0)$, observables which do not become stationary, e.g., as a superposition of projectors onto two energy levels with distinct eigenvalues.

A precise characterization of such experimentally realistic observables is far from trivial. Nevertheless, there are a few properties that appear generally reasonable and will be exploited in the following. The first of these properties concerns the range of possible measurement outcomes, i.e., the *spectral range*

$$\Delta A := \sup \mathrm{spec}(A) - \inf \mathrm{spec}(A)\,, \tag{2.18}$$

where $\sup S$ and $\inf S$ denote the supremum and infimum of the set S, respectively, and $\mathrm{spec}(A)$ is the spectrum of the operator A. Any realistic measurement instrument can only yield a finite value of the measured quantity, implying that the spectral range should be finite, $\Delta A < \infty$.

A second important property is that any realistic measurement has a finite (nonzero) *resolution*

$$\delta A := \inf\{|a - a'| : a, a' \in \mathrm{spec}(A), a \neq a'\} > 0\,, \tag{2.19}$$

limited, for example, by the number of significant digits. Combining these two properties, the observable A should thus have a finite number of distinct eigenvalues, on the order of $\Delta A/\delta A$, and even for the most precise instruments, this number will still be exceedingly small compared to the dimension of the relevant Hilbert space from (2.10).

We remark that there is no direct relation between the present notion of experimentally realistic observables and "physical observables" as defined in Sec. 2.1. One might generally expect that an experimentally realistic observable should also be physical, but the macroscopic nature of the measurement apparatus introduces additional subtleties since the latter can usually probe the microscopic degrees of freedom only on a coarse-grained level. On the other hand, physical observables in the sense of Sec. 2.1, i.e., observables composed from local and few-body operators acting on the microscopic constituents, may in principle have an unbounded or continuous spectrum. In any case, physical observables should generally be approximable by experimentally realistic ones, i.e., for a reasonable physical observable there should—in principle—exist an instrument with a finite range and resolution to measure it.

Conditions for equilibration. Equilibration in the above-defined sense (see above Eq. (2.16)) can be shown to occur very generically in systems with many degrees of freedom by bounding the fraction of time spent out of equilibrium and thereby demonstrating that it becomes exceedingly small if one waits long enough. Besides the number of distinguishable measurement outcomes $\Delta A/\delta A$ (see Eqs. (2.18) and (2.19)), important quantifiers involved in such bounds are the maximal population of an individual energy level $p_{\max}$ from (2.14) (or, equivalently, the effective dimension N_{eff} from (2.15)) and the largest degeneracy g_H of an *energy gap* (i.e., the difference between two distinct energy levels) of H. For generic interacting systems, g_H will be close to unity. The appearance of this quantity can be understood by inspection of the general time evolution (2.4), which reveals that the energy gaps determine the frequencies at which the individual terms in the sum oscillate. In case of a degenerate gap, the oscillations from two (or more) different terms will always be in phase and thus do not cancel out in the time average. Likewise, a stronger suppression of fluctuations around the time average can be expected if more levels are involved in the dynamics, i.e., if $p_{\max}$ is small or N_{eff} is large.

For a preset time $\tau > 0$, we now consider the accumulated duration $\tau_{\delta A}$ of time periods within the interval $[0, \tau]$ for which $|\langle A\rangle_{\rho(t)} - \langle A\rangle_{\bar{\rho}}| \geq \delta A$, i.e., for which deviations of the current expectation value from the time average are larger than the measurement resolution. Equilibration of experimentally realistic observables can then be established by means of the following bound as shown by Reimann and Kastner in Ref. [20], extending related previous works by Reimann [17, 18] and Short and Farrelly [23, 79]: There exists $\hat{\tau} > 0$ such that for all $\tau \geq \hat{\tau}$,

$$\frac{\tau_{\delta A}}{\tau} \leq 6\, g_H \left(\frac{\Delta A}{\delta A}\right)^2 p_{\max}\,. \tag{2.20}$$

Recalling that $\Delta A/\delta A$ will be on the order of 10^D for a measurement with D digits precision (see below Eq. (2.19)), whereas $p_{\max} = 10^{-\mathcal{O}(f)}$ according to (2.14), the fraction of times with exceptional, detectably nonstationary expectation values is indeed extremely small. In addition, it should be pointed out that the bound can actually be tightened upon replacing $p_{\max}$ by the second-largest occupation $\rho_{nn}(0)$ and in fact holds even if H exhibits degenerate energy levels [20].

A precursor of the result (2.20) and its relatives can already be found in the early work [89] by Tasaki, whose seminal character is evident from the fact that it also touches upon other important pillars of our present understanding of equilibration and thermalization such as eigenstate thermalization and typicality (see Secs. 2.2.2 and 2.3.2 below). It studies the canonical setup of a small system S of interest with Hamiltonian H_{S} coupled to a large bath B with Hamiltonian H_{B} (see above Eq. (2.8)), assuming an *ad hoc* coupling between the two components in the form of an operator V that is almost diagonal in the eigenbasis of the Hamiltonian $H_0 = H_{\text{S}} + H_{\text{B}}$ of the noninteracting compound system. Regarding the question of equilibration, the key result of Ref. [89] is that, starting from a certain class of initial states, the time-dependent expectation values of observables supported on S are indistinguishable from the long-time average, quantified in a bound similar to (2.20).

We remark that the aforementioned notions of equilibration of subsystems [19, 78] and equilibration with respect to distinguishability measures [23, 79, 80] are closely related to and in some sense entailed in the presently discussed equilibration of expectation values [9, 79]. Notably, the introduction of distinguishability measures to quantify the difference between the time-evolved state $\rho(t)$ and the supposed equilibrium state $\bar{\rho}$ remedies the legitimate critique [79] that expectation values alone cannot establish (quasi)stationarity of the system because they only consider *average* measurement outcomes.

Finally, reinforcing the generality of equilibration from a complementary point of view, it can be argued that it is extremely costly and demanding to prepare a quantum many-body system such that it does *not* equilibrate [90].

Relaxation time scales. Whereas equilibration in the sense of quasistationary observable expectation values at long times can thus be taken for granted generically, the results from Refs. [17, 19, 23, 78, 79] and particularly Eq. (2.20) provide little information about the time scales on which

relaxation to equilibrium ordinarily takes place. That is to say, these results demonstrate equilibration after sufficiently late times τ, but the associated values of τ are usually much larger than the actually observed relaxation time and thus "unrealistically late" [7]. The reason is that the setup in those references is extremely general, assuming as little as possible about the systems under study, so the results incidentally apply to some "pathological" cases as well, which can exhaust the pertinent bounds, but are also hard to exclude within a reasonably general treatment [7, 9, 79, 86, 91]. More realistic estimates for the time scales thus call for physically motivated additional constraints on H, A, or $\rho(0)$, but will inevitably forfeit generality.

A first line of investigations considers certain classes of Hamiltonians [92–98], observables [91], or nonequilibrium subspaces [99] and assesses the corresponding average relaxation times within such a class. By establishing a suitable bound for the variations among the individual members, this average can then be considered "generic" for a particular class. The reasoning here follows the typicality framework, which will be introduced in more detail in Sec. 2.3 below. This approach commonly results in extremely and often unrealistically short relaxation times close to the fastest possible scales, i.e., the inverse energy range of the initial state. The main reason is that the employed classes of consolidated systems do not have a common notion of locality and other interaction features, implying that their effect is disregarded in the obtained estimates. Nevertheless, this method can yield impressive agreement with numerical and experimental examples [96, 97], particularly if the setup does not exhibit macroscopic spatial inhomogeneities ("transportless equilibration" [100]).

There are a few studies investigating equilibration time scales for still rather broad classes of systems by different methods, too. For quantum gases of noninteracting fermions or bosons, for instance, it is argued in Ref. [101] that equilibration occurs at least on a time scale polynomial in the degrees of freedom f (as opposed to the typically exponential scaling in the general bounds similar to (2.20)). Ref. [102] considers a largely general setting at first and links the equilibration time scale to the dispersion of energy gaps (i.e., the variance among them, weighted according to the observable- and state-dependent relevance of the associated frequency for the dynamics), suggesting a similar dephasing mechanism for equilibration as the (off-diagonal) *eigenstate thermalization hypothesis* (see Sec. 2.2.2 below). For local lattice Hamiltonians and local initial states, the expected equilibration time can then be related to estimates for the inverse energy ranges of the considered observable and the initial state. A quite general and supposedly tighter upper bound on the relaxation time incorporating properties of the initial state and the observable was presented recently in Ref. [103], even though some of the underlying assumptions have been challenged particularly in setups with slow exponential relaxation characteristics such as weakly coupled system-bath compounds [104].

As insinuated above, the locality of physical interactions, essentially meaning that the interaction strength of the degrees of freedom decays with their spatial separation, has presumably not been taken into account satisfactorily in estimates of equilibration times in general. Its importance, however, is unquestioned and illustrated, for instance, by so-called *Lieb-Robinson bounds* [7, 105], which quantify how information or local perturbations can propagate through a lattice system of spins or fermions. In particular, the commutator between observables supported on different regions, which is exponentially suppressed in their spatial separation initially, can grow only gradually as time evolves, at a speed determined by the local terms of the Hamiltonian and the entailed connectivity of the lattice. Interesting approaches to incorporate such a lattice geometry were developed recently using a classical version of Lieb-Robinson bounds [106] and random-graph ensembles [107], demonstrating indeed a correlation between equilibration times and indicators of locality, and leading to time scale estimates roughly on the order of the observed behavior of concrete models.

Finally, we mention that there are plenty of numerical investigations of equilibration times in concrete model systems, see Ref. [7] and references therein for examples.

In conclusion, equilibration time scales have been assessed for a variety of different scenarios and specific models, but the general mechanisms are still far from being understood entirely. We will contribute to this catalog of time-scale estimates for specific setups in Chapter 3, in particular, where typically expected relaxation times are obtained as a by-product of our analytical predictions for the relaxation dynamics of perturbed many-body quantum systems.

2.2.2 Thermalization

Definition. Having established that generic isolated many-body quantum systems equilibrate, a natural next question concerns the characterization of the equilibrium state $\bar{\rho}$. According to textbook statistical mechanics (see Sec. 2.1), we would commonly expect that $\bar{\rho}$ coincides with an appropriate thermodynamic ensemble, e.g., the microcanonical density operator ρ_{mc} from (2.5). In the same spirit as for the concept of equilibration above, however, it is more sensible to consider expectation values of experimentally realistic or physical observables A. Hence we stipulate to say that an isolated many-body quantum system *thermalizes* if it equilibrates to a state $\bar{\rho}$ and if the equilibrium expectation values $\langle A\rangle_{\bar{\rho}}$ are practically indistinguishable from or at least very close to the pertinent thermal values, e.g.,

$$\langle A\rangle_{\bar{\rho}} \simeq \langle A\rangle_{\rho_{\mathrm{mc}}} . \tag{2.21}$$

Again, slightly different notions of thermalization can be found in the literature [7, 21, 83, 84], usually resulting from slightly different notions of equilibration (see above Eq. (2.16)). We will briefly inspect one particular alternative based on suitably defined macrostates below.

The phenomenological success of statistical mechanics and thermodynamics suggests that thermalization should be the generic behavior of macroscopic systems, but a microscopic derivation is highly nontrivial and in fact hardly possible without further assumptions. A standard line of reasoning in classical mechanics involves the *ergodic hypothesis* whereby a macroscopic system is postulated to visit each accessible phase space volume (compatible with the macroscopic constraints) for a time proportional to its relative size in the long run [108]. In other words, the system is supposed to spend an equal amount of time in each accessible microstate so that the time-averaged state coincides with the microcanonical ensemble average as introduced in and below Eq. (2.5). While the ergodic hypothesis can be proven in a few special cases (see, e.g., Refs. [109–111]), it has the status of a hypothesis to date and is in fact not sufficient nor strictly necessary for equilibration or thermalization [112, 113]. Apart from that, the requirement that the time-averaged state and the microcanonical state coincide seems overly restrictive if we are ultimately interested in expectation values of experimentally realistic or physical observables only.

Furthermore, a similar hypothesis requiring $\bar{\rho} = \rho_{\mathrm{mc}}$ is meaningless for quantum systems because the unitary time evolution under a time-independent Hamiltonian implies that the occupations $\rho_{nn}(t)$ of the energy eigenstates are constant in time (see below Eq. (2.9)), hence an initial imbalance due to nonequilibrium conditions will persist *ad infinitum*. Notwithstanding, concepts like *eigenstate thermalization* [114–116], the *quantum ergodic theorem* [84, 85, 117], and *canonical typicality* [118, 119] shed some light on the emergence of thermal behavior in quantum systems, and one may in fact even argue that the general understanding is nowadays somewhat more profound in the quantum than in the classical setting. We will explain the first two of the above-mentioned concepts, eigenstate thermalization and the quantum ergodic theorem, in the remainder of this subsection. Canonical typicality will be elucidated in Sec. 2.3.2.

Macroscopic thermal equilibrium. A slightly different take on thermalization in isolated many-body quantum systems, which emphasizes the macroscopic character of our perception, was put forward by von Neumann [117] and intensely studied recently by Goldstein and co-workers [84, 85, 99, 120, 121]. Following Ref. [121], we refer to this notion as *macroscopic thermal equilibrium*. The starting point is an energy shell $\mathcal{H}_{\mathcal{E}}$ (see below Eq. (2.9)) of dimension N, a pure state $|\psi(t)\rangle = \mathrm{e}^{-\mathrm{i}Ht}|\psi(0)\rangle \in \mathcal{H}_{\mathcal{E}}$ describing the state of the isolated system at time t, and a set of mutually commuting observables $M_1, \ldots, M_K$. These observables represent macroscopic measurements that characterize a macrostate of the system, meaning that they are coarse-grained variants of microscopic physical observables. The issue of whether such a set of commuting $M_1, \ldots, M_K$ exists is mathematically highly nontrivial [122], but physically reasonable since the order of measurements should not play a role macroscopically. We also note that similar notions of equilibrium can be defined without the need for commuting macro-observables [87, 123], but for simplicity we will stick with the set of commuting operators $M_1, \ldots, M_K$. A macrostate is then characterized by a tuple $\boldsymbol{m} = (m_1, \ldots, m_K)$ of eigenvalues of the $M_1, \ldots, M_K$, and there is an associated subspace $\mathcal{H}_{\mathcal{E}}^{(\boldsymbol{m})}$ of $\mathcal{H}_{\mathcal{E}}$ spanned by the corresponding eigenvectors. The projector onto $\mathcal{H}_{\mathcal{E}}^{(\boldsymbol{m})}$ is denoted by $\Pi_{\boldsymbol{m}}$, the dimension by $N_{\boldsymbol{m}}$, and collecting all macrospaces we recover the full energy-shell Hilbert space $\mathcal{H}_{\mathcal{E}} = \bigoplus_{\boldsymbol{m}} \mathcal{H}_{\mathcal{E}}^{(\boldsymbol{m})}$ as their direct sum. It is furthermore asserted that one of the macrospaces

will vastly dominate this decomposition in the sense that $N_{\boldsymbol{m}}/N \approx 1$, which is called the *equilibrium subspace* $\mathcal{H}_{\mathcal{E}}^{(\mathrm{eq})}$ with projector Π_{eq} and dimension N_{eq} [85, 120]. Indeed, if an equilibrium macrospace exists, which can usually be taken for granted from phenomenological evidence, then its entropy should be close to the microcanonical entropy and thus its dimension must be close to the dimension of the energy shell [120]. Within this setting, the system is then said to be in (macroscopic) thermal equilibrium at time t if

$$\langle\psi(t)|\Pi_{\mathrm{eq}}|\psi(t)\rangle \approx 1\,. \tag{2.22}$$

For any macroscopic observable of the form $A = \sum_{\boldsymbol{m}} a_{\boldsymbol{m}}\Pi_{\boldsymbol{m}}$, the condition (2.22) together with $N_{\mathrm{eq}}/N \approx 1$ immediately implies $\langle\psi(t)|A|\psi(t)\rangle \approx \langle A\rangle_{\rho_{\mathrm{mc}}}$, where $\rho_{\mathrm{mc}} = \mathbb{1}/N$ on $\mathcal{H}_{\mathcal{E}}$ as usual. The conclusions drawn from the definition (2.22) of thermal equilibrium are thus similar to those of our previous definition in and above (2.21) (called "microscopic thermal equilibrium" in [121]), i.e., macroscopically feasible measurements should yield values indistinguishable from the thermal predictions. However, the two notions are not equivalent, especially because different models for the macroscopically observable quantities are employed; for a comparison, see also Refs. [9, 85].

In terms of the definition (2.22) of thermal equilibrium, the question of whether a system thermalizes then boils down to whether Eq. (2.22) holds for most times t. This question is addressed by the *quantum ergodic theorem*, originally due to von Neumann [84, 117] and recently revived and extended by Goldstein et al. [85, 120]; see Refs. [84, 85] for an account of the tragic misconception that would almost have caused von Neumann's original work [117] to fall into oblivion. The different variants of the theorem entail that, given a Hamiltonian H with nondegenerate energy gaps, most decompositions of $\mathcal{H}_{\mathcal{E}}$ into macrospaces $\mathcal{H}_{\mathcal{E}}^{(\boldsymbol{m})}$ imply that all initial states $|\psi(0)\rangle \in \mathcal{H}_{\mathcal{E}}$ satisfy $\langle\psi(t)|\Pi_{\boldsymbol{m}}|\psi(t)\rangle \approx N_{\boldsymbol{m}}/N$ for most times t. Choosing $\Pi_{\boldsymbol{m}} = \Pi_{\mathrm{eq}}$ and exploiting $N_{\mathrm{eq}}/N \approx 1$, this readily leads to thermalization in the sense of (2.22). The notion of "most decompositions" in the above statement refers to a typicality argument (see also Sec. 2.3), meaning that the fraction of decompositions for which the statement is violated is exceedingly small in the degrees of freedom. Without further knowledge about the macroscopic observables $M_1, \ldots, M_K$, one would thus expect that the actually appropriate choice for them (which is not unique either) satisfies the statement, too. The general idea of such reasoning will be outlined in more detail in Sec. 2.3.1.

This already concludes our discussion of macroscopic thermal equilibrium. For the remainder of this thesis, thermalization will be understood in the sense formulated around Eq. (2.21), without reference to macrostates or decompositions of the energy shell. Note that an adaptation of von Neumann's methodology and results to this context can be found in Ref. [124].

Eigenstate thermalization hypothesis. Returning to our initial definition of thermalization in terms of matching expectation values in the time-averaged and thermal states, the quantity of interest given an observable A is the stationary value $\langle A\rangle_{\bar{\rho}}$, which, by adopting (2.17), reads

$$\langle A\rangle_{\bar{\rho}} = \sum_n \rho_{nn}(0)A_{nn} \tag{2.23}$$

when expressed in the eigenbasis of the Hamiltonian H. To characterize thermalization, we thus need criteria for when Eq. (2.23) coincides with the thermal expectation value $\langle A\rangle_{\rho_{\mathrm{mc}}}$, which may incorporate properties of the initial state $\rho(0)$ and/or the observable A in general.

The *eigenstate thermalization hypothesis* (ETH) addresses the observable in particular. The main idea can be expressed as the conjecture that, for physical or at least few-body observables (cf. Sec. 2.1), every single energy eigenstate $|n\rangle$ represents thermal equilibrium in the sense that $\langle n|A|n\rangle \simeq \langle A\rangle_{\rho_{\mathrm{mc}}}$, where ρ_{mc} is the microcanonical density operator corresponding to the energy E_n. In an isolated system, where $\rho_{nn}(0)$ is nonnegligible only if $E_n \approx \mathcal{E}$ (see the discussion around Eq. (2.9)), Eq. (2.23) then immediately reduces to the microcanonical prediction $\langle A\rangle_{\rho_{\mathrm{mc}}}$ at energy $\mathcal{E}$.

This general mechanism was put forward and argued for in seminal works by Deutsch [114] using a random-matrix/typicality approach (see also Sec. 2.3.2) and by Srednicki [115] studying a gas of hard spheres. It also emerges in Tasaki's study [89] of certain classes of system-bath compounds (see also below Eq. (2.20)). A similar idea had actually already been formulated somewhat earlier by

Jensen and Shankar [125] based on numerical observations in an Ising-type spin-$\frac{1}{2}$ chain. Another precursor and presumable source of inspiration of those later studies is Berry's conjecture [126], which states that quantum-mechanical energy eigenfunctions of classically chaotic systems attain thermal character in the semiclassical limit.

The arguably most widely employed formulation of the ETH today is again due to Srednicki [127] and also includes a characterization of the off-diagonal matrix elements. It postulates that the $A_{mn} = \langle m|A|n\rangle$ of a physical observable A in the eigenbasis of a generic physical Hamiltonian H can be described by the ansatz

$$A_{mn} = \mathcal{A}(\hat{E})\,\delta_{mn} + \mathrm{e}^{-\mathcal{S}(\hat{E})/2k_\mathrm{B}}\, f_A(\hat{E},\omega)\, R_{mn}\,, \tag{2.24}$$

where $\hat{E} := (E_m + E_n)/2$, $\omega := E_m - E_n$, $\mathcal{A}(E)$ and $f_A(E,\omega)$ are smooth functions of their arguments, $\mathcal{S}(E)$ is Boltzmann's entropy (2.7), and the R_{mn} exhibit statistical properties similar to the entries of a random symmetric or Hermitian matrix whose independent entries have vanishing mean and unit variance. Since $\mathcal{S}(E)/k_\mathrm{B} = \ln N = \mathcal{O}(f)$ according to (2.7) and (2.10), Eq. (2.24) thus asserts that the diagonal matrix elements A_{nn} exhibit a basically smooth dependence on energy with fluctuations exponentially suppressed in the degrees of freedom. This is also known as the *diagonal ETH*. To leading order, both the long-time average $\langle A\rangle_{\bar{\rho}}$ from (2.23) and the thermal average $\langle A\rangle_{\rho_\mathrm{mc}}$ using (2.5) thus coincide with $\mathcal{A}(\mathcal{E})$ when evaluated for states living in the macroscopically small energy window $I_\mathcal{E}$ from (2.9). Furthermore, Eq. (2.24) stipulates that the off-diagonal elements behave essentially random and that their fluctuations are again exponentially suppressed in f. This is also called the *off-diagonal ETH*. The off-diagonal terms determine, in particular, the temporal fluctuations $\overline{(\langle A\rangle_{\rho(t)} - \langle A\rangle_{\bar{\rho}})^2}$: Substituting the ansatz (2.24), one finds that these fluctuations are proportional to $\mathrm{e}^{-\mathcal{S}(\mathcal{E})/k_\mathrm{B}} = N^{-1}$ [8]. In view of (2.10), temporal fluctuations thus become unobservable in the long run for sufficiently large systems. In summary, the off-diagonal ETH thus ensures equilibration, while the diagonal ETH entails thermalization. If equilibration has been ascertained by other means already (see Sec. 2.2.1), one might well content oneself with the diagonal ETH, and the latter has indeed received considerably more attention in the literature, hence speaking of "the ETH" may commonly mean either just the diagonal part or the full form (2.24).

As matters stand, the ETH is a hypothesis, and while there are heuristic arguments for its generality [8, 89, 114, 127], there is to date no proof that it holds under generic, physically reasonable circumstances. Yet the much enhanced computational capabilities since its inception have made it possible to check the hypothesis numerically in specific models. The first numerical verification of the ETH was delivered by Rigol et al. in the pioneering study [116], and many additional demonstrations of the ETH (diagonal and off-diagonal) in a variety of different models have followed since; see, e.g., Refs. [128–135]. The diversity of these models suggests that the ETH is a pretty general feature of physical Hamiltonians and observables, meaning that it can indeed be considered a fundamental mechanism of thermalization.

Then again, Ref. [116] also presents a specific example where the ETH is violated, and similar conclusions have been reached in other studies as well (see, e.g., Refs. [128, 129, 132]). Prime examples are so-called integrable systems (see Sec. 2.2.3 below), which are characterized by an extensive number of local conserved quantities. A violation of the ETH is then often accompanied by nonthermalizing behavior, meaning that the considered systems may still equilibrate, but the time-averaged expectation values of physical observables can differ from the thermal prediction. The absence of thermalization in certain scenarios will be discussed in more detail in the subsequent Sec. 2.2.3, with a special emphasis on integrable systems.

At this point, however, a word of caution regarding the precise applicability of the ETH ansatz and statements about its violation in systems with conserved quantities is in order. If Q is such a conserved quantity, the Hamiltonian decomposes into different *symmetry sectors* characterized by the eigenvalues of Q. As a consequence, these sectors can possibly evolve somewhat independently, even though the commutator $[H,Q] = 0$ obviously implies a functional dependence between H and Q. It is thus generally agreed that the ETH cannot be expected to hold unconditionally across different symmetry sectors and some conserved quantities might have to be accounted for when testing the ETH [9, 136]. Notably, one would want to consider conserved quantities that are used

to construct the pertinent thermal density operator (e.g., the total particle number, see below Eq. (2.9)). On the other hand, projection operators onto energy eigenstates or powers of H also commute with H and are thus conserved, but since there are as many of these operators as there are eigenstates, the ETH would be meaningless if *all* conserved quantities had to be regarded. Prime candidates for relevant conserved quantities are again physical observables modeled by local and few-body operators [9], and there is evidence that for other types of conserved quantities, such as nonlocal many-body operators [9, 137–139], the observable properties are often the same across the corresponding sectors, obviating the need to pay special attention to them. Yet this seemingly does not hold invariably either, and we contend that these issues are still not satisfactorily understood [139–141]. In any case, given that the relevant conserved quantities in integrable systems are local and few-body, it is arguably debatable to classify such systems as ETH-violating if symmetry sectors are disregarded. Indeed, the notion of *generalized eigenstate thermalization* [142] presumes that integrable systems obey a kind of ETH, too, if all the local conserved quantities are taken into account, and this idea has been formalized further by means of the *quench-action approach* [143, 144]. We will elaborate in more detail on the role of conserved quantities with regard to thermalization in Sec. 2.2.3 below.

An in some sense related issue is the distinction between the *strong ETH* and the *weak ETH*, which refers to the scaling of the (diagonal) matrix elements A_{nn} with the system size. The strong ETH requires that all energy eigenstates $|n\rangle$ become thermal in the thermodynamic limit (apart from tacitly ignored but possible exceptions at the edges of the spectrum). Symbolically, this means that

$$\max_n |A_{nn} - \langle A\rangle_{\rho_{\mathrm{mc}}}| \to 0 \quad \text{as} \quad f \to \infty\,, \tag{2.25}$$

where ρ_{mc} is understood to be taken in a running window of the form (2.9) around E_n. This assumption was implicit in our discussion so far. However, as long as the statistical properties of the fluctuations are not specified in detail, the ansatz (2.24) also admits a weaker interpretation. This weak ETH only demands that the fraction of exceptional $|n\rangle$ whose expectation value deviates from the thermal prediction should vanish in the thermodynamic limit [129, 145, 146]. Formally, this notion can be expressed in terms of vanishing eigenstate-to-eigenstate fluctuations within every microcanonical window $I_{\mathcal{E}}$,

$$\frac{1}{N}\sum_{n:E_n\in I_{\mathcal{E}}}\left(A_{nn} - \frac{1}{N}\sum_{m:E_m\in I_{\mathcal{E}}} A_{mm}\right)^2 \to 0 \quad \text{as} \quad f \to \infty\,, \tag{2.26}$$

where N is the number of levels within $I_{\mathcal{E}}$ (see above Eq. (2.10)). Evidently, the strong ETH implies the weak ETH, but since the latter still admits the existence of states with nonthermal A_{nn}, systems satisfying the weak ETH can still fail to thermalize if those atypical states are significantly populated by the initial state $\rho(0)$.

The weak notion of the ETH has been proven to hold rather generically for local observables in systems with translational invariance [129, 147, 148] or a clustering property of the canonical density operator [149]. Importantly, it therefore commonly applies to integrable systems, too, as exemplified in Refs. [129, 145, 146].

2.2.3 Absence of thermalization and integrability

Absence of thermalization. While thermalization is certainly the standard behavior of isolated many-body systems, there are important classes of systems for which time-averaged expectation values of physical observables may not agree with the pertinent thermal values. The large overlap between these classes and the types of systems considered to be violating the (strong) ETH reinforces the pivotal role played by the ETH in the thermalization process.

As an aside, we remark that a system may fail to thermalize for trivial reasons if some of the prerequisites discussed in the previous subsections are violated. For instance, thermalization cannot be expected if the initial state has an overly broad energy distribution or the observable is artificially tailored with respect to the initial state (and thus, in particular, "unphysical" in the sense of Sec. 2.1). We do not have in mind such cases in this subsection.

A first class of examples with nonthermalizing behavior are systems exhibiting *many-body localization* (MBL) [7, 150, 151], a rough analog of Anderson localization [152] in interacting many-body systems. It describes a many-body phase commonly characterized by suppressed transport properties and unusually slow growth of entanglement between subsystems. As a consequence, information about the initial state is not spread diffusively and can be retained to some extent by means of local observables even at late times, which entails that the time-averaged state can give rise to nonthermal expectation values. This is in accordance with the observation that MBL systems usually violate the ETH. Notably, even the weak ETH cannot be expected to hold in general because the systems are typically not translationally invariant, e.g., due to the presence of strong disorder. Moreover, as opposed to integrability (see below), the MBL phase is found to be somewhat stable against perturbations.

Another potential source of inhibited thermalization are recently discovered *quantum many-body scars* due to kinematic constraints [153–156], related to experimentally observed nonthermalizing behavior in certain arrays of Rydberg atoms [62]. As an example, such kinematic constraints could forbid the simultaneous occupation of two neighboring sites in a lattice system, giving rise to a small number of exceptional, nonthermal energy eigenstates (the scars) which are, furthermore, found to have large overlap with certain types of structurally simple product states. Hence the exceptional states introduce a mild violation of the strong ETH, whereas the weak ETH is usually still fulfilled. A frequent dynamical characteristic are pronounced revivals on short time scales. It is yet unclear whether these kinematically constrained models feature truly nonthermalizing dynamics or if the relaxation merely happens extremely slowly (so-called glassy dynamics).

Since systems exhibiting MBL or many-body scars will be of no further concern in this thesis, we content ourselves with this brief and superficial description of their phenomenology. A third class of systems commonly considered to be nonthermalizing are so-called integrable systems, whose relaxation behavior we will inspect a little closer in the following.

Integrability. Whereas integrability is a commonly adopted term to indicate that a system has some regularity, symmetry, or increased analytical tractability, there is no generally agreed definition of integrability for quantum systems [7, 157]. In *classical* Hamiltonian mechanics, integrability usually means that there exist as many independent integrals of motion as there are degrees of freedom f [158, 159], where an integral of motion is a phase-space function whose Poisson bracket with the Hamiltonian function vanishes. After a canonical transformation to so-called action-angle variables, the equations of motion merely reflect the associated conservation laws and are thus essentially solved or "integrated" upon substitution of the initial conditions.

Unfortunately, this picture cannot be carried over to quantum mechanics without further ado for several reasons. On the one hand, for instance, even if the values of f conserved quantities or integrals of motion (self-adjoint operators commuting with the Hamiltonian H) are known, this does not fix the dynamics because the Hilbert space dimension N is exponentially large in f. On the other hand, as mentioned in Sec. 2.2.2 already, there always exist N linearly independent conserved quantities, e.g., the projection operators onto the eigenstates of H. In that sense, all quantum systems are "integrable," but—misquoting Orwell [160]—some systems are "more integrable" than others.

The common notions of quantum integrability [7, 157] thus adopt only some aspect of the classical definition, such as an extensive number of *local* integrals of motion, existence of an integrable classical limit, solvability by the Bethe ansatz [161] and fulfillment of the Yang-Baxter equation [162] or, more generally, exact diagonalizability by analytical means. Other common definitions consider the statistics of the energy levels or their consecutive gaps, respectively, again based on example systems with a classical limit. Systems with Poissonian gap statistics are then classified as integrable, whereas those with Wigner-Dyson statistics are declared nonintegrable [163]. Extending this approach, integrability may also be associated with the absence of level repulsion (like in Poissonian and as opposed to Wigner-Dyson statistics). While there is some overlap between these definitions, meaning that systems which classify as integrable by one criterion often also satisfy some of the others, the different notions are by no means equivalent.

In this thesis, we will adopt the first of the aforementioned approaches, which is arguably the most widely employed definition when it comes to questions of thermalization in integrable systems,

as reviewed, for instance, in Refs. [83, 164]. Hence we call a many-body quantum system with Hamiltonian H *integrable* if there exists an extensive number of independent physical observables $Q^{(n)}$ which commute with H, $[H, Q^{(n)}] = 0$, as well as with each other, $[Q^{(m)}, Q^{(n)}] = 0$. In fact, the Hamiltonian itself should be considered as just one of the $Q^{(n)}$, e.g., $Q^{(0)} := H$. The $Q^{(n)}$ are also called *local integrals of motion* (LIOM). By an "extensive number" of such LIOM, we mean that their amount is roughly proportional to the degrees of freedom f and grows with the system size in particular. "Independent" is to be understood somewhat vaguely in an operational sense: If $Q^{(1)}$ and $Q^{(2)}$ are conserved quantities, then also the linear combination $Q' := \alpha_1 Q^{(1)} + \alpha_2 Q^{(2)}$ with $\alpha_1, \alpha_2 \in \mathbb{R}$ is conserved. Similarly, powers of some $Q^{(n)}$ lead in principle to new conserved quantities. Both of these constructions are not considered to be independent LIOM. Yet there will always be some functional dependence between any two $Q^{(m)}$ and $Q^{(n)}$ implied by the commutator $[Q^{(m)}, Q^{(n)}] = 0$. Finally, "physical observable" is essentially taken as a synonym for a self-adjoint operator composed from (quasi)local and few-body terms (see Sec. 2.1). In summary, this definition is certainly not airtight, but suffices for all practical aspects to be discussed in the following. Notably, as mentioned before, the main focus of this thesis is on the dynamics of the relaxation process. In this context, integrable systems arise as interesting reference cases due to the possibility of nonthermalizing behavior, but we will not intend to classify specifically the dynamical behavior of integrable models.

Generalized Gibbs ensembles. In the beginning of this subsection, integrable models were identified as one class of systems for which physical observables may fail to thermalize. Observing that all the different conserved quantities $Q^{(n)}$ are such physical observables, this is immediately comprehensible since the expectation value $\langle Q^{(n)} \rangle_{\rho(t)} = \langle Q^{(n)} \rangle_{\rho(0)}$ cannot possibly approach the thermal prediction $\langle Q^{(n)} \rangle_{\rho_{\mathrm{mc}}}$ if it is nonthermal initially. Furthermore, it seems plausible that also other physical observables A, which could, for example, be combined in parts from some of the $Q^{(n)}$, may not thermalize either, even though they need not commute with the Hamiltonian themselves, $[H, A] \neq 0$, and can thus show a nontrivial (time-dependent) relaxation behavior. Since equilibration can still be expected quite generally (see Sec. 2.2.1), an interesting question is whether the equilibrium state $\bar{\rho}$ can be characterized in some "generic" way also in integrable systems.

The answer is positive and usually given in terms of the *generalized Gibbs ensemble* (GGE) associated with the set of LIOM $\{Q^{(n)}\}$ as introduced by Rigol et al. [165],

$$\rho_{\mathrm{GGE}} := \frac{\exp\bigl(-\sum_n \beta_n Q^{(n)}\bigr)}{Z_{\mathrm{GGE}}} \tag{2.27}$$

with $Z_{\mathrm{GGE}} := \mathrm{tr}[\exp(-\sum_n \beta_n Q^{(n)})]$, where the β_n are fixed such that $\langle Q^{(n)} \rangle_{\rho_{\mathrm{GGE}}} = \langle Q^{(n)} \rangle_{\rho(0)}$ for all n. The GGE thus generalizes the usual (canonical) Gibbs ensemble (2.8) by incorporating, besides the Hamiltonian $H = Q^{(0)}$, also the other local conserved quantities into the equilibrium density operator in a way similar to the grand canonical ensemble in traditional thermodynamics. It is then asserted that observable expectation values in integrable systems relax to the GGE prediction, similarly to their approach to thermal values in nonintegrable systems, i.e., $\langle A \rangle_{\bar{\rho}} \simeq \langle A \rangle_{\rho_{\mathrm{GGE}}}$ (see also Sec. 2.2.2).

We emphasize that this assertion is far from trivial despite the extensive number of constraints entailed via (2.27). Even if there are, in principle, on the order of f parameters β_n to be fixed to set up ρ_{GGE}, the associated Hilbert space is exponentially large in f (see Eq. (2.10)), meaning that there is still plenty of room for the occupation probabilities $\rho_{nn}(0)$ and thus the diagonal ensemble $\bar{\rho}$ to vary. That is to say, 10^{23} is still unimaginably smaller than $10^{10^{23}}$. Moreover, in practice it is expected that a few of the most local $Q^{(n)}$ are sufficient for a satisfactory approximation of the GGE [166].

The validity of the GGE to describe the equilibrium properties of integrable systems has been confirmed numerically in a variety of models, including bosonic [165, 167], fermionic [168, 169], and spin [170, 171] degrees of freedom (see also the reviews [33, 164] for more complete lists of examples). Furthermore, it has also been demonstrated experimentally in a one-dimensional Bose gas [172].

Role of conserved quantities. The GGE is obtained from the canonical ensemble of textbook statistical mechanics by including constraints for an additional set of conserved quantities. Importantly, the perception that conserved quantities have to be respected is already a building block of ordinary, "traditional" thermodynamics as sketched in Sec. 2.1. For example, both the microcanonical and the canonical ensembles for a system with a fixed number of particles $\mathcal{N}$ should be restricted to states exhibiting $\mathcal{N}$ particles, possibly within a macroscopically small uncertainty $[\mathcal{N}, \mathcal{N} + \Delta_{\mathcal{N}}]$, and the same applies to other macroscopically conserved quantities such as the magnetization. Alternatively, one may adopt a grand canonical ensemble, which includes states of variable particle number (magnetization), but fixes the average by means of a suitable Lagrange multiplier, in this case the chemical potential (external magnetic field). Due to ensemble equivalence (see below Eq. (2.8)), all those approaches lead to the same characterization of thermal equilibrium.

The traditional framework thus offers two ways to incorporate conserved quantities: either by restricting the set of considered microstates to the ones compatible with the conserved quantity's initial value or by introducing Lagrange multipliers to constrain the average value over all microstates accordingly. In this sense, the GGE merely extends the traditional ensembles by including the additional local conserved quantities $Q^{(n)}$ of integrable systems in the pertinent equilibrium ensemble. In particular, one could equivalently adopt a "generalized microcanonical ensemble" [142, 143] and choose the density operator proportional to the projector onto the subspace of joint eigenstates of all $Q^{(n)}$ whose eigenvalues match with the initial expectation values $\langle Q^{(n)} \rangle_{\rho(0)}$ within some macroscopically small margin of error.

Taking this reasoning one step further, one might even argue that the common declaration that "integrable systems do not thermalize" could be somewhat misleading because they seemingly *do* thermalize if the equilibrium density operator respects all relevant conservation laws. For the concept of thermalization and its absence to be meaningful, we thus have to determine which of the conservation laws should be included in the thermal ensemble and which should not. In other words, what distinguishes the LIOM in integrable systems from "traditional" conserved quantities like energy, particle number, magnetization, etc.?

There are indeed crucial differences. On the one hand, the traditional conserved quantities are reasonably generic in the sense that similar concepts can be defined for a large class of different system types. Furthermore, their composition from the physical microscopic degrees of freedom in a specific setup can be explained verbally in relatively simple terms and such that it generalizes in an obvious way to other systems. The LIOM of integrable systems, in contrast, are highly specific to a particular system at hand (apart from the traditional conserved quantities that may, of course, be part of the $Q^{(n)}$) and cannot be carried over to other types of integrable models. In addition, their construction is usually far from trivial, and their emergence from the physical degrees of freedom often hardly evident, let alone explicable in simple words. Incidentally, the same applies to the aforementioned example of projection operators onto energy eigenstates.

On the other hand, the number of traditional conservation laws pertaining to a certain system is finite and independent of the system size. Put differently, in the thermodynamic limit, nonintegrable systems usually only exhibit a finite number of physical observables commuting with the Hamiltonian. On the contrary, the number of LIOM of an integrable system grows with the system size by definition and is thus infinite in the thermodynamic limit.

In essence, the difference between traditional conserved quantities (which are to be considered when characterizing thermal equilibrium) and others such as LIOM (which should be disregarded) is therefore a conceptual one. Albeit somewhat difficult to formalize, it should usually be clear in practice whether or not a given conservation law has "traditional character." Already the question whether we have a name other than "$Q^{(n)}$" for a given operator is presumably a good indicator.

A related issue concerns conserved quantities which are arguably simple and generic, but are represented by nonlocal or many-body operators. A prominent example is the (quasi)momentum arising from translational invariance. Since such operators are not accounted for in the traditional thermodynamic examples, one would tend to disregard them, meaning that the associated decomposition of the energy spectrum into invariant subsectors should not matter with respect to thermal expectation values. As briefly mentioned in the context of ETH violations above (see Sec. 2.2.2)

it is often observed that such nonlocal many-body conserved quantities do not lead to different eigenstate expectation values of local observables in different sectors [9, 137–139], so for questions of thermalization, too, it is indeed irrelevant whether or not the conservation law is taken into account.

Role of initial states. Finally, one cannot ignore the role of initial states entirely either: In order for a conserved quantity to induce nonthermalizing behavior, the initial state must populate its nonthermal eigenstates significantly. It is thus important to understand for which combinations of observables and realistic preparation procedures this can occur at all. For instance, the fact that integrable systems normally still satisfy the weak ETH (see Sec. 2.2.2) implies that nonthermal states are exceedingly rare among all states. For an uninformed "random" choice of the initial state, one will thus typically observe thermalization also in integrable systems [173]. A popular method to prepare systems (integrable or nonintegrable) out of equilibrium are *quantum quenches* [7, 83, 174]. Here the initial state is a thermal state (often the ground state, i.e., a zero-temperature thermal state) of some Hamiltonian H', which is then suddenly changed to the actual Hamiltonian of interest H at time $t = 0$. Provided that effects of a finite switching time are negligible, this method is indeed experimentally feasible and frequently used (see, e.g., Refs. [30, 33]). In particular, quenches are the standard procedure to obtain nonequilibrium initial states which lead to manifestly nonthermalizing behavior in integrable systems while still remaining reasonably narrow in energy for suitable choices of H'. The precise mechanism of how such a quench entails significant occupations for the rare nonthermal states, however, is still not fully understood.

Another class of popular and viable initial states for lattice systems (both numerical and experimental) invokes the product basis associated with the individual lattice sites such that the number of particles, the magnetization, etc. are fixed for every individual site. The resulting pure states, however, may exhibit a relatively broad energy distribution, at least for accessible system sizes, which in turn precludes the use of standard thermodynamics for trivial reasons since the energy is not well-defined even from a "macroscopic" perspective.

A third class to generate nonequilibrium states with some control over their macroscopic properties arises within the framework of dynamical typicality, which will be introduced in Sec. 2.3.2 below.

Prethermalization. An interesting two-stage relaxation process can occur if a nonthermalizing system is subject to a reasonably weak perturbation. As explained above, such nonthermalizing systems usually still equilibrate, but the stationary expectation values of most observables do not correspond to the thermal prediction. When adding a perturbation which lifts the constraints preventing thermalization in the original system, however, the perturbed system is generally expected to thermalize. Nevertheless, if the perturbation is sufficiently weak, the system will still follow the unperturbed dynamics for some time and expectation values may even settle down to the nonthermal stationary value of the unperturbed system before eventually departing again and relaxing to the thermal state. This behavior, which has also been observed experimentally, for example, in cold-atom setups [31, 33], is known as *prethermalization* [9, 33, 175–178]. The prime example are integrable reference systems subject to weak integrability-breaking perturbations. In Sec. 3.5 we will establish that prethermalization can indeed be expected generically in the perturbed relaxation setting outlined above. The results of Sec. 3.6 will then even give a more quantitative account of the expectable dynamics.

2.3 Typicality and concentration of measure

The buzz word "typicality" has already been dropped several times in the previous section since it describes one fruitful approach to address questions of equilibration and thermalization in many-body systems. As it represents the main theoretical framework for the key results of this thesis, we devote an entire section to it, motivating and explaining the general methodology in Sec. 2.3.1 and collecting important applications from the statistical-mechanics context in Sec. 2.3.2. In fact, von Neumann's quantum ergodic theorem, briefly presented in the previous Sec. 2.2.2, constituted one such application already.

2.3.1 General considerations

Background. The basic reasoning of the *typicality* method [6, 179, 180] is founded on the same phenomenological observation that is also at the heart of statistical mechanics: Despite the vastly complicated dynamics of its microscopic degrees of freedom, the macroscopically observable behavior of a many-body system can generally be described by only a few characteristic parameters. An obvious manifestation is the fact that macroscopic experiments are reproducible at all, even though it is practically impossible to replicate a certain setup microscopically [181, 182]. Based on the relevant macroscopic characteristics, one can thus compile classes of microscopically distinct systems or setups which will nonetheless display the same macroscopic phenomenology. Note that these classes in general still depend on the observed quantity, hence different classifications arise for different phenomena or macroscopic properties of interest.

The origin of this insensitivity of many-body systems to microscopic details is a sort of "self-averaging" mechanism, whereby the complex interplay of many largely independent constituents appears practically random at larger scales. A macroscopic observable can thus be understood as a random variable whose value depends on a large number of (quasi)random and essentially independent microscopic degrees of freedom. The distribution of such a composite random variable is often extremely sharply peaked, meaning that fluctuations around the average become exceedingly rare as the number of microscopic contributions increases. This effect is known as *concentration of measure* [183–185]: Quoting Talagrand [183], "[a] random variable that depends (in a 'smooth' way) on the influence of many independent variables (but not too much on any of them) is essentially constant." Note that this rationale is akin to concepts like the law of large numbers, the central limit theorem, or large deviation theory [186].

To arrive at a prediction for the behavior of an actual physical system, we can exploit such concentration-of-measure properties as follows: Instead of the single system or setup of interest, we temporarily consider an entire class or *ensemble* of similar setups. Perhaps surprisingly, it is then often possible to calculate the average of the examined macroscopic quantity over all members of such an ensemble, even though the actual system of interest alone is not tractable. If the examined quantity exhibits concentration of measure with respect to the considered ensemble, nearly all individual members of the ensemble will be excellently described by the average behavior, i.e., their behavior is "typical." Provided that the ensemble models the relevant characteristics of the actual system of interest reasonably well, we can thus expect that also this true physical system is one of the typical members, meaning that the ensemble average becomes the desired prediction for the behavior of the true system of interest.

General procedure. To explain the general procedure, assume that we are given a certain *setup* s_0 (e.g., a Hamiltonian, observable, and initial state) as well as a macroscopic property $q(s_0)$ we wish to predict or understand theoretically (e.g., the time-dependent expectation values of the observable). The initial step towards establishing a typicality argument is to identify the decisive physical mechanisms or characteristics $\chi(s_0)$ determining the observed property $q(s_0)$. Obviously, there is no standard recipe to do so and one would want to be guided by physical intuition as well as practicality of the mathematical modeling. Once a candidate for the macroscopically relevant properties is available, we can construct classes of microscopic setups that share those characteristics. Formally, this amounts to specifying an *ensemble* S of setups (usually an ensemble of Hamiltonians, observables, or initial states) equipped with a probability distribution $\mathbb{P}$ such that the individual members of the ensemble emulate the relevant characteristics of the true setup of interest with high fidelity, $\mathbb{P}(\chi(S) = \chi(s_0)) \approx 1$.

The actual algorithm to establish typicality then consists of three steps:

1. Calculate the *ensemble average* $\hat{q} := \mathbb{E}[q(S)]$ of the quantity of interest, i.e., the expectation value of $q(S)$ with respect to the distribution $\mathbb{P}$.

Here we introduced the third type of important averaging procedure employed in this thesis, namely the average over a given typicality ensemble, denoted by the symbol $\mathbb{E}[\cdots]$. Note that these *ensemble averages* should not be confused with *thermal averages* such as $\langle\cdots\rangle_{\rho_{\mathrm{mc}}}$, $\langle\cdots\rangle_{\rho_{\mathrm{can}}}$, etc., which are also sometimes called ensemble averages in the literature. The latter are a special form of the "quantum averages" from (2.1).

2. Demonstrate concentration of measure of $q(S)$. This may be achieved in various ways. A standard approach is to calculate the variance $\mathbb{E}[(q(S) - \hat{q})^2]$ and then adopt Chebyshev's inequality (see Eq. (2.29) below) to bound the probability for $q(S)$ to deviate from the average $\hat{q}$ by more than a (macroscopically unresolvable) threshold κ, symbolically $\mathbb{P}(|q(S) - \hat{q}| \geq \kappa) \ll 1$.

3. Argue that the true setup of interest is one of the typical members, i.e., $q(s_0) = \hat{q}$.

This last step of the derivation is naturally the least rigorous and therefore also the most debated one. A careful and informed choice of the typicality ensemble is vital to convincingly argue in this direction: If the previously identified characteristic $\chi(s_0)$ is indeed the driving mechanism for the property of interest $q(s_0)$, and if the ensemble reproduces this characteristic with high probability, then it is extremely likely that the true behavior $q(s_0)$ agrees with the typical $\hat{q}$.

However, the probabilistic nature of the typicality method inevitably implicates the existence of "exceptions to the rule." Moreover, it is usually impossible to determine with certainty whether a given system will behave like a typical member of the chosen ensemble *a priori*. Put differently, there is no systematic way to check whether the considered ensemble, which unavoidably and in fact deliberately ignores many microscopic details, captures all essential characteristics relevant for the observed behavior. This unfortunate indefiniteness of the typicality method is the price to pay for reaching far into regimes inaccessible to rigorously systematic approaches, a property it shares with many other "uncontrolled" approximation schemes like density functional theory, WKB approximations, truncations of functional renormalization group equations, etc. Ultimately, the success of the method can only be judged by comparison with explicit examples. Such comparisons therefore form an integral part of this thesis (see, in particular, Secs. 3.7, 4.3, 4.4, and 5.4).

Markov's and Chebyshev's inequalities. As mentioned above, concentration of measure of a certain property can often be demonstrated conveniently by exploiting standard inequalities from probability theory that relate probabilities to low-order moments of random variables. *Markov's inequality* states that, for any nonnegative random variable X and any real number $\kappa > 0$,

$$\mathbb{P}(X \geq \kappa) \leq \frac{\mathbb{E}[X]}{\kappa} \,. \tag{2.28}$$

Given instead an arbitrary real-valued random variable X with expected value μ and variance σ^2, we can apply Markov's inequality to the (nonnegative) random variable $(X - \mu)^2$ to conclude that

$$\mathbb{P}(|X - \mu| \geq \kappa) \leq \frac{\sigma^2}{\kappa^2} \tag{2.29}$$

for all $\kappa > 0$, which is known as *Chebyshev's inequality*. Because of their generality, the bounds obtained from these two inequalities are often rather weak, but the favorable scaling of fluctuations with the degrees of freedom in many-body systems more than compensates this drawback in practice.

2.3.2 Applications

As motivated above, from a conceptual point of view, a prime example of typicality arguments at work is textbook statistical mechanics itself, which is based on the notion that fluctuations of macroscopic quantities are unobservable due to extremely sharply peaked distributions, which also underlies the equivalence of thermodynamic ensembles (see below Eq. (2.8)). Moreover, especially as a numerical tool, arguments of the type described in Sec. 2.3.1 have been utilized and rediscovered in a variety of contexts to obtain explicit quantitative estimates of otherwise extremely costly or impractical computations. Examples include stochastic trace estimators [187] (with the connection to typicality highlighted in Refs. [188, 189]), finite-temperature Lanczos methods [190], density-of-state estimations [191, 192], or thermal pure quantum states [193]. In the following, we will concentrate on those applications from research on equilibration and thermalization of many-body quantum systems and the foundations of statistical mechanics which have immediate relevance for the subsequent parts of this thesis.

Canonical typicality. As expounded in Sec. 2.1, textbook statistical mechanics characterizes thermal equilibrium of large isolated systems by means of the microcanonical ensemble ρ_{mc} from (2.5), whereas it ascribes the canonical ensemble ρ_{can} from (2.8) to systems in contact with a large bath. In the case of weak (more precisely, negligible) interactions between the system and the bath in the latter setting, the two formalisms can be connected by considering the system-bath compound as an isolated supersystem. It can then be shown that, if the supersystem is in the microcanonical state ρ_{mc}, the state of the smaller system of interest, which is obtained by tracing out the bath degrees of freedom from ρ_{mc}, is indeed given by ρ_{can} with the temperature determined by the bath as defined above Eq. (2.8) [4] (see also Ref. [18] for considerations about when a similar reduction occurs for the time-averaged state $\bar{\rho}$ of the system-bath compound). While such a relationship is generally reassuring, it is still somewhat dissatisfactory from a foundational point of view because in a purely quantum mechanical treatment, a perfectly isolated system should generally be in a pure state rather than in the maximally mixed state ρ_{mc}.

The by now classic results on *canonical typicality* by Goldstein et al. [118] and Popescu et al. [119], which were in fact first derived in the PhD thesis of Lloyd [194] almost two decades earlier, address precisely this issue: They prove that nearly all pure states of the isolated compound system with a fixed total energy lead to a reduced density matrix of the subsystem which is indistinguishable in practice from the canonical density operator ρ_{can} with the inverse temperature β of the bath. To be precise, Ref. [119] actually considers a much more general setting by showing indistinguishability between the reduced density matrix of pure compound states and the generalized canonical ensemble obtained from the maximally mixed compound state under arbitrary constraints (with a fixed total energy as in the standard setting as a special case). Again, ideas of this kind had already been sketched in Ref. [194], too.

In a related study [195], Gemmer and Mahler use a similar line of reasoning to show that the accessible Hilbert space of a system-bath compound in a pure state features a "dominant region" (essentially the typical states) for which the energy distribution upon tracing out the bath assumes the canonical form. Another related result can also be found in the previously introduced study [89] by Tasaki (see below Eq. (2.20)), which investigates compounds with a specific model for the system-bath coupling. The canonical ensemble then emerges again as the proper state to describe equilibrium expectation values of system observables. The somewhat heuristic methodology there bears resemblance to typicality arguments, albeit it is never explicitly stated this way. Not least, the general idea of canonical typicality is already discernible in Schrödinger's early work [196].

Example for the typicality method. The canonical typicality results [118, 119, 194] are all based on the fundamental observation that, for any observable A, the vast majority of pure states $|\psi\rangle$ from a Hilbert space $\mathcal{H}$ (e.g., the energy shell of the system-bath compound; see also below Eq. (2.9)) of sufficiently large (finite) dimension N yield expectation values $\langle\psi|A|\psi\rangle$ very close to the microcanonical value $\mathrm{tr}(\rho_{\mathrm{mc}}A) = \mathrm{tr}\,A/N$. As an illustrational example for the typicality method outlined in Sec. 2.3.1, we briefly demonstrate this property.

To this end, we choose an arbitrary orthonormal basis $\{|\phi_\alpha\rangle\}$ of $\mathcal{H}$. To connect to the language of Sec. 2.3.1, let us assume that we are given an observable A and a system in a pure state $|\psi\rangle = \sum_\alpha c_\alpha|\phi_\alpha\rangle \in \mathcal{H}$ with $\boldsymbol{c} := (c_1, \ldots, c_N) \in \mathbb{C}^N$, normalized such that $\langle\psi|\psi\rangle = \|\boldsymbol{c}\|^2 = \sum_\alpha |c_\alpha|^2 = 1$. The setup s_0 thus consists of the Hilbert space $\mathcal{H}$, the observable A, and the state $|\psi\rangle$. As our quantity of interest, we would like to evaluate the expectation value $q(\boldsymbol{c}) := \langle\psi|A|\psi\rangle = \sum_{\alpha,\beta} c_\alpha^* c_\beta A_{\alpha\beta}$ by means of the typicality method, where $A_{\alpha\beta} := \langle\phi_\alpha|A|\phi_\beta\rangle$ denotes the matrix elements of A in the chosen basis $\{|\phi_\alpha\rangle\}$. As indicated by the notation, the class of similar but different setups to be considered will consist of the same Hilbert space $\mathcal{H}$ and observable A, but different coefficients $\boldsymbol{c}$ and thus states $|\psi\rangle$. Of course, this may seem like a needlessly complicated approach because it is usually a simple task to calculate expectation values of pure states, but it nevertheless illustrates the general idea. Not least, as explained above, the final result still entails nontrivial physical insights.

The pertinent typicality ensemble contains all normalized pure states $|\psi'\rangle := \sum_\alpha c'_\alpha|\phi_\alpha\rangle \in \mathcal{H}$ with $\|\boldsymbol{c}'\| = 1$, and we intend to assign an equal weight (or probability) to each of them. In other words, the vector $\boldsymbol{c}'$ is to be distributed uniformly on the N-dimensional complex unit sphere. In practice, this ensemble of states can be generated, for example, by sampling random numbers c_α^{R} and c_α^{I} ($\alpha = 1, \ldots, N$) independently from a standard normal distribution (vanishing mean,

unit variance). Defining $c'_\alpha := (c^{\mathrm{R}}_\alpha + \mathrm{i}\, c^{\mathrm{I}}_\alpha)/N_c^{1/2}$ with $N_c := \sum_\alpha [(c^{\mathrm{R}}_\alpha)^2 + (c^{\mathrm{I}}_\alpha)^2]$ leads to the desired uniform distribution on the complex unit sphere for $\boldsymbol{c}'$. For the ensuing analysis, however, it is more convenient to start from the "true" $\boldsymbol{c}$ and generate $\boldsymbol{c}' := U\boldsymbol{c}$ by drawing a unitary transformation matrix $U = (U_{\mu\alpha}) \in \mathrm{U}(N)$ uniformly at random, meaning that $\mathbb{P}$ is the *Haar measure* of the unitary group $\mathrm{U}(N)$. If the Hilbert space $\mathcal{H}$ is chosen to be an energy shell $\mathcal{H}_\mathcal{E}$ (see below Eq. (2.9)), the common characteristic $\chi(\boldsymbol{c}')$ of all setups (i.e., all different states $|\psi'\rangle$) is that they have the same macroscopic energy, i.e., $\mathbb{P}(|\langle\psi'|H|\psi'\rangle - \mathcal{E}| \leq \Delta_\mathcal{E}) = 1$, and this is essentially the only characteristic fixed in the classical textbook scenario via ρ_{mc}, too.

Following the algorithm described in Sec. 2.3.1, we first compute the ensemble average

$$\mathbb{E}[q(\boldsymbol{c}')] = \sum_{\mu,\nu} \mathbb{E}\big[c'^*_\mu c'_\nu\big]\, A_{\mu\nu} = \sum_{\mu,\nu}\sum_{\alpha,\beta} c^*_\alpha\, \mathbb{E}\big[U^*_{\mu\alpha} U_{\nu\beta}\big]\, c_\beta\, A_{\mu\nu}\,. \tag{2.30}$$

Averages over factors of Haar-distributed unitary matrix elements of the type occurring on the right-hand side have been calculated in Ref. [197], yielding $\mathbb{E}[U^*_{\mu\alpha}U_{\nu\beta}] = \delta_{\mu\nu}\delta_{\alpha\beta}/N$, which in this case is also apparent due to symmetry and normalization. Substituting into (2.30), we obtain

$$\mathbb{E}[q(\boldsymbol{c}')] = \frac{1}{N}\sum_\mu A_{\mu\mu} = \mathrm{tr}(\rho_{\mathrm{mc}} A)\,. \tag{2.31}$$

To establish concentration of measure for $q(\boldsymbol{c}')$, the second step of the procedure from Sec. 2.3.1, we evaluate the variance $\mathrm{var}\, q(\boldsymbol{c}') = \mathbb{E}[q(\boldsymbol{c}')^2] - \mathbb{E}[q(\boldsymbol{c}')]^2$. Exploiting that [197]

$$\begin{aligned}\mathbb{E}\big[U_{\nu_1\beta_1}U_{\nu_2\beta_2}U^*_{\mu_1\alpha_1}U^*_{\mu_2\alpha_2}\big] = {}& \frac{1}{N^2-1}\left[\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\delta_{\alpha_1\beta_2}\delta_{\alpha_2\beta_1}\right] \\ & - \frac{1}{N(N^2-1)}\left[\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\delta_{\alpha_1\beta_2}\delta_{\alpha_2\beta_1} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2}\right],\end{aligned} \tag{2.32}$$

we obtain

$$\begin{aligned}\mathbb{E}[q(\boldsymbol{c}')^2] &= \sum_{\substack{\mu_1,\mu_2,\\ \nu_1,\nu_2}} \sum_{\substack{\alpha_1,\alpha_2,\\ \beta_1,\beta_2}} c^*_{\alpha_1} c^*_{\alpha_2} c_{\beta_1} c_{\beta_2}\, \mathbb{E}\big[U_{\nu_1\beta_1}U_{\nu_2\beta_2}U^*_{\mu_1\alpha_1}U^*_{\mu_2\alpha_2}\big]\, A_{\mu_1\nu_1}A_{\mu_2\nu_2} \\ &= \frac{1}{N^2-1}\left(1-\frac{1}{N}\right)\left[(\mathrm{tr}\, A)^2 - \mathrm{tr}(A^2)\right].\end{aligned} \tag{2.33}$$

Combined with the square of (2.31), we thus find that the variance is given by

$$\mathrm{var}\, q(\boldsymbol{c}') = \frac{1}{N+1}\left[\frac{\mathrm{tr}(A^2)}{N} - \frac{(\mathrm{tr}\, A)^2}{N^2}\right] \leq \frac{(\Delta A)^2}{4(N+1)}\,, \tag{2.34}$$

where ΔA is the spectral range of A from (2.18). Note that we exploited in the last step that $\mathrm{var}\, q(\boldsymbol{c}')$ is invariant upon adding a constant to A so that we can take the operator norm $\|A\|$ of A (largest eigenvalue in modulus) as $\Delta A/2$ without loss of generality. Substituting $q(\boldsymbol{c}') = \langle\psi'|A|\psi'\rangle$ and adopting Chebyshev's inequality (2.29) with $\kappa = \delta A$ (cf. Eq. (2.19)), we observe that

$$\mathbb{P}(|\langle\psi'|A|\psi'\rangle - \mathrm{tr}(\rho_{\mathrm{mc}}A)| \geq \delta A) \leq \frac{1}{4(N+1)}\left(\frac{\Delta A}{\delta A}\right)^2, \tag{2.35}$$

i.e., the probability that the expectation value of A for a randomly chosen pure state $|\psi'\rangle$ shows resolvable deviations from the microcanonical value decreases with the Hilbert space dimension N. In view of (2.10) and the limited number of realistic measurement outcomes (see below Eq. (2.19)), we thus conclude that for a many-body system with f degrees of freedom, such noticeable deviations are exponentially unlikely in f, and $\langle\psi'|A|\psi'\rangle = \mathrm{tr}(\rho_{\mathrm{mc}}A)$ is an excellent approximation for the vast majority of individual states $|\psi'\rangle \in \mathcal{H}$.

In the third step of the general typicality procedure from Sec. 2.3.1, we come back to the actual "true" state $|\psi\rangle$ of the considered system. In general, this state may be the result of a sophisticated preparation process. However, in the absence of any additional information about its origin, we cannot but conclude that $|\psi\rangle$ will *most likely* ("typically") be an equilibrium state in the sense

that expectation values coincide with the microcanonical prediction (see Sec. 2.1). The primary physical implication of this result is thus that there are exponentially more equilibrium states than nonequilibrium ones in reasonably large systems.

As an aside, we remark that the same calculation implies that for an arbitrary, fixed pure state $|\psi\rangle$, nearly all observables yield expectation values indistinguishable from the microcanonical value. This can be understood by observing that we can write $q(\boldsymbol{c}') = \langle\psi|U^\dagger AU|\psi\rangle$ (in a slight abuse of notation, using the same symbol U to denote the unitary operator associated with the transformation matrix between $\boldsymbol{c}$ and $\boldsymbol{c}'$). Hence, instead of considering $|\psi'\rangle = U|\psi\rangle$ as random and A as fixed, we may regard $|\psi\rangle$ as fixed and use an ensemble of observables $A' = U^\dagger AU$, comprising all Hermitian operators with the same spectrum as the true A, but randomly chosen eigenvectors.

Approximation of arbitrary density operators by pure states. The above calculation demonstrates that a single pure state can serve to approximate the microcanonical density operator ρ_{mc} with extremely high fidelity in a sufficiently large Hilbert space. Interestingly, the same idea can be employed to emulate largely arbitrary density operators ρ by means of randomly sampled pure states [181, 182, 193, 198–200], provided that ρ is of low purity, $\mathrm{tr}(\rho^2) \ll 1$. Similar conclusions can actually be drawn for ρ of very high purity, but then the result is evidently less surprising and insightful.

The method utilizes that any density operator ρ is by definition nonnegative, implying that all eigenvalues are nonnegative and there exists a unique Hermitian, nonnegative operator R such that $RR^\dagger = \rho$ (the "square root" of ρ). Following Ref. [182], the typicality ensembles of pure states to imitate ρ are then constructed in three steps: First, we generate a random pure state $|\psi_1\rangle$ by sampling its coefficients in an arbitrarily chosen orthonormal basis from a standard complex normal distribution (see above Eq. (2.30)). Second, we apply the operator R to get $|\psi_2\rangle := R\,|\psi_1\rangle$. Like $|\psi_1\rangle$, this state $|\psi_2\rangle$ is generally not normalized, but has a norm very close to unity with very high probability, to wit, the variance of the norm is given by the purity $\mathrm{tr}(\rho^2)$. More generally, the so-obtained ensemble of pure states $|\psi_2\rangle$ satisfies $\mathbb{E}[\langle\psi_2|A|\psi_2\rangle] = \mathrm{tr}(\rho A)$ and

$$\mathbb{E}[(\langle\psi_2|A|\psi_2\rangle)^2] - \mathbb{E}[\langle\psi_2|A|\psi_2\rangle]^2 \leq \Delta A\,\mathrm{tr}(\rho^2)/2 \tag{2.36}$$

for arbitrary observables A, similarly to the results (2.31) and (2.34) for the ensemble of $|\psi'\rangle$, which are closely related to the $|\psi_1\rangle$. Thus the expectation values of experimentally realistic observables (see Sec. 2.2.1) in the states $|\psi_2\rangle$ display concentration of measure for sufficiently low purity. Third, we can normalize by defining $|\psi_3\rangle := |\psi_2\rangle/\sqrt{\langle\psi_2|\psi_2\rangle}$ without changing the statistical properties significantly, so the vast majority of randomly generated, normalized pure state $|\psi_3\rangle$ approximates the expectation value of the observable A in the mixed state ρ extremely well,

$$\langle\psi_3|A|\psi_3\rangle \simeq \mathrm{tr}(\rho A) = \langle A\rangle_\rho\,. \tag{2.37}$$

Moreover, the approximation can be improved systematically by generating multiple states $|\psi_3\rangle$ and averaging over the obtained proxies for $\langle A\rangle_\rho$, even though it may sometimes be advantageous to split off parts of $\langle A\rangle_\rho$ (e.g., the partition function for the canonical ensemble $\rho = \rho_{\mathrm{can}}$, see below) and estimate them separately [188, 189].

The previous results for $\rho = \rho_{\mathrm{mc}}$ are recovered for $R = \mathbb{1}/\sqrt{N}$ by choosing $|\psi_1\rangle$ from the relevant energy shell $\mathcal{H}_\mathcal{E}$ of dimension N. For practical numerical calculations of thermal expectation values, however, this is still a bit inconvenient because one needs to know the energy eigenstates of the full Hilbert space $\mathcal{H}$ to select a basis for the energy shell $\mathcal{H}_\mathcal{E}$. More efficiently, we can exploit the equivalence of ensembles and simulate the canonical density operator ρ_{can} from (2.8) by using $R = \mathrm{e}^{-\beta H/2}$ such that $\langle\psi_3|\psi_3\rangle \simeq \mathrm{tr}(\mathrm{e}^{-\beta H}) = Z$ approximates the partition function and $\langle A\rangle_{\rho_{\mathrm{can}}} \simeq \langle\psi_3|A|\psi_3\rangle/Z$. In practice, the unnormalized $|\psi_2\rangle$ can thereby be calculated by imaginary-time propagation of $|\psi_1\rangle$ up to the inverse temperature β, which in turn is either known from the considered setup or—for an isolated system—determined as usual from the constraint $\langle H\rangle_{\rho_{\mathrm{can}}} = \mathcal{E}$. In the latter case, the functional dependence of $\langle H\rangle_{\rho_{\mathrm{can}}}$ on β has to be determined approximately, but this can be achieved in the same way from an imaginary-time series of H for a randomly generated pure state $|\psi_1\rangle$. In the same vein, one can implement other density operators that are functions of the Hamiltonian, using, for example, a Taylor expansion of the associated function

$R(H)$ to obtain $|\psi_2\rangle$ from $|\psi_1\rangle$. A particular example are Gaussian filters that pick out random states from a narrow interval of energies (or some other observable) [133, 193, 201, 202].

Likewise, it is possible to generate ensembles of pure states with approximately prescribed expectation values or measurement statistics of a few observables $A_1, \ldots, A_M$ and otherwise unbiased statistical properties [181, 182, 200] (see also Refs. [203, 204] for related ideas), modeling the preparation of an experiment with a few tunable macroscopic parameters. Denoting the target expectation values of the $A_1, \ldots, A_M$ by $a_1, \ldots, a_M$, respectively, the pertinent density operator takes the form [182]

$$\rho = \frac{1}{N}\left[\mathbb{1} - \sum_{k=1}^{M} y_k (A_k - a_k)\right]^{-1} \tag{2.38}$$

with the M parameters y_k determined by the $M+1$ constraints $\operatorname{tr}\rho = 1$ and $\operatorname{tr}(\rho A_k) = a_k$, one of which is redundant. This construction comes with the proviso that Eq. (2.38) yields a well-defined density operator, i.e., the y_k are additionally constrained to render ρ nonnegative.

The actually required operator $R = \rho^{1/2}$ is obtained by taking the square root of Eq. (2.38). In practice, determining the y_k and evaluating the square root can become computationally expensive if not impossible. Hence one might resort to discerning the y_k approximately, for instance by means of a Taylor expansion [173, 200]. This relates to the construction of nonequilibrium initial states for a single observable A from Ref. [73], which implicitly adopts a similar idea. Namely, a nonequilibrium pure state $|\psi\rangle$ is generated from a randomly Haar-distributed pure state $|\phi\rangle$ as

$$|\psi\rangle \propto \Pi(1 + \kappa A)\Pi|\phi\rangle \tag{2.39}$$

up to normalization, where κ is a real parameter used to tune the desired expectation value $\langle A\rangle_\rho \simeq \langle\psi|A|\psi\rangle$ and Π is a projector onto a suitable subspace of the considered Hilbert space (e.g., the energy shell $\mathcal{H}_\mathcal{E}$) or an approximation thereof, e.g., by means of the aforementioned Gaussian filters. Note that we introduced new notation in Eq. (2.39) for later convenience; the connection to the previous labeling is established via $|\psi_1\rangle = \Pi|\phi\rangle$, $R = \Pi(1 + \kappa A)\Pi$, and thus $|\psi_3\rangle = |\psi\rangle$.

Dynamical typicality. So far, we considered only static combinations of density operators ρ and observables A. Further insights can be gained by considering the family of time-evolved observables $A(t) := \mathcal{U}(t)^\dagger A \mathcal{U}(t)$ with $\mathcal{U}(t) = \mathcal{U}(t, 0)$ being the propagator from (2.3), i.e., we assume a Heisenberg-picture perspective on the system dynamics and explicitly allow time-dependent Hamiltonians. This way, the time-dependent expectation value $\langle A\rangle_{\rho(t)} = \operatorname{tr}[\rho(t)A]$, where $\rho(t)$ is the time-evolved state as introduced above Eq. (2.3) with $t_0 = 0$, can equivalently be written as $\langle A(t)\rangle_{\rho(0)}$. We now choose the initial state $\rho(0) = \rho$ to be one of the generic density operators considered in the previous paragraph and examine the expectation values $\langle\psi_2|A(t)|\psi_2\rangle$ of the time-evolved observable in the pure state $|\psi_2\rangle$ constructed as above. We recall that the result (2.36) entails that nearly all $|\psi_2\rangle$ generated this way exhibit very similar expectation values for $A = A(0)$. The crucial observation is that the bound (2.36) is independent of the time t when substituting $A = A(t)$ because the eigenvalues of A and thus its spectral range ΔA are unaffected by the unitary transformation $\mathcal{U}(t)$. Consequently, concentration of measure is also retained at any later time $t > 0$, implying that the overwhelming majority of pure states $|\psi_2\rangle$ or $|\psi_3\rangle$ sampled according to the above described procedure exhibit very similar expectation values not only at the initial time t, but also at any later time $t > 0$. Put differently, the time series $\langle\psi_2|A(t)|\psi_2\rangle$ closely follows the average $\mathbb{E}[\langle\psi_2|A(t)|\psi_2\rangle] = \langle A\rangle_{\rho(t)}$ for nearly all pure states $|\psi_2\rangle$, and similarly for $|\psi_3\rangle$.

This remarkable *dynamical typicality* of observable expectation values was first demonstrated by Bartsch and Gemmer in Ref. [73] for ensembles of states generated according to (2.39) and subsequently extended in various works [181, 200] with the present formulation due to Reimann and Gemmer [182]. As a particular application, one can employ ensembles of states as in Eqs. (2.38) or (2.39) with prescribed and possibly nonequilibrium expectation values $a_1, \ldots, a_M$ for observables $A_1, \ldots, A_M$. Dynamical typicality then assures that nearly all initial pure states $|\psi_3\rangle$ with these approximately fixed expectation values will lead to similar relaxation behavior of the observables A_k (or any other A) apart from very rare exceptional combinations of $|\psi_3\rangle$, A_k, and t. In the absence of further knowledge about the state, the behavior observed under a randomly generated $|\psi_3\rangle$ can thus be expected generically. If the $a_1, \ldots, a_M$ correspond to the initial values of a few

macroscopic, experimentally controllable observables $A_1, \ldots, A_M$, this offers an explanation for why the precisely prepared microstate is largely irrelevant as far as the observable dynamics is concerned. Consequently, Eqs. (2.38) and (2.39) can be utilized to construct states emulating a macroscopic preparation procedure. As another application, dynamical typicality can be exploited to compute excellent approximations of time-correlation functions [74, 75].

Predictions for typical relaxation dynamics. While dynamical typicality expresses that the relaxation of experimentally realistic observables is very similar for nearly all pure states constructed as explained above, it does not make any concrete predictions about the actually observed expectation values. Generically, these values will coincide with the ensemble average $\mathbb{E}[\langle\psi_2|A(t)|\psi_2\rangle]$, but there is no obvious way to calculate this average other than using numerical simulations for a specific setup. Yet typicality methods can be used to devise general analytical predictions for relaxation processes of isolated quantum many-body systems, and the subsequent chapters of the present thesis will in fact be mostly concerned with doing precisely this.

Particularly relevant predecessor works are Refs. [96, 97], where such predictions of the relaxation dynamics were established in a setting that is in some sense complementary to the one employed in this thesis. Focusing on an energy shell $\mathcal{H}_\mathcal{E}$, the adopted typicality approach of those references considers the dynamics generated by ensembles of Hamiltonians H' for a fixed initial state $\rho(0)$ and observable A (however, see also Ref. [96] for an alternative interpretation of the underlying randomization). More precisely, starting from the true Hamiltonian H of the actual system of interest, the considered ensembles consist of all H' obtained from H by taking the same eigenvalues, but randomly rotating [96] or shuffling [97] the eigenvectors. Incidentally, the ensemble of Hamiltonians from Ref. [96] is also the one employed for establishing von Neumann's quantum ergodic theorem (see Sec. 2.2.2) and its successors [85, 117, 120, 124].

The derivations in Refs. [96, 97] then essentially proceed as outlined in Sec. 2.3.1, resulting in a prediction for the time-dependent expectation values of the form

$$\langle A\rangle_{\rho(t)} = \langle A\rangle_{\bar{\rho}} + |\hat{d}(t)|^2 \left[\langle A\rangle_{\rho(0)} - \langle A\rangle_{\bar{\rho}}\right] . \tag{2.40}$$

Here $\bar{\rho}$ denotes the time-averaged state from (2.17) as usual, with the approach from Ref. [96] actually predicting $\bar{\rho} \simeq \rho_{\mathrm{mc}}$ in addition. Furthermore, $\hat{d}(t) := \frac{1}{N}\sum_n \mathrm{e}^{\mathrm{i}E_n t}$ is the Fourier transform of the relative density of states $D(E)/N$ (see Eqs. (2.10) and (2.11)). Crucially, the prediction (2.40) has been verified for a variety of experimental and numerical examples [96, 97]. We also remark that these works assumed an approximately uniform spreading of the initial state $\rho(0)$ across the relevant energy shell. However, pertinent extensions that include the energy distributions of $\rho(0)$ [100] or allow for macroscopic populations of a single energy level [22] have been devised, too. A characteristic feature of (2.40) is that the predicted time scale of equilibration is extremely fast since features like locality of interactions are not shared among the Hamiltonians of any given ensemble (see also the discussion towards the end of Sec. 2.2.1). Hence Eq. (2.37) cannot be expected to describe the relaxation in situations where macroscopic transport is the driving mechanism for equilibration [100].

Our investigations in Chapters 3 through 5 will apply the same idea of investigating the dynamics of an ensemble of Hamiltonians for fixed initial state and observable. The considered ensembles of Hamiltonians, however, will be markedly different (except for Sec. 4.3, where a class similar to the one from [96] is employed). Namely, we will mostly start from a given reference Hamiltonian H_0 with known dynamics and examine how the relaxation behavior changes under the influence of generic time-independent (Chapter 3) or time-dependent (Chapter 5) perturbations V. The total Hamiltonian thus takes the form $H_\lambda = H_0 + \lambda V$ with H_0 fixed and V chosen from a suitable ensemble of perturbation operators. In essence, we will thus adopt a similar setting as in Deutsch's groundbreaking work [114] which, besides foreshadowing the ETH (see Sec. 2.2.2), used such an approach to argue that generic perturbations will thermalize an isolated many-body system in the long run (see also Sec. 3.8). For the special case that H_0 describes a small system and a large bath, independent of each other, and V mediates an interaction between them, such an approach was also utilized to discern features of the system's reduced density operator, namely a dynamical-typicality property [205], a characterization of its time dependence [206], and its approximate form and fluctuations at late times [207, 208].

Random matrix theory. Given that operators are represented by matrices once a basis has been fixed, the mathematical foundations of our intention to investigate statistical ensembles of Hamiltonians are provided by *random matrix theory* [163, 209–211]. The idea to study properties of random matrices as a proxy for complicated many-body Hamiltonians goes back to Wigner [212–215], who suggested and successfully employed such an approach to explain features of atomic nuclei (see also Ref. [216], which highlights the typicality spirit of this approach). Since then, a flurry of different random matrix ensembles have been studied in the mathematical physics community (see, for example, the books and reviews [163, 209–211]), sometimes more, sometimes less closely related to concrete physical models. Special attention is usually devoted to the statistical properties of eigenvalues and eigenvectors, and the latter in particular will play a key role in all our investigations as well.

The most basic and most intensely studied of these ensembles are the family of Gaussian matrices comprising the *Gaussian Orthogonal Ensemble* (GOE), the *Gaussian Unitary Ensemble* (GUE), and the *Gaussian Symplectic Ensemble* (GSE). The GOE and GUE as well as generalizations thereof will be of particular interest to us. In the case of the GOE (GUE), the ensemble consists of real symmetric (complex Hermitian) matrices with independent—apart from the symmetry (Hermiticity) constraint—and identically distributed entries drawn from an unbiased real-valued (complex-valued) normal distribution. (For the GUE, diagonal entries are naturally also real-valued.) In fact, the adopted ensembles of perturbations V in our class of Hamiltonians $H_\lambda = H_0 + \lambda V$ will consist of similar matrices, but will include more general, possibly non-Gaussian distributions with an optional banded and sparse structure (see Sec. 3.3 for a proper definition). Early variants of such random matrices were again introduced and examined by Wigner [213, 216]. Of particular relevance from a methodological point of view are studies by Fyodorov, Mirlin and co-workers [217–219], who investigated structurally similar ensembles of banded and/or sparse matrices using *supersymmetry methods* [163, 220–225] (see also Appendices B and C). These methods will form an essential part of our derivations, too.

* * *

This concludes our survey of relevant concepts and results related to the intended investigations of this thesis. Provided with a solid understanding of the rules and history of the game as well as an ambitious strategy at hand, we are all set to take to the field and see how things play out.

3 Perturbed relaxation

The by now well-established results presented in the previous chapter provide a quite profound understanding of why isolated many-body quantum systems equilibrate in the long run, and they furthermore assert that the eventually reached equilibrium state can generically be described in terms of the traditional thermodynamic ensembles. As mentioned before, much less is known about the time scales on which this relaxation of a system initially out of equilibrium takes place, and even scarcer are general principles or predictions describing the dynamics (e.g., the time dependence of expectation values) of the relaxation process.

In this first main part of the thesis, we will develop such a prediction for the relaxation of isolated many-body quantum systems under the influence of time-independent perturbations. The starting point is a reference system whose dynamics is considered to be known, for example, because the system is noninteracting, integrable, or its dynamics has been measured. The theoretical prediction will describe the modifications of the dynamics arising from a perturbation of this reference system by establishing concentration of measure for suitable ensembles of perturbations and by calculating their concomitant average ("typical") effect.

The results presented in this chapter have been published in parts in Refs. [177, 226, 227]. In Sec. 3.1, we will specify the setup, formalize the goals, and sketch the steps of the derivation. In Sec. 3.2, we will then collect properties of the considered systems that are exploited during the derivation and thus required for the theory to be applicable. The perturbation ensembles inherent to the intended typicality approach will be introduced in Sec. 3.3. The key properties of these ensembles determining the relaxation behavior will be derived in Sec. 3.4. The main physical results are contained in the ensuing two sections: a bound for the deviations of the dynamics induced by the perturbations in Sec. 3.5 and an analytical prediction for the time evolution in the perturbed systems in Sec. 3.6. This latter result, in particular, will be illustrated by means of various examples in Sec. 3.7. Finally, in Sec. 3.8, we relate the result to similar findings in the literature and suggest possible ways to overcome some shortcomings of the approach and to extend it to setups that will be excluded in the present derivations.

3.1 Aims and setup

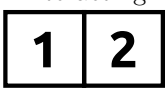
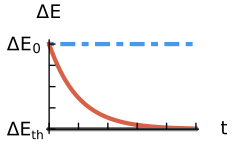
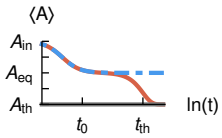
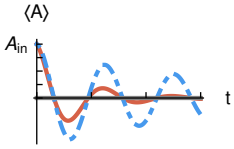
Stating the goal. The general setting to be investigated in the following involves many-body quantum systems whose Hamiltonian can be split into two parts: an unperturbed *reference system* described by the Hamiltonian H_0 and a *perturbation* V, whose strength is controlled by some (time-independent) parameter λ. Hence the perturbed Hamiltonian takes the form

$$H \equiv H_\lambda := H_0 + \lambda V\,. \tag{3.1}$$

The principal question we aim to attack is then: Provided that we know the dynamics of the reference system H_0 in a given setup, can we predict the behavior of the perturbed system under similar circumstances? This problem of characterizing the response of a given system to a perturbation is ubiquitous in physics, but the many-body nature of the systems under study poses some additional challenges that rule out standard approaches such as elementary (Rayleigh-Schrödinger) perturbation theory. The reason is that the density of states of the many-body Hamiltonian H grows exponentially with the system's degrees of freedom, cf. Sec. 2.1. The standard perturbation series is thus plagued by exceedingly small denominators, restricting its applicability to time scales far below the typically observed relaxation times of macroscopic systems. In the following, we will therefore tackle this problem by means of a nonperturbative typicality approach as sketched in Sec. 2.3.

To be more specific, we assume that the system is prepared in some *initial state* at time $t = 0$, which may be pure or mixed and is generally given by a density operator $\rho(0)$. Since the Hamiltonians considered in this section are time independent, the state at any later time $t > 0$ immediately follows as $\rho_\lambda(t) := \mathrm{e}^{-\mathrm{i}H_\lambda t}\,\rho(0)\,\mathrm{e}^{\mathrm{i}H_\lambda t}$ (see above Eq. (2.4)).

Table 3.1: Three example scenarios that can be modeled in terms of Hamiltonians of the form (3.1) and where the reference dynamics may be known explicitly: weakly coupled subsystems, integrability-breaking perturbations, and quantum quenches. In the right column, the dash-dotted blue and solid red lines sketch possible observable dynamics of the unperturbed and perturbed systems, respectively.

H_0	$H = H_0 + \lambda V$	example dynamics
isolated subsystems 	interacting 	
integrable	nonintegrable	 "prethermalization"
isolated system	quench	

A physical quantity whose dynamics we can record is described by an experimentally relevant observable A, i.e., a self-adjoint operator with a finite measurement range ΔA and resolution δA (cf. Sec. 2.2). Our focus will be on the time-dependent expectation values of this operator A,

$$\langle A \rangle_{\rho_\lambda(t)} = \mathrm{tr}[\rho_\lambda(t) A] \tag{3.2}$$

as defined in (2.1). In more technical terms, the overall goal can then be stated as follows: Assuming that the reference dynamics $\langle A \rangle_{\rho_0(t)}$ of the unperturbed system H_0 and some essential properties of the perturbation V are known, we wish to characterize the perturbed dynamics $\langle A \rangle_{\rho_\lambda(t)}$.

Applications. Generally speaking, there are a variety of physical situations that can potentially be modeled by such an approach. Three scenarios are sketched in Tab. 3.1. For example, the reference system may consist of two isolated subsystems, which includes the canonical system-plus-bath setting, but the subsystems may also be of similar size. For the sake of the argument, assume that each of them is in a thermal equilibrium state (in the sense of Sec. 2.2) such that the expectation value of, say, their energy difference is a constant. After bringing them into contact by means of a sufficiently weak coupling interaction λV, the joint system will usually relax to a new (thermal) equilibrium state, generally implying a relaxation of the energy difference between the two compartments to the new equilibrium value.

Another interesting case arises if the reference system is integrable. The dynamical behavior can then often be computed analytically, and as discussed briefly at the end of Sec. 2.2.3, adding a weak integrability-breaking perturbation commonly leads to prethermalization, where the perturbed dynamics initially resembles the unperturbed behavior before eventually departing towards thermal equilibrium.

Finally, the reference Hamiltonian may describe some system for which one happens to know the dynamics. We can then ask how this behavior is modified when changing some parameter of that Hamiltonian, called a *quantum quench* [7, 83, 174] in the literature. Note that this term is sometimes used in a more restrictive sense to refer to situations in which the initial state $\rho(0)$ is the ground state or at least a thermal state of H_0 (see also Sec. 2.2.3). For our purposes, however, the initial state need not be of this kind.

It should be emphasized that the results obtained in the following will not describe all situations that fall into these three classes of example scenarios. The prerequisites for the Hamiltonian, the initial state, and the observables will be specified more precisely in Sec. 3.2. Before collecting those, we briefly sketch the steps of our calculation to predict $\langle A\rangle_{\rho_\lambda(t)}$ from $\langle A\rangle_{\rho_0(t)}$.

Outline of the derivation. Our method of choice for finding a prediction for the perturbed dynamics is based on the notion of typicality as explained in Sec. 2.3. In the present setting, where we consider the reference dynamics generated by H_0 in (3.1) as given, the idea is to study the effect of an entire class or *ensemble* of similar perturbations simultaneously instead of the true perturbation of interest alone. As usual, this is motivated by common phenomenological evidence that the observable behavior of macroscopic systems is somewhat insensitive to the microscopic details of its constituents. In other words, only a few macroscopic characteristics of the true perturbation V are expected to actually matter with respect to the time-dependent expectation values $\langle A\rangle_{\rho_\lambda(t)}$ of experimentally relevant observables. If all members of the chosen ensemble share those key characteristics, then they will all lead to the same macroscopically observable relaxation behavior. The typicality argument is slightly weaker in that it shows that the overwhelming majority of perturbations result in practically indistinguishable dynamics. Perturbations belonging to this vast majority are then called "typical," and it remains to be argued why the true perturbation is one of these typical members. Naturally, the key here is a proper choice of the ensemble so that the typical perturbations exhibit the aforementioned essential characteristics regarding the perturbed dynamics.

The ensembles we consider will be motivated and defined in the subsequent Secs. 3.2 and 3.3. For now, we simply assume that an appropriate ensemble has been chosen, and we denote ensemble averages over V-dependent quantities by the symbol $\mathbb{E}[\cdots]$. Following the recipe from Sec. 2.3.1, the derivation then consists of three crucial steps: First, we compute the ensemble-averaged perturbed expectation values $\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}]$. Second, we consider the deviations

$$\xi_V(t) := \langle A\rangle_{\rho_\lambda(t)} - \mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}] \tag{3.3}$$

of one particular realization from the average behavior and calculate the variance $\mathbb{E}[\xi_V(t)^2]$. By showing that this variance is exceedingly small for reasonably large system sizes, we establish the above-mentioned typicality property. Third, we demonstrate agreement with concrete model systems to suggest that the obtained prediction is indeed of somewhat generic nature.

Denoting the eigenvalues and eigenvectors of the Hamiltonian H_λ from (3.1) by E_n^λ and $|n\rangle_\lambda$, respectively, the perturbed time-dependent expectation values (3.2) can be written as

$$\langle A\rangle_{\rho_\lambda(t)} = \sum_{m,n} \mathrm{e}^{\mathrm{i}(E_n^\lambda - E_m^\lambda)t}\, \rho_{mn}^\lambda(0)\, A_{nm}^\lambda \tag{3.4}$$

with $\rho_{mn}^\lambda(0) := {}_\lambda\langle m|\rho(0)|n\rangle_\lambda$ and $A_{nm}^\lambda := {}_\lambda\langle n|A|m\rangle_\lambda$. We recall that the operators $\rho(0)$ and A are independent of the perturbation V, whereas their matrix elements $\rho_{mn}^\lambda(0)$ and A_{nm}^λ are not. For computational reasons, it is advantageous to split off the V dependence by introducing the unitary transformation

$$U_{n\mu} := {}_\lambda\langle n|\mu\rangle_0 \tag{3.5}$$

between the eigenbases of the unperturbed Hamiltonian H_0 and the perturbed H_λ. Note that the dependence of the $U_{n\mu}$ on λ is notationally suppressed. Furthermore, we remark that we will use Latin indices to refer to the perturbed eigenbasis $\{|n\rangle_\lambda\}$ and Greek indices for the unperturbed eigenbasis $\{|\mu\rangle_0\}$ throughout the rest of this thesis. When dealing with the unperturbed Hamiltonian H_0, we will also drop the superscript $\lambda = 0$ for the energy levels $E_\mu := E_\mu^0$ and matrix elements such as $A_{\mu\nu} := A_{\mu\nu}^0 = {}_0\langle\mu|A|\nu\rangle_0$.

Employing the eigenvector overlaps (3.5) in (3.4), we then obtain

$$\langle A\rangle_{\rho_\lambda(t)} = \sum_{m,n} \mathrm{e}^{\mathrm{i}(E_n^\lambda - E_m^\lambda)t} \sum_{\substack{\mu_1,\mu_2,\\ \nu_1,\nu_2}} U_{m\mu_1} U_{n\mu_2} U_{m\nu_1}^* U_{n\nu_2}^*\, \rho_{\mu_1\nu_2}(0)\, A_{\mu_2\nu_1}\,. \tag{3.6}$$

Thus the random quantities in this relation are the energies E_n^λ and eigenvector overlaps $U_{n\mu}$, which inherit their distributions from the perturbation V, whereas the matrix elements $\rho_{\mu_1\nu_2}(0)$ and $A_{\mu_2\nu_1}$ in the H_0 eigenbasis are fixed. The ensemble average can therefore be written as

$$\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}] = \sum_{\substack{\mu_1,\mu_2,\\ \nu_1,\nu_2}} \rho_{\mu_1\nu_2}(0)\, A_{\mu_2\nu_1}\, \mathbb{E}\left[\sum_{m,n} \mathrm{e}^{\mathrm{i}(E_n^\lambda - E_m^\lambda)t}\, U_{m\mu_1} U_{n\mu_2} U^*_{m\nu_1} U^*_{n\nu_2}\right]. \tag{3.7}$$

Hence the key task in the first step of the typicality recipe is to calculate the average on the right-hand side of this equation involving four factors of eigenvector overlaps $U_{n\mu}$. As for the second step, which is the computation of the variance

$$\mathbb{E}\big[\xi_V(t)^2\big] = \mathbb{E}\Big[\big(\langle A\rangle_{\rho_\lambda(t)}\big)^2\Big] - \big(\mathbb{E}\big[\langle A\rangle_{\rho_\lambda(t)}\big]\big)^2\,, \tag{3.8}$$

we consequently need to evaluate a similar ensemble average comprising eight factors of $U_{n\mu}$ and two additional energy summands in the exponent.

3.2 Prerequisites

Having formulated the goals and laid out our method to tackle the question of perturbed relaxation in many-body quantum systems, we will now specify more concretely the classes of physical systems that may be amenable to our theoretical approach. In the subsequent derivations, it is assumed implicitly or explicitly that the systems under study satisfy the following prerequisites. Nevertheless, it should also be said that the fulfillment of all those requirements does not guarantee the successful theoretical prediction of a system's behavior. As explained in Sec. 2.3, the difficulty to determine *a priori* whether a chosen ensemble emulates all relevant properties of a given system is inherent to the typicality approach. Hence the following list of requirements should rather be understood as a guideline of conditions under which a successful modeling becomes more likely (and whose violation diminishes these chances). The eventual justification can and will only be given *a posteriori* in Sec. 3.7, where the theoretical prediction is compared to concrete numerical and experimental results.

The first prerequisite addresses the reference Hamiltonian H_0 and the initial state $\rho(0)$:

Prerequisite (i): Homogeneous density of states. Recalling the properties of isolated systems collected at the end of Sec. 2.1, the reference system H_0 prepared in the state $\rho(0)$ should have a well-defined macroscopic energy. Hence there exists an energy window $I_\mathcal{E}$ (Eq. (2.9)) such that the level populations $\rho_{\mu\mu}(0)$ are negligible if $E_\mu \notin I_\mathcal{E}$. For this window $I_\mathcal{E}$, we require that the density of states $D(E)$ from (2.11) should be uniform to a good approximation, $D(E) \approx \varepsilon^{-1} = \text{const}$ (cf. Eq. (2.13)). Observing that $I_\mathcal{E}$ is a macroscopically small interval by definition, this assumption is usually fulfilled rather generically.

Since the state $\rho(0)$ (and thus $\rho(t)$ for all $t > 0$) is essentially contained in $I_\mathcal{E}$, we can technically replace the "true" Hamiltonian H_0 by an auxiliary one that is identical to H_0 on $I_\mathcal{E}$, but exhibits similar properties *ad infinitum* outside of $I_\mathcal{E}$, too, notably a constant $D(E)$. The reason is that $|\rho_{\mu\nu}(0)|^2 \leq \rho_{\mu\mu}(0)\,\rho_{\nu\nu}(0)$ due to the Cauchy-Schwarz inequality, hence any contributions to time-dependent expectation values such as (3.6) from outside of $I_\mathcal{E}$ are also negligible. Such a natural continuation of the energy window will sometimes be mathematically convenient.

We emphasize that the initial state $\rho(0)$ does not define the window $I_\mathcal{E}$. The only condition is that $\rho(0)$ should be contained in that window up to negligibly small populations. Apart from that, the initial state is essentially arbitrary and may thus, in particular, populate only a small subset of states in $I_\mathcal{E}$ or may even be a single energy eigenstate. As will be discussed in detail in Sec. 3.6.3, the nature of the considered perturbation ensembles will entail that the initial state should not exhibit any macroscopic *spatial* inhomogeneities in order for the eventually obtained theoretical prediction of the dynamics to be applicable. However, this is not a formal requirement for the derivation and thus not demanded here.

As far as the admissible perturbations in (3.1) are concerned, we assume that they exhibit the following five key properties:

Prerequisite (ii): Sufficiently weak perturbations. For our intention to relate the behavior of the perturbed system (3.1) to the dynamics of the reference system H_0 to be meaningful, it should be clear that admissible perturbations cannot become arbitrarily strong. Specifically, we therefore require that the perturbation should not change the system's thermodynamic properties. Phase transitions induced by V are thus explicitly ruled out. The principal consequence of interest is that also for the perturbed system and its levels E_n, there exists an energy window similar to $I_\mathcal{E}$ from (2.9) which contains the initial state $\rho(0)$. Notably, it is taken for granted that this energy window exhibits the same uniform density of states $D(E) \approx \varepsilon^{-1}$ since $D(E) = \mathrm{e}^{\mathcal{S}(E)/k_\mathrm{B}}\mathcal{S}'(E)$ is directly related to the system's thermodynamic properties via Boltzmann's entropy $\mathcal{S}(E)$ from (2.6).

Given this intimate relation between thermodynamic properties and the distribution of energy levels, the present weak-perturbation prerequisite then essentially means that Prerequisite (i) should apply similarly to H_λ and H_0. Away from the edges, a substantial degree of "rigidity" or "stiffness" of the spectrum can actually be expected quite generically if H_λ is nonintegrable due to a so-called level-repulsion mechanism [163]: The "evolution" of the energy levels with λ can be described as a one-dimensional system of particles at positions E_n^λ interacting via a repulsive logarithmic two-body potential, the so-called *Pechukas-Yukawa gas* [228, 229]. As a consequence, the levels will fluctuate with λ, but tend to avoid forming clusters or crossing each other. Away from the borders, the levels thus feel an effective pressure, which stabilizes their distribution and in fact furthermore promotes an approximately homogeneous spreading (cf. Prerequisite (i)).

Prerequisite (iii): Sufficiently strong perturbations. Our intention to model the influence of perturbations within a typicality framework necessitates that many degrees of freedom are involved in mediating the effect of V in order for self-averaging to take place (cf. Sec. 2.3.1). The perturbation should therefore significantly mix a large number of unperturbed levels. More precisely, this means that the overlaps (3.5) between the unperturbed and perturbed eigenvectors should extend across an energy scale Γ_v with $\Gamma_v \gg \varepsilon$. In a coarse-grained sense that will become clear below, $U_{n\mu}$ should thus be nonnegligible for $|E_n^\lambda - E_\mu| \lesssim \Gamma_v$. Together with Prerequisite (ii), which implies $\Gamma_v \ll \Delta_\mathcal{E}$ with $\Delta_\mathcal{E}$ being the width of the energy window $I_\mathcal{E}$ from (2.9), we thus need $\varepsilon \ll \Gamma_v \ll \Delta_\mathcal{E}$. Because of the extreme level density of typical many-body systems (see also Sec. 2.1), there still exists a wide range of parameters λ satisfying these bounds. In particular, we can and will take for granted that the number of levels N_v that get mixed by the perturbation is exponentially large in the system's degrees of freedom f [226], i.e.,

$$N_v := \Gamma_v/\varepsilon = 10^{\mathcal{O}(f)}\,. \tag{3.9}$$

Note that perturbations violating $\varepsilon \ll \Gamma_v$ (or $N_v \gg 1$) are actually so weak that they do not lead to notable deviations from the unperturbed behavior on any reasonable time scale [226]: If Γ_v were on the order of ε, meaning that the perturbation mixes only levels within a few multiples of the mean level spacing, it could only induce modifications on time scales ε^{-1}, too, because it only affects the corresponding frequencies $E_n^\lambda - E_m^\lambda \sim \varepsilon$ in (3.6). Considering the extremely small level spacing in many-body systems, however, the time scale associated with ε^{-1} is unimaginably large and typically exceeds the age of the universe by many orders of magnitude. For all times of interest, the perturbed dynamics would thus not be distinguishable from the unperturbed behavior. As our final result will turn out to entail this limit correctly, the requirement $\varepsilon \ll \Gamma_v$ and thus the present Prerequisite (iii) may be dropped again *a posteriori*.

As usual in statistical physics, we can expect that the observable behavior of the considered many-body system is characterized well by a few macroscopic parameters. Prerequisites (ii) and (iii) did not incorporate any more specific properties of the perturbation other than its overall strength. As we will see in Sec. 3.6, this can indeed be sufficient to predict the behavior under weak perturbations. However, the theoretical prediction and its applicability can be improved significantly if additional information about the perturbation is included. The following assumption therefore asks for a more detailed, energy-resolved account of the perturbation strength.

Prerequisite (iv): Perturbation profile. On a coarse-grained level, the magnitude of the perturbation matrix elements $V_{\mu\nu} := {}_0\langle\mu|V|\nu\rangle_0$ within the energy window $I_\mathcal{E}$ should only depend on the

energy difference $E_\mu - E_\nu$ of the coupled levels. We express this symbolically as

$$|V_{\mu\nu}|^2 \simeq \sigma_v^2(E_\mu - E_\nu)\,, \tag{3.10}$$

where '$\simeq$' here is supposed to mean an asymptotic equality in sufficiently large systems when the matrix elements on the left-hand side are locally averaged over eigenstates $|\mu\rangle_0$ and $|\nu\rangle_0$ that are close-by in energy. The quantity $\sigma_v^2(E)$ will be called the *perturbation profile* henceforth. Apart from a potential discontinuity at $E = 0$, which will be addressed in the subsequent Prerequisite (v), it will be—basically by definition—a smooth and, compared to the mean level spacing ε, slowly varying function of E. In addition, it is an even function by definition, $\sigma_v^2(-E) = \sigma_v^2(E)$, because V is Hermitian.

Supported by semiclassical arguments [219, 230], analytical studies of lattice systems [102, 231], and many numerical examples [130, 134, 232–234] (see also Fig. 3.2 below), realistic perturbations indeed commonly exhibit a well-defined perturbation profile $\sigma_v^2(E)$. Moreover, the function $\sigma_v^2(E)$ is often found to decay to zero as $E \to \infty$, leading to a so-called *banded* matrix structure. As a simple example, consider a reference Hamiltonian H_0 that is noninteracting, meaning that its eigenstates $|\mu\rangle_0$ are momentum modes. For a perturbation consisting of local interactions, meaning that the interaction strength decays on a length scale ξ with the spatial separation of the degrees of freedom, standard Fourier analysis then suggests that interactions fall off on a scale proportional to ξ^{-1} in momentum space (recalling $\hbar = 1$). Hence locality of interactions, which is a generic physical assumption, is expected to translate into a banded matrix structure of $V_{\mu\nu}$ in this scenario. However, we point out that $\sigma_v^2(E)$ is not required to vanish for large E. For example, in the above mentioned weak-perturbation limit, it often suffices to consider $\sigma_v^2(E) = \text{const}$.

Besides being possibly banded, the matrix $V_{\mu\nu}$ is also often found to be *sparse* [209, 219, 233, 235], which is to say that the majority of the matrix elements $V_{\mu\nu}$ vanish. For instance, in our example of a noninteracting reference system H_0, the eigenbasis $\{|\mu\rangle_0\}$ is composed of single-particle states. If the perturbation adds few-body interactions, meaning that only a limited number of particles interact simultaneously, the resulting matrix $V_{\mu\nu}$ will indeed exhibit a sparse structure.

Associated with the perturbation profile $\sigma_v^2(E)$ there are two important quantifiers which encode characteristics of V on an even more coarse-grained level. The first one, defined by

$$\sigma_v := \lim_{E\to 0+} \sqrt{\sigma_v^2(E)}\,, \tag{3.11}$$

sets an overall energy scale of the operator V and will be called the *intrinsic strength* of the perturbation in the following. Note that we only distinguish the squared intrinsic strength σ_v^2 from the perturbation profile $\sigma_v^2(E)$ by omission of the argument. The total perturbation strength in (3.1) can thus be estimated as $\lambda\sigma_v$. In fact, it will turn out later that the more prominent characteristic is

$$\alpha_v := \frac{\sigma_v^2}{\varepsilon}\,, \tag{3.12}$$

which relates the intrinsic perturbation strength to the mean level spacing of the unperturbed H_0 and thereby connects the energy scales of H_0 and V. Being slightly sloppy in the naming conventions, we will also refer to α_v as the intrinsic perturbation strength.

The second quantifier, defined by

$$\Delta_v := \frac{1}{\sigma_v^2}\int_0^\infty \mathrm{d}E\,\sigma_v^2(E)\,, \tag{3.13}$$

measures the energy range across which unperturbed eigenstates are coupled by the perturbation and will be called its *band width*. This band width Δ_v should not be confused with the eigenvector mixing scale Γ_v from above (3.9). The former quantifies the absolute energy range of directly coupled unperturbed levels via the perturbation, whereas the latter measures the effective mixing of the perturbed and unperturbed eigenvectors, i.e., the relative energy range from which unperturbed states contribute significantly (compared to all others) to a given perturbed one (and vice versa). In a perturbation-theoretic picture, Γ_v can thus receive "higher-order corrections" and will generally deviate substantially from Δ_v. Furthermore, we emphasize that the perturbation profile $\sigma_v^2(E)$ is

not required to decay as $E \to \infty$; hence Δ_v may be infinite if the matrix $V_{\mu\nu}$ is not or only very weakly banded.

Up until now, the demanded properties can be regarded as a specification of our model of isolated many-body quantum systems and admissible perturbations, guided by physical principles and insights about the setting to be investigated. In contrast, the following final two prerequisites are included mostly for technical reasons.

Prerequisite (v): Sufficiently small level fluctuations. The fluctuations of the energy levels induced by the perturbation should be sufficiently small so that their effect on the relaxation dynamics is negligible. We observe that this assumption is related to Prerequisite (ii), which demanded, in particular, that the energy window $I_\mathcal{E}$ and the mean level spacing ε should not change under the influence of the perturbation. As mentioned before, the generic level repulsion mechanism is a first reason why this prerequisite is usually a sensible assumption. More specifically, its most prominent exploitation occurs in connection with ensemble averages of the form (3.7): We will assume that in such averages, we can replace the difference $E_n^\lambda - E_m^\lambda$ between perturbed energy levels in the exponential by its mean value $\mathbb{E}[E_n^\lambda - E_m^\lambda] = E_n - E_m$, where E_n denotes the energy levels of the unperturbed H_0 as before. Indeed, we will verify *a posteriori* that even considerable fluctuations extending over large multiples of the level spacing ε have no noticeable effect on the observable dynamics within the typical relaxation time scales. Technically speaking, this means that eigenvector mixing as encoded in the factors of $U_{n\mu}$ in Eq. (3.7) vastly dominates over level fluctuations as far as the relaxation process is concerned.

To estimate the time scale associated with the level fluctuations, we can employ a typicality approach similar to the one outlined in Sec. 3.1. We define the difference $\epsilon_n^\lambda := E_n^\lambda - E_n$ between the perturbed and unperturbed energy levels and consider how the reference dynamics $\langle A\rangle_{\rho_0(t)}$ is changed when replacing all energies E_μ by $E_\mu + \epsilon_\mu^\lambda$. In fact, since this amounts to choosing a diagonal perturbation matrix $V_{\mu\nu}$, it is just a special (and particularly simple) case of the more general setting to be considered below,

$$\lambda V_{\mu\nu} := \delta_{\mu\nu}\, \epsilon_\mu^\lambda \,. \tag{3.14}$$

For the sake of the argument, we assume that the ϵ_μ^λ are independent and follow a normal distribution of mean zero and variance $\lambda^2\sigma_0^2$; the precise distribution is irrelevant for our present purposes because we are only interested in a rough estimate of the time scales. Note that a vanishing mean, in particular, can be assumed without loss of generality because we can always add a trivial constant to the perturbation that shifts all energies, but does not alter the dynamics, which only depends on energy differences. Substituting into (3.7) and exploiting $U_{n\mu} = \delta_{n\mu}$ yields

$$\begin{aligned}\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}] &= \sum_{\mu\neq\nu} \mathrm{e}^{\mathrm{i}(E_\nu - E_\mu)t}\, \mathbb{E}\Big[\mathrm{e}^{\mathrm{i}(\epsilon_\nu^\lambda - \epsilon_\mu^\lambda)t}\Big]\, \rho_{\mu\nu}(0)\, A_{\nu\mu} + \sum_\mu \rho_{\mu\mu}(0)\, A_{\mu\mu} \\ &= \langle A\rangle_{\overline{\rho}_0} + \mathrm{e}^{-\lambda^2\sigma_0^2 t^2}\, \big[\langle A\rangle_{\rho_0(t)} - \langle A\rangle_{\overline{\rho}_0}\big]\,,\end{aligned} \tag{3.15}$$

where $\overline{\rho}_0$ denotes the time-averaged state of the unperturbed dynamics (cf. Eqs. (2.16) and (2.17)). Averaged over the ensemble, the level fluctuations thus start to noticeably affect the relaxation only for times $t \gtrsim (\lambda\sigma_0)^{-1}$. Moreover, the variance (3.8) can be bounded from above as detailed in Appendix E.1 by

$$\mathbb{E}\big[\xi_V(t)^2\big] \leq 3\Delta A\sqrt{p_{\max}} \tag{3.16}$$

with the measurement range ΔA of A as defined in (2.18) and $p_{\max} := \max_\mu \rho_{\mu\mu}(0)$ as in (2.14). As long as the initial state is spread out sufficiently broadly across the pertinent energy window $I_\mathcal{E}$ such that Eq. (2.14) holds, the average effect (3.15) is then also observed typically, i.e., for the vast majority of combined level fluctuations ϵ_μ^λ (cf. Eq. (2.29) for quantitative details). Consequently, the deviations induced by fluctuating energy levels scale like

$$\frac{\big|\langle A\rangle_{\rho_\lambda(t)} - \langle A\rangle_{\rho_0(t)}\big|}{\Delta A} \leq \frac{\big|\langle A\rangle_{\rho_\lambda(t)} - \langle A\rangle_{\rho_0(t)}\big|}{|\langle A\rangle_{\rho_0(t)} - \langle A\rangle_{\overline{\rho}_0}|} \simeq (1 - \mathrm{e}^{-\lambda^2\sigma_0^2 t^2}) = (1 - \mathrm{e}^{-\mathbb{E}[(\epsilon_\mu^\lambda)^2]t^2}) \leq t^2\, \mathbb{E}[(\epsilon_\mu^\lambda)^2]\,, \tag{3.17}$$

and if t_R is the observed relaxation time scale of the system, the effect of level fluctuations is expected to be negligible provided that

$$t_\mathrm{R} \ll (\lambda\sigma_0)^{-1} = \left(\mathbb{E}[(\epsilon_\mu^\lambda)^2]\right)^{-1/2} . \tag{3.18}$$

Of course, a more quantitative assessment of the admissible fluctuations will only be possible once we know the actual relaxation time scale associated with the perturbed Hamiltonian (3.1) and is therefore postponed to Sec. 3.6.3. Furthermore, we remark that the deviations between the reference dynamics $\langle A\rangle_{\rho_0(t)}$ and the dynamics $\langle A\rangle_{\rho_\lambda(t)}$ observed upon applying the diagonal perturbation (3.14) can also be bounded rigorously [177] by

$$\frac{\left|\langle A\rangle_{\rho_\lambda(t)} - \langle A\rangle_{\rho_0(t)}\right|}{\Delta A} \leq |t| \max_\mu |\epsilon_\mu^\lambda| . \tag{3.19}$$

Finally, the fact that the time scale associated with the level fluctuations is usually much larger than the observed relaxation time is also the reason why we can allow the diagonal elements $V_{\mu\mu}$ of admissible perturbations to fluctuate significantly stronger than the off-diagonal $V_{\mu\nu}$ as indicated in Prerequisite (iv) below Eq. (3.10), where the perturbation profile $\sigma_v^2(E)$ was permitted to be discontinuous at $E = 0$. In combination with the results from Secs. 3.4 and 3.6, the present Prerequisite (v) and particularly Eqs. (3.18) and (3.19) specify how strong this discontinuity may become in practice. Generically, at least if H_0 satisfies the ETH and V is a physical observable, the fluctuations of diagonal and off-diagonal matrix elements should still be of similar order (cf. the ETH ansatz (2.24)). Somewhat larger fluctuations of the $V_{\mu\mu}$ compared to the $V_{\mu\nu}$ for $\mu \neq \nu$ are, however, not uncommon (see also the example in Fig. 3.2 below), and we do not require H_0 to satisfy the ETH either.

The prerequisites collected so far are quite specific and could—at least in principle—be readily tested for a given model using, for example, exact diagonalization. In fact, we will discuss one such example shortly. Before that, however, we shall complete our list of requirements by including one additional technical (and unfortunately rather vague) prerequisite about the perturbation (see also Secs. 3.3 and 3.4).

Prerequisite (vi): Sufficiently uncorrelated perturbations. Besides the trivial correlations due to the constraint $V_{\mu\nu} = V_{\nu\mu}^*$, the true perturbation will inevitably exhibit correlations due to functional interdependencies between the matrix elements $V_{\mu\nu}$ for different μ and ν. These correlations should be sufficiently weak so that their influence on the dynamics is of subleading importance. Put differently, the individual terms in H_0 and V should be "orthogonal" to each other in some operational sense, such that V adds a new aspect or characteristic to the system. An immediate consequence is that in a given setup, the splitting into H_0 and V according to (3.1) cannot simply be reversed by defining $H_0' := H_\lambda = H_0 + \lambda V$ and $\lambda V' := H_\lambda - H_\lambda' = -\lambda V$ because H_0' and V' will now typically be strongly correlated.

Example. While the above catalog of requirements may appear to impose utterly restrictive constraints on the admissible systems, we contend that these properties are in fact widely observed in realistic physical systems with sufficiently many degrees of freedom. As already mentioned above, Prerequisites (i) through (iv) in particular merely formalize physical properties of the considered systems. Incidentally, if not stated explicitly, similar assumptions are taken for granted implicitly in many works on related questions.

To illustrate these properties in an explicit example, we examine the corresponding characteristics of a two-dimensional spin-$\frac{1}{2}$ system by means of exact diagonalization, for which we employ the *Eigen3* library [236] here and in any other later examples. The reference Hamiltonian H_0 couples nearest neighbors on an $L\times L$ square lattice with open boundary conditions via isotropic Heisenberg terms,

$$H_0 = \sum_{i=1}^{L-1}\sum_{j=1}^{L} \boldsymbol{\sigma}_{i,j}\cdot\boldsymbol{\sigma}_{i+1,j} + \sum_{i=1}^{L}\sum_{j=1}^{L-1} \boldsymbol{\sigma}_{i,j}\cdot\boldsymbol{\sigma}_{i,j+1} , \tag{3.20a}$$

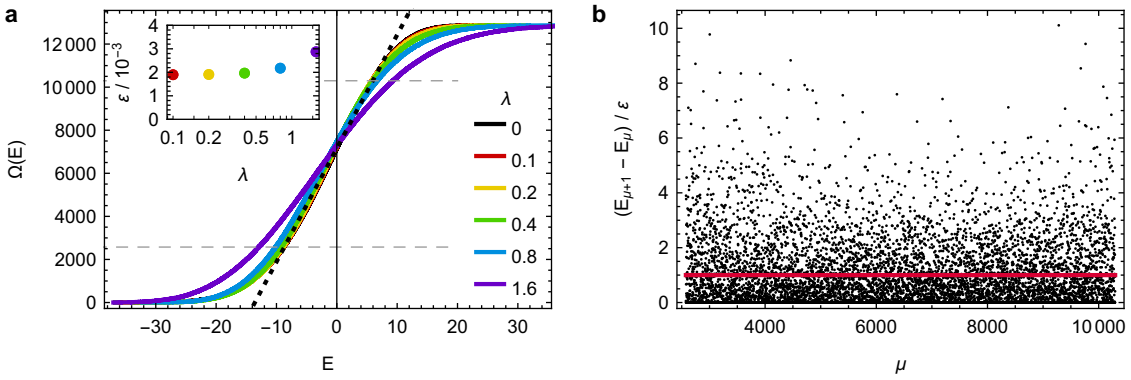


Figure 3.1: Level distribution of the spin system (3.20a) and (3.20b) on a 4×4 lattice in the zero-magnetization subsector of dimension $\binom{16}{8} = 12\,870$. **a.** Integrated density of states $\Omega(E) := \sum_m \Theta(E - E_m^\lambda)$, counting the number of levels with energy below E, for various values of the perturbation strength λ. The dotted black line corresponds to a constant density of states $D(E) = \varepsilon^{-1}$. The horizontal dashed lines indicate the window $I_\mathcal{E}$ comprising the central 60 % of levels. Inset: mean level spacing ε in $I_\mathcal{E}$ as a function of λ. **b.** Distribution of consecutive energy gaps of H_0 in $I_\mathcal{E}$, measured in units of the mean level spacing. The range of the y-axis comprises all spacings in the considered regime, i.e., no gap in $I_\mathcal{E}$ exceeds 11ε.

where $\boldsymbol{\sigma}_{i,j} := (\sigma^x_{i,j}, \sigma^y_{i,j}, \sigma^z_{i,j})$ with $\sigma^\alpha_{i,j}$ denoting the Pauli matrices acting on site (i,j). The perturbation adds additional spin-flip terms between next-nearest neighbors, i.e.,

$$V = \sum_{i,j=1}^{L-1} \left(\sigma^x_{i,j}\sigma^x_{i+1,j+1} + \sigma^y_{i,j}\sigma^y_{i+1,j+1} + \sigma^x_{i+1,j}\sigma^x_{i,j+1} + \sigma^y_{i+1,j}\sigma^y_{i,j+1} \right). \tag{3.20b}$$

Note that the total magnetization $M^z := \frac{1}{L}\sum_{i,j} \sigma^z_{i,j}$ is conserved for all H_λ. For the following analysis, we choose $L = 4$, restrict to the $M^z = 0$ subsector, and select the central 60 % of its energy levels to define the energy window $I_\mathcal{E}$ from (2.9). In the unperturbed system, this window consists of 7722 states with energies between $E = -8.8$ and $E = 5.8$ and a mean level spacing $\varepsilon = 1.90 \times 10^{-3}$.

The distribution of the energy levels for various values of λ is shown in Fig. 3.1. The figure illustrates that Prerequisites (i) and (ii) are satisfied well, at least for values of $\lambda \leq 0.8$: As apparent from Fig. 3.1a, the density of states within the selected window $I_\mathcal{E}$ is indeed homogeneous and does not vary much between different values of λ apart from the strongest perturbation with $\lambda = 1.6$. Furthermore, Fig. 3.1b shows that the fluctuations of the level spacings are rather mild with a standard deviation of 1.3ε and the largest gap of order 10ε.

To investigate the further assumptions about the perturbation, Fig. 3.2a shows the coarse-grained matrix $V_{\mu\nu}$ for $E_\mu, E_\nu \in I_\mathcal{E}$ averaged over blocks of 100×100 levels. The color gradient clearly illustrates the banded structure. A more detailed view of $V_{\mu\nu}$ for the central 60×60 levels is shown in Fig. 3.2b, where every pixel corresponds to one matrix element. Here the *sparsity* of $V_{\mu\nu}$ becomes apparent. Quantified by the fraction s of vanishing matrix elements, we estimate it to be $s = 0.855$ using a threshold of $|V_{\mu\nu}|^2 < 10^{-16}$ to count as zero. In other words, 14.5 % of the matrix elements are nonvanishing.

The resulting perturbation profile (3.10) is plotted in Fig. 3.2c using a bin width of 0.01 (approximately 5ε) for the local averages. The plot also includes a fit to the exponential form

$$\sigma_v^2(E) = \sigma_v^2 \, \mathrm{e}^{-|E|/\Delta_v} \tag{3.21}$$

with $\sigma_v^2 = 5.02 \times 10^{-3}$ (such that $\alpha_v = \sigma_v^2/\varepsilon \approx 2.64$) and $\Delta_v = 7.32$, which offers a good approximation of the empirical profile and demonstrates that the functional relation can indeed be considered as smooth and slowly varying with respect to ε despite the rather small system size, confirming Prerequisite (iv). Fluctuations of the diagonal elements are considerably larger, with the empirical variance yielding an estimate of $\sigma_0^2 \approx 6.3$ for $\sigma_0^2 := \mathbb{E}[(V_{\mu\mu} - \mathbb{E}V_{\mu\mu})^2]$ (cf. Eqs. (3.14) and (3.18)). The ratio of diagonal to off-diagonal fluctuations is thus $\sigma_0/\sigma_v \approx 35$.

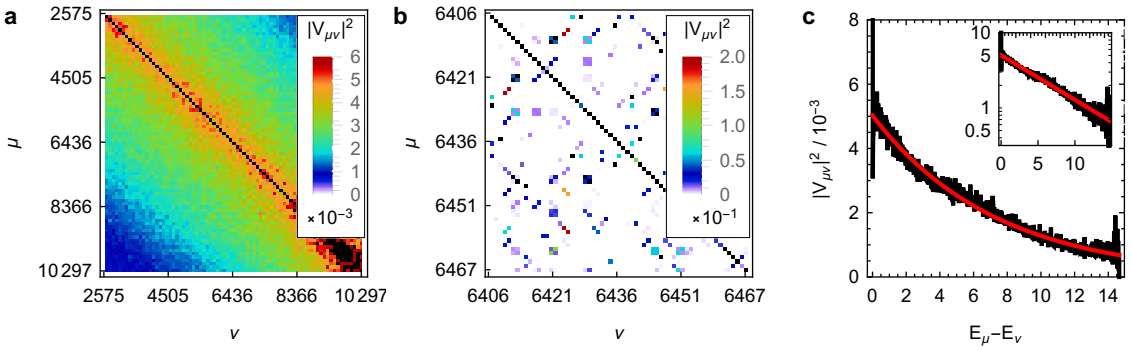


Figure 3.2: Structure of the perturbation matrix $V_{\mu\nu}$ for the spin system (3.20a) and (3.20b) on a 4×4 lattice in the zero-magnetization subsector. **a.** Squared matrix elements $|V_{\mu\nu}|^2$ of the perturbation (3.20b) in the eigenbasis of the reference Hamiltonian (3.20) in a central energy window $I_\mathcal{E}$ of 7722 states (60 % of the total Hilbert space dimension), averaged over blocks of 100×100 levels, illustrating the overall banded structure of the matrix. **b.** Squared matrix elements $|V_{\mu\nu}|^2$ in a central segment of 60×60 states with each pixel representing one matrix element, showing the sparse character of the matrix with a fraction $s = 0.855$ of vanishing entries. **c.** Coarse-grained perturbation profile (3.10) (black, bin width 0.01) and fit to the exponential form (3.21) with $\sigma_v^2 = 5.02 \times 10^{-3}$ and $\Delta_v = 7.32$ (red). The inset shows the same data with a logarithmically scaled y axis.

3.3 Perturbation ensembles

Within our typicality approach, for any given system of the form (3.1), we intend to embed the "true" perturbation of interest into an ensemble of self-adjoint operators sharing with the true V those properties which are essential for the observable dynamics. Obviously, a proper choice of the ensemble is crucial for the success of the method and the chosen class of operators should neither be too general nor too specific: In case our description is too generic, we risk to oversimplify the problem and may not be able to make any useful predictions about the perturbed behavior at all because the decisive characteristics of the perturbation are lost. On the other hand, an overly restrictive choice limits the predictive power, too, because it requires very detailed knowledge about the structure of the true perturbation, and it also obscures the physical principles taking microscopic complexity to macroscopic regularity. (The extreme example is the specification of all the exact matrix elements $V_{\mu\nu}$, which would obviously allow an excellent prediction in principle, but does not teach us anything about the mechanism behind the observed behavior.) Moreover, the necessity for a large enough ensemble of perturbations is also inherent to the intended typicality method as the underlying concentration-of-measure principle only becomes meaningful in sufficiently high-dimensional spaces.

Distribution of matrix elements. Our working hypothesis is that the five Prerequisites (ii)–(vi) required for the operator V (see Sec. 3.2) restrict the class of admissible perturbation just enough to successfully model a large variety of physical systems and settings. Hence a proper V ensemble should conform with these five assumptions, but should otherwise be unbiased and general. Using $\mathbb{E}[\cdots]$ to indicate ensemble averages as before, the most important properties of any given ensemble thus are the first two moments of the matrix elements $V_{\mu\nu}$,

$$\mathbb{E}[V_{\mu\nu}] = 0 \qquad \text{and} \qquad \mathbb{E}[|V_{\mu\nu}|^2] = \sigma_v^2(E_\mu - E_\nu)\,, \tag{3.22}$$

along with the constraint $V_{\mu\nu}^* = V_{\nu\mu}$ to ensure Hermiticity. Here $\sigma_v^2(E)$ is the perturbation profile (3.10) of the true V.

Asking for a vanishing mean is a consequence of taking the ensemble to be unbiased with respect to the choice of phase factors of the basis vectors, i.e., the eigenstates $|\mu\rangle_0$ of H_0. Instead of the $\{|\mu\rangle_0\}$, we could equally well adopt a basis $\{e^{i\phi_\mu}|\mu\rangle_0\}$ with arbitrary $\phi_\mu \in \mathbb{R}$ because $|\mu\rangle_0$ and $e^{i\phi_\mu}|\mu\rangle_0$ describe the same state and are thus physically equivalent. Hence it is natural to require that the distribution of the matrix elements $V_{\mu\nu}$ should be invariant upon changing from the $|\mu\rangle_0$ to the $e^{i\phi_\mu}|\mu\rangle_0$, i.e., $V_{\mu\nu}$ and $e^{i(\phi_\nu - \phi_\mu)}V_{\mu\nu}$ should have identical statistical properties. For $\mu \neq \nu$, this implies $\mathbb{E}[V_{\mu\nu}] = 0$, in particular. As far as the diagonal matrix elements $V_{\mu\mu}$ are concerned,

we recall that adding a constant to the perturbation does not change the dynamics (see below Eq. (3.14)). Moreover, the expected value of $|V_{\mu\mu}|$ cannot depend on the corresponding eigenvalue E_μ due to Prerequisite (iv). Hence requiring $\mathbb{E}[V_{\mu\mu}] = 0$ does not entail any loss of generality.

Choosing the variance according to (3.22) incorporates the perturbation profile (3.10) into the ensemble and thus implements Prerequisite (iv) in an ergodic sense, i.e., if the local average in (3.10) is replaced by the ensemble average. We thereby ensure that nearly all members of the perturbation ensemble share this characteristic with the true perturbation.

By analogy with the central limit theorem, it will turn out below (see Sec. 3.4, especially the discussion below Eq. (3.48)) that the distribution of the $V_{\mu\nu}$ is otherwise rather arbitrary. Therefore, we stipulate that the probability density functions of the $V_{\mu\nu}$ follow the general form

$$p_{\mu\nu}(v) := \mathbb{E}[\delta(V_{\mu\nu} - v)] = f_{|E_\mu - E_\nu|}(v)\,, \tag{3.23}$$

where $\{f_E(v)\}_{E>0}$ is a family of probability densities on $\mathbb{R}$ or $\mathbb{C}$ with mean zero and variance $\sigma_v^2(E)$, and $f_0(v)$ is a probability density on $\mathbb{R}$ having vanishing mean and variance $\sigma_v^2(0)$. Respecting the invariance property of the perturbation ensemble under phase rotations of the unperturbed eigenvectors discussed below (3.22), we can furthermore take it for granted that $f_E(v)$ only depends on the absolute value $|v|$. Notably, the form (3.23) of the distribution admits the paradigms of banded and sparse $V_{\mu\nu}$ matrices discussed in the previous section: The bandedness can be controlled in terms of the energy-dependent variance $\sigma_v^2(E)$, whereas sparsity may be accounted for by including in $f_E(v)$ a term proportional to $\delta(v)$.

The question of whether ensembles of complex Hermitian or real symmetric matrices should be preferred is known to relate to the time-reversal symmetry of the Hamiltonian of interest [163]. We will mostly focus on the case of complex Hermitian matrices ($V_{\mu\nu}$) and comment on pertinent modifications for real symmetric ones when appropriate. Eventually, the choice turns out to be basically irrelevant for the questions we are interested in here since both options lead to the same conclusions. In particular, the final prediction for the perturbed relaxation behavior will be symmetric in time also for complex Hermitian perturbation matrices.

So far, we only considered the marginal distributions $p_{\mu\nu}(v)$ of the individual matrix elements $V_{\mu\nu}$. In the following, we will exclusively deal with ensembles of random matrices with statistically independent elements (apart from $V_{\mu\nu}^* = V_{\nu\mu}$). In essence, this is a technical requirement enabling the evaluation of ensemble averages such as $\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}]$ in (3.7) at all. However, it is generally hard to identify (let alone quantify) potential correlations in any given system, so independence may also be understood as unbiasedness in the absence of more detailed information. Moreover, mild correlations will not noticeably impact the properties relevant for the observable dynamics [177]. Nevertheless, the fact that correlations do exist in physical systems is undeniable and may in principle restrict the applicability considerably [237, 238]; Prerequisite (vi) is meant to formalize to some extent under which circumstances the approach can be expected to work nonetheless. For the rest, we will come back to the question of how to improve the modeling to incorporate common types of correlations and how these may affect our results in Sec. 3.8.

Combining the assumption of statistical independence with the marginal distributions (3.23), the probability density

$$p(V) := \prod\nolimits_{\mu\leq\nu} p_{\mu\nu}(V_{\mu\nu}) \tag{3.24}$$

thus characterizes the distribution of the entire perturbation matrix. The ensemble average $\mathbb{E}[\cdots]$ can then be written explicitly as

$$\mathbb{E}[\cdots] \equiv \int [\mathrm{d}V]\, \cdots p(V)\,, \tag{3.25}$$

where $[\mathrm{d}V]$ denotes the Lebesgue measure of all independent entries of V, i.e.,

$$[\mathrm{d}V] := \prod\nolimits_{\mu\leq\nu} \mathrm{d}V_{\mu\nu} \quad \text{or} \quad [\mathrm{d}V] := \Big[\prod\nolimits_{\mu} \mathrm{d}V_{\mu\mu}\Big] \Big[\prod\nolimits_{\mu<\nu} \mathrm{d}V_{\mu\nu}\, \mathrm{d}V_{\mu\nu}^*\Big] \tag{3.26}$$

for real symmetric or complex Hermitian matrices, respectively. Moreover, for integrals over complex-valued variables v, we generally employ the definition $\mathrm{d}v\, \mathrm{d}v^* := 2\, \mathrm{d}(\mathrm{Re}\, v)\, \mathrm{d}(\mathrm{Im}\, v)$ to map them onto integrals over $\mathbb{R}^2$ (see also Appendix A.2).

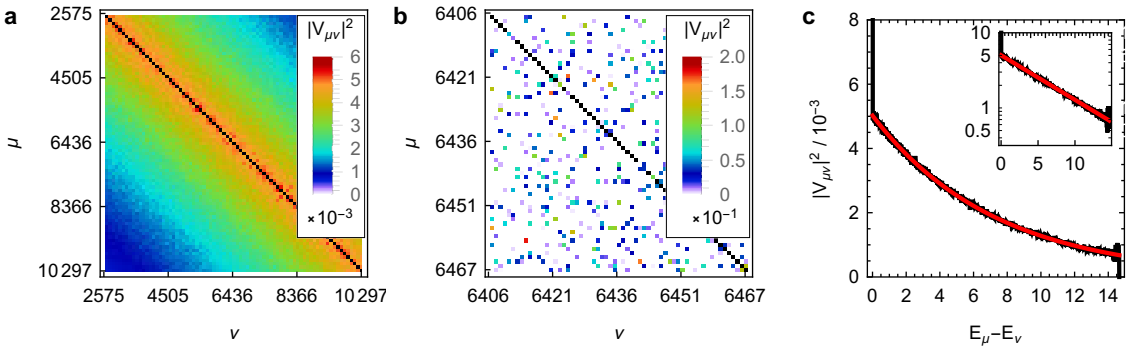


Figure 3.3: Structure of the perturbation matrix for an imitation of the spin model (3.20) from Fig. 3.2. The values of $\varepsilon = 1.90 \times 10^{-3}$, $s = 0.855$, $\sigma_v^2 = 5.02 \times 10^{-3}$, and $\Delta_v = 7.32$ are chosen identical to those of the spin system, but the unperturbed system has equally spaced energy levels, Eq. (3.27), and the matrix V is generated according to the distribution (3.28). All details regarding the visualization are as in Fig. 3.2.

Example. To showcase the embedding of a particular "true" perturbation into a random matrix ensemble with a distribution of the form (3.24), we return to the example (3.20) of a lattice spin system from the previous section. The relevant parameters extracted from the true perturbation were the level spacing $\varepsilon = 1.90 \times 10^{-3}$, the sparsity $s = 0.855$, and an approximately exponential perturbation profile (3.21) with strength parameter $\sigma_v^2 = 5.02 \times 10^{-3}$ and band width $\Delta_v = 7.32$. Based on these empirical values, we design a random matrix ensemble whose typical members exhibit the same properties. In accordance with Prerequisite (ii) of an approximately constant density of states, we take the energy levels of the unperturbed Hamiltonian to be equally spaced, meaning that

$$H_0 = \varepsilon \sum_\mu \mu \, |\mu\rangle_{00}\langle\mu| \,. \tag{3.27}$$

The perturbation matrix elements $V_{\mu\nu}$ for $\mu < \nu$ are distributed according to (3.23) with $v \in \mathbb{C}$ and

$$f_E(v) = s\,\delta(v) + (1-s)^2 \, \frac{\mathrm{e}^{-(1-s)|v|^2/\sigma_v^2(E)}}{\pi \sigma_v^2(E)} \qquad (E > 0) \,, \tag{3.28}$$

meaning that the nonvanishing off-diagonal entries follow an unbiased complex normal distribution whose variance decays as prescribed by the perturbation profile. The diagonal elements are sampled from a real normal distribution of mean zero and variance $\sigma_0^2 = 6.3$ (no sparsity).

For a single realization drawn from this ensemble, we display in Fig. 3.3 the same quantities that were previously presented for the spin model (3.20) in Fig. 3.2. Qualitatively, the resulting characteristics look quite similar. Perhaps the most striking difference is found when comparing the middle panels of both figures. Since the excerpt from the true matrix in Fig. 3.2b has more white spots below the resolution of the color scale than the emulation in Fig. 3.3b, the tails of the true distribution seemingly differ from the imitation's normal distribution. As will become clear in Sec. 3.4, such details of the distribution beyond the first two moments are expected to be basically irrelevant provided that the system is sufficiently large.

3.4 Eigenvector overlap moments

As outlined at the end of Sec. 3.1, the crucial steps in deriving our prediction for the perturbed dynamics—after exploiting Prerequisite (v) from Sec. 3.2—are calculations of ensemble averages over products of eigenvector overlaps $U_{n\mu} = {}_\lambda\langle n|\mu\rangle_0$ of the perturbed and unperturbed Hamiltonians (see Eq. (3.5)), i.e., objects of the form

$$\mathbb{E}[\, U_{n_1\mu_1} U^*_{n_1\nu_1} U_{n_2\mu_2} U^*_{n_2\nu_2} \cdots] \,. \tag{3.29}$$

The present section is devoted exclusively to this task. It is therefore of a rather technical nature and does not contain any new physical insights.

We begin in Sec. 3.4.1 with a brief introduction to the computational approach, which is based on the resolvent of the Hamiltonian H_λ and exploits so-called supersymmetry methods to extract overlap moments like (3.29) from it. In Secs. 3.4.2 and 3.4.3, we apply these techniques to evaluate moments of the form (3.29) involving two and four factors of $U_{n\mu}$, respectively. In Sec. 3.4.4, we sketch a scheme to approximate moments of fourth and higher order in terms of the results for the second moment from Sec. 3.4.2. These approximation can then be used to extend the results from Sec. 3.4.3 for the fourth moment to more general cases and to assess the eighth moment needed for the variance (3.8) of the time-dependent expectation values.

To spare the impatient reader the hassle of going through the derivations in all their glorious details, we briefly summarize those main findings of the four subsections which are most important for the subsequent analysis regarding the typical time evolution of expectation values.

Some conclusions about the general structure of expressions like (3.29) can already be drawn by exploiting the statistical invariance of the V ensemble under phase rotations of the unperturbed eigenvectors (see below Eq. (3.22)). Considering instead of $\{|\mu\rangle_0\}$ the basis vectors $\{\mathrm{e}^{\mathrm{i}\phi_\mu}|\mu\rangle_0\}$ with arbitrary phase factors $\phi_\mu \in [0, 2\pi)$, the matrix elements $V_{\mu\nu}$ transform into $\mathrm{e}^{\mathrm{i}(\phi_\nu-\phi_\mu)}V_{\mu\nu}$ and the overlap products $U_{n\mu}$ from (3.5) turn into $\mathrm{e}^{\mathrm{i}\phi_\mu}U_{n\mu}$. But as the statistical properties of the $\mathrm{e}^{\mathrm{i}(\phi_\nu-\phi_\mu)}V_{\mu\nu}$ agree with those of the $V_{\mu\nu}$ by definition, the same must hold for $\mathrm{e}^{\mathrm{i}\phi_\mu}U_{n\mu}$ and $U_{n\mu}$. Since the phases ϕ_μ may be chosen arbitrarily and independently for all μ, the average in (3.29) thus vanishes unless all factors $U_{n_k\mu_k}$ have a "partner" $U^*_{n_l\nu_l}$ such that $\mu_k = \nu_l$, much like in the Isserlis or Wick theorem [239, 240] for complex Gaussian random variables. In fact, this structure will emerge in the explicit calculations below (see, e.g., the step from Eq. (3.105) to (3.106)) as a consequence of the Isserlis-Wick theorem for (supersymmetric) Gaussian integrals (see also Appendix C). In particular, this implies that averages over products of an unbalanced number of U and U^* factors always vanish.

The first crucial insight of Sec. 3.4.2 is that the second moment $\mathbb{E}[U_{n\mu_1}U^*_{n\nu_1}]$ takes the form

$$\mathbb{E}[U_{n\mu}U^*_{n\nu}] = \delta_{\mu\nu}\, u(E_n - E_\mu)\,, \tag{3.30}$$

see Eq. (3.61), where the *overlap distribution* $u(E)$ essentially arises as the imaginary part of the ensemble-averaged resolvent $\mathbb{E}[(E-\mathrm{i}\eta-H_\lambda)^{-1}]$ in the limit $\eta \to 0$. More precisely, the latter average can be obtained from the complex-valued function $G(z)$ as $\mathbb{E}[(z-H_\lambda)^{-1}] = \sum_\mu G(z-E_\mu)|\mu\rangle_{00}\langle\mu|$ (see Eqs. (3.36) and (3.58)), where $G(z)$ in turn satisfies the nonlinear integral equation

$$G(z)\left[z - \lambda^2 \int \frac{\mathrm{d}E}{\varepsilon}\, G(z-E)\, \sigma_v^2(E)\right] = 1\,, \tag{3.31}$$

see Eq. (3.59). From the solution of this equation, the overlap distribution $u(E)$ in (3.30) can then be extracted as

$$u(E) = \frac{\varepsilon}{\pi} \lim_{\eta\to 0+} \operatorname{Im} G(E - \mathrm{i}\eta)\,, \tag{3.32}$$

see Eq. (3.60). Crucially, the integral equation (3.31) thus connects the perturbation profile $\sigma_v^2(E)$ from (3.10), which we announced as the decisive characteristic of the perturbation in Prerequisite (iv) from Sec. 3.2, to the second moment (3.30) of the eigenvector overlaps. As an aside, we remark that the overlap distribution $u(E)$ is closely related to the so-called *strength function* or *local density of states* $D_\mu(E) := \sum_n |U_{n\mu}|^2 \delta(E-E_n)$, a commonly studied characteristic in nuclear physics and random matrix theory [210, 233]. Indeed, $\mathbb{E}[D_\mu(E)] \simeq u(E-E_\mu)/\varepsilon$ if the sum is approximated by an integral (cf. the discussion below Eq. (2.12)).

An important step of the calculation thus consists in solving (3.31). Analytical solutions are obtained in three special cases. The first one basically covers weak perturbations in the sense that $\Gamma := 2\pi\lambda^2\alpha_v \ll \Delta_v$, where α_v and Δ_v were defined in (3.12) and (3.13), respectively. In this regime, the overlap distribution assumes the Breit-Wigner form

$$u(E) = \frac{\varepsilon}{2\pi} \frac{\Gamma}{E^2 + \Gamma^2/4}\,, \tag{3.33}$$

see Eq. (3.67). The second special case essentially corresponds to reasonably strong perturbations satisfying $\gamma := \sqrt{8\Delta_v \alpha_v}\,\lambda \gg \Delta_v$. Here $u(E)$ becomes a semicircle distribution,

$$u(E) = \frac{2\varepsilon}{\pi\gamma^2}\sqrt{\gamma^2 - E^2}\,\Theta(\gamma^2 - E^2)\,, \tag{3.34}$$

see Eq. (3.72). Third, an exact solution $G(z)$ in the form of a continued-fraction expansion (see Eq. (3.77)) can be given for a perturbation profile of Breit-Wigner form, $\sigma_v^2(E) = \sigma_v^2/[1 + (\pi E/2\Delta_v)^2]$. From this exact solution, approximations for $u(E)$ of arbitrary accuracy can be computed.

Not least, we also describe a method to solve the integral equation (3.31) numerically for arbitrary perturbation profiles $\sigma_v^2(E)$. Comparing such numerically exact solutions to the special cases (see Figs. 3.4 and 3.5), we observe that the function $G(z)$ (and hence also $u(E)$) is largely unaffected by the precise details of $\sigma_v^2(E)$ and well characterized by the two parameters $\alpha_v\lambda^2$ (overall perturbation strength, cf. Eq. (3.12)) and Δ_v (perturbation band width, cf. Eq. (3.13)) alone. Thus the analytic solutions derived for those three special cases in fact serve as excellent approximations in much more general settings, too.

As far as the fourth-order moment of eigenvector overlaps is concerned, the main result of Sec. 3.4.3 is the following leading-order approximation for $N_v \gg 1$, where N_v quantifies the number of levels mixed by the perturbation and is—according to Prerequisite (iii)—exponentially large in the system's degrees of freedom (cf. Eq. (3.9)). Namely, for weak perturbations such that $u(E)$ is given by (3.33), we find that

$$\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}] = \delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\,d^{n_1n_2}_{\mu_1\mu_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\left(\delta_{n_1n_2}d^{n_1n_2}_{\mu_1\mu_2} + f^{n_1n_2}_{\mu_1\mu_2}\right), \tag{3.35a}$$

where

$$d^{n_1n_2}_{\mu_1\mu_2} = u(E_{n_1} - E_{\mu_1})\,u(E_{n_2} - E_{\mu_2})\,, \tag{3.35b}$$

$$f^{n_1n_2}_{\mu_1\mu_2} = \left(\frac{\Gamma\varepsilon}{2\pi}\right)\frac{u(E_{n_1} - E_{\mu_2})\,u(E_{n_2} - E_{\mu_1}) - u(E_{n_1} - E_{\mu_1})\,u(E_{n_2} - E_{\mu_2})}{(E_{n_1} - E_{n_2})(E_{\mu_1} - E_{\mu_2})}\,, \tag{3.35c}$$

see Eqs. (3.120), (3.113b), and (3.121), respectively. Moreover, we verify this result numerically (see Fig. 3.6), establishing, in particular, that the leading-order approximation already applies to rather small values of $N_v \gtrsim 10$.

Finally, we sketch a method to approximately determine fourth and higher-order moments of eigenvector overlaps by reducing them to second-order expressions in Sec. 3.4.4, exploiting that the $U_{n\mu}$ are almost Gaussian distributed (see also the Supplemental Material of Ref. [226]).

3.4.1 Resolvent approach and supersymmetry methods

We first lay the foundation of the concrete calculations in the subsequent Secs. 3.4.2 and 3.4.3 by introducing the general framework. All calculations are formally carried out in a large, but finite Hilbert space of dimension $N \gg 1$, e.g., the number of levels in the energy window $I_\mathcal{E}$ from (2.9). Eventually, we will let $N \to \infty$ by naturally extending the system properties within $I_\mathcal{E}$ *ad infinitum* as sketched in Prerequisite (i), keeping, in particular, the perturbation strength and the density of states fixed.

Eigenvector overlaps from resolvents. The spectral properties of the Hamiltonian $H_\lambda = H_0 + \lambda V$ from (3.1) are encoded in its *resolvent* or Green's function, defined as the operator

$$\mathcal{G}(z) := (z - H_\lambda)^{-1} \equiv \sum_n \frac{1}{z - E_n^\lambda}\,|n\rangle_{\lambda\,\lambda}\langle n|\,. \tag{3.36}$$

For example, the eigenvalues $E_n^\lambda \in \mathbb{R}$ of the Hermitian operator H_λ are the poles of this $\mathcal{G}(z)$, and except for these poles $\mathcal{G}(z)$ is analytic for all $z \in \mathbb{C}$. Products of eigenvector overlaps such

as $U_{n\mu}U^*_{n\mu}$ can be expressed in terms of the matrix elements $\mathcal{G}_{\nu\mu}(z) := {}_0\langle\nu|\mathcal{G}(z)|\mu\rangle_0$ by observing that

$$\mathcal{G}_{\nu\mu}(E_n^\lambda \pm \mathrm{i}\eta) = \sum_m \frac{E_n^\lambda - E_m^\lambda \mp \mathrm{i}\eta}{(E_n^\lambda - E_m^\lambda)^2 + \eta^2}\, U_{m\mu}U^*_{m\nu} \tag{3.37}$$

for arbitrary $\eta > 0$. Taking the difference between the expressions for the lower and upper signs in front of $\mathrm{i}\eta$, we obtain

$$\mathcal{G}_{\nu\mu}(E_n^\lambda - \mathrm{i}\eta) - \mathcal{G}_{\nu\mu}(E_n^\lambda + \mathrm{i}\eta) = 2\mathrm{i}\sum_m \frac{\eta}{(E_n^\lambda - E_m^\lambda)^2 + \eta^2}\, U_{m\mu}U^*_{m\nu}\,. \tag{3.38}$$

Up to a factor of π, the fraction on the right-hand side is a Breit-Wigner or Cauchy distribution of width η for the energy difference $E_n^\lambda - E_m^\lambda$. For small η, this fraction thus approaches a Dirac-δ distribution $\pi\delta(E_n^\lambda - E_m^\lambda)$. Rewriting the sum $\sum_m \cdots$ as an integral $\int \mathrm{d}E\, D(E) \cdots$ (cf. Eq. (2.11)) and exploiting Prerequisite (i) of a homogeneous density of states $D(E) = \varepsilon^{-1}$, we can thus conclude that

$$\mathbb{E}[U_{n\mu}U^*_{n\nu}] = \frac{\varepsilon}{2\pi\mathrm{i}} \lim_{\eta\to 0+} \mathbb{E}\big[\mathcal{G}_{\nu\mu}(E_n^\lambda - \mathrm{i}\eta) - \mathcal{G}_{\nu\mu}(E_n^\lambda + \mathrm{i}\eta)\big] \tag{3.39}$$

upon averaging over the ensemble of perturbations. There are three remarks in order. First, we note that it is implicitly assumed here that the average $\mathbb{E}[U_{n\mu}U^*_{n\nu}]$ is slowly varying with n compared to the mean level spacing ε so that the approximation of a uniform density of states is indeed justified. This smoothness condition on $\mathbb{E}[U_{n\mu}U^*_{n\nu}]$, which is working standard in random matrix theory [163, 223], will turn out to hold self-consistently in the final result obtained below and can also be confirmed numerically. Second, we mention that the step from (3.38) to (3.39) can be generalized straightforwardly to higher moments (3.29) as will be detailed when needed in Sec. 3.4.3. Third, we point out that the limit entailed in (3.39) to target the singularities of $\mathcal{G}(z)$ on the real line is akin to pole prescriptions for Green's functions in other contexts such as classical electrodynamics [241]. By analogy, $\mathcal{G}(E + \mathrm{i}\eta)$ and $\mathcal{G}(E - \mathrm{i}\eta)$ with $E \in \mathbb{R}$, $\eta > 0$ are therefore also called the *retarded* and *advanced* resolvents, respectively.

The usefulness of the resolvent formalism is founded in the observation that the matrix elements $\mathcal{G}_{\nu\mu}(z)$ can be computed as Gaussian integrals with kernel $\mathcal{G}^{-1}(z) = z - H_\lambda$ as we elaborate in more detail in Appendix C. Namely, introducing the abbreviation $z^\pm := E_n^\lambda \pm \mathrm{i}\eta$ for the arguments of the retarded and advanced resolvents in (3.39), those matrix elements can be written as

$$\mathcal{G}_{\nu\mu}(z^\pm) = \frac{\mp\mathrm{i}\,\det(z^\pm - H_\lambda)}{(\pm 2\pi\mathrm{i})^N} \int \Big[\prod\nolimits_\alpha \mathrm{d}x_\alpha \mathrm{d}x_\alpha^*\Big]\, x_\nu x_\mu^* \exp\Big\{\pm\mathrm{i}\sum\nolimits_{\alpha,\beta} x_\alpha^* \left[(z^\pm - E_\alpha)\delta_{\alpha\beta} - V_{\alpha\beta}\right] x_\beta\Big\}\,, \tag{3.40}$$

where the choice of sign in the exponent ensures convergence due to $\eta > 0$. The issue with this expression is the normalization factor $\det(z^\pm - H_\lambda)$ because it is generally impractical to compute this high-dimensional matrix determinant, let alone its ensemble average. A remedy is found by extending the Gaussian integral in Eq. (3.40) to anticommuting numbers.

Supersymmetry method. The supersymmetry techniques employed in the following are based on the concept of graded algebras and vector spaces. These are generated by a set of *anticommuting* or *Grassmann numbers* $\chi_1, \chi_2, \ldots$, whose defining property is that $\chi_i\chi_j = -\chi_j\chi_i$ for any two such elements. By analogy with ordinary (commuting) numbers, it is then possible to extend notions from linear algebra and calculus to these *superspaces* of commuting and anticommuting degrees of freedom. A collection of concepts relevant for our present purposes is provided in Appendix B. For a more in-depth introduction to the method and its applications, we refer to Refs. [163, 220–225].

In our present context, these Grassmann numbers turn out to be useful because they allow us to express the determinant in the prefactor of (3.40) as a Gaussian integral, too. To wit, we introduce anticommuting numbers χ_α and χ_α^* associated with the unperturbed eigenstates $|\alpha\rangle_0$ and obtain (cf. Appendix C.2)

$$\det(z^\pm - H_\lambda) = \mathrm{i}^{-N} \int \Big[\prod\nolimits_\alpha \mathrm{d}\chi_\alpha \mathrm{d}\chi_\alpha^*\Big] \exp\Big\{\mathrm{i}\sum\nolimits_{\alpha,\beta} \chi_\alpha^* \left[(z^\pm - E_\alpha)\delta_{\alpha\beta} - V_{\alpha\beta}\right] \chi_\beta\Big\}\,. \tag{3.41}$$

To combine Eqs. (3.40) and (3.41) in compact notation, we define a supervector $X := (X_1 \cdots X_N)^\mathsf{T}$ with

$$X_\alpha := \begin{pmatrix} x_\alpha \\ \chi_\alpha \end{pmatrix} \tag{3.42}$$

consisting of commuting or *bosonic* components $X_{\alpha\mathrm{B}} := x_\alpha$ and anticommuting or *fermionic* components $X_{\alpha\mathrm{F}} := \chi_\alpha$. Note that we will employ the same language to address the different sectors of a supermatrix

$$M = \begin{pmatrix} M_{\mathrm{BB}} & M_{\mathrm{BF}} \\ M_{\mathrm{FB}} & M_{\mathrm{FF}} \end{pmatrix}, \tag{3.43}$$

too, see also Appendix B.2. Finally, we also introduce the diagonal (2×2) matrices $L^\pm := \mathrm{diag}(\pm 1, 1)$ and the shorthand $[\mathrm{d}X\mathrm{d}X^*] := \prod_\alpha \mathrm{d}x_\alpha \mathrm{d}x_\alpha^* \mathrm{d}\chi_\alpha \mathrm{d}\chi_\alpha^*$. Using (3.41) and (3.42), the resolvent matrix elements from (3.40) can then be written as

$$\mathcal{G}_{\nu\mu}(z^\pm) = \mp \mathrm{i} \int \frac{[\mathrm{d}X\mathrm{d}X^*]}{(\mp 2\pi)^N}\, x_\nu x_\mu^*\, \exp\left\{\mathrm{i}X^\dagger[(z^\pm - H_0 - \lambda V)\otimes L^\pm]X\right\}. \tag{3.44}$$

We observe that the factor $(z^\pm - H_0 - \lambda V)$ in the Kronecker product on the right-hand side acts on the "Hilbert space part" of the vector X (i.e., the indices α) whereas the factor $L^\pm$ acts on the "superspace part" (the indices B and F). In the following, we will identify, in a slight abuse of notation, $(z^\pm - H_0 - \lambda V)$ with $(z^\pm - H_0 - \lambda V)\otimes \mathbb{1}$ and $L^\pm$ with $\mathbb{1}\otimes L^\pm$ whenever this is unambiguous. The same convention will also be adopted for other operators acting trivially on either Hilbert or superspace.

The ensemble average of the resolvent matrix elements (3.44) entering (3.39) is obtained via (3.25) by integrating over the distribution $p(V)$ of the perturbations from (3.24),

$$\mathbb{E}[\mathcal{G}_{\nu\mu}(z^\pm)] = \mp \mathrm{i} \int \frac{[\mathrm{d}X\mathrm{d}X^*]}{(\mp 2\pi)^N}\, x_\nu x_\mu^* \int [\mathrm{d}V]\, p(V)\, \exp\left[\mathrm{i}X^\dagger L^\pm (z^\pm - H_0 - \lambda V)X\right]. \tag{3.45}$$

This expression provides the starting point for computing the second moment (3.39) within the intended supersymmetry approach. Representations of higher-order moments can be set up in a similar way as will be detailed directly in Sec. 3.4.3.

Modifications for real-valued perturbation ensembles. So far, we have implicitly assumed that the matrix $(V_{\mu\nu})$ is complex Hermitian. The case of real symmetric perturbation matrices can be treated similarly in principle, but the enhanced symmetry (see also the comment in the second paragraph after Eq. (3.23)) necessitates a doubling of the number of supersymmetric parameters compared to the number of random variables [221–223, 242]. The pertinent equivalent of Eq. (3.44) for real-valued V ensembles is thus

$$\mathcal{G}_{\nu\mu}(z^\pm) = \mp \mathrm{i} \int \left[\prod_\alpha \frac{\mathrm{d}x_\alpha^{(1)}\mathrm{d}x_\alpha^{(2)}\mathrm{d}\chi_\alpha \mathrm{d}\chi_\alpha^*}{\mp 2\pi}\right] x_\nu^{(1)} x_\mu^{(1)}\, \exp\left\{\frac{\mathrm{i}}{2}\sum\nolimits_{\alpha,\beta}\left[\pm x_\alpha^{(1)} J_{\alpha\beta}^\pm x_\beta^{(1)} \pm x_\alpha^{(2)} J_{\alpha\beta}^\pm x_\beta^{(2)} + \chi_\alpha^* J_{\alpha\beta}^\pm \chi_\beta - \chi_\alpha J_{\alpha\beta}^\pm \chi_\beta^*\right]\right\} \tag{3.46}$$

with real-valued $x_\alpha^{(1)}$ and $x_\alpha^{(2)}$, anticommuting χ_α and χ_α^*, and where $J_{\alpha\beta}^\pm := (z^\pm - E_\alpha)\delta_{\alpha\beta} - \lambda V_{\alpha\beta}$. In total, there are thus $N(N+1)$ commuting and $N(N+1)$ anticommuting integration variables as opposed to $\frac{N(N+1)}{2}$ independent random variables $V_{\alpha\beta}$ $(\alpha \le \beta)$.

For convenience, we will solely discuss the case of complex Hermitian perturbations in Secs. 3.4.2 and 3.4.3. In case of the second moment, the actual calculation for real symmetric V matrices proceeds in essentially the same way and yields the same result (see also Ref. [219]). For the fourth- and higher-order moments, additional terms arise due to the enhanced symmetry, and we will briefly comment on the expected modifications compared to the complex Hermitian case at the end of Sec. 3.4.3.

Outline of the calculation. The general scheme to evaluate expressions like (3.45) for the second and fourth moments is the same and proceeds in four steps:

1. Average over the perturbation ensemble. As a result, the integral over the supervector X will no longer be of Gaussian type.
2. Perform a supersymmetric Hubbard-Stratonovich transformation (see Appendix C.4 and Refs. [220, 243, 244]), which introduces an auxiliary supermatrix to render the X integral Gaussian again.

3. Integrate over the supervector X.
4. Employ a saddle-point approximation (cf. Appendix D) to evaluate the remaining integral over the auxiliary Hubbard-Stratonovich matrix.

The first step is straightforward because the corresponding integrals are Gaussian after invoking a generalized central limit theorem. The second step serves, as indicated, to remove terms quartic in X introduced during the first step, and thereby enables the third step. The saddle-point approximation in the fourth step exploits the large Hilbert space dimension N and becomes asymptotically exact as $N \to \infty$. Finally, the ensemble-averaged resolvents can be combined according to Eq. (3.39) and similar relations to find the overlap moments of interest.

3.4.2 Second moment and overlap distribution

As far as the mathematical foundations of our subsequent physical predictions are concerned, this present subsection is presumably the most important one because we will encounter several key quantities and relations which will play prominent roles in the final results. The overall goal is to evaluate the second-order eigenvector overlap moment $\mathbb{E}[U_{n\mu}U_{n\nu}^*]$. Due to (3.39), the principal object of study is thus the ensemble-averaged resolvent $\mathbb{E}[\mathcal{G}_{\nu\mu}(z^\pm)]$ with $z^\pm = E_n \pm \mathrm{i}\eta$, $\eta > 0$. Note that we redefined $z^\pm$ compared to how it was introduced above Eq. (3.40) by replacing E_n^λ by its ensemble average $\mathbb{E}[E_n^\lambda] = E_n$ since we intend to calculate the ensemble average of $\mathcal{G}(z)$ for fixed z, in the spirit of the definition (3.36).

Asymptotic distribution and ensemble average. Starting from Eq. (3.45) and following the recipe outlined at the end of Sec. 3.4.1, the first step consists in evaluating the integral over the V ensemble for a given distribution $p(V)$, i.e., the ensemble average

$$\begin{aligned}\mathbb{E}\big[\exp\big\{-\mathrm{i}\lambda\, X^\dagger L^\pm V X\big\}\big] \\ = \mathbb{E}\Big[\exp\Big\{-\mathrm{i}\lambda\Big(\sum\nolimits_\alpha X_\alpha^\dagger L^\pm X_\alpha V_{\alpha\alpha} + \sum\nolimits_{\alpha<\beta}(X_\alpha^\dagger L^\pm X_\beta + X_\beta^\dagger L^\pm X_\alpha)(\mathrm{Re}\, V_{\alpha\beta}) \\ + \sum\nolimits_{\alpha<\beta}(X_\alpha^\dagger L^\pm X_\beta - X_\beta^\dagger L^\pm X_\alpha)(\mathrm{i}\,\mathrm{Im}\, V_{\alpha\beta})\Big)\Big\}\Big]\,.\end{aligned} \tag{3.47}$$

Introducing the symbol $\sigma_{\alpha\beta}^2 := \mathbb{E}[|V_{\alpha\beta}|^2] = \sigma_v^2(E_\alpha - E_\beta)$ for the variance of the matrix element $V_{\alpha\beta}$ (see Eqs. (3.22)–(3.24)) and the definitions $Y_\alpha := \sigma_{\alpha\alpha}X_\alpha^\dagger L^\pm X_\alpha/2$ and $Y_{\alpha\beta} := \sigma_{\alpha\beta}X_\alpha^\dagger L^\pm X_\beta$ (with the same fixed choice of L^+ or L^- for all α), we can abbreviate $X^\dagger L^\pm V X = Z + Z^*$ with

$$Z := \sum\nolimits_\alpha Y_\alpha \frac{V_{\alpha\alpha}}{\sigma_{\alpha\alpha}} + \sum_{\alpha<\beta} Y_{\alpha\beta}\frac{V_{\alpha\beta}}{\sigma_{\alpha\beta}}\,. \tag{3.48}$$

Taking into account the general form (3.24) of the considered perturbation ensembles, this quantity Z is a weighted sum of N^2 independent random variables of zero mean and unit variance, namely $V_{\alpha\alpha}/\sigma_{\alpha\alpha}$, $\mathrm{Re}\, V_{\alpha\beta}/\sigma_{\alpha\beta}$, and $\mathrm{Im}\, V_{\alpha\beta}/\sigma_{\alpha\beta}$ ($\alpha < \beta$). Recalling the central limit theorem, we expect that Z approaches a Gaussian distribution with mean zero and variance $Y^2 := \sum_\alpha Y_\alpha^2 + \sum_{\alpha<\beta}|Y_{\alpha\beta}|^2$ for large N, and similarly for $X^\dagger L^\pm V X = Z + Z^*$. For any given X, the asymptotic distribution of the exponent on the right-hand side of (3.47) for large N is therefore determined by the first two moments (3.22) of the matrix elements $V_{\alpha\beta}$. Since we will eventually let $N \to \infty$, we can thus approximate any given $p(V)$ by any other distribution whose first two moments (3.22) coincide with those of $p(V)$. A particularly convenient choice is the Gaussian distribution

$$p_{\alpha\alpha}(v) = \frac{\mathrm{e}^{-v^2/2\sigma_{\alpha\alpha}^2}}{\sqrt{2\pi}\sigma_{\alpha\alpha}} \quad \text{and} \quad p_{\alpha\beta}(v) = \frac{\mathrm{e}^{-|v|^2/\sigma_{\alpha\beta}^2}}{\pi(\sigma_{\alpha\beta})^2} \quad (\alpha < \beta)\,. \tag{3.49}$$

Assuming this distribution in (3.47), the right-hand side factors into a product of N^2 one-dimensional Gaussian integrals in the variables $V_{\alpha\alpha}$, $\mathrm{Re}\, V_{\alpha\beta}$, and $\mathrm{Im}\, V_{\alpha\beta}$ ($\alpha < \beta$). Performing these integrals and substituting into (3.45), the ensemble-averaged resolvent takes the form

$$\begin{aligned}\mathbb{E}[\mathcal{G}_{\nu\mu}(z^\pm)] = \mp\mathrm{i}\int \frac{[\mathrm{d}X\mathrm{d}X^*]}{(\mp 2\pi)^N}\, x_\nu x_\mu^*\, \exp\Big\{-\tfrac{\lambda^2}{2}\sum\nolimits_{\alpha,\beta}(\sigma_{\alpha\beta})^2\, \mathrm{str}\Big[X_\alpha X_\alpha^\dagger L^\pm X_\beta X_\beta^\dagger L^\pm\Big] \\ + \mathrm{i}\sum\nolimits_\alpha (z^\pm - E_\alpha) X_\alpha^\dagger L^\pm X_\alpha\Big\}\,,\end{aligned} \tag{3.50}$$

where $\operatorname{str} M := \operatorname{tr} M_{\mathrm{BB}} - \operatorname{tr} M_{\mathrm{FF}}$ denotes the supertrace of the supermatrix M (cf. Eq. (3.43) and Appendix B.2). For later reference, we note that the integrand possesses a (pseudo)unitary symmetry, i.e., it is invariant under transformations $X \mapsto TX$, $X^\dagger \mapsto X^\dagger T^\dagger$ satisfying $T^\dagger L^\pm T = L^\pm$.

Hubbard-Stratonovich transformation. Proceeding with the second step, we employ a supersymmetric generalization (see Appendix C.4 and Refs. [163, 220]) of the Hubbard-Stratonovich transformation [243, 244] to rewrite the exponential of the fourth-order term in X as a superintegral involving only quadratic terms in X, namely

$$\begin{aligned}&\exp\Big[-\tfrac{\lambda^2}{2}\sum\nolimits_{\alpha,\beta}(\sigma_{\alpha\beta})^2\,\operatorname{str}(X_\alpha X_\alpha^\dagger L^\pm X_\beta X_\beta^\dagger L^\pm)\Big]\\ &\quad=\int\frac{[\mathrm{d}R]}{(2\pi)^N}\exp\Big[-\tfrac{1}{2\lambda^2}\sum\nolimits_{\alpha,\beta}(\sigma^{-2})_{\alpha\beta}\operatorname{str}(R_\alpha R_\beta)+\mathrm{i}\sum\nolimits_\alpha\operatorname{str}(R_\alpha X_\alpha X_\alpha^\dagger L^\pm)\Big]\,.\end{aligned}\tag{3.51}$$

Here σ^{-2} denotes the inverse of the Hilbert-space matrix σ^2 with $(\sigma^2)_{\alpha\beta} = (\sigma_{\alpha\beta})^2$ (see below Eq. (3.47)). The auxiliary (2×2) supermatrices R_α are parametrized as

$$R_\alpha := \begin{pmatrix} r_{1\alpha} & \rho_\alpha \\ \rho_\alpha^* & \mathrm{i}r_{2\alpha}\end{pmatrix}\tag{3.52}$$

with real numbers $r_{1\alpha}$, $r_{2\alpha}$ and anticommuting ρ_α, ρ_α^*, and $[\mathrm{d}R] := \prod_\alpha \mathrm{d}R_\alpha$ with $\mathrm{d}R_\alpha := \mathrm{d}r_{1\alpha}\mathrm{d}r_{2\alpha}\mathrm{d}\rho_\alpha\mathrm{d}\rho_\alpha^*$ for short. Upon substitution of (3.51) into (3.50), the remaining integral over the supervector X is Gaussian and can be evaluated straightforwardly (see also Appendix C). Thus, after the third step of the algorithm from the end of Sec. 3.4.1, we are left with

$$\begin{aligned}\mathbb{E}[\mathcal{G}_{\nu\mu}(z^\pm)] = \delta_{\mu\nu}\int\frac{[\mathrm{d}R]}{(2\pi)^N}\,&\big[(R_\mu + z^\pm - E_\mu)^{-1}\big]_{\mathrm{BB}}\\ &\times\exp\Big\{-\operatorname{str}\Big[\tfrac{1}{2\lambda^2}\sum\nolimits_{\alpha,\beta}(\sigma^{-2})_{\alpha\beta}R_\alpha R_\beta+\sum\nolimits_\alpha\ln(R_\alpha+z^\pm-E_\alpha)\Big]\Big\}\,.\end{aligned}\tag{3.53}$$

Saddle-point approximation. For the final step in the calculation we exploit that the exponent of the integrand in (3.53) is extensive in the Hilbert space dimension N, which in turn is exponentially large in the system's degrees of freedom. As a result, the integral is dominated by the highest saddle points of the exponent in the complex, multidimensional R plane, where the integrand becomes sharply peaked along suitably chosen integration contours. This justifies to evaluate the integral by means of a saddle-point approximation (see Appendix D and Refs. [163, 221, 223, 245, 246]). To find the stationary points of the exponent, we look for supermatrices R_μ such that the first variation of the exponent in (3.53) with respect to R vanishes, i.e.,

$$R_\mu + \lambda^2\sum\nolimits_\alpha(\sigma_{\mu\alpha})^2\,(R_\alpha + z^\pm - E_\alpha)^{-1} = 0\,.\tag{3.54}$$

From the solutions of this saddle-point equation, we have to select the dominant one that can be reached by a deformation of the original integration contour without crossing any singularities. The saddle-point approximation of (3.53) is then obtained as the product of the integrand and the inverse square root of the superdeterminant corresponding to the second variation of the exponent in (3.53), where both are evaluated at the dominating saddle point (cf. Appendix D.3). If there are several such dominating saddles, we have to sum their contributions.

To solve (3.54), it suffices to search for diagonal solutions $\hat{R}_\mu$ because all further solutions can be generated from diagonal ones by exploiting the (pseudo)unitary symmetry of the integral observed below (3.50) [163]. In view of the Hubbard-Stratonovich transformation (3.51), which effectively identifies $R_\mu \sim X_\mu X_\mu^\dagger L^\pm$, the auxiliary matrix $R = (R_\mu)$ transforms as $R \mapsto TRT^{-1}$ under that symmetry. Once the dominant diagonal solution $\hat{R}$ has been identified, we thus need to substitute $R = T\hat{R}T^{-1}$ into the integrand in (3.53) and average over the (pseudo)unitary group of transformation matrices T satisfying $T^\dagger L^\pm T = L^\pm$.

Assuming the matrix R_μ to be diagonal, the matrix equation (3.54) decouples into two identical equations for its entries. Consequently, any diagonal solution will be of the form $\hat{R}_\mu = \hat{r}(E_\mu, z^\pm)\mathbb{1}$,

where we explicitly indicated the dependence of the solution on both the unperturbed and the (infinitesimally shifted) perturbed energies E_μ and $z^\pm$, respectively. Since $\hat{R}_\mu$ is proportional to the unit matrix, all equivalent solutions obtained via the transformation T are identical, so the group average is trivial. Moreover, since the superdeterminant of any matrix proportional to $\mathbb{1}$ is unity (see Appendix B.2), the contribution involving the second variation of the exponent in (3.53) amounts to a trivial factor of one. Eventually, we therefore find that the ensemble-averaged resolvent from (3.53) takes the form

$$\mathbb{E}[\mathcal{G}_{\nu\mu}(z^\pm)] = \frac{\delta_{\mu\nu}}{z^\pm - E_\mu + \hat{r}(E_\mu, z^\pm)} \tag{3.55}$$

after the saddle-point approximation, where $\hat{r}(E_\mu, z^\pm)$ solves

$$\hat{r}(E_\mu, z^\pm) + \lambda^2 \sum\nolimits_\alpha \frac{(\sigma_{\mu\alpha})^2}{z^\pm - E_\alpha + \hat{r}(E_\alpha, z^\pm)} = 0\,. \tag{3.56}$$

Ensemble-averaged resolvent and overlap distribution. The aim of this paragraph is to recast (3.56) into a computationally more accessible form. To this end, we exploit Prerequisite (i) of a homogeneous density of states and assume that the summands are slowly varying with α (which they will be as a consequence of Prerequisite (iii), see below), such that the sum can be approximated by an integral, i.e., $\sum_\alpha \cdots \approx \int \mathrm{d}E/\varepsilon \cdots$. We also revert to the perturbation profile $\sigma_v^2(E)$ from (3.10) by substituting $(\sigma_{\mu\alpha})^2 = \sigma_v^2(E_\mu - E_\alpha)$ (see below (3.47)) and utilize that the solution $\hat{r}(E_\mu, z^\pm)$ will only depend on the difference $z^\pm - E_\mu$ due to Prerequisites (i) and (ii) as well as the property (3.22) of the perturbation matrix elements. In view of Eq. (3.55), it is therefore natural to introduce the *ensemble-averaged (scalar) resolvent*

$$G(z^\pm - E_\mu) := \frac{1}{z^\pm - E_\mu + \hat{r}(E_\mu, z^\pm)}\,, \tag{3.57}$$

implying

$$\mathbb{E}[\mathcal{G}_{\nu\mu}(z^\pm)] = \delta_{\mu\nu}\, G(z^\pm - E_\mu) \tag{3.58}$$

or, in operator form, $\mathbb{E}[\mathcal{G}(z)] = G(z - H_0)$. Adopting the above-mentioned transformation steps, Eq. (3.56) for $\hat{r}(E_\mu, z^\pm)$ then turns into the nonlinear integral equation

$$G(z^\pm)\left[z^\pm - \lambda^2 \int \frac{\mathrm{d}E}{\varepsilon}\, G(z^\pm - E)\, \sigma_v^2(E)\right] = 1 \tag{3.59}$$

for the ensemble-averaged resolvent $G(z)$.

Recalling Eq. (3.39) and observing that $\mathcal{G}(z^*) = \mathcal{G}(z)^\dagger$ according to (3.36), we understand that the second moment of eigenvector overlaps is encoded in the imaginary part of $G(z)$. Namely, defining the *overlap distribution*

$$u(E) := \frac{\varepsilon}{2\pi\mathrm{i}} \lim_{\eta\to 0+} [G(E - \mathrm{i}\eta) - G(E + \mathrm{i}\eta)] = \frac{\varepsilon}{\pi} \lim_{\eta\to 0+} \operatorname{Im} G(E - \mathrm{i}\eta)\,, \tag{3.60}$$

and exploiting (3.39) and (3.58), we immediately find

$$\mathbb{E}[U_{n\mu} U_{n\nu}^*] = \delta_{\mu\nu}\, u(E_n - E_\mu)\,. \tag{3.61}$$

Since the distribution of the perturbation V is even in the sense that V and $-V$ exhibit identical statistical properties (see below Eq. (3.22)), we readily conclude that the overlap distribution is an even function, too,

$$u(E) = u(-E)\,. \tag{3.62}$$

Indeed, since $G(E - \mathrm{i}\eta) = \mathbb{E}[(E - \mathrm{i}\eta - \lambda V)^{-1}]$ (see below Eq. (3.58)), this follows directly from the definition (3.60).

The problem of computing the second moment of eigenvector overlaps has thus been reformulated as the problem to solve the integral equation (3.59) for the ensemble-averaged resolvent $G(z)$. An equivalent equation for a related (but real-valued) random matrix ensemble had already been

reported in Ref. [219], even though the authors never published its derivation [247]. Incidentally, already Wigner's seminal studies [213, 214] of banded random matrices comprised a special case of (3.59) for perturbation matrix elements distributed according to $f_E(v) = [1 - \sigma_v^2(E)]\delta(v) + \sigma_v^2(E)[\delta(v-v_0)+\delta(v+v_0)]/2$ with $v_0 > 0$ and a step profile $\sigma_v^2(E) = \Theta(\Delta_v^2 - E^2)$. In other words, the matrix elements have magnitude v_0 but random signs within a band of width Δ_v, whereas they vanish outside of this band.

For the remainder of this subsection, we will discuss solutions of (3.59) for three special cases of the perturbation profile that are analytically tractable, before sketching a method to calculate such solutions numerically for arbitrary profile functions. Finally, we will compare all those special solutions to numerically exact ones in a concluding paragraph.

Special case 1: constant profile or weak perturbations. Let us consider eigenstates $|n\rangle_\lambda$ and $|\mu\rangle_0$ of the perturbed and unperturbed Hamiltonians, respectively. From Eq. (3.58) we understand that $\mathbb{E}[\mathcal{G}_{\mu\mu}(E_n^\lambda \pm \mathrm{i}\eta)]$ depends only on the distance of the two levels, $E_n^\lambda - E_\mu$, and similarly for the average overlap $\mathbb{E}[|U_{n\mu}|^2] = \mathbb{E}[|{}_\lambda\langle n|\mu\rangle_0|^2]$ of the state vectors from (3.61). Together with (3.37) and continuity in λ, we therefore generally expect that the function $G(z)$ from (3.57) is dominated in magnitude by the region around $z = 0$, corresponding to states $|n\rangle_\lambda$ and $|\mu\rangle_0$ that are close-by in energy, and decays to zero as $|z| \to \infty$.

For our first approximate approach to solve the integral equation (3.59), we assume that this decay of $G(z)$ happens on a scale much smaller than the typical scale of variations of the perturbation profile $\sigma_v^2(E)$ from (3.10). Consequently, the integral in (3.59) is dominated by the dominating region of $G(z-E)$ at $E \approx |z|$, and $\sigma_v^2(E)$ can be approximated by its central value $\sigma_v^2(|z|)$ in this region. With the definition $C(z) := \int \mathrm{d}E\, G(z-E)$, we then find

$$G(z) = \frac{1}{z - \lambda^2\, \sigma_v^2(|z|)\, C(z)/\varepsilon}\,. \tag{3.63}$$

Exploiting once again that the decay scale of $G(z)$ is assumed to be much smaller than the scale of variations of $\sigma_v^2(E)$, we can approximate $\sigma_v^2(|z|)$ in this expression by its value σ_v^2 around zero (cf. Eq. (3.11)) for all relevant values of $|z|$ for which $G(z)$ significantly deviates from zero. Recalling that $z = E' \pm \mathrm{i}\eta$ with $E' \in \mathbb{R}$ and $\eta \to 0+$ in the cases of interest (cf. Eq. (3.39)), the quantity $C(z)$ defined above (3.63) becomes a constant $C^\pm := \lim_{\eta\to 0+} C(\pm\mathrm{i}\eta)$ depending only on the sign of the imaginary part of z. Substituting (3.63) with these simplifications into the definition of $C(z)$ and evaluating the integral in the principal-value sense, we obtain $C^\pm = \mp\mathrm{i}\pi$ as the only consistent solution. As a result, we find the approximate solution

$$G(z) := \frac{1}{z + \mathrm{i}\,\mathrm{sgn}(\mathrm{Im}\, z)\Gamma/2}\,, \tag{3.64}$$

where $\mathrm{sgn}(x)$ denotes the sign function and

$$\Gamma := \frac{2\pi\lambda^2\sigma_v^2}{\varepsilon} = 2\pi\alpha_v\lambda^2\,, \tag{3.65}$$

where the definition (3.12) of α_v was substituted in the last equality. The Hubbard-Stratonovich auxiliary supermatrix $\hat{R}_\mu = \hat{r}(E_\mu, z^\pm)\mathbb{1}$ (see above Eq. (3.55)) solving the saddle-point equation (3.54) thus takes the form $\hat{R}_\mu = [G(z^\pm - E_\mu)^{-1} - (z^\pm - E_\mu)]\mathbb{1} = \pm\mathrm{i}\Gamma\mathbb{1}/2$ according to Eqs. (3.57) and (3.64). In view of the parametrization (3.52), we notice that this solution does not lie on the original contour of integration. Nevertheless, we can (and should) adjust that contour appropriately by shifting $r_{1\alpha} \mapsto r_{1\alpha} \pm \mathrm{i}\Gamma/2$, which is allowed because the poles of the integrand in (3.53) at $r_{1\alpha} = E_\alpha - z^\pm$ lie on the opposite side of the real line (below it for '+' and above it for '−'). Similar adjustments are always possible as long as the sign of $\mathrm{Im}\, G(z)$ is opposite to the sign of $\mathrm{Im}\, z$ and will thus be tacitly understood for all further solutions discussed below. Note that this relation of signs is also entailed in the combination of Eqs. (3.36) and (3.58), which implies $G(z) = \mathbb{E}[(z - \lambda V)^{-1}]$ and thus $\mathrm{sgn}[\mathrm{Im}\, G(z)] = -\,\mathrm{sgn}(\mathrm{Im}\, z)$ since V is Hermitian.

Returning to the scalar solution (3.64), the condition of a sufficiently fast decaying function $G(z)$ (see above Eq. (3.63)) is thus satisfied self-consistently if the perturbation profile $\sigma_v^2(E)$ varies on a scale much larger than Γ. In particular, this is obviously the case for a constant profile $\sigma_v^2(E) = \sigma_v^2$.

More generally, if the perturbation profile is sufficiently regular such that the scale of variations is also the decay scale, i.e., the band width Δ_v from (3.13), then we expect (3.64) to hold for

$$\Gamma \ll \Delta_v \,. \tag{3.66}$$

Since Γ scales quadratically with the coupling strength λ according to (3.65), Eq. (3.64) is thus essentially a weak-perturbation approximation for $G(z)$. The overlap distribution (3.60) associated with (3.64) is the Breit-Wigner, Cauchy, or Lorentz distribution

$$u(E) = \frac{\varepsilon}{2\pi} \frac{\Gamma}{E^2 + \Gamma^2/4} \,. \tag{3.67}$$

In view of (3.61), we thus conclude that the mixing of eigenvectors extends across the energy scale Γ for weak perturbations, i.e., we can identify Γ here with the scale Γ_v introduced in Prerequisite (iii) from Sec. 3.2 to assess how strong the mixing should be in order for the typicality approach to be applicable.

Special case 2: narrow profile or strong perturbations. For our second approximate solution of Eq. (3.59), we adopt a similar idea, but exchange the roles of $G(z)$ and $\sigma_v^2(E)$. More precisely, we now assume that the mixing of perturbed and unperturbed eigenvectors is so strong that the scale of $G(z)$ is much larger than that of $\sigma_v^2(E)$ or, equivalently, that the perturbation profile is sharply peaked around $E = 0$. If this is the case, we can approximate $G(z - E)$ by $G(z)$ in the integrand in (3.59) and exploit (3.13), resulting in the algebraic equation

$$\gamma^2 \, G(z)^2/4 - z \, G(z) + 1 = 0 \tag{3.68}$$

with

$$\gamma := \sqrt{\frac{8\Delta_v}{\varepsilon}} \, \lambda \, \sigma_v = \sqrt{8\Delta_v \alpha_v} \, \lambda \,. \tag{3.69}$$

Next we recall that the sign of $\operatorname{Im} G(z)$ must be opposite to that of $\operatorname{Im} z$ for the contour to be adjustable such that it passes through the resulting saddle point (see below Eq. (3.65)). Of the two solutions of (3.68), we thus find

$$G(z) = \frac{2}{\gamma^2} \left[z - \mathrm{i} \operatorname{sgn}(\operatorname{Im} z) \sqrt{\gamma^2 - z^2} \right] \tag{3.70}$$

as the only mathematically consistent one. The typical scale of variations of $G(z)$ is thus γ from (3.69), hence the initial assumption of slowly varying $G(z)$ on the decay scale Δ_v of $\sigma_v^2(E)$ is verified self-consistently if

$$\gamma \gg \Delta_v \,. \tag{3.71}$$

With (3.69), we conclude that the adopted approximation is expected to hold for larger values of the coupling strength λ, i.e., for stronger perturbations. The overlap distribution (3.60) obtained from (3.70) takes the semicircular form

$$u(E) = \frac{2\varepsilon}{\pi\gamma^2} \sqrt{\gamma^2 - E^2} \, \Theta(\gamma^2 - E^2) \,. \tag{3.72}$$

In this case, the eigenvector mixing scale Γ_v from Prerequisite (iii) corresponds to the radius γ of the semicircle distribution.

Remarkably, the two solutions (3.64) and (3.70) along with their respective regimes of validity (3.66) and (3.71) (under reasonable regularity conditions on $\sigma_v^2(E)$) only depend on the two perturbation characteristics α_v from (3.12) and Δ_v from (3.13), but not on any further details of the profile $\sigma_v^2(E)$. In particular, we therefore expect the overlap distribution $u(E)$ to follow a smooth crossover from the Breit-Wigner distribution (3.67) to the semicircle distribution (3.72) as the coupling strength λ is increased for fixed α_v and Δ_v, regardless of any further details of $\sigma_v^2(E)$. The value λ_c at which this crossover occurs can be estimated by equating the rates Γ and γ from (3.65) and (3.69), yielding

$$\lambda_\mathrm{c} := \sqrt{\frac{2\varepsilon\Delta_v}{\pi^2\sigma_v^2}} = \sqrt{\frac{2\Delta_v}{\pi^2\alpha_v}} \,. \tag{3.73}$$

Special case 3: Breit-Wigner profile. The third case for which we discuss an analytic solution here deals with perturbation profiles of the Breit-Wigner form, i.e.,

$$\sigma_v^2(E) = \frac{\sigma_v^2}{1 + (\pi E/2\Delta_v)^2} \,. \tag{3.74}$$

The strategy to solve (3.59) for this choice of $\sigma_v^2(E)$ is to exploit results from complex analysis to express the occurring integral in terms of residues of the integrand. We restrict ourselves to the main steps of the derivation here; details can be found in Ref. [248].

The idea is to focus on the lower half $\mathbb{C}^- := \{z \in \mathbb{C} : \operatorname{Im} z < 0\}$ of the complex plane, which suffices to extract the physically relevant overlap distribution $u(E)$ according to (3.60). (Note that analogous calculations could also be carried out for the upper half-plane.) Under the proviso that, for all $z \in \mathbb{C}^-$, $G(z)$ is analytic and vanishes as $|z| \to \infty$, the integral (3.59) with $\sigma_v^2(E)$ from (3.74) becomes

$$\int \mathrm{d}E\, G(z^- - E)\, \sigma_v^2(E) = 2\sigma_v^2\, \Delta_v\, G(z^- - 2\mathrm{i}\Delta_v/\pi)\,, \tag{3.75}$$

picking up the residue of the sole simple pole of the integrand in the upper half-plane at $E = 2\mathrm{i}\Delta_v/\pi$. Exploiting this result in Eq. (3.59) leads to

$$G(z^-) = \frac{1}{z^- - 2\lambda^2\sigma_v^2\Delta_v\, G(z^- - 2\mathrm{i}\Delta_v/\pi)/\varepsilon} = \frac{1}{z^- - b_v\, G(z^- - \mathrm{i}a_v)} \tag{3.76}$$

with $a_v := 2\Delta_v/\pi$ and $b_v := 2\lambda^2\sigma_v^2\Delta_v/\varepsilon$. By iterated substitution of $G(z)$ in the denominator on the right-hand side, we find that $G(z^-)$ can be expressed as an infinite continued fraction,

$$G(z^-) = \cfrac{1}{z^- - \cfrac{b_v}{z^- - \mathrm{i}a - \cfrac{b_v}{z^- - 2\mathrm{i}a - b_v \cdots}}} \,. \tag{3.77}$$

Moreover, this infinite continued fraction can be shown to converge for all $z \in \mathbb{C}^-$ and $b_v > 0$ [248], hence Eq. (3.77) provides a (formally) exact solution of the integral equation (3.59) for the Breit-Wigner profile $\sigma_v^2(E)$ from (3.74). By truncating the continued fraction at a finite order M_{cf}, we find approximate solutions of improving quality as M_{cf} is increased, all of which verify the initial prerequisites of being analytic in $\mathbb{C}^-$ and vanishing for large $|z|$ [248]. For later reference, we record that, in particular, the second-order approximation reads

$$G(z^-) = \cfrac{1}{z^- - \cfrac{b_v}{z^- - \mathrm{i}a - \cfrac{b_v}{z^- - 2\mathrm{i}a}}} = \frac{(z^- - \mathrm{i}\gamma_0)(z^- - 2\mathrm{i}\gamma_0) - \Delta_v\Gamma/\pi}{(z^- - \mathrm{i}\gamma_+)(z^- - \mathrm{i}\gamma_0)(z^- - \mathrm{i}\gamma_-)} \,, \tag{3.78}$$

where we substituted Γ from (3.65) and additionally introduced

$$\gamma_n := \frac{2\Delta_v}{\pi}\left[1 + n\sqrt{1 - \frac{\pi\Gamma}{2\Delta_v}}\right] \tag{3.79}$$

with the abbreviations $\gamma_+ \equiv \gamma_{+1}$ and $\gamma_- \equiv \gamma_{-1}$. In light of (3.77), the lower-order truncations of the infinite continued fraction are expected to work better the smaller the numerical value of b_v is because every higher-order correction is effectively suppressed by an additional factor of b_v. For fixed α_v and Δ_v, the second-order expression (3.78) may thus be regarded as a weak-perturbation approximation again. Its applicability, however, extends far beyond the simpler result (3.64) for an essentially constant profile found above, as will become clear below, especially in Sec. 3.6.

Numerical solutions for general perturbation profiles. The hitherto derived solutions of the integral equation (3.59) applied to special choices of the perturbation profile and/or certain limits of the coupling strength λ. For a more general treatment, we need to resort to numerical methods. While other approaches are of course possible, we here propose to expand the function $G(z)$ in terms of Chebyshev rational functions $B_n(x)$, obtained from the Chebyshev polynomials of the first kind $T_n(x)$ according to

$$B_n(x) := T_n\!\left(\frac{x}{\sqrt{x^2 + \ell^2}}\right) \qquad (n = 0, 1, \ldots)\,. \tag{3.80}$$

The parameter $\ell > 0$, which mediates the compactification of the real line to the interval $[-1, 1]$, is arbitrary *a priori*, but ideally reflects the typical scale of the function to be expanded in order to optimize convergence [249].

As before, we are particularly interested in solutions of (3.59) in the vicinity of the real line, i.e., for arguments $z = x - \mathrm{i}\eta$ with $x \in \mathbb{R}$ and very small $\eta > 0$ (cf. Eq. (3.60)). As observed below (3.65), the sign of the imaginary part of $G(z)$ jumps when crossing the real line, implying that the solutions of (3.59) can become ambiguous for purely real z. Formally, we therefore introduce the symbol

$$G^+(x) := \lim_{\eta \to 0+} G(x - \mathrm{i}\eta) \tag{3.81}$$

to denote the continuation from the lower complex half-plane. The numerical problem thus consists of solving

$$G^+(x) \left[x - \lambda^2 \int \frac{\mathrm{d}E}{\varepsilon} \, G^+(x - E) \, \sigma_v^2(E) \right] = 1 \tag{3.82a}$$

for real-valued x with the additional constraint that

$$\operatorname{Im} G^+(x) \geq 0 \,, \tag{3.82b}$$

which is implied by the observation below (3.65). Following the above-announced route to express the solution as a series in the $B_n(x)$ from (3.80), we write

$$G^+(x) = G^{\mathrm{R}}(x) + \mathrm{i}\, G^{\mathrm{I}}(x) \tag{3.83a}$$

and expand the real-valued functions $G^{\mathrm{R}}(x)$ and $G^{\mathrm{I}}(x)$ as

$$G^{\mathrm{R}}(x) := \sum_{n=0}^{M_{\mathrm{Cheb}}} G_n^{\mathrm{R}} \, B_n(x) \quad \text{and} \quad G^{\mathrm{I}}(x) := \sum_{n=0}^{M_{\mathrm{Cheb}}} G_n^{\mathrm{I}} \, B_n(x) \tag{3.83b}$$

with $G_n^{\mathrm{R}}, G_n^{\mathrm{I}} \in \mathbb{R}$ and M_{Cheb} denoting the order of the truncation. The remaining task thus consists in determining the $2(M_{\mathrm{Cheb}}+1)$ coefficients $\boldsymbol{G} := (G_0^{\mathrm{R}}, G_0^{\mathrm{I}}, \ldots, G_{M_{\mathrm{Cheb}}}^{\mathrm{R}}, G_{M_{\mathrm{Cheb}}}^{\mathrm{I}})$ such that the violation of (3.82a) is minimal. To this end, we employ the pseudospectral method [249, 250] and require the residual

$$R(\boldsymbol{G}, x) := G^+(x) \left[x - \lambda^2 \int \frac{\mathrm{d}E}{\varepsilon} \, G^+(x - E) \, \sigma_v^2(E) \right] - 1 \tag{3.84}$$

to vanish on a discrete set of collocation points $x_m \in \mathbb{R}$ ($m = 0, 1, \ldots, M_{\mathrm{Cheb}}$). Specifically, the x_m are chosen as the roots of the $(M_{\mathrm{Cheb}} + 1)$th Chebyshev rational function $B_{M_{\mathrm{Cheb}}+1}(x)$, implying that the pseudospectral method agrees with a spectral expansion when the inner products are computed numerically using an optimal Gaussian quadrature rule [249, 250].

The conditions $\operatorname{Re} R(\boldsymbol{G}, x_m) = \operatorname{Im} R(\boldsymbol{G}, x_m) = 0$ for $m = 0, \ldots, M_{\mathrm{Cheb}}$ thus yield a set of $2(M_{\mathrm{Cheb}}+1)$ nonlinear algebraic equations for the $2(M_{\mathrm{Cheb}} + 1)$ expansion coefficients G_n^{R}, G_n^{I}. For the solution of this system of equations, we adopt a Newton-Raphson iteration [251] with the residual vector $R_m = R(\boldsymbol{G}, x_m)$ and the Jacobian $J_{mn}^{\mathrm{R,I}} = \partial R(\boldsymbol{G}, x_m)/\partial G_n^{\mathrm{R,I}}$ calculated from (3.84). For the first initial guess, we use either of the limiting distributions (3.64) or (3.70) for small or large λ, respectively, and gradually vary λ across the transition thereafter. If this initial guess is sufficiently close to the actual solutions and obeys the constraint (3.82b), then so will the resulting iteratively improved approximations of the Newton-Raphson scheme.

Comparison of special cases and numerics. With a method to compute the ensemble-averaged resolvent $G(z)$ defined via (3.59) numerically for arbitrary perturbation profiles $\sigma_v^2(E)$, we can now compare these numerically exact results to the different solutions or approximations previously obtained for special choices of the perturbation profile. We focus here on four different shapes for $\sigma_v^2(E)$: the step profile

$$\sigma_v^2(E) = \sigma_v^2 \, \Theta(\Delta_v^2 - E^2) \,, \tag{3.85}$$

the exponential profile (3.21), the Breit-Wigner profile (3.74), and a double Breit-Wigner profile

$$\sigma_v^2(E) = \frac{\sigma_v^2 \, b_1^2 (b_2^2 + d^2)}{(b_1^2 + E^2)[(b_2^2 + (E - d)^2]} \tag{3.86}$$

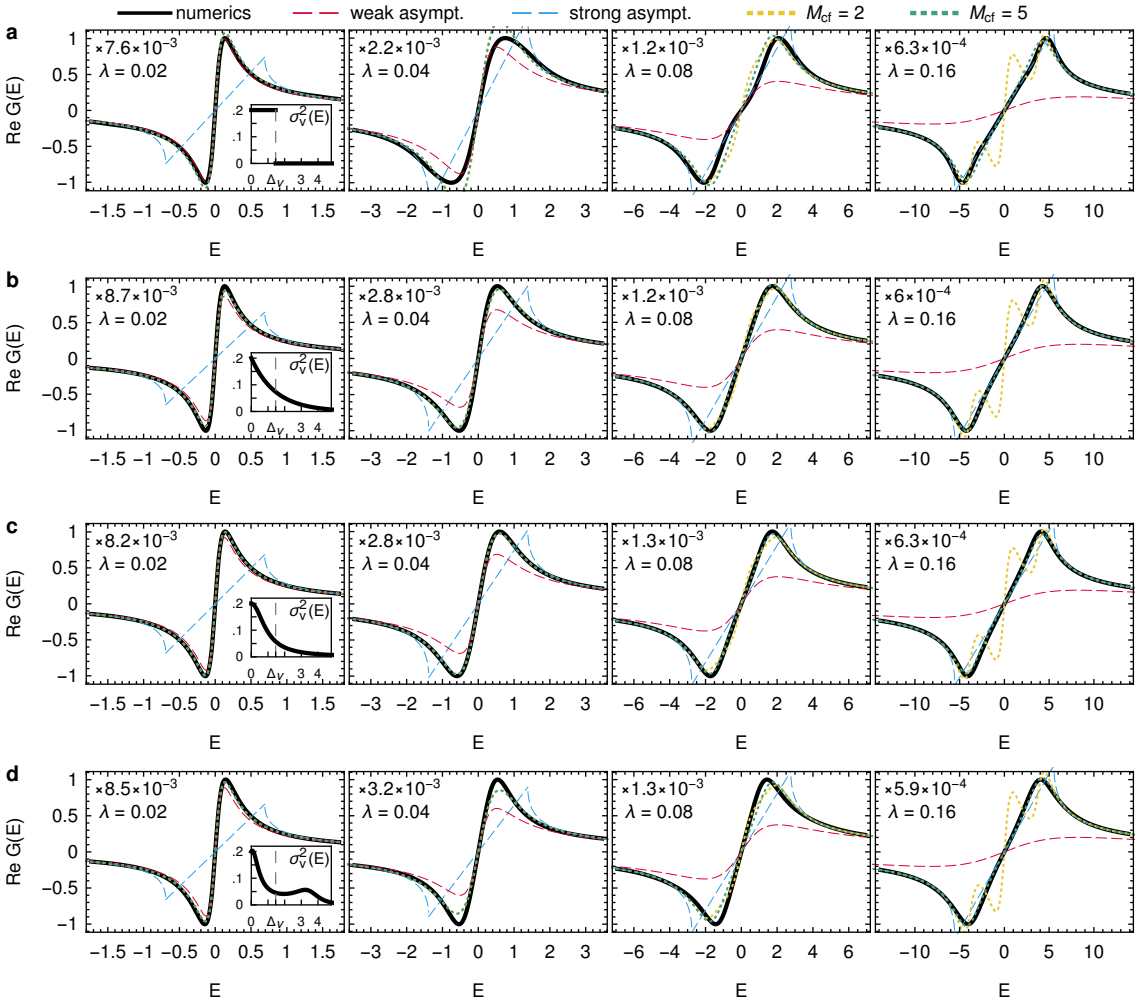


Figure 3.4: Real part $\mathrm{Re}\, G^+(E)$ of the scalar average resolvent (3.81) for various perturbation profiles $\sigma_v^2(E)$ (see insets of the left-most column) and coupling strengths λ (see top-left corner of each panel). For all configurations, $\varepsilon^{-1} = 512$, $\sigma_v^2 = 0.2$, and $\Delta_v = 750\varepsilon \approx 1.46$, yielding a crossover coupling of $\lambda_c \approx 0.05$ according to (3.73). Solid: Numerically obtained Chebyshev expansions (3.83) with $M_{\mathrm{Cheb}} = 80 \ldots 160$ and $\ell = 0.5 \ldots 8$ (increasing with λ) for **a.** the step profile (3.85); **b.** the exponential profile (3.21); **c.** the Breit-Wigner profile (3.74); **d.** the double Breit-Wigner profile (3.86) with $b_1 = 0.45$, $b_2 = 0.9$, $d = 3.5$. Values are scaled as indicated in the top-left corner of each panel. Dashed: Universal asymptotic solutions for weak (red, Eq. (3.64)) and strong (blue, Eq. (3.70)) coupling. Dotted: Truncated continued-fraction solutions for the Breit-Wigner profile (3.74) of orders $M_{\mathrm{cf}} = 2$ (yellow) and $M_{\mathrm{cf}} = 5$ (green). The corresponding imaginary parts $\mathrm{Im}\, G^+(E)$ are displayed in Fig. 3.5.

with $b_2 > b_1 > 0$ and $d > 0$. These different shapes are also sketched in the insets of Figs. 3.4 and 3.5.

We fix the mean level spacing at $\varepsilon = 1/512$. For each perturbation profile, we then choose parameters such that $\sigma_v^2 = 0.2$ and $\Delta_v = 750\varepsilon \approx 1.46$; in particular, $b_1 = 0.45$, $b_2 = 0.9$, and $d = 3.5$ for the double Breit-Wigner profile (3.86). The numerical solutions are obtained as detailed in the previous paragraph, generally using truncated Chebyshev series of order $M_{\mathrm{Cheb}} = 80$, but increasing the order up to $M_{\mathrm{Cheb}} = 160$ for the smallest λ values as the corresponding solutions $G(z)$ exhibit the fattest tails. We show the real and imaginary parts of $G^+(E)$ from (3.81) as solid black lines in Figs. 3.4 and 3.5, respectively.

In each panel, we then also plot the asymptotic result (3.64) for weak perturbations or a constant profile (dashed, red) as well as the asymptotic result (3.70) for stronger perturbations or very narrow profiles (dashed, blue). The transition from (3.64) to (3.70) is clearly visible as λ is increased from left to right in each row of both figures. Note that the estimated crossover coupling (3.73) is

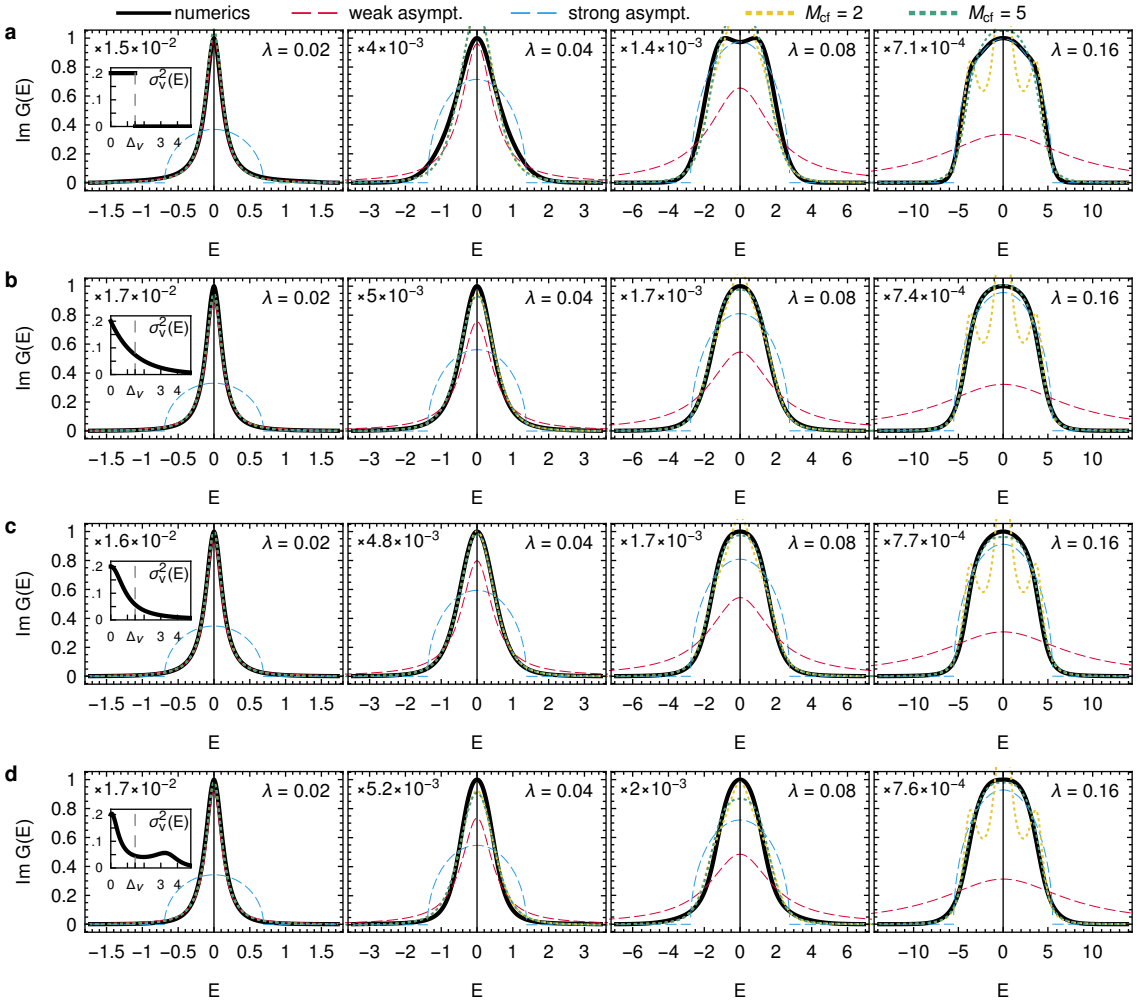


Figure 3.5: Imaginary part $\operatorname{Im} G^+(E)$ of the scalar average resolvent (3.81) for various perturbation profiles $\sigma_v^2(E)$ (see insets of the left-most column) and coupling strengths λ (see top-right corner of each panel). For all configurations, $\varepsilon^{-1} = 512$, $\sigma_v^2 = 0.2$, and $\Delta_v = 750\varepsilon \approx 1.46$, yielding a crossover coupling of $\lambda_c \approx 0.05$ according to (3.73). Solid: Numerically obtained Chebyshev expansions (3.83) with $M_{\text{Cheb}} = 80 \ldots 160$ and $\ell = 0.5 \ldots 8$ (increasing with λ) for **a.** the step profile (3.85); **b.** the exponential profile (3.21); **c.** the Breit-Wigner profile (3.74); **d.** the double Breit-Wigner profile (3.86) with $b_1 = 0.45$, $b_2 = 0.9$, $d = 3.5$. Values are scaled as indicated in the top-left corner of each panel. Dashed: Universal asymptotic solutions for weak (red, Eq. (3.64)) and strong (blue, Eq. (3.70)) coupling. Dotted: Truncated continued-fraction solutions (3.77) for the Breit-Wigner profile (3.74) of orders $M_{\text{cf}} = 2$ (yellow) and $M_{\text{cf}} = 5$ (green). The corresponding real parts $\operatorname{Re} G^+(E)$ are displayed in Fig. 3.4.

$\lambda_c \approx 0.05$ for the chosen parameters.

In addition to these universally expected limiting forms, we also display in each panel the second- (dotted, yellow) and fifth-order (dotted, green) continued-fraction approximations (3.77) for the Breit-Wigner perturbation profile (3.74). Naturally, these approximations should and do recover the numerical solutions in Figs. 3.4c and 3.5c, which were also obtained for the Breit-Wigner profile. To faithfully describe the function $G(z)$, we observe that higher-order approximations are needed for larger values of λ, whereas the lower-order approximations work equally well for small λ.

Moreover, and perhaps somewhat surprisingly, the solutions for the Breit-Wigner perturbation profile also provide quite decent approximations of the solutions for the other perturbation profiles displayed in rows a, b, and d of Figs. 3.4 and 3.5. We emphasize that the dotted and dashed curves are the same in each column of these figures because the values of ε, σ_v^2, and Δ_v are fixed, meaning that those approximations only depend on the coupling strength λ. Hence we notice a remarkable universality of the transition from weak to stronger perturbations. Minor differences are visible in

the region around $E = 0$, but the overall shape of the perturbation profile appears to be largely irrelevant once the intrinsic strength α_v and the band width Δ_v have been fixed. A first hint at the origins of this stability is the universality of the distributions for small and large λ because the smooth crossover with the coupling λ must interpolate between these two limits. Furthermore, the fact that $\sigma_v^2(E)$ only enters Eq. (3.59) under the integral sign may also explain that its details are washed out with regard to $G(z)$. In any case, the observed universality of $G(z)$ and hence also of the overlap distribution $u(E)$ from (3.60) will turn out convenient when it comes to devising predictions as general as possible for the time evolution of expectation values in Sec. 3.6.

3.4.3 Fourth moment

With solutions for the second moment of eigenvector overlaps available from the previous Sec. 3.4.2, we turn to the next nontrivial order consisting of four factors of overlap matrices $U_{n\mu}$, i.e., the average

$$\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}]\,. \tag{3.87}$$

Notably, this is the quantity required in (3.7) to specify the ensemble-averaged time evolution of perturbed expectation values after adopting Prerequisite (v).

The individual steps to evaluate (3.87) parallel those for the second-order moments. In particular, we follow again the recipe provided at the end of Sec. 3.4.1. By similar arguments as those invoked below Eq. (3.48), we can and will restrict to the normal distribution (3.49) for the perturbation matrix elements $V_{\mu\nu}$ because more general distributions are asymptotically equivalent for Hilbert space dimensions $N \gg 1$. Moreover, we will focus in this subsection on the physically most relevant case of sufficiently weak perturbations and consider a constant perturbation profile $\sigma_v^2(E) = \sigma_v^2 = 1$, whose overall strength is thus controlled by λ from (3.1). Extensions to more general profile functions are possible in principle, but require considerably more computational effort and will here instead be covered by the alternative approximation method sketched in Sec. 3.4.4 below.

Preliminary considerations. The fourth-order moment (3.87) comprises overlaps of two perturbed and four unperturbed eigenvectors ($|n_1\rangle_\lambda$, $|n_2\rangle_\lambda$ and $|\mu_1\rangle_0$, $|\mu_2\rangle_0$, $|\nu_1\rangle_0$, $|\nu_2\rangle_0$, respectively). The labels 1 and 2 here are obviously interchangeable, implying that also the final result for the fourth moment should be invariant when simultaneously exchanging $n_1 \leftrightarrow n_2$, $\mu_1 \leftrightarrow \mu_2$, and $\nu_1 \leftrightarrow \nu_2$. Moreover, when swapping $\mu_1 \leftrightarrow \nu_1$ and $\mu_2 \leftrightarrow \nu_2$ but keeping n_1 and n_2 fixed, the fourth-order moment (3.87) should be complex conjugated.

Another important property is the unitarity of the matrix $(U_{n\mu})$, which implies that (3.87) should reduce to the second moment when tracing out $|n_1\rangle_\lambda$ or $|n_2\rangle_\lambda$, e.g.,

$$\sum_{n_2} \mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}] = \delta_{\mu_2\nu_2}\,\mathbb{E}[U_{n_1\mu_1}U^*_{n_1\nu_1}]\,. \tag{3.88a}$$

Likewise, the second moment should be recovered when summing over $\mu_1 = \nu_2$ or $\mu_2 = \nu_1$, e.g.,

$$\sum_{\mu_1,\nu_2} \delta_{\mu_1\nu_2}\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}] = \delta_{n_1n_2}\,\mathbb{E}[U_{n_1\mu_2}U^*_{n_1\nu_1}]\,, \tag{3.88b}$$

and similarly when summing over $\mu_1 = \nu_1$ or $\mu_2 = \nu_2$, e.g.,

$$\sum_{\mu_1,\nu_1} \delta_{\mu_1\nu_1}\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}] = \mathbb{E}[U_{n_2\mu_2}U^*_{n_2\nu_2}]\,. \tag{3.88c}$$

These reduction properties reveal that the random variables $U_{n\mu}$ are not independent for distinct values of n and/or μ. From a mathematical point of view, naturally, this is clear because the orthogonality (or unitarity) constraint induces a functional dependence between all the $U_{n\mu}$.

The first step in our calculation of (3.87) is again to express the four factors of $U_{n\mu}$ in terms of matrix elements of the resolvent (3.36), and to set up the latter as a Gaussian superintegral. To do so, the cases of one perturbed eigenvector (i.e., $n_1 = n_2$ in (3.87)) and two distinct perturbed eigenvectors ($n_1 \neq n_2$) need to be treated differently.

Resolvent approach: single perturbed eigenvector. For a single perturbed eigenvector, i.e., $n_1 = n_2 = n$, the product $U_{n\mu_1}U_{n\mu_2}U^*_{n\nu_1}U^*_{n\nu_2}$ can be expressed as a linear combination of products of two resolvent matrix elements $\mathcal{G}_{\nu_1\mu_1}(z^\pm)$ and $\mathcal{G}_{\nu_2\mu_2}(z^\pm)$ by exploiting (3.37), where we denoted $z^\pm := E_n^\lambda \pm \mathrm{i}\eta$ as before. Averaging over the perturbation ensemble and taking the limit $\eta \to 0+$ similarly as in (3.39), we can thus write

$$\begin{aligned}\mathbb{E}[U_{n\mu_1}U_{n\mu_2}U^*_{n\nu_1}U^*_{n\nu_2}] = -\frac{\varepsilon^2}{4\pi^2}\lim_{\eta\to 0+}\mathbb{E}\big[&\mathcal{G}_{\nu_1\mu_1}(z^+)\mathcal{G}_{\nu_2\mu_2}(z^+) + \mathcal{G}_{\nu_1\mu_1}(z^-)\mathcal{G}_{\nu_2\mu_2}(z^-)\\ &-\mathcal{G}_{\nu_1\mu_1}(z^+)\mathcal{G}_{\nu_2\mu_2}(z^-) - \mathcal{G}_{\nu_1\mu_1}(z^-)\mathcal{G}_{\nu_2\mu_2}(z^+)\big]\,.\end{aligned} \tag{3.89}$$

Since we intend to express the matrix elements of $\mathcal{G}(z)$ in terms of Gaussian integrals, however, we need to symmetrize this expression in the Greek indices. For example, similarly as in (3.44), one finds that

$$\int \frac{[\mathrm{d}X\,\mathrm{d}X^*]}{(2\pi)^N} x_{\nu_1}x_{\nu_2}x^*_{\mu_1}x^*_{\mu_2}\,\mathrm{e}^{\mathrm{i}X^\dagger L^+(z^+-H_\lambda)X} = \mathcal{G}_{\nu_1\mu_1}(z^+)\mathcal{G}_{\nu_2\mu_2}(z^+) + \mathcal{G}_{\nu_1\mu_2}(z^+)\mathcal{G}_{\nu_2\mu_1}(z^+)\,, \tag{3.90}$$

a manifestation of the Isserlis-Wick theorem [239, 240], see also Appendix C. In other words, the Gaussian integral "automatically" produces a symmetrized product since the left-hand side is invariant under exchanging $\nu_1 \leftrightarrow \nu_2$ or $\mu_1 \leftrightarrow \mu_2$. The appropriate relation between the four factors of eigenvector overlaps and the resolvent matrix elements to work with is therefore

$$\begin{aligned}&-\frac{8\pi^2}{\varepsilon}\mathbb{E}[U_{n\mu_1}U_{n\mu_2}U^*_{n\nu_1}U^*_{n\nu_2}]\\ &= \lim_{\eta\to 0+}\mathbb{E}\big[\mathcal{G}_{\nu_1\mu_1}(z^+)\mathcal{G}_{\nu_2\mu_2}(z^+) + \mathcal{G}_{\nu_1\mu_2}(z^+)\mathcal{G}_{\nu_2\mu_1}(z^+) + \mathcal{G}_{\nu_1\mu_1}(z^-)\mathcal{G}_{\nu_2\mu_2}(z^-) + \mathcal{G}_{\nu_1\mu_2}(z^-)\mathcal{G}_{\nu_2\mu_1}(z^-)\\ &\quad -\mathcal{G}_{\nu_1\mu_1}(z^+)\mathcal{G}_{\nu_2\mu_2}(z^-) - \mathcal{G}_{\nu_1\mu_2}(z^+)\mathcal{G}_{\nu_2\mu_1}(z^-) - \mathcal{G}_{\nu_1\mu_1}(z^-)\mathcal{G}_{\nu_2\mu_2}(z^+) - \mathcal{G}_{\nu_1\mu_2}(z^-)\mathcal{G}_{\nu_2\mu_1}(z^+)\big]\,.\end{aligned} \tag{3.91}$$

The ensemble average of the terms in the second line, i.e., over expressions of the form (3.90), is evaluated completely analogously to the second-order case in Eq. (3.50) because the dependence on V is identical. Likewise, the subsequent Hubbard-Stratonovich transformation works the same way, except that a single auxiliary supermatrix R now suffices due to the constant variance $(\sigma_{\mu\nu})^2 = 1$ assumed in this subsection. Hence we find that

$$\begin{aligned}&\mathbb{E}\big[\mathcal{G}_{\nu_1\mu_1}(z^\pm)\mathcal{G}_{\nu_2\mu_2}(z^\pm) + \mathcal{G}_{\nu_1\mu_2}(z^\pm)\mathcal{G}_{\nu_2\mu_1}(z^\pm)\big]\\ &\quad= (\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1})\int\frac{\mathrm{d}R}{2\pi}\,(R+z^\pm-E_{\mu_1})^{-1}_{\mathrm{BB}}\,(R+z^\pm-E_{\mu_2})^{-1}_{\mathrm{BB}}\\ &\qquad\qquad\times\exp\left\{-\operatorname{str}\left[\frac{R^2}{2\lambda^2} + \sum\nolimits_\alpha \ln(R+z^\pm-E_\alpha)\right]\right\}.\end{aligned} \tag{3.92}$$

For the remaining integral over the supermatrix R, we employ a saddle-point approximation, which again works analogously to the second-order case. In particular, the solution of the saddle-point equation is of the previously found form (3.64) for the special case of a constant profile, so we are left with

$$\mathbb{E}\big[\mathcal{G}_{\nu_1\mu_1}(z^\pm)\mathcal{G}_{\nu_2\mu_2}(z^\pm) + \mathcal{G}_{\nu_1\mu_2}(z^\pm)\mathcal{G}_{\nu_2\mu_1}(z^\pm)\big] = \frac{\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}}{(z^\pm - E_{\mu_1} \pm \mathrm{i}\Gamma/2)\,(z^\pm - E_{\mu_2} \pm \mathrm{i}\Gamma/2)}\,. \tag{3.93}$$

For the terms in the third line of (3.91), the two factors of resolvent matrix elements in each product involve distinct arguments z^+ and z^-. Therefore, they have to be dealt with similarly to the case of two distinct perturbed eigenvectors, to which we will turn next.

Resolvent approach: two distinct perturbed eigenvectors. Barring degeneracies, the product of eigenvector overlaps in (3.87) for the case of two distinct ($n_1 \neq n_2$) perturbed eigenvectors $|n_1\rangle_\lambda$ and $|n_2\rangle_\lambda$ can be written as

$$\begin{aligned}\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}] = -\frac{\varepsilon^2}{4\pi^2}\lim_{\eta\to 0+}\mathbb{E}\big[&\mathcal{G}_{\nu_1\mu_1}(z_1^+)\mathcal{G}_{\nu_2\mu_2}(z_2^+) + \mathcal{G}_{\nu_1\mu_1}(z_1^-)\mathcal{G}_{\nu_2\mu_2}(z_2^-)\\ &-\mathcal{G}_{\nu_1\mu_1}(z_1^+)\mathcal{G}_{\nu_2\mu_2}(z_2^-) - \mathcal{G}_{\nu_1\mu_1}(z_1^-)\mathcal{G}_{\nu_2\mu_2}(z_2^+)\big]\end{aligned} \tag{3.94}$$

based on (3.37) as before and introducing the abbreviation $z_k^\pm := E_{n_k}^\lambda \pm \mathrm{i}\eta$. Regarding the right-hand side of this relation, there is a crucial difference between the terms in the first and second lines. In the first line, the resolvents in each product are evaluated at points shifted to the same side of the real line, above it for z_k^+ or below it for z_k^-. Hence we have a product of two retarded or two advanced resolvents, respectively. In the second line, by contrast, each product involves one retarded and one advanced resolvent.

In the first case, we essentially fall back on the second-order calculation again. Observing that $z_1^\pm \neq z_2^\pm$, we introduce two supervectors $X^{(1)}$ and $X^{(2)}$ of the form (3.42) and write

$$\begin{aligned}\mathcal{G}_{\nu_1\mu_1}(z_1^\pm)\,\mathcal{G}_{\nu_2\mu_2}(z_2^\pm) = \int \frac{[\mathrm{d}X^{(1)}\mathrm{d}X^{(1)*}]}{(\mp 2\pi)^N}\frac{[\mathrm{d}X^{(2)}\mathrm{d}X^{(2)*}]}{(\mp 2\pi)^N} x_{\nu_1}^{(1)} x_{\mu_1}^{(1)*} x_{\nu_2}^{(2)} x_{\mu_2}^{(2)*} \\ \times \exp\Big[\mathrm{i}X^{(1)\dagger}L^\pm(z_1^\pm - H_\lambda)X^{(1)} + \mathrm{i}X^{(2)\dagger}L^\pm(z_2^\pm - H_\lambda)X^{(2)}\Big]\,.\end{aligned} \tag{3.95}$$

After performing the average over the perturbation ensemble, the resulting expression factorizes into two copies of the integral (3.50). Consequently, all considerations from Sec. 3.4.2 carry over immediately, eventually yielding that the two ensemble-averaged resolvent matrix elements factorize, too, i.e.

$$\mathbb{E}[\mathcal{G}_{\nu_1\mu_1}(z_1^\pm)\,\mathcal{G}_{\nu_2\mu_2}(z_2^\pm)] = \mathbb{E}[\mathcal{G}_{\nu_1\mu_1}(z_1^\pm)]\,\mathbb{E}[\mathcal{G}_{\nu_2\mu_2}(z_2^\pm)] \tag{3.96}$$

with the single averages given in (3.58).

The situation is manifestly different—and considerably more involved—for products of one retarded and one advanced resolvent, which is the form of the remaining terms both in the second line of Eq. (3.94) and in the third line of Eq. (3.91). The reason for this is that there the solution of the corresponding saddle-point equation is no longer proportional to the unit matrix, resulting in an entire manifold of degenerate saddles that needs to be integrated over (see also the remarks below Eq. (3.54)).

Writing the product of one retarded and one advanced resolvent as a Gaussian integral, we obtain

$$\begin{aligned}\mathcal{G}_{\nu_1\mu_1}(z_1^+)\,\mathcal{G}_{\nu_2\mu_2}(z_2^-) = \int \frac{[\mathrm{d}X^{(1)}\mathrm{d}X^{(1)*}]}{(-2\pi)^N}\frac{[\mathrm{d}X^{(2)}\mathrm{d}X^{(2)*}]}{(2\pi)^N} x_{\nu_1}^{(1)} x_{\mu_1}^{(1)*} x_{\nu_2}^{(2)} x_{\mu_2}^{(2)*} \\ \times \exp\Big[\mathrm{i}X^{(1)\dagger}L^+(z_1^+ - H_\lambda)X^{(1)} + \mathrm{i}X^{(2)\dagger}L^-(z_2^- - H_\lambda)X^{(2)}\Big]\,.\end{aligned} \tag{3.97}$$

For notational convenience, we define a new collective supervector $X := (X_1 \cdots X_N)^\mathsf{T}$ with

$$X_\alpha := \begin{pmatrix} x_\alpha^{(1)} & \chi_\alpha^{(1)} & x_\alpha^{(2)} & \chi_\alpha^{(2)}\end{pmatrix}^\mathsf{T}\,. \tag{3.98}$$

Note that the components with superscript '(1)' correspond to the retarded resolvent and those with superscript '(2)' to the advanced resolvent. We will therefore refer to them as the retarded and advanced sectors, respectively, in the following. By means of the abbreviations

$$\bar{z} := \frac{z_1^+ + z_2^-}{2} \quad \text{and} \quad \Delta_z := z_1^+ - z_2^- \tag{3.99}$$

with $\operatorname{Im}\bar{z} = 0$ and $\operatorname{Im}\Delta_z = 2\eta$ as well as the diagonal matrices

$$L := \operatorname{diag}(1,1,-1,1) \quad \text{and} \quad \Lambda := \operatorname{diag}(1,1,-1,-1)\,, \tag{3.100}$$

Eq. (3.97) can then be written in the more compact form

$$\mathcal{G}_{\nu_1\mu_1}(z_1^+)\,\mathcal{G}_{\nu_2\mu_2}(z_2^-) = \int \frac{[\mathrm{d}X\mathrm{d}X^*]}{(2\pi\mathrm{i})^{2N}}\, x_{\nu_1}^{(1)} x_{\mu_1}^{(1)*} x_{\nu_2}^{(2)} x_{\mu_2}^{(2)*}\, \exp\big[\mathrm{i}X^\dagger L\left(\bar{z} + \tfrac{\Delta_z}{2}\Lambda - H_\lambda\right)X\big]\,. \tag{3.101}$$

From here on, we proceed with steps 1 through 4 of the algorithm laid out at the end of Sec. 3.4.1 to calculate the ensemble average of (3.101).

Ensemble average and symmetries. The first step of calculating the average over the perturbation ensemble in (3.101) is still structurally similar to the corresponding averaging procedure from (3.45) to (3.50) for the second moment. We obtain

$$\mathbb{E}[\mathcal{G}_{\nu_1\mu_1}(z_1^+)\,\mathcal{G}_{\nu_2\mu_2}(z_2^-)] = \int \frac{[\mathrm{d}X\mathrm{d}X^*]}{(2\pi\mathrm{i})^{2N}}\, x_{\nu_1}^{(1)}\, x_{\mu_1}^{(1)*}\, x_{\nu_2}^{(2)}\, x_{\mu_2}^{(2)*} \times \exp\left[-\tfrac{\lambda^2}{2}\sum_{\alpha,\beta}\mathrm{str}\left(X_\alpha X_\alpha^\dagger L X_\beta X_\beta^\dagger L\right) + \mathrm{i}\sum_\alpha X_\alpha^\dagger L\left(\bar{z} + \tfrac{\Delta}{2}\Delta_z - E_\alpha\right) X_\alpha\right]. \tag{3.102}$$

Similarly as for the second-order expression (3.50), it is worthwhile to examine the symmetries of the integrand in (3.102) as they will turn out to be crucial when performing the saddle-point approximation later. We observe that the character of the integrand changes depending on the relative location of the perturbed eigenvectors $|n_1\rangle_\lambda$ and $|n_2\rangle_\lambda$ as quantified by the parameter Δ_z from (3.99). On the one hand, we may intuitively expect that significant correlations due to the orthonormality constraint exist between those eigenvectors if they correspond to close-by levels, meaning that the difference Δ_z is small, $\Delta_z \sim \varepsilon \ll \Gamma$ with Γ from (3.65) (see also the discussion below Eq. (3.88c)). Since Γ is the typical scale of eigenvector correlations, we can neglect the term proportional to Δ_z in (3.102) to leading order in this case. The integrand then exhibits a pseudounitary symmetry mediated by transformations $X \mapsto TX$, $X^\dagger \mapsto X^\dagger T^\dagger$ with the matrix T satisfying $T^\dagger L T = L$. On the other hand, if the perturbed eigenvectors $|n_1\rangle_\lambda$ and $|n_2\rangle_\lambda$ are well separated such that $\Delta_z \gg \Gamma$, the Δ_z term is no longer negligible and the pseudounitary symmetry breaks down. As will become clear below, the average (3.102) then maps back onto the case of independent eigenvectors (see also Sec. 3.4.4), consistent with the intuition that the eigenvectors should not "feel" each other if they correspond to eigenvalues that lie far apart in the spectrum. In the intermediate regime, the situation is much more subtle; we will come back to this issue once we derived the explicit result for small Δ_z, i.e., below Eq. (3.116).

Hubbard-Stratonovich transformation. The Hubbard-Stratonovich transformation in the second step also takes a similar, yet simpler form compared to Eq. (3.51), because a single (4×4) supermatrix R is sufficient thanks to the constant variance $(\sigma_{\mu\nu})^2 = 1$. Hence we can write (cf. also Appendix C.4)

$$\exp\left[-\frac{\lambda^2}{2}\sum\nolimits_{\alpha,\beta}\mathrm{str}\left(X_\alpha X_\alpha^\dagger L X_\beta X_\beta^\dagger L\right)\right] = \int \frac{\mathrm{d}R}{(2\pi)^2}\,\exp\left[-\mathrm{str}\left(\frac{R^2}{2\lambda^2} + \mathrm{i}R\sum\nolimits_\alpha X_\alpha X_\alpha^\dagger L\right)\right], \tag{3.103}$$

where R is conveniently parameterized as [163, 222, 223]

$$R = T\begin{pmatrix} P_1 - \mathrm{i}\delta_0 & 0 \\ 0 & P_2 + \mathrm{i}\delta_0 \end{pmatrix} T^{-1} \tag{3.104}$$

with Hermitian (2×2) supermatrices P_1 and P_2 and $\delta_0 > 0$ to be adapted such that the integration contour passes through the saddle points [223]. Furthermore, the block-diagonalizing transformation matrix T satisfies $T^\dagger L T = L$ and thus belongs to the (approximate) pseudounitary symmetry group of the integrand in (3.102). Adopting (3.103), Eq. (3.102) transforms into

$$\mathbb{E}[\mathcal{G}_{\nu_1\mu_1}(z_1^+)\,\mathcal{G}_{\nu_2\mu_2}(z_2^-)] = \int \frac{\mathrm{d}R}{(2\pi)^2}\int \frac{[\mathrm{d}X\mathrm{d}X^*]}{(2\pi\mathrm{i})^{2N}}\, x_{\nu_1}^{(1)}\, x_{\mu_1}^{(1)*}\, x_{\nu_2}^{(2)}\, x_{\mu_2}^{(2)*} \times \exp\left[-\mathrm{str}\left(\frac{R^2}{2\lambda^2}\right) + \mathrm{i}\sum\nolimits_\alpha X_\alpha^\dagger L\left(R + \bar{z} + \frac{\Delta}{2}\Delta_z - E_\alpha\right) X_\alpha\right]. \tag{3.105}$$

The Gaussian integral over X, i.e., the third step of the recipe from the end of Sec. 3.4.1, is calculated by exploiting the Isserlis-Wick theorem (see Appendix C), yielding

$$\begin{aligned}
&\mathbb{E}[\mathcal{G}_{\nu_1\mu_1}(z_1^+)\,\mathcal{G}_{\nu_2\mu_2}(z_2^-)] \\
&= \int \frac{\mathrm{d}R}{(2\pi)^2} \exp\left\{-\operatorname{str}\left[\frac{R^2}{2\lambda^2} + \sum\nolimits_\alpha \ln\left(R + \bar{z} + \Delta_z\Lambda/2 - E_\alpha\right)\right]\right\} \\
&\quad \times \Big\{-\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\left[(R + \bar{z} + \Delta_z\Lambda/2 - E_{\nu_1})^{-1}\right]_{1\mathrm{B},1\mathrm{B}}\left[(R + \bar{z} + \Delta_z\Lambda/2 - E_{\nu_2})^{-1}\right]_{2\mathrm{B},2\mathrm{B}} \\
&\qquad -\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\left[(R + \bar{z} + \Delta_z\Lambda/2 - E_{\nu_1})^{-1}\right]_{1\mathrm{B},2\mathrm{B}}\left[(R + \bar{z} + \Delta_z\Lambda/2 - E_{\nu_2})^{-1}\right]_{2\mathrm{B},1\mathrm{B}}\Big\}\,.
\end{aligned} \tag{3.106}$$

Note that the supermatrix indices in the last two lines refer to the retarded (1) or advanced (2) components (corresponding to z_1^+ and z_2^-, respectively) of the bosonic sector (B).

Saddle-point approximation. In the fourth and final step, we evaluate the integral over the supermatrix R by means of a saddle-point approximation. The associated saddle-point equation is obtained by requiring the first variation of the exponent in (3.106) to vanish, resulting in

$$R + \lambda^2 \sum_\alpha (R + \bar{z} + \Delta_z\Lambda/2 - E_\alpha)^{-1} = 0\,. \tag{3.107}$$

By analogy with the solution strategy for the second moment (see below Eq. (3.54)), we first look for a diagonal solution $\hat{R}$. Expressing the sum over α as a principal-value integral, we straightforwardly find

$$\hat{R} = \mathrm{i}\Gamma\Lambda/2 \tag{3.108}$$

with $\Gamma = 2\pi\lambda^2/\varepsilon$ as defined in (3.65), observing that we set $\sigma_v^2 = 1$. Recalling the definition of Λ in (3.100), we notice that this diagonal solution is *not* proportional to the identity matrix, a crucial difference to the second-order case. In the regime where $\Delta_z \sim \varepsilon$ and the pseudounitary symmetry $T^\dagger L T = L$ is present, we therefore obtain additional nontrivial solutions $T\hat{R}T^{-1}$, all of which need to be accounted for in the saddle-point approximation, meaning that we need to integrate over the symmetry group of transformation matrices T.

The reduction of the R integral in Eq. (3.106) works analogously to the calculation for the second moment (see below Eq. (3.53)), hence we essentially substitute $R = T\hat{R}T^{-1}$ in the integrand in (3.106). To calculate the remaining integral over the manifold of degenerate saddles, it is convenient to introduce a new integration variable $Q := T\Lambda T^{-1}$ satisfying $Q^2 = \mathbb{1}$ and thus $\operatorname{str}(Q^2) = 0$. Altogether, we then obtain

$$\begin{aligned}
&\mathbb{E}[\mathcal{G}_{\nu_1\mu_1}(z_1^+)\,\mathcal{G}_{\nu_2\mu_2}(z_2^-)] \\
&= -\int \mathrm{d}\mu(Q)\ \exp\left[-\operatorname{str}\sum\nolimits_\alpha \ln\left(\mathrm{i}\Gamma Q/2 + \bar{z} + \Delta_z\Lambda/2 - E_\alpha\right)\right] \\
&\quad \times \Big[\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\left(\mathrm{i}\Gamma Q/2 + \bar{z} + \Delta_z\Lambda/2 - E_{\nu_1}\right)^{-1}_{1\mathrm{B},1\mathrm{B}}\left(\mathrm{i}\Gamma Q/2 + \bar{z} + \Delta_z\Lambda/2 - E_{\nu_2}\right)^{-1}_{2\mathrm{B},2\mathrm{B}} \\
&\qquad +\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\left(\mathrm{i}\Gamma Q/2 + \bar{z} + \Delta_z\Lambda/2 - E_{\nu_1}\right)^{-1}_{1\mathrm{B},2\mathrm{B}}\left(\mathrm{i}\Gamma Q/2 + \bar{z} + \Delta_z\Lambda/2 - E_{\nu_2}\right)^{-1}_{2\mathrm{B},1\mathrm{B}}\Big]\,,
\end{aligned} \tag{3.109}$$

where the integration measure $\mathrm{d}\mu(Q)$ will be given once a suitable parameterization for Q has been fixed (see Eqs. (E.4)–(E.9)). As the remaining integration in (3.109) is a rather tedious endeavor, we relegate the details to Appendix E.2.

To state the eventually obtained result, we define, by analogy with (3.81), the functions

$$G_{\mathrm{w}}^\pm(E) := \frac{1}{E \mp \mathrm{i}\Gamma/2}\,, \tag{3.110}$$

corresponding to the weak-perturbation asymptotics (3.64) of the scalar ensemble-averaged resolvent $G(z)$ with positive (+) or negative (−) imaginary part, respectively. Note that $|G_{\mathrm{w}}^\pm(E)| \equiv |G_{\mathrm{w}}(E)|$ is the same for both choices of the sign. In addition, we introduce the abbreviations $G_{n\nu}^\pm := G_{\mathrm{w}}^\pm(E_n - E_\nu)$ and $\operatorname{sinc}(x) := (\sin x)/x$.

For $N_v \gg 1$, where N_v from (3.9) quantifies the number of unperturbed levels mixed by the perturbation and is given by $N_v = \Gamma/\varepsilon$ here (see below Eq. (3.67)), the missing ensemble averages in Eqs. (3.91) and (3.94) are then found to read

$$\begin{aligned}
&\lim_{\eta\to 0+} \mathbb{E}[\mathcal{G}_{\nu_1\mu_1}(z_1^+)\,\mathcal{G}_{\nu_2\mu_2}(z_2^-) + \mathcal{G}_{\nu_1\mu_1}(z_1^-)\,\mathcal{G}_{\nu_2\mu_2}(z_2^+)] \\
&= \delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\left\{ G^+_{n_1\nu_1}\,G^-_{n_2\nu_2} + G^-_{n_1\nu_1}\,G^+_{n_2\nu_2} + \Gamma^2\,|G_{n_1\nu_1}G_{n_2\nu_2}|^2\,\mathrm{sinc}\Big(\tfrac{\pi(E_{n_1}-E_{n_2})}{\varepsilon}\Big)\right. \\
&\qquad \left. -\tfrac{\Gamma^4\varepsilon^2}{4\pi^2}|G_{n_1\nu_1}G_{n_2\nu_2}|^4\,|G_{n_1\nu_2}G_{n_2\nu_1}|^2\left[E_{\nu_1}^2+E_{\nu_2}^2+2E_{n_1}E_{n_2}-(E_{\nu_1}+E_{\nu_2})(E_{n_1}+E_{n_2})+\tfrac{\Gamma^2}{2}\right]^2\right\} \\
&\quad + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\,\tfrac{\Gamma^3\varepsilon}{2\pi}\,|G_{n_1\nu_2}G_{n_2\nu_1}|^2\left[|G_{n_1\nu_1}|^2+|G_{n_2\nu_2}|^2-\frac{(E_{n_1}-E_{n_2})\left(|G_{n_1\nu_1}|^2-|G_{n_2\nu_2}|^2\right)}{E_{\nu_1}+E_{\nu_2}-E_{n_1}-E_{n_2}}\right].
\end{aligned} \tag{3.111}$$

Collecting terms. With this, we have all terms needed to express the fourth moment of eigenvector overlaps (3.87) explicitly at our disposal: For $n_1 = n_2$, the ensemble average is given by (3.91), where the terms in the second line were found in (3.93) and those in the third follow from (3.111) by setting $n_1 = n_2$. This leads to

$$\mathbb{E}[U_{n\mu_1}U_{n\mu_2}U^*_{n\nu_1}U^*_{n\nu_2}] = (\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1})\,u(E_n - E_{\mu_1})\,u(E_n - E_{\mu_2}) + \mathcal{O}(N_v^{-3}) \tag{3.112}$$

with $u(E)$ as given in (3.67), i.e., the overlap distribution (3.60) in the weak-perturbation limit, and where we omitted terms of order N_v^{-3} or higher in the large parameter $N_v = \Gamma/\varepsilon$ (cf. Eq. (3.9)).

For $n_1 \neq n_2$, the ensemble average is expressed as (3.94), where the terms in the first line were given in (3.96) and the just-derived Eq. (3.111) provides the missing terms in the second line. We obtain

$$\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}]\big|_{n_1\neq n_2} = \delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\,d^{n_1n_2}_{\mu_1\mu_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\,\tilde{f}^{n_1n_2}_{\mu_1\mu_2} + \mathcal{O}(N_v^{-4}) \tag{3.113a}$$

to next-to-leading order in N_v, where

$$d^{n_1n_2}_{\mu_1\mu_2} := u(E_{n_1} - E_{\mu_1})\,u(E_{n_2} - E_{\mu_2}) \tag{3.113b}$$

$$\begin{aligned}
\tilde{f}^{n_1n_2}_{\mu_1\mu_2} &:= -\left(\tfrac{\Gamma\varepsilon}{4\pi}\right)\,u(E_{n_1} - E_{\mu_1})\,u(E_{n_2} - E_{\mu_2})\,, \\
&\quad \times \frac{\Gamma^2+E_{\mu_1}^2/2+E_{\mu_2}^2/2+E_{n_1}E_{n_2}-(E_{\mu_1}+E_{\mu_2})(E_{n_1}+E_{n_2})/2}{[(E_{n_1}-E_{\mu_2})^2+(\Gamma/2)^2][(E_{n_2}-E_{\mu_1})^2+(\Gamma/2)^2]}\,.
\end{aligned} \tag{3.113c}$$

Combining (3.112) and (3.113a), we are left with

$$\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}] \simeq \delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\,d^{n_1n_2}_{\mu_1\mu_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\left(\delta_{n_1n_2}d^{n_1n_2}_{\mu_1\mu_2} + \tilde{f}^{n_1n_2}_{\mu_1\mu_2}\right). \tag{3.114}$$

We point out that we truncated the expression for a single perturbed eigenvector (3.112) at one order lower than that for two distinct eigenvectors. The reason for this is the additional prefactor $\delta_{n_1n_2}$ they receive in the combined expression (3.114), which effectively reduces their order by a factor of N_v^{-1} when summing over n_1 and n_2 as we will eventually do to evaluate, for instance, Eq. (3.7).

Symmetry restoration. To conclude, we return to the remarks about general symmetry and reduction properties of the fourth moment from the beginning of this subsection. The preliminary result (3.114) is obviously symmetric in the labels 1 and 2, and, being real-valued, it also validates the condition of complex conjugation when swapping $\mu_1 \leftrightarrow \nu_1$ and $\mu_2 \leftrightarrow \nu_2$. Moreover, it verifies the reduction property (3.88a) for perturbed eigenvectors as can be tested straightforwardly by approximating the sum $\sum_{n_1}\cdots$ by an integral $\int \mathrm{d}E/\varepsilon\cdots$ as usual. The reduction property (3.88c) for unperturbed eigenvectors is not satisfied exactly, but the violations are of subleading order as will become clear below (see the discussion below Eq. (3.124)). However, the reduction property (3.88b) is violated relevantly because

$$\begin{aligned}
&\sum_{\mu_1\nu_2}\delta_{\mu_1\nu_2}\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}] \\
&= \delta_{\mu_2\nu_1}\left[u(E_{n_1} - E_{\mu_2}) + \tilde{u}(E_{n_1} - E_{n_2})\,u(E_{n_1} - E_{\mu_2})\left(\frac{(E_{n_1}-E_{n_2})(E_{n_1}-3E_{n_2}+2E_{\mu_2})}{2(E_{n_2}-E_{\mu_2})^2+\Gamma^2/2} - 1\right)\right]
\end{aligned} \tag{3.115}$$

with the convolution

$$\tilde{u}(E) := \int \frac{\mathrm{d}E'}{\varepsilon}\, u(E-E')\, u(E')\,. \tag{3.116}$$

While the first term on the right-hand side of (3.115) is precisely the expected result according to (3.88b), the nonnegligible second term spoils the symmetry. We emphasize that this violation is not an artifact of truncating the ensemble-averaged resolvents at leading order in N_v (cf. Eqs. (3.112) and (3.113a)). Instead, the prime suspect is the approximate character of the saddle-point degeneracy observed below Eq. (3.102). Taking the symmetry for granted, we integrated in (3.109) (see also Appendix E.2) over all saddle points $T\hat{R}T^{-1}$ with $\hat{R}$ from (3.108) and T verifying $T^\dagger L T = L$. This presumes a perfect pseudounitary symmetry, reflecting the situation for $\Delta_z \simeq \varepsilon$. Then again, we observe that also the case of $|\Delta_z| \gg \Gamma$, where the pseudounitary symmetry breaks down, is retrieved correctly in (3.114) because $\tilde{f}^{n_1 n_2}_{\mu_1\mu_2}$ becomes negligible compared to $d^{n_1 n_2}_{\mu_1\mu_2}$ then. Unfortunately, the intermediate regime with $\varepsilon \ll |\Delta_z| \ll \Gamma$, where the symmetry is neither perfect nor completely broken, is hardly accessible by analytical means. Yet we can restore the reduction property (3.88b) *a posteriori*, leading to a fully consistent expression for the fourth moment (3.87).

This is achieved by devising a correction term $c^{n_1 n_2}_{\mu_1\mu_2}$ for $\tilde{f}^{n_1 n_2}_{\mu_1\mu_2}$ that is of the same order in N_v, fixes (3.88b), and simultaneously preserves (3.88a) as well as the aforementioned symmetry properties upon exchanging indices. Due to these symmetries, the only admissible dependencies of the correction term are the five invariants

$$\Gamma,\ (E_{\mu_1}+E_{\mu_2}),\ (E_{n_1}+E_{n_2}),\ (E_{\mu_1}-E_{\mu_2})^2,\ \text{and}\ (E_{n_1}-E_{n_2})^2\,. \tag{3.117}$$

Anticipating a structural similarity to (3.113c), we therefore make an ansatz of the form

$$c^{n_1 n_2}_{\mu_1\mu_2} = \left(\tfrac{\Gamma\varepsilon}{4\pi}\right) u(E_{n_1}-E_{\mu_1})\, u(E_{n_2}-E_{\mu_2}) \frac{A(c_0,c_1,c_2,c_3,c_4,c_5)}{[(E_{n_1}-E_{\mu_2})^2+(\Gamma/2)^2][(E_{n_2}-E_{\mu_1})^2+(\Gamma/2)^2]} \tag{3.118a}$$

with

$$\begin{aligned} A(c_0,c_1,c_2,c_3,c_4,c_5) = {} & c_0\Gamma^2 + c_1(E_{\mu_1}+E_{\mu_2})(E_{n_1}+E_{n_2}) + c_2(E_{\mu_1}-E_{\mu_2})^2 + c_3(E_{n_1}-E_{n_2})^2 \\ & + c_4(E_{\mu_1}+E_{\mu_2})^2 + c_5(E_{n_1}+E_{n_2})^2\,. \end{aligned} \tag{3.118b}$$

Requiring (3.88a) and restricting to constant solutions for the c_i yields $c_5 = -c_4 = c_3 = c_1/2$ and $c_2 = c_0 = 0$, hence there is only one free parameter remaining, for example, the coefficient c_1. Substituting into (3.88b), we find that $c_1 = -1$. The correction term then becomes

$$c^{n_1 n_2}_{\mu_1\mu_2} = \left(\tfrac{\Gamma\varepsilon}{2\pi}\right) u(E_{n_1}-E_{\mu_1})\, u(E_{n_2}-E_{\mu_2}) \frac{(E_{\mu_1}+E_{\mu_2}-2E_{n_1})(E_{\mu_1}+E_{\mu_2}-2E_{n_2})}{[(E_{n_1}-E_{\mu_2})^2+(\Gamma/2)^2][(E_{n_2}-E_{\mu_1})^2+(\Gamma/2)^2]}. \tag{3.119}$$

Altogether, our leading-order approximation for the fourth moment of eigenvector overlaps in the weak-perturbation limit thus reads

$$\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}] = \delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\, d^{n_1 n_2}_{\mu_1\mu_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\left(\delta_{n_1 n_2} d^{n_1 n_2}_{\mu_1\mu_2} + f^{n_1 n_2}_{\mu_1\mu_2}\right) \tag{3.120}$$

with $d^{n_1 n_2}_{\mu_1\mu_2}$ from (3.113b) and $f^{n_1 n_2}_{\mu_1\mu_2} := \tilde{f}^{n_1 n_2}_{\mu_1\mu_2} + c^{n_1 n_2}_{\mu_1\mu_2}$, i.e.,

$$\begin{aligned} f^{n_1 n_2}_{\mu_1\mu_2} &= -\left(\tfrac{\Gamma\varepsilon}{4\pi}\right) u(E_{n_1}-E_{\mu_1})\, u(E_{n_2}-E_{\mu_2}) \frac{\Gamma^2+(E_{\mu_1}-E_{\mu_2})^2+(E_{n_1}-E_{n_2})^2-(E_{\mu_1}+E_{\mu_2}-E_{n_1}-E_{n_2})^2}{[(E_{n_1}-E_{\mu_2})^2+(\Gamma/2)^2][(E_{n_2}-E_{\mu_1})^2+(\Gamma/2)^2]} \\ &= \left(\tfrac{\Gamma\varepsilon}{2\pi}\right) \frac{u(E_{n_1}-E_{\mu_2})\, u(E_{n_2}-E_{\mu_1}) - u(E_{n_1}-E_{\mu_1})\, u(E_{n_2}-E_{\mu_2})}{(E_{n_1}-E_{n_2})(E_{\mu_1}-E_{\mu_2})}\,. \end{aligned} \tag{3.121}$$

For consistency, we check the influence of the correction term depending on the relative locations of the perturbed and unperturbed eigenvectors to each other. Introducing $\Delta := E_{\mu_1} - E_{\mu_2}$ for the difference of the unperturbed energies, $\Delta' := E_{n_1} - E_{n_2}$ for the difference of the perturbed ones, and $\hat{\Delta} := \frac{E_{n_1}+E_{n_2}}{2} - \frac{E_{\mu_1}+E_{\mu_2}}{2}$ for the difference between the mean perturbed and the mean unperturbed energies, respectively, the ratio of the correction (3.119) to the corrected term (3.121) reads

$$\frac{c^{n_1 n_2}_{\mu_1\mu_2}}{f^{n_1 n_2}_{\mu_1\mu_2}} = \frac{2\,(4\hat{\Delta}^2-\Delta'^2)/\Gamma^2}{1-(4\hat{\Delta}^2-\Delta^2-\Delta'^2)/\Gamma^2}\,. \tag{3.122}$$

From this expression, we understand that the correction is small when $|\Delta'| \ll \Gamma$ and $|\hat{\Delta}| \ll \Gamma$ such that all eigenvectors involved correspond to levels that are close in energy. This is precisely the regime where the pseudounitary symmetry of the integrand in (3.102) is intact. Furthermore, the correction is of minor relevance if all of $|\Delta|$, $|\Delta'|$ and $|\hat{\Delta}|$ are large compared to Γ, i.e., if the associated levels lie far apart. Again, this is consistent with our observations below Eqs. (3.102) and (3.116) that the overlap factors $U_{n\mu}$ then essentially become independent of each other. In short, we indeed observe that the correction term (3.119) principally targets the intermediate regime of an approximate pseudounitary symmetry with $\varepsilon \ll |\Delta'| \ll \Gamma$ or $\varepsilon \ll |\hat{\Delta}| \ll \Gamma$.

Finally, we come back to the second reduction property (3.88c) for unperturbed eigenvectors. Substituting the final result (3.120), we find that

$$\begin{aligned}
&\sum_{\mu_1\nu_2} \delta_{\mu_1\nu_1} \mathbb{E}[U_{n_1\mu_1} U_{n_2\mu_2} U^*_{n_1\nu_1} U^*_{n_2\nu_2}] \\
&= \delta_{\mu_2\nu_2} \Bigg[u(E_{n_2} - E_{\mu_2}) + u(E_{n_1} - E_{\mu_2})\, u(E_{n_2} - E_{\mu_2}) \\
&\qquad\qquad \times \left(\delta_{n_1 n_2} - \frac{\Gamma^2 - 4(E_{n_1} - E_{\mu_2})(E_{n_2} - E_{\mu_2})}{\Gamma\varepsilon/\pi}\, u(E_{n_1} - E_{\mu_2})\, u(E_{n_2} - E_{\mu_2}) \right) \Bigg] .
\end{aligned} \tag{3.123}$$

The first term corresponds to the expected result, whereas the second term is adverse. Assessing orders of N_v, however, we observe that

$$\sum_{\mu_1\nu_2} \delta_{\mu_1\nu_1} \mathbb{E}[U_{n_1\mu_1} U_{n_2\mu_2} U^*_{n_1\nu_1} U^*_{n_2\nu_2}] = \delta_{\mu_2\nu_2} u(E_{n_2} - E_{\mu_2}) \left[1 + \mathcal{O}(N_v^{-2})\right] , \tag{3.124}$$

i.e., the infraction is doubly suppressed in the number of mixed levels N_v from (3.9) and thus insignificant. Note that the violation of (3.88b) observed in (3.115) before symmetry restoration was of order N_v^{-1} instead and therefore potentially relevant in sums of a large number of terms, e.g., Eq. (3.7). Altogether, our final leading-order approximation (3.120) for the fourth moment of eigenvector overlaps is thus consistent with the symmetry considerations from the beginning of this subsection.

Numerical verification. To test our analytical result (3.120) for the fourth-order eigenvector overlap moment and to verify that the leading-order approximation should be sufficient for all practical purposes, we compute these fourth moments numerically in an explicit example system with a Hilbert space dimension of $N = 512$. The Hamiltonian has the structure (3.1), where the unperturbed part is chosen like in (3.27) with level spacing $\varepsilon = 1$, such that $E_\mu = \mu$. The diagonal matrix elements $V_{\mu\mu}$ of the perturbation are drawn from a standard (real-valued) normal distribution, whereas the off-diagonal $V_{\mu\nu}$ with $\mu < \nu$ are sampled from a complex normal distribution and those with $\mu > \nu$ follow as $V_{\mu\nu} = V^*_{\nu\mu}$ due to Hermiticity. In both cases, the mean is zero and the variance is unity, hence the perturbation ensemble coincides with the Gaussian Unitary Ensemble (GUE). Finally, the coupling strength is $\lambda = 1.33$, implying $\Gamma \approx 11$ according to (3.65) and thus $N_v = \Gamma/\varepsilon \approx 11$, too.

The resulting empirical moments (3.87) for various index combinations are given by the black dots in Fig. 3.6. The comparison with the theoretical result (3.120), shown as solid red lines, reveals very good agreement despite the rather small dimension N and number of mixed levels N_v. Recalling that we are eventually interested in values of N and N_v that are exponentially large in the system's degrees of freedom, the leading-order approximation (3.120) will most certainly be sufficiently accurate.

For later reference, we also include in Fig. 3.6 the approximation (3.128), which will result from the alternative approach to be laid out in the subsequent Sec. 3.4.4. While this method yields identical results for the $\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}$ branch, the $\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}$ branch is reproduced only to a somewhat lesser extent with apparent quantitative deviations.

We remark that estimates and approximations of the fourth moment (3.87) in similar settings have been studied in the literature, too. In particular, Ithier and Ascroft [254] used an approach via Lippmann-Schwinger-type equations combined with a self-averaging conjecture for products of the

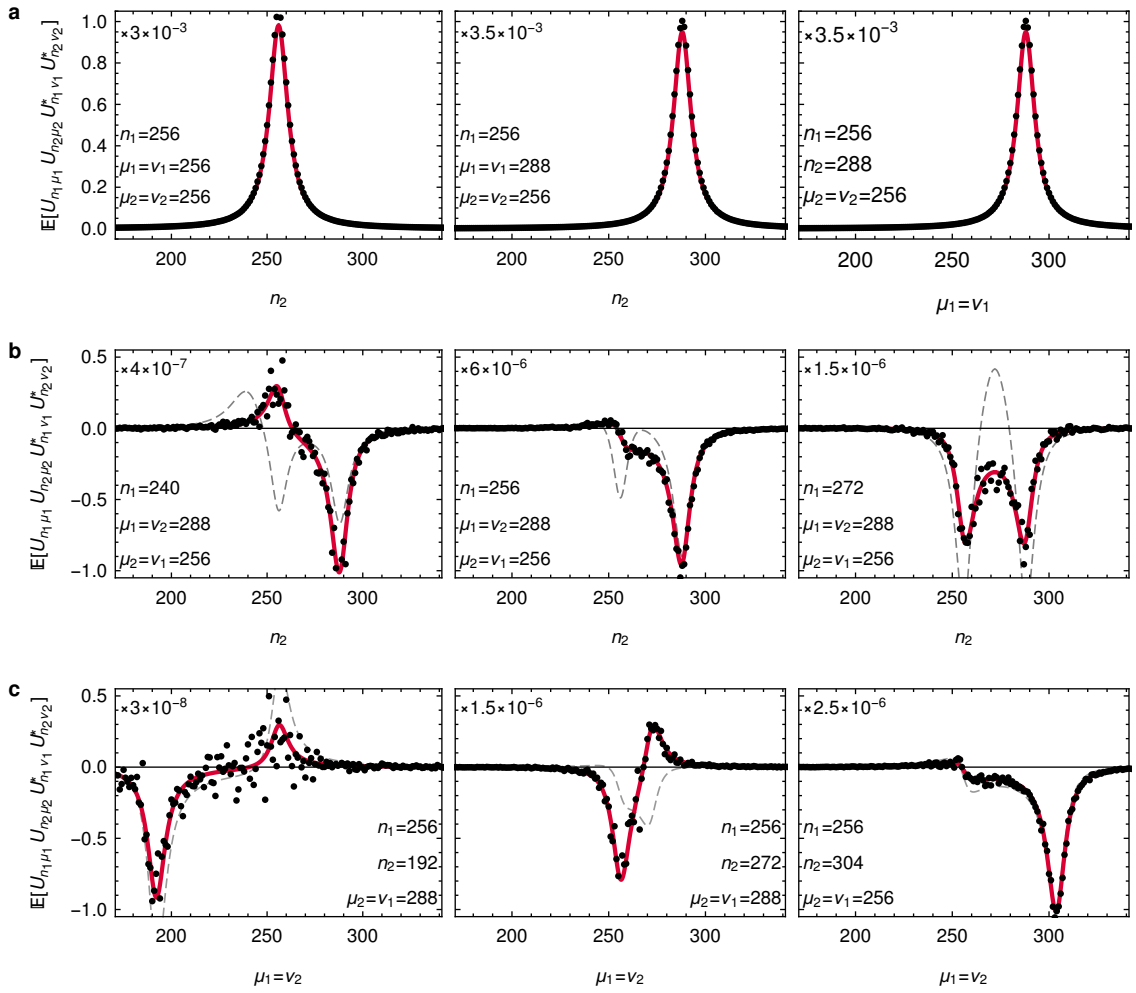


Figure 3.6: Fourth moment of eigenvector overlaps $\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}]$ for various combinations of fixed and variable indices as indicated in the respective panels. Dots: Numerical averages over 10^5 randomly sampled perturbation matrices $V_{\mu\nu}$ from the Gaussian Unitary Ensemble with mean zero and variance $\sigma_v^2 = 1$. The unperturbed Hamiltonian is (3.27) with a level spacing of $\varepsilon = 1$, and the coupling strength is $\lambda = 1.33$, so $\Gamma = N_v \approx 11$ (cf. Eqs. (3.9) and (3.65)). The Hilbert space dimension is $N = 512$. Solid: Analytical approximation (3.120) from supersymmetry methods. Dashed: Analytical approximation (3.128) obtained by the alternative approach from Sec. 3.4.4. **a.** Probe of the $\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}$ branch in (3.120) as a function of n_2 or $\mu_1 = \nu_1$. **b.** Probe of the $\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}$ branch as a function of n_2. **c.** Probe of the $\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}$ branch as a function of $\mu_1 = \nu_2$. The y axes are scaled as specified in the top-left corner of each panel.

resolvent $\mathcal{G}(z)$ with the covariance tensor of the $V_{\mu\nu}$ to evaluate (3.87) for different index combinations which essentially correspond to the individual $\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}$ or $\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}$ branches in (3.120). Their results coincide with ours in these cases.

A related problem was also studied by Nation and Porras in Ref. [252]. Their setup is identical to the one in Deutsch's 1991 paper [114], i.e., the perturbation matrices are sampled from the Gaussian Orthogonal Ensemble (GOE), meaning that all entries are real-valued. Methodologically, they extended the approach from [253], approximating the $U_{n\mu}$ by Gaussian random variables with an additional orthogonality constraint, resulting in an approximation that violates the reduction property (3.88b) (see Eq. (48) in Ref. [252]).

Perhaps surprisingly, it appears that our result (3.120), explicitly derived for complex-valued perturbation ensembles, can be generalized straightforwardly to the real-valued case. Since the Isserlis-Wick theorem takes a slightly different form (the matrices $U_{\mu\nu}$ are now real-valued and orthogonal), one should add an additional $\delta_{\mu_1\mu_2}\delta_{\nu_1\nu_2}$ branch with weight $f^{n_1n_2}_{\mu_1\nu_1}$, observing that the $\delta_{\mu_1\mu_2}\delta_{\nu_1\nu_2}$ and $\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}$ branches must share the same dependence on the "uncontracted" indices

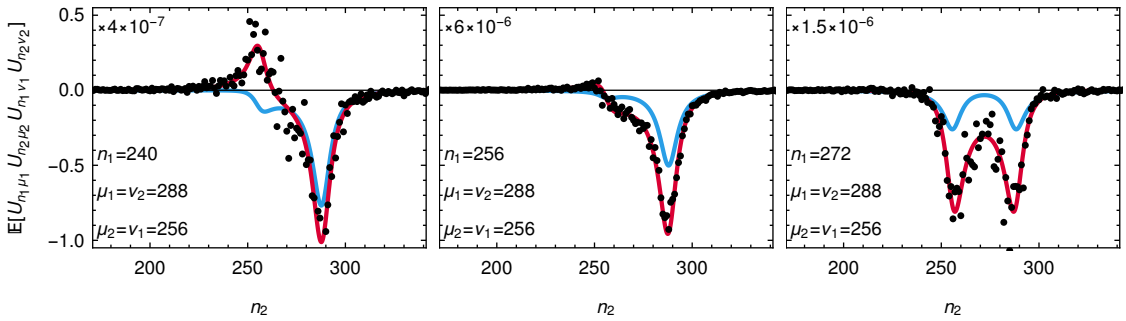


Figure 3.7: Fourth moment of eigenvector overlaps $\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U_{n_1\nu_1}U_{n_2\nu_2}]$, adopting real-valued perturbations, for various combinations of fixed and variable indices as indicated in the respective panels. Dots: Numerical averages over 10^5 randomly sampled perturbation matrices $V_{\mu\nu}$ from the Gaussian Orthogonal Ensemble with mean zero and variance $\sigma_v^2 = 1$. The unperturbed Hamiltonian is (3.27) with a level spacing of $\varepsilon = 1$, and the coupling strength is $\lambda = 1.33$, so $\Gamma = N_v \approx 11$ (cf. Eqs. (3.9) and (3.65)). Red lines: Analytical approximation (3.120) from supersymmetry methods (originally obtained for complex-valued perturbation ensembles). Blue lines: Analytical approximation from Ref. [252] obtained by an extension of Deutsch's method from [114, 253]. The y axes are scaled as specified in the top-left corner of each panel.

by symmetry. Apart from this modification, the overall structure is seemingly similar at leading order, as is illustrated by means of example in Fig. 3.7. Namely, we choose a similar setup for the system Hamiltonian as in Fig. 3.6, but sample the perturbation from the GOE instead. As before, the dots show the empirical moments (note that the $\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}$ branch is probed), while the solid red lines correspond to (3.120) and again agree very well with the numerics. For comparison, we also display the approximation from Ref. [252] as solid blue lines, unveiling that these expressions reproduce the dependence qualitatively, but cannot faithfully recover all quantitative details.

3.4.4 Alternative approach and higher-order moments

The supersymmetry methods employed in the previous subsections are a powerful approach to compute asymptotically exact expressions for eigenvector overlap moments like (3.29). Yet the calculations quickly become technically involved and eventually unsustainable as foreshadowed by the need for an amending symmetry restoration in the fourth-order calculation in Sec. 3.4.3. On the other hand, the final result (3.120) with (3.113b) and (3.121) for the fourth moment suggests that higher-order moments may ultimately be well approximated in terms of combinations of the second moments from (3.61). In this subsection, we therefore briefly sketch a generalized approximation scheme which allows to reduce correlators like (3.29) of arbitrary order to combinations of the second moment (3.61). Details can be found in the Supplemental Material of Ref. [226].

Independent Gaussian eigenvectors. It is known that the distribution of the $U_{n\mu}$ is approximately Gaussian [114, 253], and for Gaussian random variables, higher-order moments collapse to combinations of second moments as a consequence of the Isserlis-Wick theorem. Approximating the $U_{n\mu}$ as Gaussian and additionally independent random variables was precisely the approach from Refs. [114, 253]. Under these premises, one readily obtains, for example, a decomposition of the fourth moment (3.87) into second-moment expressions,

$$\begin{aligned}\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}] \approx{}& \delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\,u(E_{n_1}-E_{\mu_1})\,u(E_{n_2}-E_{\mu_2}) \\ &+ \delta_{n_1n_2}\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\,u(E_{n_1}-E_{\mu_2})\,u(E_{n_2}-E_{\mu_1})\,.\end{aligned} \tag{3.125}$$

We observe that the $\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}$ branch coincides with the leading-order approximation (3.120) found in the previous subsection by supersymmetry methods, but the crucial $f^{n_1n_2}_{\mu_1\mu_2}$ term in the $\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}$ branch is missing entirely. Consequently, also the reduction properties (3.88) are violated. While this approximation was sufficient for the purposes of Ref. [114], namely to establish thermalization for the majority of perturbed systems similar to (3.1), it is unfortunately too simplistic for our aim of predicting the dynamical relaxation behavior [226]. More precisely, as detailed in [226], assuming an approximately Gaussian distribution for the $U_{n\mu}$ turns out to be still tenable (see also Refs. [208, 252]), whereas the independence assumption has to be abandoned because the

concomitant violations of the unitarity constraint (particularly the properties (3.88a) and (3.88b)) are too severe.

Orthogonal eigenvector approximation. Exploiting those insights, the main idea of the alternative approach is to maintain a set of "raw" perturbed eigenvectors (or overlap products $U_{n\mu}$) that are independent and Gaussian distributed, but to enforce the unitarity constraint in expressions like (3.29) explicitly by an additional *ad hoc* orthogonalization.

Consider a product (3.29) of eigenvector overlaps $U_{n\mu} = {}_\lambda\langle n|\mu\rangle_0$ comprising K distinct perturbed eigenvectors $|n_1\rangle_\lambda, \ldots, |n_K\rangle_\lambda$. For each of the $|n_k\rangle_\lambda$, we introduce a random vector $v^{n_k} = (v_1^{n_k}, \ldots, v_N^{n_k})$, independent of v^{n_l} for $k \neq l$, whose components $v_\mu^{n_k}$ are independent random variables following a Gaussian distribution of mean zero and variance $u(E_{n_k} - E_\mu)$. Denoting averages over this ensemble of random vectors by the symbol $\mathrm{E}[\cdots]$, we thus have $\mathrm{E}\big[v_\mu^{n_k}\big] = 0$ and $\mathrm{E}\big[v_\mu^{n_k} v_\nu^{n_k *}\big] = \delta_{\mu\nu}\, u(E_{n_k} - E_\mu)$. Hence the $v_\mu^{n_k}$ exhibit the same first two moments as the $U_{n_k\mu}$ for all n_k and μ, but contrary to these $U_{n_k\mu}$, the auxiliary $v_\mu^{n_k}$ are all mutually independent by construction. In other words, the $v_\mu^{n_k}$ are an explicit implementation of the assumption adopted in the previous paragraph, where we temporarily considered the $U_{n\mu}$ as independent, Gaussian-distributed random variables. The $v_\mu^{n_k}$ thus represent the aforementioned set of raw perturbed eigenvectors.

In general, an individual realization of these raw vectors will violate the unitarity conditions $\sum_\mu v_\mu^{n_k *} v_\mu^{n_l} = \delta_{kl}$ and (if $K = N$) $\sum_k v_\mu^{n_k *} v_\nu^{n_k} = \delta_{\mu\nu}$ satisfied by the $U_{n\mu}$, even though both relations still hold on average. To overcome that deficiency, we define a second set of random vectors $\{\hat{v}^{n_k}\}$ by orthogonalizing the independent v^{n_k} via the Gram-Schmidt process, i.e.,

$$\hat{v}_\mu^{n_k} := a_k \left(v_\mu^{n_k} - \sum_{j=1}^{k-1} \hat{v}_\mu^{n_j} \sum_\nu \hat{v}_\nu^{n_j *} v_\nu^{n_k} \right), \tag{3.126}$$

with a_k chosen such that $\sum_\mu |\hat{v}_\mu^{n_k}|^2 = 1$. The $\hat{v}_\mu^{n_k}$ now form an orthonormal set by construction, i.e., $\sum_\mu \hat{v}_\mu^{n_k *} \hat{v}_\mu^{n_l} = \delta_{kl}$. The suggested scheme to compute eigenvector overlap moments like (3.29) then consists in approximating

$$\mathbb{E}[\, U_{n_1\mu_1} U^*_{n_1\nu_1} U_{n_2\mu_2} U^*_{n_2\nu_2} \cdots] \approx \mathrm{E}[\, \hat{v}_{\mu_1}^{n_1} \hat{v}_{\nu_1}^{n_1 *} \hat{v}_{\mu_2}^{n_2} \hat{v}_{\nu_2}^{n_2 *} \cdots], \tag{3.127}$$

i.e., we take the $\hat{v}_\mu^{n_k}$ as proxies for the true $U_{n_k\mu}$. Finally, we observe that the normalization factors $a_k \approx 1$ because the raw v^{n_k} are already orthonormalized on average. It turns out [226] that the deviations of the a_k from unity are subleading in the final approximations for the overlap moments (3.29), meaning that they are accompanied by inverse powers of the number N_v of mixed levels that are beyond the leading order N_v^{-3} we are interested in (see also the discussion around Eqs. (3.112) and (3.113)). Hence we will directly set $a_k = 1$ for all k in the following.

Computation of overlap moments. Upon substitution of (3.126) into (3.127), the eigenvector overlap moments from (3.29) can be approximated in terms of moments of the $v_\mu^{n_k}$. Since the latter are independent and unbiased Gaussian random variables by definition, the Isserlis-Wick theorem [239, 240] provides a straightforward algorithm to reduce these expressions to linear combinations of products of their second moments, which in turn are expressed in terms of the overlap distribution $u(E)$ of the "true" overlap factors $U_{n\mu}$ (cf. Eq. (3.60)). This way we accomplish the aforementioned reduction of higher-order moments to the exactly known second-order result from Sec. 3.4.2. For a moment involving K factors of $U_{n\mu}$, we thus need, according to (3.126) and (3.127), the moments of orders $K, K+2, \ldots, 2K$ of the K independent raw vectors v^{n_k} with components $v_\mu^{n_k}$.

As an explicit example, we present the derivation of the fourth-order approximation in Appendix E.3. This calculation yields

$$\mathrm{E}[\, \hat{v}_{\mu_1}^{n_1} \hat{v}_{\mu_2}^{n_2} \hat{v}_{\nu_1}^{n_1 *} \hat{v}_{\nu_2}^{n_2 *}] = \delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\, d_{\mu_1\mu_2}^{n_1 n_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1} \left(\delta_{n_1 n_2} d_{\mu_1\mu_2}^{n_1 n_2} + \hat{f}_{\mu_1\mu_2}^{n_1 n_2} \right) \tag{3.128}$$

to leading order in N_v for every branch, respectively, with $d_{\mu_1\mu_2}^{n_1 n_2}$ from (3.113b) and

$$\hat{f}_{\mu_1\mu_2}^{n_1 n_2} := u(E_{n_1} - E_{\mu_1})\, u(E_{n_1} - E_{\mu_2}) \left[\tilde{u}(E_{n_1} - E_{n_2}) - u(E_{n_2} - E_{\mu_1}) - u(E_{n_2} - E_{\mu_2}) \right], \tag{3.129}$$

where $\tilde{u}(E)$ is given by Eq. (3.116) as before. An obvious shortcoming of this approximation is that it is not symmetric upon exchanging all labels 1 and 2 (see the preliminary considerations at the beginning of Sec. 3.4.3), traced back to the unequal treatment of the perturbed eigenvectors $|n_1\rangle_\lambda$ and $|n_2\rangle_\lambda$ or, equivalently, the asymmetric construction (3.126) of $\hat{v}^{n_1}$ and $\hat{v}^{n_2}$ from v^{n_1} and v^{n_2}. As a remedy, we could either symmetrize (3.129) *a posteriori* or employ a symmetric orthogonalization procedure (e.g., Löwdin orthogonalization [255, 256]) right from the beginning. However, we avoid such technical complications here since this missing symmetry is eventually unimportant when it comes to computing ensemble averages of time-dependent expectation values as in Eq. (3.7), basically because the average in question is already symmetrized by means of the sum over n_1 and n_2 there. More importantly, the reduction properties from Eqs. (3.88a) and (3.88b) are verified straightforwardly.

The approximation (3.127) with (3.128) for the fourth moment in the weak-perturbation limit, such that $u(E)$ is given by (3.67), had already been included as dashed gray lines in Fig. 3.6 above. The comparison of this approximate result with numerical simulations or the supersymmetry solution (3.120) does not reveal overly convincing agreement at first glance. Nonetheless, the qualitative reproduction of the essential features of the correlator $\mathbb{E}[U_{n_1\mu_1}U_{n_2\mu_2}U^*_{n_1\nu_1}U^*_{n_2\nu_2}]$ along with the faithful incorporation of the reduction properties (3.88a) and (3.88b) will turn out to be sufficient for the purpose of computing the ensemble-averaged time evolution (3.7) in Sec. 3.6, and the same holds true for the higher-order moments when bounding the variance (3.8). In spite of the substantial technical simplifications compared to the supersymmetry approach from the previous subsections, the actual calculations of higher-order moments quickly become quite cumbersome, too, but can at least be aided by computer algebra software.

3.5 Prethermalization

Having established their mathematical background, we finally begin to state the main physical results of this chapter, which characterize the relaxation of the observable expectation values $\langle A\rangle_{\rho_\lambda(t)}$. Despite the fourfold dependence of these expectation values on the $U_{n\mu}$ (see Eq. (3.7)), it is possible to gain interesting insights regarding the *difference*

$$\Delta_\lambda(t) := \langle A\rangle_{\rho_\lambda(t)} - \langle A\rangle_{\rho_0(t)} \tag{3.130}$$

of the perturbed and unperturbed expectation values by exploiting just the second moment (3.30) of eigenvector overlaps (overlap distribution), i.e., without resorting to the considerably more involved fourth and eighth moments. Notably, we will obtain an analytical argument for why prethermalization (cf. Sec. 2.2.3) can be commonly expected when an integrable system is subject to weak integrability-breaking perturbations [177]. More generally, a similar scenario potentially arises whenever a nonthermalizing (e.g., many-body localized) system exhibiting sufficiently fast equilibration [97] is weakly perturbed such that the constraints inhibiting thermalization are removed.

Bound on expectation-value differences. The first step in the derivation establishes a rigorous bound on the difference $\Delta_\lambda(t)$ from (3.130). With the definition $W(t) := \mathrm{e}^{-\mathrm{i}H_\lambda t}\mathrm{e}^{\mathrm{i}H_0 t}$, $\Delta_\lambda(t)$ can be written as

$$\Delta_\lambda(t) = \mathrm{tr}[\rho_0(t)\,W^\dagger(t)\,A\,W(t)] - \mathrm{tr}[\rho_0(t)A]\,. \tag{3.131}$$

Next it is helpful to introduce the two "moments" $\hat{w}_1(t) := 2\,\mathrm{Re}\,\mathrm{tr}\{\mathbb{E}[W(t)]\rho_0(t)\}$ and $\hat{w}_2(t) := \mathrm{tr}\{\mathbb{E}[W(t)]\,\rho_0(t)\,\mathbb{E}[W^\dagger(t)]\}$ involving the ensemble-averaged (nonunitary) operator $\mathbb{E}[W(t)]$. Exploiting these definitions, the ensemble average of the difference (3.130) can then be upper-bounded in modulus by a series of manipulations detailed in the Supplemental Material of Ref. [177], yielding

$$\mathbb{E}[|\Delta_\lambda(t)|] \leq \Delta A\, d(t) \quad \text{with} \quad d(t) := \frac{1}{2}\left[3\sqrt{1-\hat{w}_2(t)} + \sqrt{1-5\hat{w}_2(t)+\hat{w}_1(t)^2}\right]. \tag{3.132}$$

As before, ΔA is the spectral range of the observable A (see Eq. (2.18)). So far, no assumptions about the ensemble of perturbations V in (3.1) were necessary, i.e., the bound (3.132) is completely general and independent of the distribution (3.24).

Moreover, by means of Markov's inequality (2.28), the bound (3.132) on the ensemble average of differences between the unperturbed and perturbed dynamics can be used to bound the probability to observe such differences in an individual realization of the perturbation ensemble, leading to

$$\mathbb{P}(|\Delta_\lambda(t)| \geq \kappa) \leq \frac{\Delta A\, d(t)}{\kappa} \tag{3.133}$$

for arbitrary $\kappa > 0$. Hence the probability to observe differences above the resolution $\kappa = \delta A$ of the macroscopic operator A is small as long as $d(t)$ from (3.132) is much smaller than the measurement precision $\Delta A/\delta A$.

From a technical point of view, it is remarkable that the otherwise commonly required step of computing the variance of the considered quantity to establish its typicality property (see Sec. 2.3.1) can be circumvented here. This is because the quantity $|\Delta_\lambda(t)|$, whose expectation value is bounded in (3.132), is nonnegative, so Markov's inequality (2.28) can be applied in place of Chebyshev's inequality (2.29). Ultimately, this and the fact that averages of order $W(t)^2$ have "magically" disappeared when passing from (3.131) to (3.132) explain why the second moment of eigenvector overlaps will suffice to evaluate $d(t)$ for the considered perturbation ensembles and thus to provide a more quantitative interpretation of (3.132) and (3.133).

Exploiting ensemble properties. As our next step, we successively incorporate certain properties of the allowed perturbation ensembles (see also Secs. 3.2 and 3.3) to simplify the right-hand side of (3.132) and extract the physical implications of this general bound. The details of these derivations can be found in Appendix E.4 (or in the Supplemental Material of Ref. [177], too). The primary objects of study in this context are the matrix elements

$$W_{\mu\nu}(t) := {}_0\langle\mu|W(t)|\nu\rangle_0 = \sum_n \mathrm{e}^{\mathrm{i}(E_\mu - E_n^\lambda)t}\, U_{n\mu}^*\, U_{n\nu} \tag{3.134}$$

of the operator $W(t)$ in the unperturbed eigenbasis $\{|\mu\rangle_0\}$. First, we assume that the statistical properties of the matrix elements $V_{\mu\nu}$ are invariant upon multiplication of the basis vectors $\{|\mu\rangle_0\}$ by arbitrary sign factors $s_\mu \in \{-1, 1\}$, a special case of phase invariance as discussed below Eq. (3.22) with $\phi_\mu \in \{0, \pi\}$. It follows that the ensemble-averaged operator $\mathbb{E}[W(t)]$ is diagonal in the unperturbed eigenbasis, i.e.,

$$\mathbb{E}[W_{\mu\nu}(t)] = \delta_{\mu\nu}\, \mathbb{E}[w_\mu(t)] \tag{3.135}$$

for some well-defined function $w_\mu(t)$ (see Eq. (E.35)). Second, the ensemble average of this function $w_\mu(t)$ is independent of μ provided that Prerequisites (i) and (iv) from Sec. 3.2 are satisfied, such that

$$\mathbb{E}[W_{\mu\nu}(t)] = \delta_{\mu\nu}\, \hat{w}(t) \tag{3.136}$$

for another well-defined function $\hat{w}(t)$. Hence $\mathbb{E}[W(t)]$ is actually proportional to the identity operator. If one additionally requires that the distribution of all diagonal matrix elements $V_{\mu\mu}$ be equivalent to the one of $-V_{\mu\mu}$, which is essentially contained in Prerequisite (iv) again, potentially after adding a trivial constant to V (see the remarks below Eq. (3.22)), it can be inferred that the function $\hat{w}(t)$ is real-valued. Exploiting all this to evaluate the functions $\hat{w}_1(t)$ and $\hat{w}_2(t)$ defined below Eq. (3.131) yields $\hat{w}_1(t) = 2\hat{w}(t)$ and $\hat{w}_2(t) = \hat{w}(t)^2$. Substituting into (3.132), we finally obtain the bound

$$\mathbb{E}[|\Delta_\lambda(t)|] \leq \Delta A\, d(t) \quad \text{with} \quad d(t) = 2\sqrt{1 - \hat{w}(t)^2} \tag{3.137}$$

for the ensemble-averaged difference between the perturbed and unperturbed dynamics. The question regarding the smallness of $d(t)$ thus boils down to the time dependence of $\hat{w}(t)$. Closer inspection of (3.134) in conjunction with (3.136) reveals that $\hat{w}(0) = 1$ and $0 \leq |\hat{w}(t)| \leq 1$. Provided that the ensemble is well-behaved, such that averages over eigenvector overlaps as in (3.134) are reasonably smooth, the function $\hat{w}(t)$ can also be expected to depend smoothly on t, so there is indeed room for an appreciable time span up to which $\hat{w}(t) \approx 1$, and consequently the perturbed dynamics will typically resemble the unperturbed one during this interval (see also Eq. (3.133)).

Quantitative estimate. So far, only abstract properties of possible perturbation ensembles were taken into account to make the general bound (3.132) more concrete, leading to (3.137) under still rather general assumptions about the V operators. Notably, it was *not* necessary to assume small level fluctuations (cf. Prerequisite (v)) or small correlations or even statistical independence of the matrix elements $V_{\mu\nu}$ (cf. Prerequisite (vi)). Hence the bound (3.137) applies to the specific perturbation ensembles introduced in Sec. 3.3, but also to considerably more general situations. For more quantitative insights regarding the smallness of $d(t)$, however, we will now employ results from Sec. 3.4 for those more specific ensembles from Sec. 3.3.

By requiring also Prerequisite (v), we can approximate $E_n^\lambda \approx E_n$ in the exponent of (3.134) and apply our results from Sec. 3.4.2 for the second moment of eigenvector overlaps to compute $\hat{w}(t)$ in (3.135). Combining Prerequisite (v) with Eqs. (3.32), (3.134), and (3.135), we then find that

$$\hat{w}(t) = \sum_n \mathrm{e}^{-\mathrm{i}(E_n - E_\mu)t}\, u(E_n - E_\mu) = [g_\lambda(t)]^* \,, \tag{3.138}$$

where we defined

$$g_\lambda(t) := \sum_n \mathrm{e}^{\mathrm{i}E_n t}\, u(E_n) = \int \mathrm{d}E\, D(E)\, \mathrm{e}^{\mathrm{i}Et}\, u(E) \,, \tag{3.139}$$

and $D(E)$ is the density of states, see Eq. (2.11). This Fourier transform (3.139) of the overlap distribution $u(E)$ from (3.32) will turn out to play a crucial role in characterizing the response of the system to typical perturbations in the subsequent Sec. 3.6, too. Since we already established that $\hat{w}(t)$ is real-valued (see above Eq. (3.137)), we actually find the same property for $g_\lambda(t)$, implying $\hat{w}(t) = g_\lambda(t)$.

In the present setup, we are naturally most interested in the case of weak perturbations because only then can we possibly expect that marked signatures of the special (nonthermalizing) reference dynamics survive also in the perturbed system. Exploiting that the overlap distribution universally assumes the Breit-Wigner form (3.33) for the considered ensembles in this regime, we find $\hat{w}(t) = \mathrm{e}^{-\Gamma|t|/2}$ via (3.138) and (3.139). Substituting into (3.137), we are then left with

$$\mathbb{E}[|\Delta_\lambda(t)|] \leq \Delta A\, d(t) \quad \text{with} \quad d(t) = 2\sqrt{1 - \mathrm{e}^{-\Gamma|t|}} \leq 2\sqrt{\Gamma|t|} \,. \tag{3.140}$$

Focusing on $t > 0$, the energy range Γ of the overlap distribution (3.33) thus determines the time scale up to which deviations between the unperturbed and perturbed dynamics remain small for nearly all perturbations of the considered ensemble: For any given $t \ll \kappa^2/(\Delta A)^2\Gamma$, Eqs. (3.133) and (3.140) imply that $\langle A\rangle_{\rho_\lambda(t)}$ and $\langle A\rangle_{\rho_0(t)}$ are indistinguishable on the preset precision level κ with overwhelming probability when randomly drawing a perturbation from the considered ensemble. Moreover, similarly as in Sec. 3.6.1 below, this argument is readily extended from single time points to intervals, meaning that we can typically expect resemblance of the unperturbed and perturbed dynamics during the entire initial relaxation period.

Interpretation and discussion. We finally come back to the scenario outlined in the beginning of this section, namely when the reference system H_0 exhibits certain constraints (such as integrability or many-body localization) which inhibit thermalization. In spite of such restrictions, as time progresses, a nonequilibrium initial state is still commonly expected to relax towards some nonthermal equilibrium state under rather general assumptions (cf. Sec. 2.2.1). We denote the time scale of this unperturbed relaxation by τ_0.

Turning to the perturbed systems, on the one hand, the vast majority of perturbations in any considered ensemble will cause the system to thermalize eventually [114, 252, 257, 258] because they usually break the aforementioned constraints. On the other hand, the bound (3.140) implies a close resemblance between the unperturbed and perturbed dynamics up to times of order Γ^{-1}. As long as $\tau_0 \ll \Gamma^{-1}$, the perturbed system will therefore also display the initial relaxation towards a nonequilibrium, quasistationary state found in the reference dynamics. The ultimate approach of thermal equilibrium will only set in at times beyond Γ^{-1}. This is precisely the type of relaxation known as *prethermalization* in the literature [9, 31, 33, 175, 176, 178], and the results of this section and Ref. [177] establish that it is the generic behavior of nonthermalizing systems subject to sufficiently weak perturbations which remove obstructions for thermalization. The most important

and paradigmatic case arises when the reference system is integrable whereas the perturbed one is not.

It remains to be clarified what "sufficiently weak" in this context means. Taking the condition $\tau_0 \ll \Gamma^{-1}$ and substituting the definition (3.65) of Γ, we find that the perturbation strength $\lambda\sigma_v$ should satisfy

$$\lambda^2\sigma_v^2 \ll \frac{\varepsilon}{\tau_0}\,. \tag{3.141}$$

For the entire argument to be valid, we then have to verify that this condition is compatible with Prerequisites (ii) and (iii) from Sec. 3.2, which essentially delimit Γ as $\varepsilon \ll \Gamma \ll \Delta_{\mathcal{E}}$, i.e., between the mean level spacing ε and the total extent $\Delta_{\mathcal{E}}$ of the microcanonical energy window $I_{\mathcal{E}}$ (cf. Eq. (2.9)). In terms of the perturbation strength, this implies $\varepsilon^2 \ll \lambda^2\sigma_v^2 \ll \varepsilon\Delta_{\mathcal{E}}$. As discussed around Eq. (3.9), the lower bound is a natural condition for the perturbation to have any noticeable effect at all. Combined with (3.141), we thus have $\varepsilon^2 \ll \lambda^2\sigma_v^2 \ll \varepsilon/\tau_0$, meaning that τ_0 should be very much smaller than the time scale set by the inverse level spacing (sometimes called the "Heisenberg time" in the literature [98]) to leave enough room for visible effects of the perturbation. Since the level spacing decreases exponentially with the degrees of freedom, this Heisenberg time usually exceeds the age of the universe by many orders of magnitude in macroscopic systems, hence $\tau_0 \ll \varepsilon^{-1}$ will always hold in practice.

For a given perturbation λV satisfying $\lambda\sigma_v \gg \varepsilon$, Eq. (3.141) therefore essentially states how fast the relaxation in the unperturbed system must be in order that prethermalization can occur. The in some sense optimal situation arises when the system relaxes as fast as $\tau_0 \sim \Delta_{\mathcal{E}}^{-1}$. In this case, the right-hand side of (3.141) coincides with the upper bound $\Gamma \ll \Delta_{\mathcal{E}}$ entailed in Prerequisite (ii) for the perturbation to leave the system's thermodynamic properties unchanged. Such fast relaxation times can indeed arise under quite generic circumstances, especially if equilibration happens without macroscopic transport currents [9, 96, 97, 100].

To sum up, we demonstrated that a nonthermalizing many-body system prepared out of equilibrium and relaxing on a time scale τ_0 will commonly exhibit prethermalization when subject to generic perturbations whose strength satisfies $\lambda^2\sigma_v^2 \ll \varepsilon/\tau_0$, provided that they are strong enough to have an effect on the relaxation behavior at all. This conclusion is based on the bound (3.140) for the difference between perturbed and unperturbed expectation values as well as the results from Refs. [114, 252, 257] ensuring that these systems will typically thermalize in the long run. The time scale on which this ultimate thermalization takes place, however, is not accessible from those results. The quantity Γ^{-1} from (3.65) merely provides a lower bound according to (3.140). The results presented in the next section, in turn, will allow us to obtain a direct estimate of this time scale of thermalization and will in fact even yield a prediction for the dynamical details of the entire relaxation process.

3.6 Typical time evolution

After the somewhat technical elaborations on prerequisites, perturbation ensembles, and overlap moments as well as a digression on prethermalization, we are finally ready to tackle the principal goal of this chapter as formulated in Sec. 3.1, namely, to find a prediction for the observable expectation values (3.2) of perturbed isolated many-body quantum systems. As outlined towards the end of Sec. 3.1, the intended typicality approach (see also Sec. 2.3) boils down to computing the ensemble-averaged dynamics (3.7) and bounding the variance (3.8). These steps will be carried out in Sec. 3.6.1. A major role in the final prediction is played by the so-called response profile function, which we already met briefly in the previous section. The properties of this response profile will be analyzed extensively in Sec. 3.6.2, before we will close this section with a more detailed discussion of the result and its implications in Sec. 3.6.3. Another important aspect of any theoretical investigation is the comparison of predictions with concrete physical systems, to which we will turn in the subsequent Sec. 3.7.

3.6.1 Expectation-value dynamics

Ensemble-averaged dynamics. As outlined at the end of Sec. 3.1, our first step to derive a prediction for the perturbed time-dependent expectation values (3.2) consists in calculating the average over all perturbations of a suitably chosen ensemble from Sec. 3.3. To this end, we resume with Eq. (3.7) for this ensemble-averaged dynamics. Adopting Prerequisite (v) of sufficiently small level fluctuations, we can write (3.7) as

$$\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}] = \sum_{\substack{\mu_1,\mu_2,\\ \nu_1,\nu_2}} \rho_{\mu_1\nu_2}(0)\, A_{\mu_2\nu_1} \sum_{m,n} \mathrm{e}^{\mathrm{i}(E_n-E_m)t}\, \mathbb{E}[U_{m\mu_1} U_{n\mu_2} U^*_{m\nu_1} U^*_{n\nu_2}]\,. \tag{3.142}$$

We will first focus on the case of sufficiently weak perturbations, for which we evaluated the required fourth-order moment of eigenvector overlaps in Sec. 3.4.3. Utilizing the result (3.35) from there (see also Eq. (3.120)), we find that

$$\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}] = \sum_{\mu,\nu} \left\{ \sum_n d^{nn}_{\mu\nu}\, \rho_{\mu\mu}(0)\, A_{\nu\nu} + \sum_{m,n} \mathrm{e}^{\mathrm{i}(E_n-E_m)t} \left[d^{mn}_{\mu\nu}\, \rho_{\mu\nu}(0)\, A_{\nu\mu} + f^{mn}_{\mu\nu}\, \rho_{\mu\mu}(0)\, A_{\nu\nu} \right] \right\} \tag{3.143}$$

with $d^{mn}_{\mu\nu}$ and $f^{mn}_{\mu\nu}$ as specified in Eqs. (3.35b) and (3.35c), respectively. To evaluate the sums over m and n, we adopt our usual convention to replace sums over energy eigenstates by integrals with a constant density of states ε^{-1}, justified by Prerequisite (iii), which implies that both $d^{mn}_{\mu\nu}$ and $f^{mn}_{\mu\nu}$ are slowly varying functions of the associated energies compared to the mean level spacing ε. For the first term in (3.143), we then obtain

$$\sum_n d^{nn}_{\mu\nu} = \int \frac{\mathrm{d}E}{\varepsilon}\, u(E-E_\mu)\, u(E-E_\nu) = \tilde{u}(E_\mu - E_\nu) \tag{3.144}$$

with the overlap distribution $u(E)$ assuming the Breit-Wigner form (3.33), and where we exploited (3.62) and the definition (3.116) of $\tilde{u}(E)$ in the last equality. For the second term, we proceed similarly to find

$$\sum_{m,n} \mathrm{e}^{\mathrm{i}(E_m-E_n)t} d^{mn}_{\mu\nu} = \mathrm{e}^{\mathrm{i}(E_\nu-E_\mu)t}\, |g_\lambda(t)|^2\,. \tag{3.145}$$

Here $g_\lambda(t)$ is the above announced *response profile*, defined in accordance with (3.139) as the Fourier transform of the overlap distribution $u(E)$ from (3.32), i.e.,

$$g_\lambda(t) = \int \frac{\mathrm{d}E}{\varepsilon}\, \mathrm{e}^{\mathrm{i}Et}\, u(E)\,. \tag{3.146}$$

Note that we explicitly indicate the dependence on the coupling strength λ here again, which is inherited from $u(E)$ but notationally suppressed there. For the presently studied case with $u(E)$ given by (3.33), we thus find $g_\lambda(t) = \mathrm{e}^{-\Gamma|t|/2}$ with Γ from (3.65). For the third term in (3.143), we observe that the summation over m and n also leads to a Fourier transformation. It can be readily evaluated by means of residue techniques, yielding

$$\sum_{m,n} \mathrm{e}^{\mathrm{i}(E_m-E_n)t} f^{mn}_{\mu\nu} = r(t, E_\mu - E_\nu) - \mathrm{e}^{-\Gamma|t|}\, \tilde{u}(E_\mu - E_\nu) \tag{3.147}$$

with

$$r(t,\omega) := \mathrm{e}^{-\Gamma|t|}\, \tilde{u}(\omega) \left[1 - \cos(\omega t) - \frac{\Gamma\, \sin(\omega|t|)}{\omega} \right]. \tag{3.148}$$

Collecting (3.144), (3.145), and (3.147), we thus conclude that the ensemble-averaged perturbed expectation values in the weak-perturbation limit take the form

$$\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}] = \langle A\rangle_{\tilde{\rho}_\lambda} + |g_\lambda(t)|^2 \left[\langle A\rangle_{\rho_0(t)} - \langle A\rangle_{\tilde{\rho}_\lambda} \right] + R(t)\,, \tag{3.149}$$

where the state $\tilde{\rho}_\lambda$ is defined in terms of its matrix elements in the eigenbasis of the unperturbed Hamiltonian as

$${}_0\langle\mu|\tilde{\rho}_\lambda|\nu\rangle_0 := \delta_{\mu\nu} \sum_\alpha \tilde{u}(E_\mu - E_\alpha)\, \rho_{\alpha\alpha}(0) \tag{3.150}$$

and the last term in (3.149) is given by

$$R(t) := \sum_{\mu,\nu} \rho_{\mu\mu}(0)\, A_{\nu\nu}\, r(t, E_\mu - E_\nu) \tag{3.151}$$

with $r(t,\omega)$ from (3.148). This latter term satisfies $R(t \to 0) = R(t \to \infty) = 0$ and turns out to be negligible in essentially all situations of practical relevance (see also Sec. 3.6.3 below). Taking this into account in (3.149), we arrive at our final result for the ensemble-averaged dynamics,

$$\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}] = \langle A\rangle_{\tilde{\rho}_\lambda} + |g_\lambda(t)|^2 \left[\langle A\rangle_{\rho_0(t)} - \langle A\rangle_{\tilde{\rho}_\lambda}\right]. \tag{3.152}$$

Until now, we explicitly worked in the limit of weak perturbations where $u(E)$ is the Breit-Wigner distribution (3.33), implying that $\tilde{u}(E)$ in (3.150) is also of the Breit-Wigner form with the width parameter Γ from (3.65) replaced by 2Γ, and $g_\lambda(t)$ from (3.146) describes an exponential decay, $g_\lambda(t) = \mathrm{e}^{-\Gamma|t|/2}$. However, we eventually reach the same final result (3.152) when substituting in (3.142) the generalized approximation for the fourth moment of eigenvector overlaps from (3.128), except that $u(E)$ now is the general overlap distribution (3.32) associated with the perturbation profile $\sigma_v^2(E)$ from (3.10) via (3.31) and encoding the corresponding second moment of eigenvector overlaps via (3.30). In other words, Eq. (3.152) also describes the ensemble-averaged observable expectation values for general $\sigma_v^2(E)$ and λ provided that Eqs. (3.116), (3.150), and (3.146) are evaluated for the overlap distribution (3.32) obtained from the corresponding solution $G(z)$ of (3.31). In particular, the resulting expression for $\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}]$ is also structurally similar to Eq. (3.149), but with a slightly different remnant term $R(t)$, which is nonetheless found to be negligible again.

Variance. The crucial second step of the adopted typicality approach concerns the deviations $\xi_V(t) = \langle A\rangle_{\rho_\lambda(t)} - \mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}]$ between one particular realization V from the perturbation ensemble and the average behavior (see also Eq. (3.3)). The idea here is to bound the variance $\mathbb{E}[\xi_V(t)^2]$, which leads to a bound for the probability to observe such deviations via Chebyshev's inequality (2.29). In view of (3.8), we thus need to compute the ensemble average of the square of (3.6),

$$\begin{aligned}\mathbb{E}\Big[\big(\langle A\rangle_{\rho_\lambda(t)}\big)^2\Big] = &\sum_{\substack{\mu_1,\mu_2,\mu_3,\mu_4\\ \nu_1,\nu_2,\nu_3,\nu_4}} \rho_{\mu_1\nu_2}(0)\, \rho_{\mu_3\nu_4}(0)\, A_{\mu_2\nu_1}\, A_{\mu_4\nu_3} \\ &\times \sum_{\substack{n_1,n_2\\ n_3,n_4}} \mathrm{e}^{\mathrm{i}(E_{n_4}-E_{n_3}+E_{n_2}-E_{n_1})t}\, \mathbb{E}[U_{n_1\mu_1} U_{n_2\mu_2} U_{n_3\mu_3} U_{n_4\mu_4} U^*_{n_1\nu_1} U^*_{n_2\nu_2} U^*_{n_3\nu_3} U^*_{n_4\nu_4}],\end{aligned} \tag{3.153}$$

where we exploited Prerequisite (v) similarly as in Eq. (3.142). Hence we have to determine the eighth moment of eigenvector overlaps (3.29). By means of the approximation scheme from Sec. 3.4.4, this leads to a quite extensive linear combination of pairwise Kronecker-δ contractions between the μ_i and ν_j indices, weighted by factors involving the second moments $\mathbb{E}[|U_{n\mu}|^2] = u(E_n - E_\mu)$. The crucial observation is that all terms of order unity in the number of mixed levels N_v from (3.9) cancel exactly against the terms obtained by squaring (3.142) [226]. With (3.8), one therefore eventually finds that

$$\mathbb{E}[\xi_V(t)^2] \leq \frac{c\,(\Delta A)^2}{N_v}, \tag{3.154}$$

where c is a positive constant of order 10^3 or less, independent of any system details (particularly H_0, λ, $\sigma_v^2(E)$, $\rho(0)$, or A), and ΔA denotes the measurement range of A as before (see the Supplemental Material of Ref. [226] for details). Therefore, considering that N_v grows exponentially with the system size (cf. Eq. (3.9)), the variance (3.154) indeed turns out to be unimaginably small in macroscopic systems or even moderately large mesoscopic ones.

The relation (3.154) thus establishes the crucial concentration-of-measure property of the time-dependent expectation values in any of the admitted perturbation ensembles from Sec. 3.3. We illustrate this measure concentration by an explicit example in Fig. 3.8, comparing fluctuations in systems of different dimensions N and different numbers of mixed levels N_v (cf. Eqs. (3.9) and (3.154)). The unperturbed Hamiltonian H_0 is taken as in (3.27) with a fixed level spacing $\varepsilon = 512/N$. The perturbations are drawn from the Gaussian Unitary Ensemble (GUE), i.e., the entries follow a complex normal distribution with mean zero and variance $\sigma_v^2 = 1$. The coupling

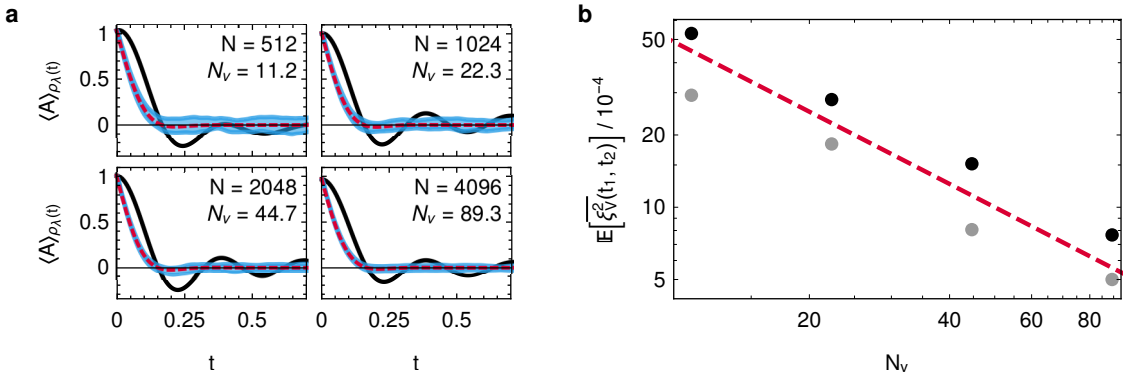


Figure 3.8: Numerical illustration of measure concentration of the perturbed time-dependent expectation values for systems of different dimensions N with H_0 from (3.27) and perturbations from the GUE with $\sigma_v^2 = 1$. Observable and initial state are constructed according to (3.155) and (2.39), respectively, with $h(t) = \mathrm{sinc}(t/0.05)$ and $\kappa = 1$. Parameters are scaled as $\varepsilon = 512/N$, $\lambda = 1.33\sqrt{512/N}$, implying $\Gamma = 11.2 = \text{const}$ and $N_v = 11.2 \times 512/N$. For each value of N, we sample 50 realizations of V and compute the resulting average time evolution $\mathbb{E}[\langle A \rangle_{\rho_\lambda(t)}]$ and variance $\mathbb{E}[\xi_V(t)^2]$ empirically. **a.** Time evolution of the unperturbed (solid black line) and perturbed (shaded blue range) systems along with the theoretically predicted average (3.152) (dashed red line) with $g_\lambda(t) = \mathrm{e}^{-\Gamma|t|/2}$ (see below Eq. (3.146)) and $\langle A \rangle_{\bar{\rho}_\lambda} = 0$. For the perturbed dynamics, the shaded ranges span one standard deviation around the mean. **b.** Time- and ensemble-averaged fluctuations (3.158) as a function of the number N_v of levels mixed by the perturbation (cf. Eqs. (3.9) and (3.154)). Gray dots correspond to time averages over the relaxation phase, $t \in [0, 0.05\pi]$, black dots additionally include the stationary phase, $t \in [0, 1]$. The dashed red line is given by $0.05/N_v$ as a guide to the eye.

strength is scaled as $\lambda = 1.33\sqrt{512/N}$. For $N = 512$, this is the same setup as in Fig. 3.6. The particular scaling of λ and ε with N is chosen to ensure that the width $\Gamma = 11.2$ of the overlap distribution $u(E)$ (see Eqs. (3.33) and (3.65)) is independent of N, implying that also the relaxation time scale mediated by the perturbations via $g_\lambda(t) = \mathrm{e}^{-\Gamma|t|/2}$ (see below Eq. (3.146)) is the same for all N. At the same time, the number of mixed levels scales as $N_v = \Gamma/\varepsilon = 11.2\,(512/N)$, so the fluctuations are expected to decrease as N is increased according to (3.154).

For the observable A, we adopt the construction from Ref. [259] of an ETH-compatible observable with prescribed reference dynamics $\langle A \rangle_{\rho_0(t)} = h(t)$ by choosing

$$A_{\mu\nu} := \hat{h}(E_\mu - E_\nu) M_{\mu\nu}\,. \tag{3.155}$$

Here $\hat{h}(\omega)$ is the Fourier transform of the desired dynamics $h(t)$ and $M_{\mu\nu} = M_{\nu\mu}$ are real-valued, independent, Gaussian-distributed random variables of vanishing mean and variance $1/N$ for $\mu \leq \nu$. Within the dynamical-typicality framework, starting from an initial state $\rho(0) = |\psi\rangle\langle\psi|$ generated according to (2.39) with $\kappa = 1$, this indeed yields $\langle A \rangle_{\rho_0(t)} = h(t)$ as the typical dynamics found in sufficiently large systems [259]. In the following example, we choose Π in (2.39) as a projector onto the central 20 % of states.

In Fig. 3.8a, we show the so-obtained reference dynamics for $h(t) = \mathrm{sinc}(20t)$ and various N as black solid lines. We then generate 50 random perturbations V from the GUE for each N and compute the empirical mean and variance of the resulting $\langle A \rangle_{\rho_\lambda(t)}$ numerically for all t. The blue shaded areas in Fig. 3.8 show the ranges of one standard deviation around the mean. We also depict the analytically calculated ensemble average (3.152) with $g_\lambda(t) = \mathrm{e}^{-\Gamma|t|/2}$ (see below Eq. (3.146)) and $\langle A \rangle_{\bar{\rho}_\lambda} = 0$ as red dashed lines. Reassuringly, these predictions lie well in the middle of the shaded areas. Moreover, we observe that the fluctuations indeed decrease qualitatively as N_v is increased. To quantify this decrease, we look at the time average of the squared fluctuations $\xi_V(t)^2$ from (3.8) during a fixed interval $[t_1, t_2]$,

$$\overline{\xi_V^2}(t_1, t_2) := \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \mathrm{d}t\; \xi_V(t)^2\,. \tag{3.156}$$

The empirical average of this quantity over all V samples, which is an estimate of the ensemble average $\mathbb{E}[\overline{\xi_V^2}(t_1, t_2)]$, is displayed in Fig. 3.8b for two different intervals: The gray dots focus on the

initial relaxation period with $t_1 = 0$ and $t_2 = \pi/20$ (the first zero of $h(t)$), whereas the black dots incorporate the stationary phase as well by employing $t_2 = 1$ instead. In both cases, we observe that the integrated fluctuations (3.156) approximately decrease like $1/N_v$, reinforcing the scaling obtained in (3.154) and thus verifying measure concentration in this particular example.

Typical dynamics. The smallness of the variance (3.154) entails that deviations from the average behavior (3.152) are exceedingly rare when randomly sampling a perturbation from any of the considered ensembles. As mentioned in Sec. 2.3.1, this notion can be cast into a more quantitative form by applying Chebyshev's inequality (2.29) to the random variable $\langle A\rangle_{\rho_\lambda(t)}$. Choosing, for example, $\kappa = \Delta A/N_v^{1/3}$ in (2.29), we obtain that

$$\mathbb{P}\left(\left|\langle A\rangle_{\rho_\lambda(t)} - \mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}]\right| \geq \frac{\Delta A}{N_v^{1/3}}\right) \leq \frac{c}{N_v^{1/3}} \tag{3.157}$$

for any preset time point t. Given that N_v increases exponentially with the degrees of freedom (see Eq. (3.9)), the probability that an individual perturbation induces deviations from the average behavior above any practically reasonable measurement precision is thus unimaginably small for the considered many-body systems. Moreover, it is straightforward to extend the bound (3.157) to entire time intervals [96]. To this end, we consider again the time-averaged squared fluctuations $\overline{\xi_V^2}(t_1, t_2)$ from (3.156). Averaging over the ensemble of perturbations and exploiting (3.154) leads to

$$\mathbb{E}\left[\overline{\xi_V^2}(t_1, t_2)\right] \leq \frac{c\,(\Delta A)^2}{N_v}\,, \tag{3.158}$$

and since $\overline{\xi_V^2}(t_1, t_2) \geq 0$, Markov's inequality (2.28) implies that

$$\mathbb{P}\left(\overline{\xi_V^2}(t_1, t_2) \geq \frac{(\Delta A)^2}{N_v^{1/2}}\right) \leq \frac{c}{N_v^{1/2}}\,. \tag{3.159}$$

Consequently, the probability to observe noticeable deviations from the average dynamics during an extended time interval $[t_1, t_2]$, and thus for any given time $t \in [t_1, t_2]$, is also exceedingly small. Altogether, we therefore conclude that the ensemble average $\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}]$ from (3.152) is an excellent approximation of the actual dynamics $\langle A\rangle_{\rho_\lambda(t)}$ for the overwhelming majority of times t and perturbations V from any admissible ensemble. Put differently, Eq. (3.152) is thus turned into a prediction for the relaxation behavior of an individual perturbed system,

$$\boxed{\langle A\rangle_{\rho_\lambda(t)} = \langle A\rangle_{\tilde{\rho}_\lambda} + |g_\lambda(t)|^2 \left[\langle A\rangle_{\rho_0(t)} - \langle A\rangle_{\tilde{\rho}_\lambda}\right]}, \tag{3.160}$$

where exceptions from this typical behavior for a single perturbation are exponentially suppressed in the system's degrees of freedom. This relation (3.160) between the perturbed and unperturbed time-dependent expectation values constitutes the main result of the present chapter. Evidently the principle characteristic governing the modifications induced by the perturbation is the response profile $g_\lambda(t)$ from (3.146). We will therefore analyze this quantity in more detail in the ensuing Sec. 3.6.2, before discussing the implications of the result for concrete physical systems in Sec. 3.6.3.

At last, we remark that the prediction (3.160) bears some structural resemblance to the typical relaxation prediction (2.40) for a rather different set of Hamiltonians (see Sec. 2.3.2). We therefore highlight that the present relation (3.160) is in fact quite distinct: Instead of the initial value $\langle A\rangle_{\rho(0)}$, the time-dependent expectation values $\langle A\rangle_{\rho_0(t)}$ of the unperturbed reference system occur in the bracket on the right-hand side, implying that this bracket is generally time dependent, too, contrary to the one in (2.40). Furthermore, also the response profile $g_\lambda(t)$ is distinct from and somewhat more complicated than the function $\hat{d}(t)$ in (2.40) since the former is the Fourier transform of the product $D(E)\,u(E)$ whereas $\hat{d}(t)$ is obtained by Fourier transformation of just $D(E)$. Apart from that, the similarity essentially arises because the structure of Eqs. (2.40) or (3.160) itself is kind of generic, entailing a long-time limit and a characterization of the time-dependent deviations from it for finite times.

3.6.2 Response profile

According to the theoretical prediction (3.160), the response profile $g_\lambda(t)$ from (3.146) encodes how the perturbed expectation values typically deviate from the reference dynamics when randomly sampling perturbations from an ensemble as introduced in Sec. 3.3. This function $g_\lambda(t)$ is defined as the Fourier transform of the overlap distribution $u(E)$ from (3.32) describing the second moment of eigenvector overlaps between the perturbed and unperturbed Hamiltonians. This overlap distribution, in turn, was discussed extensively in Sec. 3.4.2 as part of a more general exploration of the ensemble-averaged resolvent $G(z)$ from (3.58) solving the integral equation (3.31), whose imaginary part entails $u(E)$ via (3.32).

Most of the subsequently established properties of the response profile $g_\lambda(t)$ can therefore be understood as translations of corresponding properties of the overlap distribution $u(E)$ from the energy to the time domain. Notably, the (approximate) universality of $G(z)$ and thus $u(E)$ will carry over to $g_\lambda(t)$, whose indifference towards the detailed structure of the perturbation (i.e., the perturbation profile $\sigma_v^2(E)$) will turn out to be even more striking in some sense.

General properties. We begin by collecting a few general properties and observations independent of any particular choice of the perturbation ensemble or strength. The normalization of $u(E)$, $\int \mathrm{d}E\, u(E)/\varepsilon = 1$, immediately implies that

$$g_\lambda(0) = 1\,. \tag{3.161}$$

Moreover, if $u(E)$ is a regular function, which it is found to be in all examples of practical relevance, we also have

$$\lim_{t\to\pm\infty} g_\lambda(t) = 0\,. \tag{3.162}$$

Next, the fact that $u(E)$ is real-valued in combination with the symmetry (3.62) entails that $g_\lambda(t)$ is also real-valued and symmetric in time, i.e.,

$$g_\lambda(-t) = g_\lambda(t)\,. \tag{3.163}$$

Finally, the energy scale Γ_v associated with the width of the overlap distribution (see the discussion around Eq. (3.9)) provides an estimate of the typical relaxation time scale Γ_v^{-1} of $g_\lambda(t)$. All these properties will be seen explicitly in the following examples, too.

Special cases. As our first concrete set of examples, we discuss the three special cases for the overlap distribution considered in Sec. 3.4.2. In the first of these examples, pertaining to the limit of weak perturbations such that $\Gamma \ll \Delta_v$ (see Eqs. (3.13), (3.65), and (3.66)), the perturbation profile is effectively constant and $u(E)$ is the Breit-Wigner distribution (3.33). Hence the response profile (3.146) describes an exponential decay,

$$g_\lambda(t) = \mathrm{e}^{-\Gamma|t|/2}\,. \tag{3.164}$$

In the opposite limit of an effectively narrow perturbation profile, corresponding to stronger perturbations satisfying $\gamma \gg \Delta_v$ (see Eqs. (3.13), (3.69), and (3.71)), $u(E)$ assumes a semicircle distribution (3.34), whose Fourier transform is

$$g_\lambda(t) = \frac{2\,J_1(\gamma t)}{\gamma t}\,. \tag{3.165}$$

Here $J_1(x)$ is the Bessel function of the first kind of order 1. Observing that $u(E)$ generically approaches these two limits for sufficiently small or large values of λ, respectively, while keeping all other parameters fixed, the same must hold for $g_\lambda(t)$ as λ is varied. Accordingly, we can expect an essentially exponential response profile (3.164) for weak perturbations, which gradually crosses over to the Bessel-like behavior (3.165) with increasing λ, provided the prerequisites from Sec. 3.2 remain satisfied. An estimate for the coupling λ_c at which the crossover approximately occurs was given in Eq. (3.73).

In the third special case considered in Sec. 3.4.2, the perturbation profile was taken to exhibit the Breit-Wigner form (3.74). Approximations for $g_\lambda(t)$ can be extracted from the exact solution (3.77)

for $G(z)$ by truncating the continued fraction at some finite order M_{cf} and computing the Fourier transform of its imaginary part. In particular, the second-order approximation (3.78) yields

$$g_\lambda(t) = \frac{(\gamma_+ - \frac{\Gamma}{2})\,\mathrm{e}^{-\gamma_-|t|} - \Gamma\,\mathrm{e}^{-\gamma_0|t|} + (\gamma_- - \frac{\Gamma}{2})\,\mathrm{e}^{-\gamma_+|t|}}{2(\gamma_0 - \Gamma)} \tag{3.166}$$

with $\gamma_\pm \equiv \gamma_{\pm 1}$ and γ_n as defined in (3.79), ultimately depending only on the overall perturbation strength $\alpha_v\lambda^2$ and the perturbation band width Δ_v. Since the expansion in the order M_{cf} of the continued fractions basically amounts to a weak-perturbation expansion, the approximation (3.166) is expected to apply primarily to the regime of moderately small λ. However, as we will see below (e.g., in Fig. 3.9), the approximation works remarkably well across a large range of λ values and for perturbation profiles other than the Breit-Wigner one from (3.74), too.

Integro-differential equation and numerical solutions. The ensemble-averaged resolvent $G(z)$ solves the integral equation (3.31). Observing that $g_\lambda(t)$ is related to the Fourier transform of its imaginary part, it is possible to transform the equation (3.31) into an equation directly for $g_\lambda(t)$. Notably, this will relate the response profile $g_\lambda(t)$ to the Fourier transform

$$\hat{\sigma}_v^2(t) := \int \frac{\mathrm{d}E}{\varepsilon}\,\mathrm{e}^{\mathrm{i}Et}\,\sigma_v^2(E) \tag{3.167}$$

of the perturbation profile $\sigma_v^2(E)$ from (3.10). To this end, we let $z = x - \mathrm{i}\eta$ for fixed $\eta > 0$, multiply both sides of (3.31) by $\mathrm{e}^{\mathrm{i}xt}$, and integrate over x. With the definition

$$h_\eta(t) := \int \mathrm{d}x\;\mathrm{e}^{\mathrm{i}xt}\,G(x - \mathrm{i}\eta)\,, \tag{3.168}$$

this procedure results in

$$\frac{1}{\mathrm{i}}\left(\frac{\partial}{\partial t} + \eta\right) h_\eta(t) - \lambda^2 \int \frac{\mathrm{d}E}{\varepsilon}\,\sigma_v^2(E) \int \mathrm{d}x \int \mathrm{d}y\;\mathrm{e}^{-\mathrm{i}xt} G(x - \mathrm{i}\eta)\,G(y - \mathrm{i}\eta)\,\delta(x - y - E) = 2\pi\,\delta(t)\,. \tag{3.169}$$

Expressing the δ distribution on the left-hand side by a Fourier integral, $2\pi\,\delta(x - y - E) = \int \mathrm{d}s\,\mathrm{e}^{\mathrm{i}(x-y-E)s}$, exploiting the definitions (3.167) and (3.168), and taking the limit $\eta \to 0+$, we obtain

$$\dot{h}_{0+}(t) + \frac{\lambda^2}{2\pi\mathrm{i}} \int \mathrm{d}s\;h_{0+}(t - s)\,h_{0+}(s)\,\hat{\sigma}_v^2(s) = 2\pi\mathrm{i}\,\delta(t)\,, \tag{3.170}$$

where the dot in the first term denotes the derivative with respect to time t. In view of the δ inhomogeneity on the right-hand side, we now make an ansatz of the form $h_{0+}(t) = 2\pi\mathrm{i}\,\Theta(t)\,g(t)$, where $g(t)$ is assumed to be a bounded and sufficiently smooth function of t. Substituting into (3.170), we are left with

$$\Theta(t)\,\dot{g}(t) + \lambda^2 \int_0^t \mathrm{d}s\;g(t - s)\,g(s)\,\hat{\sigma}_v^2(s) = 0\,. \tag{3.171}$$

We observe that if $g(0) \in \mathbb{R}$, then the function $g(t)$ is real-valued for all $t > 0$, too. Taking into account (3.32), (3.146) and (3.168), we furthermore notice that

$$g_\lambda(t) = \frac{1}{2\pi\mathrm{i}}\,[h_{0+}(t) - h_{0+}(-t)^*] = \Theta(t)\,g(t) + \Theta(-t)[g(-t)]^*\,. \tag{3.172}$$

For $t > 0$, we therefore conclude that $g_\lambda(t)$ satisfies the integro-differential equation

$$\dot{g}_\lambda(t) = -\lambda^2 \int_0^t \mathrm{d}s\;g_\lambda(t - s)\,g_\lambda(s)\,\hat{\sigma}_v^2(s)\,. \tag{3.173}$$

The response profile for $t < 0$ is then obtained via (3.163). Moreover, it also solves (3.173) as can be verified by noticing that $\sigma_v^2(E) = \sigma_v^2(-E)$ is a real-valued and even function, so the same applies to its Fourier transform $\hat{\sigma}_v^2(t)$ from (3.167). In conclusion, the response profile $g_\lambda(t)$ thus satisfies (3.173) for all t.

Solving (3.173) together with the initial condition (3.161) thus offers a direct way to numerically determine the response profile for any given perturbation profile $\sigma_v^2(E)$, provided that its Fourier

transform $\hat{\sigma}_v^2(t)$ exists. Another way to obtain $g_\lambda(t)$ for general perturbation profiles $\sigma_v^2(E)$ is to solve the integral equation (3.31) numerically using, for instance, a pseudospectral expansion in terms of Chebyshev rational functions as suggested in Sec. 3.4.2. Extracting the overlap distribution $u(E)$ from such a numerical solution via (3.32), we can then perform a (numerical) Fourier transformation to find $g_\lambda(t)$.

Crossover in the time domain. We already explored above that $g_\lambda(t)$ undergoes a transition from the exponential decay (3.164) for small coupling λ to the Bessel-like behavior (3.165) for large λ, at least under the rather mild condition that $\sigma_v^2(E)$ is sufficiently slowly varying compared to the level spacing ε (see also the discussion around Eq. (3.63)). Interestingly, a similar crossover is observed universally in the time domain. To investigate this crossover in more detail, the integral equation (3.173) provides a convenient starting point.

To determine the asymptotic behavior as $t \to \infty$, we use the ansatz $g_\lambda(t) \sim \mathrm{e}^{-\kappa t}$. Inserting it into (3.173) and employing the inverse transformation of (3.167) together with $\hat{\sigma}_v^2(-t) = \hat{\sigma}_v^2(t)$ (see below Eq. (3.173)), we obtain

$$\kappa = \lambda^2 \int_0^t \mathrm{d}s\; \hat{\sigma}_v^2(s) \xrightarrow{t\to\infty} \frac{\pi\lambda^2\sigma_v^2}{\varepsilon} = \frac{\Gamma}{2}\,. \tag{3.174}$$

Here Γ was defined in (3.65). Hence $g_\lambda(t)$ approaches for large t the limiting expression (3.164), which was previously found to govern the regime of small λ or large Δ_v. Hence the approximation (3.165) for large λ will eventually break down at large times for any fixed, finite value of λ because the associated asymptotic behavior there is $g_\lambda(t) \sim t^{-3/2}$ as $t \to \infty$.

For small times t, in turn, the integral equation (3.173) can be used to find a Taylor series representation of $g_\lambda(t)$ by evaluating the derivatives $g_\lambda^{(k)}(t) := \mathrm{d}^k g_\lambda(t)/\mathrm{d}t^k$ recursively at $t = 0$, provided that the corresponding integrals on the right-hand side exist. More precisely, the value of $g_\lambda^{(k)}(0)$ can thereby be related to the *moments*

$$\Sigma_n := \frac{1}{\sigma_v^2} \int \mathrm{d}E\; E^n \sigma_v^2(E) \qquad (n \in \mathbb{N}_0) \tag{3.175}$$

of the perturbation profile of order $n \leq k-2$. Hence the kth derivative of $g_\lambda(t)$ at $t = 0$ exists if the $(k-2)$th moment of $\sigma_v^2(E)$ is finite. Moreover, we note that—if they exist—the odd moments of $\sigma_v^2(E)$ vanish because it is an even function (see below Eq. (3.10)).

To establish the relationship between the derivatives of $g_\lambda(t)$ and the moments of $\sigma_v^2(E)$, we first notice that $\dot{g}_\lambda(0) = 0$ if the integrand in (3.173) is sufficiently regular. Differentiating (3.173) with respect to t, we find that

$$\ddot{g}_\lambda(t) = -\lambda^2\, g_\lambda(t)\, \hat{\sigma}_v^2(t) - \int_0^t \mathrm{d}s\; \dot{g}_\lambda(t-s)\, g_\lambda(s)\, \hat{\sigma}_v^2(s) \tag{3.176}$$

and hence

$$\ddot{g}_\lambda(0) = -\lambda^2 \hat{\sigma}_v^2(0) = -\lambda^2\, \sigma_v^2\, \Sigma_0/\varepsilon = -2\lambda^2\sigma_v^2\Delta_v/\varepsilon = -\gamma^2/4\,, \tag{3.177}$$

where we adopted the definition (3.69) in the last equality. If the band width Δ_v of $\sigma_v^2(E)$ is finite, the curvature of $g(t)$ is thus also universally determined by just two characteristics of the perturbation profile, its overall strength $\alpha_v\lambda^2$ and its band width Δ_v (see Eqs. (3.12) and (3.13)). Moreover, for finite band width and finite but small λ, we therefore conclude that the weak-perturbation asymptotics (3.164) will eventually break down for very small times t because neither does its first derivative vanish nor does the second derivative agree with (3.177). Put differently, the limits $\lambda \to 0$ and $t \to 0$ do not commute.

Upon repeated differentiation of (3.173) or (3.176), it is now straightforward to obtain expressions for higher-order derivatives $g_\lambda^{(k)}(t)$, depending recursively on derivatives of lower order. Specifically, we can readily verify by induction that

$$g_\lambda^{(2n)}(t) = -\lambda^2 \sum_{r=0}^{n-1} g_\lambda^{(2n-2r-2)}(0) \sum_{k=0}^{2r} \binom{2r}{k} \frac{\mathrm{d}^k \hat{\sigma}_v^2(t)}{\mathrm{d}t^k}\, g_\lambda^{(2r-k)}(t) - \lambda^2 \int_0^t \mathrm{d}s\; g_\lambda^{(2n-1)}(t-s)\, g_\lambda(s)\, \hat{\sigma}_v^2(s)\,, \tag{3.178a}$$

$$g_\lambda^{(2n+1)}(t) = -\lambda^2 \sum_{r=0}^{n-1} g_\lambda^{(2n-2r-2)}(0) \sum_{k=0}^{2r} \binom{2r+1}{k} \frac{\mathrm{d}^k \hat{\sigma}_v^2(t)}{\mathrm{d}t^k}\, g_\lambda^{(2r+1-k)}(t) - \lambda^2 \int_0^t \mathrm{d}s\; g_\lambda^{(2n)}(t-s)\, g_\lambda(s)\, \hat{\sigma}_v^2(s) \tag{3.178b}$$

for all $n \in \mathbb{N}$. Evaluated at $t = 0$ (and still assuming regularity of the integrands), this implies

$$g_\lambda^{(2n)}(0) = -\lambda^2 \alpha_v \sum_{r=0}^{n-1} g_\lambda^{(2n-2r-1)}(0) \sum_{k=0}^{r} \binom{2r}{2k} (-1)^k \Sigma_k\, g_\lambda^{(2r-2k)}(0)\,, \tag{3.179a}$$

$$g_\lambda^{(2n+1)}(0) = 0\,. \tag{3.179b}$$

For small t, in particular, we thus observe that

$$g_\lambda(t) = 1 - \frac{\gamma^2 t^2}{8} + \mathcal{O}(t^4)\,. \tag{3.180}$$

Altogether, we thus found that the short-time dynamics is governed by the quadratic polynomial (3.180), whereas the long-time limit is dominated by the exponential decay (3.164). The time scale at which the crossover between these two regimes occurs can be estimated by equating (3.180) with the leading order of (3.164), $\mathrm{e}^{-\Gamma|t|/2} \simeq 1 - \Gamma|t|/2$, yielding

$$t_\mathrm{c} := \frac{\pi}{\Delta_v}\,. \tag{3.181}$$

To leading order, the crossover time is thus solely determined by the perturbation's band width Δ_v, irrespective of its strength $\lambda\sigma_v$. This is consistent with the findings from Ref. [206], where a transition from a Gaussian behavior to an exponential decay around the time Δ_v^{-1} was predicted based on a related random matrix model and numerical observations (see also Sec. 3.8).

Notably, the Bessel-type asymptotic expression (3.165) for $g_\lambda(t)$ in case of large λ or small Δ_v agrees with (3.180) to third order and can thus serve as a general approximation for the regime of small times, too.

Graphical illustration and universality. Summarizing the asymptotic analysis of this subsection, we found that, for any sufficiently regular perturbation profile with fixed α_v and Δ_v, the response profile $g_\lambda(t)$ approaches universal functions for small and large coupling strengths λ as well as for short and long times t. Universal here means that the functions solely depend on those two parameters α_v and Δ_v, regardless of further details of the perturbation profile $\sigma_v^2(E)$.

For small $\lambda \ll \lambda_\mathrm{c}$ (cf. Eq. (3.73)), the response profile approaches (3.164) and is thus predominantly exponential on the relevant time scales. For large $\lambda \gg \lambda_\mathrm{c}$, in turn, $g_\lambda(t)$ turns into the Bessel-like form (3.165) for those relevant time scales where it noticeably deviates from zero. For intermediate values, we expect a smooth crossover between these two limiting cases.

Perhaps surprisingly, the same limiting expressions govern the behavior of $g_\lambda(t)$ at fixed λ for long and short times, respectively. For large times $t \gg t_\mathrm{c}$ (cf. Eq. (3.181)), the decay of $g_\lambda(t)$ eventually becomes exponential as in (3.164), whereas for small times $t \ll t_\mathrm{c}$, it is well approximated by the Bessel-like behavior (3.165). At intermediate times, we again expect a crossover between these two limits. We remark that the relaxation time scale corresponding to λ_c, which is obtained by evaluating either Γ^{-1} from (3.65) or γ^{-1} from (3.69) for $\lambda = \lambda_\mathrm{c}$, is given by $t_\mathrm{c}/4$. Hence the transition in the time domain will typically only be visible for moderately small couplings $\lambda \lesssim \lambda_\mathrm{c}$.

All findings of the present subsection are illustrated in Fig. 3.9. The four rows in this figure show numerical solutions (solid black lines) of the squared magnitude $|g_\lambda(t)|^2$ of the response profile (which is the relevant quantity for the dynamics, see Eq. (3.160)) for four different perturbation profiles: the step profile (3.85), the exponential profile (3.21), the Breit-Wigner profile (3.74), and the double Breit-Wigner profile (3.86). In all cases, $\varepsilon^{-1} = 512$, $\sigma_v^2 = 0.2$ (hence $\alpha_v \approx 102$), and $\Delta_v = 1.46$, and the perturbation profiles are also sketched in the left-most panel of each row. These response profile functions are thus in direct correspondence to the overlap distributions displayed for the same four perturbation profiles in Fig. 3.5 (up to the scaling factor ε/π). The coupling strength λ is varied from $\lambda = 0.02$ to $\lambda = 0.16$ from left to right, tuning across the transition around $\lambda_\mathrm{c} \approx 0.05$ according to (3.73).

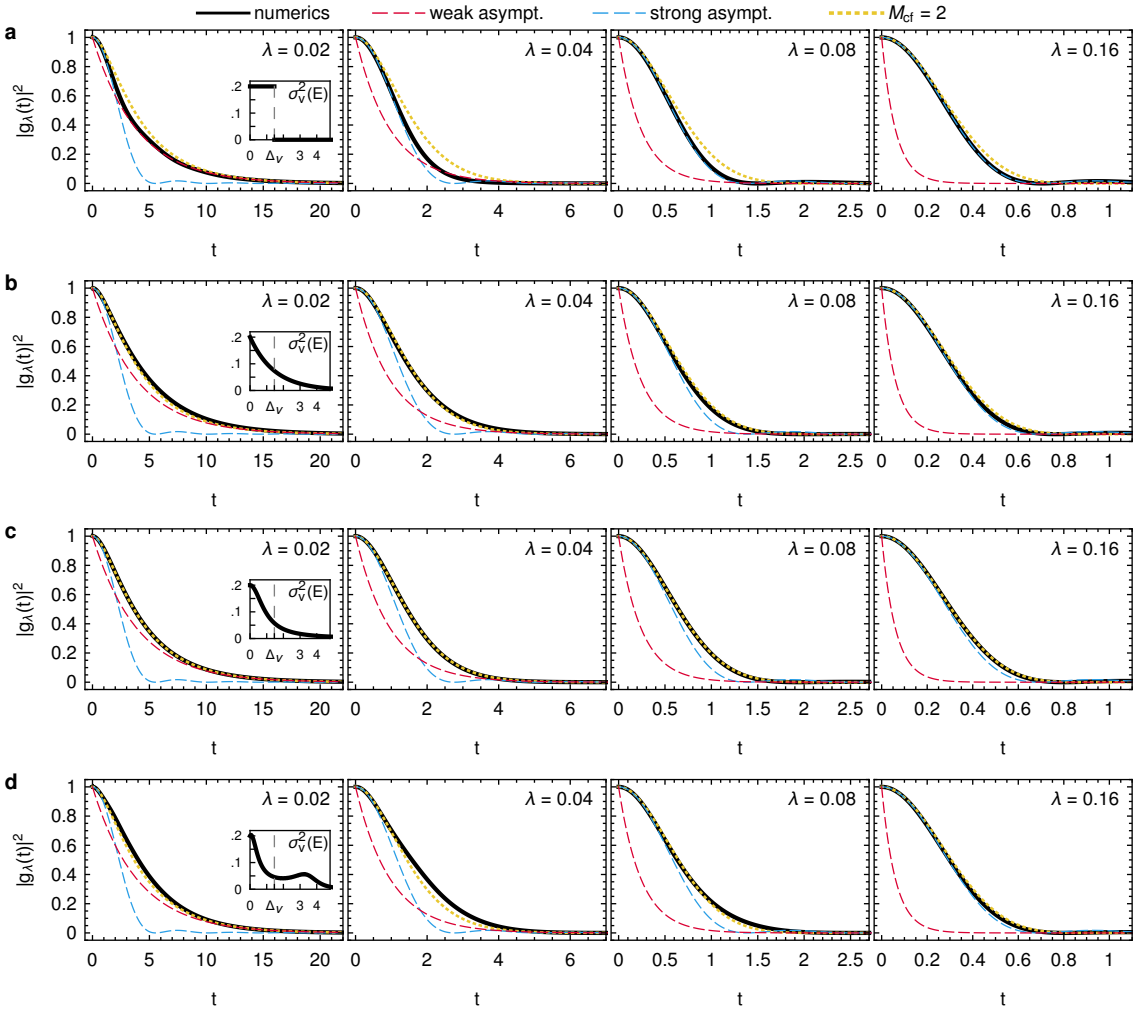


Figure 3.9: Squared response profile $|g_\lambda(t)|^2$ for various perturbation profiles $\sigma_v^2(E)$ (see insets of the left-most column) and coupling strengths λ (see top-right corner of each panel). For all configurations, $\varepsilon^{-1} = 512$, $\sigma_v^2 = 0.2$, and $\Delta_v = 750\varepsilon \approx 1.46$, yielding a crossover coupling of $\lambda_c \approx 0.05$ according to (3.73) and a crossover time $t_c \approx 2.2$ according to (3.181). Solid: Numerical solutions obtained by Fourier transformation of the respective overlap distributions $u(E) = \varepsilon \operatorname{Im} G^+(E)/\pi$ from Fig. 3.5 for **a.** the step profile (3.85); **b.** the exponential profile (3.21); **c.** the Breit-Wigner profile (3.74); **d.** the double Breit-Wigner profile (3.86) with $b_1 = 0.45$, $b_2 = 0.9$, $d = 3.5$. Dashed: Universal asymptotic solutions for weak (red, Eq. (3.164)) and strong (blue, Eq. (3.165)) coupling. Dotted: $g_\lambda(t)$ from (3.166) for the Breit-Wigner profile (3.74), obtained from the second-order continued fraction truncation (3.78) for $G(z)$.

With each of these numerically exact solutions for $g_\lambda(t)$, we plot the corresponding asymptotic expressions for weak (dashed red lines) and strong (dashed blue lines) perturbations from Eqs. (3.164) and (3.165), respectively. Recalling that α_v and Δ_v are fixed, these dashed curves are thus identical in every column. The transition between the two limiting shapes is clearly visible as λ is varied. Moreover, this also demonstrates the aforementioned crossover in the time domain, which is best observed in the left-most column of each row, noting that the expected crossover time is $t_c \approx 2.2$ according to (3.181).

In addition, we show in each panel the approximation (3.166) (dotted yellow lines) obtained from the second-order continued-fraction truncation of $G(z)$ for the Breit-Wigner perturbation profile (3.74). Naturally, this approximation is therefore expected to work best in Fig. 3.9c. Having said this, it turns out that the agreement is also amazingly good for the other three perturbation profiles. This illustrates directly the (approximate) universality of the response profile, which was already foreshadowed by the universal limiting expressions for small and large coupling strengths and short and long times. As visualized in Fig. 3.9, the behavior of $g_\lambda(t)$ for any reasonably regular

perturbation profile is largely determined by the values of α_v and Δ_v.

For completeness, we mention that there are some deviations between the second-order continued-fraction solution (3.166) for even larger times t and values of the coupling λ not shown here [248]. However, for such large values of λ, the premises from Sec. 3.2, notably Prerequisite (ii), will presumably break down, too, meaning that this situation was essentially excluded *a priori*. Otherwise, this scenario will still be covered well by the strong-perturbation asymptotics (3.165). Furthermore, we recall that a related universality was observed in Sec. 3.4.2 (see Figs. 3.4 and 3.5 in particular) for the ensemble-averaged resolvent $G(z)$ from (3.57) already.

Altogether, we therefore conclude that the analytic expressions (3.164), (3.165), and (3.166), which were originally obtained for special cases of the perturbation profile, in fact cover a very broad (if not comprehensive) regime of physically relevant coupling strengths and perturbation profiles. Moreover, the necessary information about the perturbation has been reduced even further. Instead of the full perturbation profile, which was the starting point of our modeling in Sec. 3.2, we infer that, in essence, only the intrinsic strength α_v from (3.12) and the band width Δ_v from (3.13) are required for a decent description of the response profile $g_\lambda(t)$.

3.6.3 Discussion

The main result of this chapter, stated in Sec. 3.6.1, is the prediction (3.160) of how the time-dependent expectation values of the reference system H_0 change under the influence of most perturbations. The principal characteristic is the response profile $g_\lambda(t)$, whose properties we studied extensively in the previous subsection. Here, we will first elaborate on the other quantities appearing in (3.149) and (3.160), namely $R(t)$ and $\langle A\rangle_{\tilde{\rho}_\lambda}$. We also discuss the validity of Prerequisite (v) about negligibly small level fluctuations. Thereafter, we will comment on what the result (3.160) entails for the real physical system of interest in (3.1) and its "true" perturbation.

Bound for the remnant term. We had already commented briefly on the remnant term $R(t)$ from (3.151) below that equation, arguing that $R(t=0) = R(t\to\infty) = 0$ and even $R(t) \simeq 0$ for all practical purposes. As a first step to fortify the latter claim, we notice that $\int \mathrm{d}\omega\, r(t,\omega) = 0$ for all t, where $r(t,\omega)$ was defined in (3.148). If the reference system H_0 satisfies the ETH (see Sec. 2.2.2), the relevant diagonal matrix elements $A_{\nu\nu}$ of A in the eigenbasis of H_0, corresponding to levels $E_\nu \in I_\mathcal{E}$ (see Eq. (2.9)) and appearing in (3.151), are well approximated by the microcanonical expectation value, i.e., $A_{\nu\nu} \approx \langle A\rangle_{\rho_\mathrm{mc}}$. Employing this approximation in (3.151) and rewriting the sum over ν as an integral as usual, we then readily find that $R(t) = 0$ in this case.

More generally, the magnitude of $R(t)$ can be bounded on the relevant relaxation time scales in terms of the violation of the ETH as measured by summing the deviations between $A_{\nu\nu}$ and $\langle A\rangle_{\rho_\mathrm{mc}}$ across the energy window $I_\mathcal{E}$ from (2.9), i.e.,

$$\Delta_\mathrm{mc}(A) := \sum_{\nu\,:\,E_\nu\in I_\mathcal{E}} |A_{\nu\nu} - \langle A\rangle_{\rho_\mathrm{mc}}|\,. \tag{3.182}$$

To this end, we first rewrite the result (3.149) in the form

$$\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}] = \langle A\rangle_{\tilde{\rho}_\lambda} + |g_\lambda(t)|^2 \left[\langle A\rangle_{\rho_0(t)} - \langle A\rangle_{\tilde{\rho}_\lambda} + \tilde{R}(t)\right] \tag{3.183}$$

with

$$\tilde{R}(t) := \sum_{\mu,\nu} \rho_{\mu\mu}(0)\, A_{\nu\nu}\, \tilde{r}(t, E_\mu - E_\nu)\,, \quad \tilde{r}(t,\omega) := \tilde{u}(\omega)\left[1 - \cos(\omega t) - \frac{\Gamma\sin(\omega|t|)}{\omega}\right], \tag{3.184}$$

meaning that we factored out the response-profile contribution $|g_\lambda(t)|^2 = \mathrm{e}^{-\Gamma|t|}$ in Eqs. (3.148) and (3.151). As detailed in Appendix E.5, we can then show that

$$|\tilde{R}(t)| \leq \frac{C_\kappa\, \Delta_\mathrm{mc}(A)}{N_v} \quad \text{for all } t \leq t_{\kappa *}\,, \tag{3.185}$$

where $t_{\kappa *} = -\ln(\kappa)/\Gamma$ is the time up to which deviations of $\mathbb{E}[\langle A\rangle_{\rho_\lambda(t)}]$ from $\langle A\rangle_{\tilde{\rho}_\lambda}$ are resolvable with precision κ, i.e., $|g_\lambda(t)|^2 = \mathrm{e}^{-\Gamma|t|} \geq \kappa$. Furthermore, C_κ is a constant that roughly corresponds

to the number of digits precision (e.g., $C_\kappa \approx 20$ for $\kappa = 10^{-20}$) and N_v is the number of mixed levels from (3.9). Given that the latter number scales exponentially in the degrees of freedom f, mild violations of the ETH as quantified by (3.182) are thus indeed negligible, even though the scaling of $\Delta_{\mathrm{mc}}(A)$ has to be taken into account in any concrete example, too.

The derivation of an even stronger bound applying to the original $R(t)$ and for all times t can be found in Ref. [260], yielding

$$R(t)^2 \leq \frac{\Delta_{\mathrm{mc}}^2(A)}{50N_v}\,, \tag{3.186}$$

where

$$\Delta_{\mathrm{mc}}^2(A) := \sum_{\nu\,:\,E_\nu \in I_{\mathcal{E}}} \left(A_{\nu\nu} - \langle A\rangle_{\rho_{\mathrm{mc}}}\right)^2 \tag{3.187}$$

also measures the violation of the ETH, albeit in terms of the squared magnitude of deviations between the diagonal matrix elements and the microcanonical prediction. Even if the strong ETH is violated, the reference system may still often satisfy the weak ETH (cf. Sec. 2.2.2 and particularly Eq. (2.26)), e.g., if the Hamiltonian has translational symmetry. As a consequence of Eq. (2.26), the violation quantified by (3.187) will then be subextensive in N, whereas N_v from (3.9) is expected to grow extensively with N. The bound (3.186) thus entails that $R(t)$ is indeed negligible if the system is sufficiently large.

Similar estimates are also expected to apply to the pertinent remnant terms for larger perturbations, arising from the approximation (3.128) of the fourth-order overlap moment, as well as due to (subleading) corrections to any such approximation. Moreover, these bounds are quite conservative, considering that the individual terms in the sum in (3.151) are oscillating for any fixed t, meaning that cancellations unaccounted for in the above bounds will arise naturally. Notably, we have so far never encountered a particular example where $R(t)$ in (3.149) entailed visible deviations from (3.152).

Long-time limit and prethermalization. In light of $R(t)$ being negligible, we will concentrate exclusively on Eq. (3.160) for the typical expectation values. Taking into account the property (3.162) of $g_\lambda(t)$, we understand that the perturbed systems are predicted to equilibrate (even if the unperturbed system does not), and that the corresponding stationary expectation value approached for long times is $\langle A\rangle_{\tilde{\rho}_\lambda}$ with the state $\tilde{\rho}_\lambda$ from (3.150). This quantity was already at the focus of Deutsch's early studies on thermalization [114, 253] (see also Ref. [257]). According to its definition (3.150), the state $\tilde{\rho}_\lambda$ is based on the diagonal ensemble $\overline{\rho}_0$ (see Sec. 2.2.1) of the reference system H_0, but the level occupations are additionally averaged over the energy scale Γ_v that measures the mixing of unperturbed eigenvectors caused by the perturbation. If the reference system satisfies the (strong) ETH, then already $\overline{\rho}_0$ can be well approximated by the microcanonical density operator ρ_{mc}, and the same will hold for $\tilde{\rho}_\lambda$. Moreover, even if the unperturbed system is, for example, integrable and fulfills just the weak ETH, expectation values with respect to $\tilde{\rho}_\lambda$ will usually still coincide with those obtained from the microcanonical ensemble ρ_{mc} [7, 114, 252, 257], unless the perturbations are very weak or still exhibit conservation laws disregarded in ρ_{mc}. Quite generally, the prediction (3.160) can therefore be simplified further by setting $\langle A\rangle_{\tilde{\rho}_\lambda} = \langle A\rangle_{\rho_{\mathrm{mc}}}$ and thus

$$\langle A\rangle_{\rho_\lambda(t)} = \langle A\rangle_{\rho_{\mathrm{mc}}} + |g_\lambda(t)|^2 \left[\langle A\rangle_{\rho_0(t)} - \langle A\rangle_{\rho_{\mathrm{mc}}}\right]. \tag{3.188}$$

Hence the perturbed dynamics will typically resemble the behavior of the unperturbed system initially, notably even if the latter does not thermalize or even equilibrate, but will eventually approach a thermal state as time progresses. At this point we thus highlight that the prediction (3.188) includes, as a special case, a description of the prethermalization process discussed in Sec. 3.5, as announced towards the end of that subsection (see also Sec. 2.2.3 and Table 3.1).

Adopting the results from Ref. [97] (see also Sec. 2.3.2 and Eq. (2.40) in particular), if applicable, one can even devise an additional prediction for the unperturbed dynamics $\langle A\rangle_{\rho_0(t)}$. In this case, the entire prethermalization scenario is captured by an analytical theory that merely depends on the initial value $\langle A\rangle_{\rho_0(0)}$, the nonthermal stationary value $\langle A\rangle_{\overline{\rho}_0}$, and the temperature (for the prediction of $\langle A\rangle_{\rho_0(t)}$, see Eq. (2.40) and Ref. [97]) as well as the thermal value $\langle A\rangle_{\rho_{\mathrm{mc}}}$, the intrinsic perturbation strength α_v, and (possibly) the perturbation band width Δ_v (for the present prediction of $\langle A\rangle_{\rho_\lambda(t)}$ from $\langle A\rangle_{\rho_0(t)}$).

Fermi's golden rule. Another interesting and important special case entailed in the result (3.160) is a form of Fermi's golden rule [1]. If we choose the initial state to be an eigenstate of the reference system, $\rho(0) = |\nu_\mathrm{i}\rangle\langle\nu_\mathrm{i}|$, and take the projector $A = |\nu_\mathrm{f}\rangle\langle\nu_\mathrm{f}|$ onto another unperturbed eigenstate as our observable, the time-dependent expectation value $\langle A\rangle_{\rho_\lambda(t)}$ is just the probability $p_{\nu_\mathrm{i}\to\nu_\mathrm{f}}(t)$ to observe the transition from $|\nu_\mathrm{i}\rangle$ to $|\nu_\mathrm{f}\rangle$ after time t. Upon substitution into (3.160), we find that

$$p_{\nu_\mathrm{i}\to\nu_\mathrm{f}}(t) = \tilde{u}(E_{\nu_\mathrm{f}} - E_{\nu_\mathrm{i}})\left[1 - |g_\lambda(t)|^2\right] + \delta_{\nu_\mathrm{i}\nu_\mathrm{f}}|g_\lambda(t)|^2\,. \tag{3.189}$$

For sufficiently weak perturbations, in particular, the response profile $g_\lambda(t)$ assumes the exponential form (3.164), such that the transition probability according to (3.189) approaches its predicted equilibrium value $\tilde{u}(E_{\nu_\mathrm{f}} - E_{\nu_\mathrm{i}})$ at the rate $\Gamma = 2\pi\lambda^2\sigma_v^2/\varepsilon$ (see Eq. (3.65)). Observing that $\sigma_v^2 \simeq \mathbb{E}[|V_{\mu\nu}|^2]$ in this case, Eq. (3.189) is the ensemble average of Fermi's golden rule in disguise. It is noteworthy that this relation was derived here by nonperturbative methods, contrary to the "traditional" approach. This is particularly reflected by the fact that (3.189) is expected to hold beyond the traditional (exponential) golden-rule regime when $g_\lambda(t)$ is given by the solutions of (3.173) in general. Yet the analysis from Sec. 3.6.2 suggests that the traditional rule with the exponential $g_\lambda(t)$ will eventually apply at late times.

Insignificance of level fluctuations. As promised in Sec. 3.2 (see below Eq. (3.18)), we finally come back to the question of when the fluctuations of energy levels are negligible for the dynamics so that we can, in particular, approximate $E_n^\lambda - E_m^\lambda$ in (3.7) by $E_n - E_m$ as required in Prerequisite (v) (see also Eq. (3.142)). As recorded in (3.18), the influence of level fluctuations on the relaxation dynamics is expected to be negligible as long as the relaxation time t_R is much larger than the inverse perturbation strength, $t_\mathrm{R} \ll (\lambda\sigma_0)^{-1}$.

According to our main result (3.160), the relaxation time t_R is set by the characteristic time scale of the response profile $g_\lambda(t)$. Focusing on sufficiently weak perturbations first, such that $g_\lambda(t)$ is given by (3.164), we associate t_R with Γ^{-1} from (3.65). The condition $t_\mathrm{R} \ll (\lambda\sigma_0)^{-1}$ is then equivalent to

$$\frac{\sigma_0}{\sigma_v} \ll 2\pi\lambda\frac{\sigma_v}{\varepsilon} \approx \sqrt{N_v} \tag{3.190}$$

with N_v from (3.9), and where we identified $\Gamma_v = \Gamma$ (see above Eq. (3.9) and Eq. (3.33)) in the last step. In view of (3.9), the fluctuations σ_0 of the diagonal matrix elements $V_{\mu\mu}$ may exceed the fluctuations σ_v of the off-diagonal $V_{\mu\nu}$ by many orders of magnitude without violating the condition (3.190) and hence (3.18).

For stronger perturbations, such that $g_\lambda(t)$ assumes the form (3.165), the typical relaxation time is γ^{-1} from (3.69). The condition (3.18) then leads to

$$\frac{\sigma_0}{\sigma_v} \ll \sqrt{\frac{\Delta_v}{\varepsilon}}\,. \tag{3.191}$$

Since the level spacing ε decreases exponentially in the degrees of freedom f, whereas the band width Δ_v (i.e., the energy range of the perturbation) should be roughly independent of f, we eventually reach the same conclusion that the fluctuations of the diagonal matrix elements of V remain insignificant with respect to the relaxation dynamics even if they exceed the off-diagonal fluctuations by many orders of magnitude.

Finally, we remark that we would have arrived at the same conclusions, too, if the rigorous bound (3.19) had been employed in lieu of (3.18). Observing that the time scale t_R decreases gradually as the response profile crosses over from (3.164) to (3.165), we have thus justified the generic validity of Prerequisite (v) *a posteriori*.

Applicability to real systems. The original goal formulated at the beginning of this chapter was to describe the dynamics of an actual physical system with Hamiltonian (3.1). So far, we showed that the vast majority of perturbations within any of the admitted ensembles from Sec. 3.3 result in the relaxation behavior (3.160) or (3.188). It thus remains to be argued that the true perturbation of the system of interest is a typical member of one of those considered ensembles. Unfortunately, it is virtually impossible to prove this for any concrete given system, so we can only collect evidence supporting this conjecture and investigate disqualifying properties (see also Sec. 2.3). In the

subsequent considerations, we take the prerequisites collected in Sec. 3.2 for granted. That is to say, we exclude violations of these requirements, which were explicitly exploited in the derivation of the prediction (3.160), from the following discussion since a system may fail to follow this prediction for obvious reasons in that case.

Eventually, the decisive question is whether or not the key features of the true perturbation with regard to the relaxation behavior are shared by the majority of perturbations in a suitable ensemble. Summarizing the preceding subsections, we established that the relaxation dynamics induced by a large variety of perturbations is essentially determined by the perturbation profile (3.10), i.e., the coarse-grained squared magnitude of the perturbation matrix elements $V_{\mu\nu}$ in the unperturbed basis. In fact, the discussion of the response profile from Sec. 3.6.2 suggests that an even stronger reduction to just the parameters α_v from (3.12) and Δ_v from (3.13) is legitimate. At the same time, minor fluctuations of the $V_{\mu\nu}$ around their "true" values will not entail any noticeable deviations since, from a mathematical point of view, the propagator $\mathrm{e}^{-\mathrm{i}H_\lambda t}$ is continuous in these variables, and from a physical point of view, no experiment (real or numerical) would be reproducible otherwise.

Thereby, the large freedom to choose the precise distribution (3.24) of the perturbation operator leaves room to tailor an ensemble in such a way that the true perturbation of interest is realized with reasonably high probability. Moreover, the structure of the considered ensembles was designed to emulate common features of real perturbations such as bandedness, sparsity, etc. (see Sec. 3.3), reinforcing that real perturbations can be faithfully modeled or embedded in such an ensemble. Yet the fact that the true perturbation is sampled with reasonable probability within a single ensemble does not necessarily imply that it leads to the typical behavior of that ensemble when it comes to the relaxation behavior of observable expectation values. Indeed, if this behavior depended on some subtle details of the perturbation (for instance, the value of a single matrix element $V_{\mu\nu}$ as an extreme example), the observed dynamics could still deviate from the typical behavior of the majority in case that this subtle feature happens to assume a rare value. Fortunately, again, the predicted behavior (3.160) is remarkably robust and—as demonstrated in Sec. 3.6.2 in particular—independent of any fine-tuned perturbation characteristics.

Nevertheless, there are certain possible features of physical perturbations that are not explicitly accounted for. Such features are, for example, special operators commuting with the perturbation (notably the observable A [238, 259]) or quenches from systems with few symmetries (such as nonintegrable ones) to systems with more symmetries (such as integrable ones). Interestingly, it may still be possible to model such setups within the present approach if the considered observable and initial state do not explicitly "probe" these features. In any case, those examples constitute rather special situations.

In contrast, a potentially severe shortcoming is the only rudimentary modeling of the local and few-body character of interactions in common physical systems. While the considered ensembles explicitly allow for a banded and sparse matrix structure as it is often found as a result of local and few-body perturbations (e.g., if the reference system is noninteracting, see also below Eq. (3.10)), there is no geometry or notion of physical entities ("particles" or "bodies") incorporated, meaning that sparsity and bandedness arise randomly in the considered perturbations and not in the way implied by the physical structure. It is known that such issues may render physical systems atypical with respect to certain random matrix ensembles [87, 107, 237], i.e., certain aspects of the true system may therefore indeed behave atypically compared to a selected ensemble, even though all considered characteristics (that is, the perturbation profile (3.10)) formally agree. Moreover, the outcome of our theoretical prediction (3.160) that essentially all experimentally relevant observables behave similarly can certainly not be upheld if the initial state exhibits macroscopic inhomogeneities. For instance, Lieb-Robinson bounds [7, 105] (see also Sec. 2.2.1) limit the speed at which a local perturbation can spread across a locally interacting lattice system, implying that observables probing regions far away from the perturbation will notice the change later than those monitoring the vicinity of the perturbation, and thus the two will relax on different time scales. On the other hand, these issues are again not expected to matter if the considered setup does not explicitly probe them. In particular, the *ad hoc* modeling of locality and sparsity should therefore still be satisfactory if the initial state (or the observable) are sufficiently homogeneous on a macroscopic scale.

This raises an interesting point about admissible initial states and observables in general. The only explicit requirement on the initial state for the derivation was a well-defined macroscopic energy (cf. Prerequisite (i)). Whereas the prediction (3.160) and especially the response profile (3.146) do not exhibit an explicit dependence on any more specific properties of the state, there is an implicit dependence mediated by the density of states (or rather, the mean level spacing ε), which sets a basic energy and thus also time scale for the reference system as well as a scale to gauge the perturbation strength. Since the mean level spacing may change with the state's energy (see also Fig. 3.1), initial states pertaining to different energy windows will generally also lead to different relaxation characteristics in (3.160). On the other hand, besides a finite spectral range and resolution (cf. Eqs. (2.18) and (2.19)), there are no obvious restrictions as far as the observable A is concerned. However, we remark that there will always be special combinations of initial state and observable correlated in such a way that the resulting dynamics is atypical with respect to a given perturbation ensemble. Put differently, every combination of an observable A and an initial state $\rho(0)$ entails a set of atypical perturbations (which could in principle contain the true perturbation of interest), and furthermore these atypical perturbations will generally differ for different A and $\rho(0)$. *A priori*, however, it is unfortunately not immediately obvious whether a given combination of A, $\rho(0)$, and V is correlated such that it behaves atypically compared to the remaining members of a certain ensemble.

In summary, there is compelling evidence to believe that real perturbations can be modeled in terms of the ensembles considered here, unless there are specific reasons to the contrary. These reasons may be explicitly comprehensible in a given setup (macroscopically inhomogeneous initial state, quench from general to special case, ...) or they may be rooted more subtly in correlations between the Hamiltonian, the observable, and/or the initial state. Consequently, it is imperative to verify the prediction (3.160) in explicit numerical or experimental examples, and this will be the subject of the next subsection.

3.7 Examples

The purpose of our present theory was to explain the perturbed relaxation of isolated many-body quantum systems, and the main result (3.160) or (3.188) provides a prediction for this relaxation in the form of a relation between the time-dependent expectation of the perturbed and unperturbed systems. A natural next step is thus to test this prediction against numerical and experimental data for concrete example systems, for which we partly employ numerical studies on our own and partly adopt pertinent results from the literature. The investigated systems will usually still be (effectively) small from a macroscopic point of view since it is clearly impossible to simulate exactly (numerically or experimentally) a quantum system of, say, 10^{23} degrees of freedom. Yet the analytical result (3.160) will turn out to reproduce the prevailing features of the dynamics remarkably well. Looking ahead, the large variety of models and setups which are amenable to and described well by the theory underpins its broad applicability and the considerable generality of the prediction (3.160).

General remarks. In the spirit of the setup and applications described in Sec. 3.1, we generally take the reference dynamics $\langle A\rangle_{\rho_0(t)}$ of the unperturbed system H_0 as given, i.e., as an input to the theory. The essential second input then is the perturbation profile $\sigma_v^2(E)$ from (3.10) or at least the intrinsic perturbation strength α_v from (3.12) and, ideally, the band width Δ_v from (3.13). In numerical examples, these quantities can be determined in principle by exact diagonalization or approximately identified, for example, by means of typicality methods and imaginary time evolution. From these characteristics, the response profile $g_\lambda(t)$ can be computed numerically for general $\sigma_v^2(E)$ or estimated by means of the analytical approximations (3.164), (3.165), and (3.166), depending on the information available.

Similarly, if detailed numerical information is available, for example, from exact diagonalization, also the state $\tilde{\rho}_\lambda$ from (3.150) can be calculated directly from the initial occupations $\rho_{\mu\mu}(0)$ of the unperturbed levels and the function $\tilde{u}(E)$ from (3.116), which is derived from the overlap distribution (3.32) and thus the perturbation profile, too. However, employing the result (3.188),

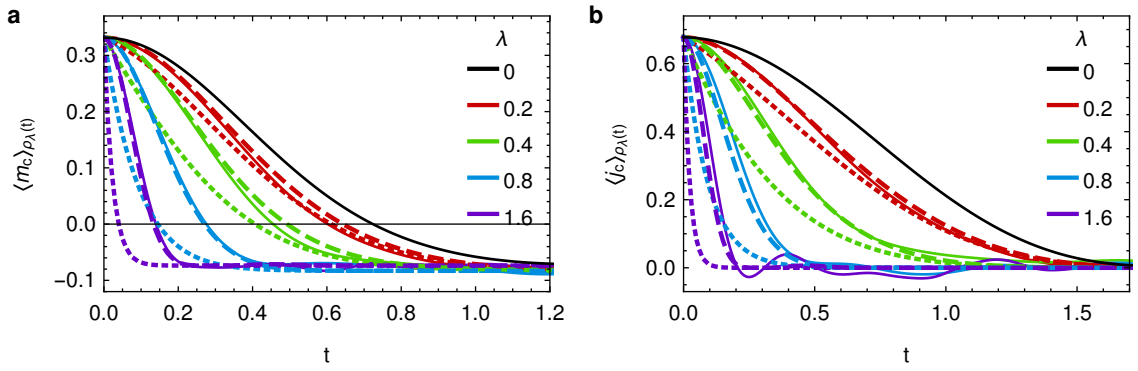


Figure 3.10: Time evolution in the two-dimensional spin-$\frac{1}{2}$ model $H_\lambda = H_0 + \lambda V$ with H_0 from (3.20a) and V from (3.20b) for various perturbation strengths. Solid: Numerical simulation results obtained by exact diagonalization. Dashed: Analytical prediction (3.160), employing the numerical solution for $\lambda = 0$ (solid black curve) for the reference dynamics $\langle A\rangle_{\rho_0(t)}$, the second-order continued-fraction result (3.166) for the response profile $g_\lambda(t)$ with the empirically determined intrinsic perturbation strength $\alpha_v = 2.64$ and band width $\Delta_v = 7.32$ (cf. Figs. 3.1 and 3.2), and $\langle A\rangle_{\bar{\rho}_\lambda}$ as indicated below. Dotted: Analytical prediction (3.160) adopting the weak-perturbation and large-time asymptotics (3.164) instead (same parameters). **a.** Time-dependent expectation values of the magnetization correlation m_c from (3.192), starting from the state (3.194). The long-time limiting values are computed directly via Eq. (3.150), yielding $\langle m_\mathrm{c}\rangle_{\bar{\rho}_\lambda} = -0.0896, -0.0820, -0.0830, -0.0738$ for $\lambda = 0.2, 0.4, 0.8, 1.6$, respectively. **b.** Time-dependent expectation values of the spin-flip correlation j_c from (3.195), starting from the dynamical-typicality state (2.39) with $A = j_\mathrm{c}$, $\kappa = 2$, and Π projecting onto the central 2048 energy levels. The long-time limiting values are the thermal expectation values, $\langle j_\mathrm{c}\rangle_{\bar{\rho}_\lambda} = \langle j_\mathrm{c}\rangle_{\rho_\mathrm{mc}} = 0$.

the thermal expectation value $\langle A\rangle_{\rho_\mathrm{mc}}$ is often obtained even simpler by symmetry arguments or other methods.

Unfortunately, such detailed information about a certain system is not always available, and it may sometimes not even be possible to decide unequivocally whether all of the prerequisites from Sec. 3.2 are satisfied in the setup in question, hence this will be tacitly assumed in these cases. Moreover, the perturbation profile $\sigma_v^2(E)$ or the parameters α_v and Δ_v are often not assessed or quoted along with time series of observable expectation values in the literature, not least because determining these quantities quickly becomes unfeasible for larger system sizes. In this case, Eqs. (3.160) and (3.188) still amount to powerful predictions of the relaxation behavior since one or two fit parameters are sufficient to forecast the dynamics for a broad range of coupling strengths λ and times t. Depending on the specific setting and the information on hand, the various analytical approximations for $g_\lambda(t)$ provide different degrees of control and precision. If nothing or only the rough strength of the perturbation is known and presumed to be weak, then $g_\lambda(t)$ may be approximated by (3.164) with α_v as the only free parameter according to (3.65). If the perturbation strength and band width are accessible or the regime of stronger perturbations is reached, the approximation (3.166) can be used with Δ_v as an additional parameter (see also Eq. (3.79)). For even stronger perturbations, one might also adopt (3.165) instead.

Two-dimensional spin-1/2 lattice. As our first example, we take the spin-$\frac{1}{2}$ system defined on a 4×4 lattice as introduced in Sec. 3.2. The unperturbed Hamiltonian (3.20a) consists of isotropic Heisenberg interactions between nearest neighbors and the perturbation (3.20b) comprises spin-flip terms (with respect to the z direction) between next-nearest neighbors. The properties of this system were analyzed in detail in Figs. 3.1 and 3.2 using exact diagonalization. In particular, we found that $\varepsilon = 1.90 \times 10^{-3}$, $\sigma_v^2 = 5.02 \times 10^{-3}$, and $\Delta_v = 7.32$, hence $\alpha_v = 2.64$. For the transition between the limiting forms (3.164) and (3.165) of $g_\lambda(t)$, the corresponding crossover coupling (3.73) is thus $\lambda_\mathrm{c} \approx 0.75$, and the crossover time (3.181) is $t_\mathrm{c} \approx 0.43$.

We consider two different combinations of observables and initial states, both focusing on the central spins at sites $(2, 2)$ and $(3, 3)$, which could thus be regarded as the "system" with the remaining spins around them serving as a "bath." The first observable measures the magnetization

correlation in the z direction between these two sites,

$$m_{\mathrm{c}} := \sigma^z_{2,2}\,\sigma^z_{3,3}\,. \tag{3.192}$$

To achieve an appreciable nonequilibrium initial value, we prepare the two "system spins" in the "up" state, whereas the bath is supposed to be in equilibrium, modeled by a random (Haar-distributed) pure state on the corresponding subspace. Finally, to satisfy Prerequisite (i) of a well-defined macroscopic energy, we filter the so-obtained state by means of a Gaussian projector

$$\mathit{\Pi}_{\mathcal{E},\Delta_\mathcal{E}} := \frac{\mathrm{e}^{-(H_0-\mathcal{E})^2/2\Delta_\mathcal{E}^2}}{\sqrt{2\pi\Delta_\mathcal{E}^2}} \tag{3.193}$$

of mean energy $\mathcal{E} = 0$ and standard deviation $\Delta_\mathcal{E} = 2$ [133, 201, 202]. Roughly speaking, this simulates a macroscopic measurement of the system energy that yielded $\mathcal{E} = 0$. Altogether, the initial state is thus given by $\rho(0) = |\psi\rangle\langle\psi|$ with

$$|\psi\rangle \propto \mathit{\Pi}_{\mathcal{E},\Delta_\mathcal{E}}\,\sigma^+_{2,2}\,\sigma^+_{3,3}|\phi\rangle\,, \tag{3.194}$$

where $\sigma^+_{i,j} := \sigma^x_{i,j} + \mathrm{i}\sigma^y_{i,j}$ and $|\phi\rangle$ is a Haar-distributed random vector on the full $\binom{16}{8}$-dimensional Hilbert space of the zero-magnetization subsector (see also Sec. 3.2). The resulting dynamics is shown for various coupling strengths λ by the solid lines in Fig. 3.10a.

The second observable we consider is the spin-flip correlation between the two "system" sites,

$$j_{\mathrm{c}} := \sigma^x_{2,2}\,\sigma^y_{3,3} - \sigma^y_{2,2}\,\sigma^x_{3,3} = \tfrac{1}{2\mathrm{i}}\left(\sigma^-_{2,2}\,\sigma^+_{3,3} - \sigma^+_{2,2}\,\sigma^-_{3,3}\right), \tag{3.195}$$

with $\sigma^\pm_{i,j} := \sigma^x_{i,j} \pm \mathrm{i}\sigma^y_{i,j}$. The initial state $\rho(0) = |\psi\rangle\langle\psi|$ is chosen in the spirit of the dynamical-typicality construction according to (2.39) with $\mathit{\Pi}$ projecting onto the central 2048 states of the zero-magnetization sector (ensuring Prerequisite (i)), and $\kappa = 2$. For several choices of λ, this leads to the dynamics depicted by the solid lines in Fig. 3.10b.

Since all parameters entering the analytical prediction (3.160) with the second-order continued-fraction expression (3.166) for the response profile $g_\lambda(t)$ are explicitly available, we can directly compare that prediction to the numerical simulation without any free parameters, resulting in the dashed curves in Fig. 3.10. For the reference dynamics $\langle A\rangle_{\rho_0(t)}$, we thereby adopted the corresponding numerical results from the solid black curves in the respective panels. The long-time limit $\langle A\rangle_{\tilde{\rho}_\lambda}$ is explicitly computed via (3.150) for $A = m_{\mathrm{c}}$ in Fig. 3.10a, exploiting the known occupations $\rho_{\mu\mu}(0)$ and calculating the overlap distribution $u(E)$ numerically as a Chebyshev series (cf. Sec. 3.4.2), assuming an exponential perturbation profile (3.21) with $\sigma_v^2 = 5.02 \times 10^{-3}$ and $\Delta_v = 7.32$ (cf. Fig. 3.2c). The explicit values are stated in the caption of Fig. 3.10. Note that the associated microcanonical expectation value $\langle m_{\mathrm{c}}\rangle_{\rho_{\mathrm{mc}}} = -0.0805$ as obtained from the energy shell of the central 60 % of states (see below Eq. (3.20b)) could be employed as well without significant loss of accuracy. This route is adopted for $A = j_{\mathrm{c}}$ in Fig. 3.10b, where we effectively use the prediction (3.188) instead, taking $\langle A\rangle_{\tilde{\rho}_\lambda} = \langle A\rangle_{\rho_{\mathrm{mc}}} = 0$ (by symmetry). The agreement between theory and numerics is very good in both settings and for all values of λ, i.e., across the transition from weak to stronger perturbations with the crossover around $\lambda_{\mathrm{c}} \approx 0.75$.

For completeness, we also show the analytical prediction obtained by utilizing the approximation (3.164), expected to apply for weak perturbations $\lambda \ll \lambda_{\mathrm{c}} \approx 0.75$ or late times $t \gg t_{\mathrm{c}} \approx 0.43$. In these limits, we likewise find good agreement between the prediction and the numerical results. At the same time, it becomes apparent that this approximation cannot faithfully describe stronger perturbations at short times.

Bosonic Hubbard chain. As a second example, we consider the bosonic Hubbard chain as studied by Flesch and co-workers in Ref. [261], defined by the Hamiltonian

$$H = -J\sum_{i=1}^{L}\left(b^\dagger_{i+1}b_i + b^\dagger_i b_{i+1}\right) + \frac{U}{2}\sum_{i=1}^{L} n_i(n_i-1) \tag{3.196}$$

with periodic boundary conditions. Here $b^\dagger_i$ and b_i are the bosonic creation and annihilation operators on site i, and $n_i := b^\dagger_i b_i$. In Ref. [261], time-dependent expectation values of the odd-site

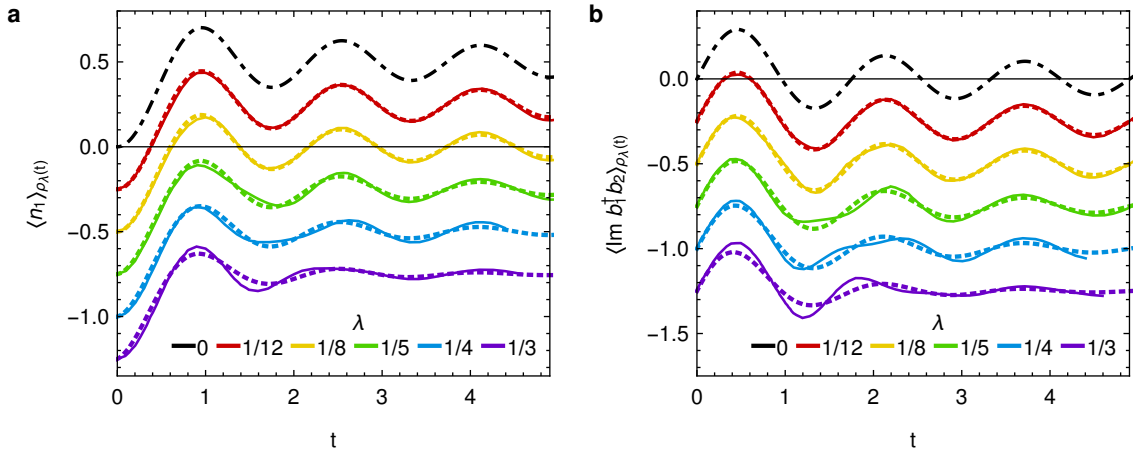


Figure 3.11: Time evolution in the bosonic Hubbard chain with Hamiltonian H from (3.196) for $J = 1$, various coupling strengths $\lambda = 1/U$ and **a.** the odd single-site occupation n_1 or **b.** the nearest-neighbor hopping correlation $\mathrm{Im}\, b_1^\dagger b_2$ (see below Eq. (3.196)). The initial state $\rho(0)$ consists of singly occupied even sites and empty odd sites. Dash-dotted: Analytical solutions $\langle n_1\rangle_{\rho_0(t)} = [1 - J_0(4t)]/2$ and $\langle \mathrm{Im}\, b_1^\dagger b_2\rangle_{\rho_0(t)} = J_1(4t)/2$ for the reference dynamics with $\lambda = 0$ in an infinite chain. Solid: Numerical tDMRG results from Ref. [261] for $L = 32$, vertically shifted in steps of -0.25 for better visibility. Dotted: Analytical prediction (3.188), employing the analytic result (dash-dotted black curve) for the reference dynamics $\langle A\rangle_{\rho_0(t)}$, the weak-perturbation response profile (3.164) with $\alpha_v = 0.79$, the thermal expectation values $\langle n_1\rangle_{\rho_{\mathrm{mc}}} = \frac{1}{2}$ and $\langle \mathrm{Im}\, b_1^\dagger b_2\rangle_{\rho_{\mathrm{mc}}} = 0$, and the same vertical shifts as for the numerics.

occupation n_1 and of the nearest-neighbor hopping correlation $\mathrm{Im}\, b_1^\dagger b_2 = (b_1^\dagger b_2 - b_2^\dagger b_1)/2\mathrm{i}$ were computed for a chain of length $L = 32$ with $J = 1$ and various values of the interaction strength U using time-dependent density-matrix renormalization group (tDMRG) methods. The considered initial state $\rho(0)$ has one particle on every even site and empty odd sites. The obtained results are reproduced by the solid lines in Fig. 3.11.

To compare to the analytic prediction (3.188), we choose the limit $U \to \infty$ of large interactions as our reference case. In this limit, the Hamiltonian (3.196) becomes equivalent to a spin-$\frac{1}{2}$ XX model, which thus comprises our reference Hamiltonian H_0. The time-dependent expectation values are known analytically in this case [261] and are given by $\langle n_1\rangle_{\rho_0(t)} = [1 - J_0(4t)]/2$ and $\langle \mathrm{Im}\, b_1^\dagger b_2\rangle_{\rho_0(t)} = J_1(4t)/2$, respectively, where $J_\alpha(x)$ are Bessel functions of the first kind. Corrections for large but finite U and the given initial state can be obtained by means of an expansion in the inverse interaction strength U^{-1} [262]. The leading-order contribution $U^{-1}V$ consists of nearest and next-nearest neighbor interactions as well as three-spin terms, and serves as the perturbation in our setup (3.1). The coupling strength is thus $\lambda = U^{-1}$.

The values of the parameters α_v and Δ_v are not available from Ref. [261]. Since the coupling values of the data presented in Fig. 3.11 turn out to lie well inside the weak-perturbation regime[2], we choose the expression (3.164) for the response profile, so α_v is the only fit parameter required. With $\alpha_v = 0.79$ and the microcanonical expectation values $\langle n_1\rangle_{\rho_{\mathrm{mc}}} = \frac{1}{2}$ and $\langle \mathrm{Im}\, b_1^\dagger b_2\rangle_{\rho_{\mathrm{mc}}} = 0$, the prediction (3.188) then leads to the dotted curves in Fig. 3.11. The agreement with the numerical data is again very pleasing.

Cold-atom experiments. A great and quite recent technological advancement is the ability to emulate lattice systems of interacting bosons or fermions experimentally by confining ultracold atoms using optical traps [26–28, 30] (see also the discussion at the end of Sec. 2.1). In particular, an effective bosonic Hubbard chain (3.196) was realized experimentally by Trotzky et al. [32], deliberately aiming at reproducing the setup from Ref. [261] (see Fig. 3.11) as closely as possible.

In Fig. 3.12, we compare their experimentally measured expectation values of the single-site occupation n_1 with our theory (3.188), employing the same parameter values as in Fig. 3.11, i.e., there are no additional fit parameters. The analytical prediction describes the experimental data very

[2] The estimated crossover coupling (3.73) is $\lambda_c \approx 1.5$ when using the more refined response profile (3.166) and fitting both α_v and Δ_v.

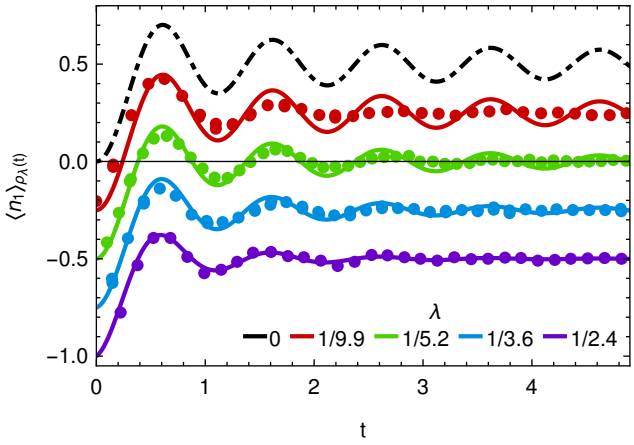


Figure 3.12: Time evolution of the single-site occupation in an optical lattice of ultracold, repulsively interacting Rb atoms. Dots: Experimental data adopted from Fig. 2 of Ref. [32]. The experimental setup emulates the bosonic Hubbard model (3.196) with the same initial conditions as in Fig. 3.11. Dash-dotted: Analytical solution $\langle n_1 \rangle_{\rho_0(t)} = [1 - J_0(4t)]/2$ for the reference dynamics with $\lambda = 0$ in an infinite bosonic Hubbard chain. Solid: Analytical prediction (3.188), employing the analytical result (dash-dotted black curve) for the reference dynamics $\langle A \rangle_{\rho_0(t)}$, the weak-perturbation response profile (3.164) with $\alpha_v = 0.79$, and the thermal expectation value $\langle n_1 \rangle_{\rho_{\mathrm{mc}}} = \frac{1}{2}$, i.e., the same parameters as in Fig. 3.11. Like there, the perturbed data are shifted vertically in steps of -0.25 for better visibility.

well, especially for the larger values of λ. In view of the numerical simulations shown in Fig. 3.11 and their good agreement with the theory, the model (3.196) presumably misses some relevant features of the cold-atom experiment for very small λ. Indeed, the authors of Ref. [32] speculate about possible origins of these deviations themselves, identifying tunneling between different copies of the chain as well as nonadiabatic heating as the primary sources of deviations. Since the regime of small λ corresponds to small values of the tunneling constant J in (3.196) in the way parameters are controlled in the experiment [32], the relative influence of those effects is arguably strongest there.

Spin-1/2 XXZ chain. As another example, we turn to the antiferromagnetic spin-$\frac{1}{2}$ XXZ chain, whose Hamiltonian reads

$$H = J \sum_i \left(\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \Delta \, \sigma_i^z \sigma_{i+1}^z \right), \tag{3.197}$$

where σ_i^α are Pauli matrices acting on site i as before. Note that this system is integrable for all values of the parameters $J > 0$ and $\Delta \geq 0$ [162]. Furthermore, it exhibits a phase transition from a gapless Luttinger liquid for $\Delta \leq 1$ to a gapped, Ising-ordered antiferromagnetic phase for $\Delta > 1$.

In Fig. 3.13a, we compare our prediction (3.160) to numerically obtained results by Barmettler et al. [263], who studied the dynamics of the system (3.197) for $J = 1/4$ and various values of the anisotropy parameter Δ in a formally infinite chain using a time-evolving block-decimation (TEBD) algorithm, which effectively keeps track of up to 7000 dominant states. The system was prepared in a Néel state of alternating up and down spins, i.e., $\rho(0) = |\psi\rangle\langle\psi|$ with $|\psi\rangle = |\uparrow\downarrow\uparrow\downarrow\cdots\rangle$, and the monitored observable was the staggered magnetization

$$M_{\mathrm{s}}^z := \frac{1}{2L} \sum_i (-1)^i \sigma_i^z \, . \tag{3.198}$$

The most natural choice in view of the available data is to take the case $\Delta = 0$ as the reference system H_0 and identify the coupling strength λ with the anisotropy parameter Δ. The unperturbed system is thus an XX model, effectively similar to the example from Fig. 3.11, but the perturbation, which couples neighboring spins in the z direction, is different. Notably, the unperturbed dynamics is again known exactly, $\langle M_{\mathrm{s}}^z \rangle_{\rho_0(t)} = J_0(2t)/2$ [263]. The long-time limit is $\langle A \rangle_{\bar{\rho}_\lambda} = 0$ by symmetry and is thus equal to the thermal expectation value $\langle A \rangle_{\rho_{\mathrm{mc}}}$ despite the system's integrability. We

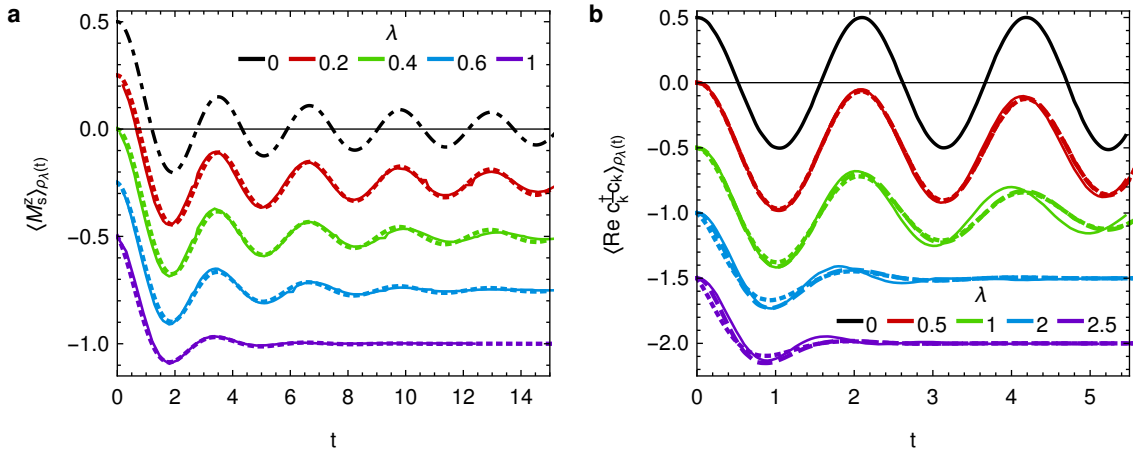


Figure 3.13: a. Time evolution of the staggered magnetization (3.198) in the spin-$\frac{1}{2}$ chain (3.197), starting from a Néel initial state, for various values of the anisotropy parameter $\lambda = \Delta$. Solid: Numerical infinite-size TEBD results adopted from Figs. 1 and 2 of Ref. [263] and vertically shifted in steps of -0.25. Dash-dotted: Analytical solution $\langle M_s^z \rangle_{\rho_0(t)} = J_0(2t)/2$ for the unperturbed dynamics ($\lambda = 0$). Dotted: Analytical prediction (3.160), employing the analytically known reference dynamics (dash-dotted black curve), the weak-perturbation response profile (3.164) with $\alpha_v = 0.073$, and $\langle A \rangle_{\bar{\rho}_\lambda} = 0$ by symmetry. **b.** Time evolution of the correlation (3.200) between conjugated momentum modes k and $\bar{k}$ in the fermionic Hubbard model (3.199) on a Bethe lattice of infinite coordination number, starting from a state with one particle per site with alternating spins between nearest neighbors, for various interaction strengths $\lambda = U$. Solid: Numerical DMFT results, adopted from Fig. 3 of Ref. [264] and vertically shifted in steps of -0.5. Dashed: Analytical prediction (3.188) using the unperturbed result ($\lambda = 0$, black solid curve) for $\langle \mathrm{Re}\, c_{\bar{k}}^\dagger c_k \rangle_{\rho_0(t)}$, $\langle \mathrm{Re}\, c_{\bar{k}}^\dagger c_k \rangle_{\rho_{\mathrm{mc}}} = 0$, and the two-parameter approximation (3.166) for the response profile with fitted $\alpha_v = 0.045$ and $\Delta_v = 4.9$. Dotted: Analytical prediction employing instead the weak-perturbation approximation (3.164) for the response profile with the same parameters.

content ourselves with the weak-perturbation asymptotics (3.164) again and treat the unknown ratio α_v as the only fit parameter, yielding $\alpha_v = 0.073$.

The resulting theoretical predictions (dotted lines in Fig. 3.13a) are in excellent agreement with the numerical results (solid lines) for all values of λ up to the critical point at $\lambda = \Delta = 1$.

Fermionic Hubbard model on Bethe lattice. As our next example, we consider the fermionic Hubbard model on a Bethe lattice of infinite coordination number in the form studied by Balzer et al. in Ref. [264]. Denoting pairs of connected sites i and j by $i \leftrightarrow j$ and introducing the creation and annihilation operators $c_{i\sigma}^\dagger$ and $c_{i\sigma}$, respectively, for a fermion of spin $\sigma \in \{\uparrow, \downarrow\}$ on site i, the Hamiltonian takes the form

$$H = -\sum_{i \leftrightarrow j, \sigma} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i \left(n_{i\uparrow} - \tfrac{1}{2}\right)\left(n_{i\downarrow} - \tfrac{1}{2}\right) \tag{3.199}$$

with $n_{i\sigma} := c_{i\sigma}^\dagger c_{i\sigma}$. The first term here describes hopping between neighboring sites and the second term is a repulsive ($U > 0$) on-site interaction. The system is integrable for $U = 0$ and nonintegrable for $0 < U < \infty$ [264]. We select the interaction as the perturbation in our setup (3.1), i.e., $\lambda = U$. The unperturbed Hamiltonian H_0 is diagonalized essentially by Fourier transformation, introducing (in a slight abuse of notation) creation and annihilation operators $c_k^\dagger$ and c_k, respectively, of momentum modes k. Choosing an initial state with one particle per lattice site and alternating spins between nearest neighbors, Balzer et al. then calculated time-dependent expectation values for various values of λ using dynamical mean-field theory (DMFT), investigating, in particular, the correlation

$$\mathrm{Re}\, c_{\bar{k}}^\dagger c_k = \tfrac{1}{2}\left(c_{\bar{k}}^\dagger c_k + c_k^\dagger c_{\bar{k}}\right) \tag{3.200}$$

between conjugated momentum modes k and $\bar{k}$ whose single-particle energies agree up to a sign factor [264]. The obtained dynamics is shown by the solid lines in Fig. 3.13b.

To compare to the prediction (3.188), we adopt $\langle \mathrm{Re}\, c_k^\dagger c_k \rangle_{\rho_{\mathrm{mc}}} = 0$ and the expression (3.166) for the response profile. Since the explicit values of α_v and Δ_v are unfortunately not known, we estimate them by a fit and find $\alpha_v = 0.045$ and $\Delta_v = 4.9$. This yields the dashed curves in Fig. 3.13b for the theoretical prediction, which again agree well with the numerics

We note that the estimates for α_v and Δ_v imply a crossover coupling (3.73) of $\lambda_{\mathrm{c}} \approx 4.7$ and a crossover time (3.181) of $t_{\mathrm{c}} \approx 0.64$, hence the data displayed in Fig. 3.13b all satisfy $\lambda < \lambda_{\mathrm{c}}$. Therefore, we also show the corresponding weak-perturbation approximation as dotted lines in the figure, using (3.164) for the response profile. This reveals that the improved two-parameter approximation (3.166) performs better for larger λ values and times $t \lesssim t_{\mathrm{c}}$, indicating that the perturbation matrix $V_{\mu\nu}$ could indeed exhibit a decisive banded structure.

As a last remark on this example, we point out that the reference dynamics here exhibits persistent oscillations, i.e., the unperturbed system does not equilibrate (cf. Sec. 2.2). Hence the example demonstrates that such systems are equally covered by the present theoretical approach, as they should since there was no formal restriction necessary that would exclude them.

Two coupled spin-1/2 chains. For our last example, we consider an isolated system consisting of two antiferromagnetic Heisenberg spin-$\frac{1}{2}$ chains of length L with periodic boundary conditions, such that the respective Hamiltonians $H^{(1)}$ and $H^{(2)}$ are given by

$$H^{(s)} := \sum_i \boldsymbol{\sigma}_{s,i} \cdot \boldsymbol{\sigma}_{s,i+1} \,. \tag{3.201}$$

As usual, $\boldsymbol{\sigma}_{s,i} = (\sigma^x_{s,i}, \sigma^y_{s,i}, \sigma^z_{s,i})$ is a vector of Pauli matrices acting on the ith site of chain $s \in \{1, 2\}$. In the unperturbed system, the chains are completely isolated from each other such that

$$H_0 := H^{(1)} + H^{(2)} \,. \tag{3.202}$$

By means of the perturbation V, we now bring them into contact, coupling the ith sites of both chains via Heisenberg terms, too, such that

$$V := \sum_i \boldsymbol{\sigma}_{1,i} \cdot \boldsymbol{\sigma}_{2,i} \,. \tag{3.203}$$

The perturbed system $H_\lambda = H_0 + \lambda V$ thus constitutes a so-called spin ladder. We remark that this setup implements the first two example scenarios from Table 3.1 since the perturbation couples two isolated subsystems and breaks the integrability of the simple one-dimensional chains.

We focus on the zero-magnetization sector in the (arbitrarily chosen) z direction. The initial state $\rho(0) = |\psi\rangle\langle\psi|$ is based on an infinite-temperature state modeled by a state $|\phi\rangle$ drawn uniformly at random from that sector, but we subsequently align two neighboring spins in the first chain in the "up" state and the corresponding spins in the second chain in the "down" state. Consequently,

$$|\psi\rangle \propto \sigma^+_{1,1}\sigma^+_{1,2}\sigma^-_{2,1}\sigma^-_{2,2}|\phi\rangle \tag{3.204}$$

with $\sigma^\pm_{s,i} = \sigma^x_{s,i} \pm \mathrm{i}\sigma^y_{s,i}$. Since this state already exhibits a relatively narrow energy distribution for the system sizes employed in the following, there is no need for additional filtering (i.e., applying a reasonably narrow filter does not have a significant effect on the dynamics).

Using a second-order Suzuki-Trotter decomposition with time step $\Delta t = 0.01$, we simulate the time evolution for chains of length $L = 12$, amounting to a total of $2L = 24$ lattice sites. The solid lines in Fig. 3.14a show the resulting dynamics of the single-site magnetization $\sigma^z_{1,1}$ for the uncoupled chains ($\lambda = 0$, black) as well as for various coupling strengths $\lambda > 0$ (color-coded as indicated in the legend of subfigure b). Since the magnetization is conserved within every chain individually for $\lambda = 0$, the equilibrium expectation value is $\langle \sigma^z_{1,1} \rangle_{\bar{\rho}_0} = 2/L = 0.17$ for the unperturbed dynamics. Upon coupling the two chains, conservation of the subchain magnetizations is broken, so $\langle \sigma^z_{1,1} \rangle_{\rho_{\mathrm{mc}}} = 0$. According to Sec. 3.5 (see also Sec. 2.2.3), we thus expect prethermalization for sufficiently weak perturbations, and this is indeed observed as highlighted in particular in the log-linear plot of the inset.

To compare to the theory (3.188), we display the predictions obtained by adopting both the second-order continued-fraction approximation (3.166) (dashed lines) and the weak-perturbation

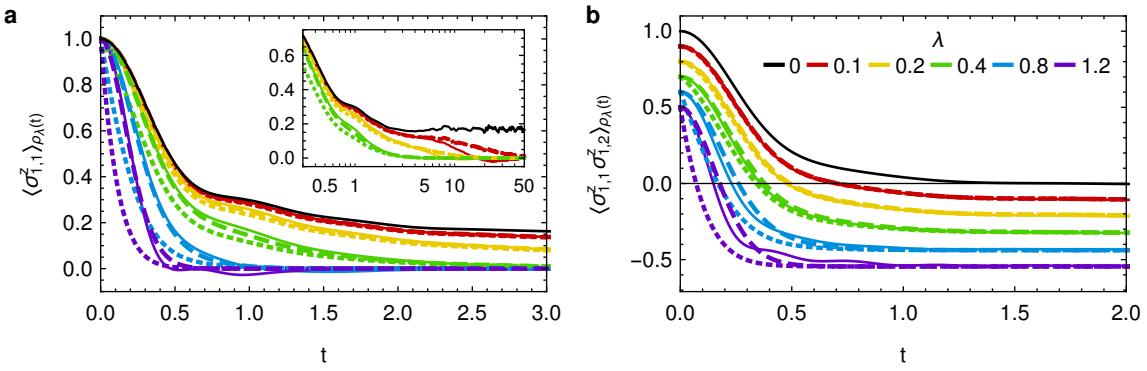


Figure 3.14: Time evolution of two coupled spin-$\frac{1}{2}$ chains with $H_\lambda = H_0 + \lambda V$, where H_0 and V are given by (3.202) and (3.203), respectively, for various coupling strengths, starting from the state $\rho(0) = |\psi\rangle\langle\psi|$ with $|\psi\rangle$ from (3.204). Solid: Numerical simulation results obtained by Suzuki-Trotter propagation. Dashed: Analytical prediction (3.160), employing the numerical solution for $\lambda = 0$ (solid black curve) for the reference dynamics $\langle A\rangle_{\rho_0(t)}$, the second-order continued-fraction result (3.166) for the response profile $g_\lambda(t)$ with the intrinsic perturbation strength $\alpha_v = 0.95$ and band width $\Delta_v = 5.8$, and $\langle A\rangle_{\bar\rho_\lambda}$ as indicated below. Dotted: Analytical prediction (3.160) adopting the weak-perturbation and large-time asymptotics (3.164) instead (same parameters). **a.** Time-dependent expectation values of the single-site magnetization $A = \sigma^z_{1,1}$, adopting $\langle\sigma^z_{1,1}\rangle_{\bar\rho_\lambda} = \langle\sigma^z_{1,1}\rangle_{\rho_{\rm mc}} = 0$. **b.** Time-dependent expectation values of the magnetization correlation $A = \sigma^z_{1,1}\sigma^z_{1,2}$, utilizing the long-time limiting values $\langle\sigma^z_{1,1}\sigma^z_{1,2}\rangle_{\bar\rho_\lambda} = -0.014, -0.019, -0.026, -0.038, -0.045$, respectively, for $\lambda = 0.1, 0.2, 0.4, 0.8, 1.2$. Curves for $\lambda > 0$ are shifted in steps of -0.1 for better visibility.

asymptotics (3.164) for the response profile $g_\lambda(t)$. Since the system size is beyond our numerical capabilities for exact diagonalization, we treat the intrinsic strength α_v and the band width Δ_v of the perturbation as fit parameters, yielding $\alpha_v = 0.95$ and $\Delta_v = 5.8$. For the crossover coupling (3.73) between the weak- and strong-perturbation regimes, this yields an estimate of $\lambda_{\rm c} \approx 1.1$. Note that exact diagonalization of a smaller system with $L = 9$ suggests that the actual response profile is asymptotically exponential with a minor plateau at small energy differences. Moreover, an exponential fit to that profile yields estimates on the order of $\alpha_v \approx 2$ and $\Delta_v \approx 7\ldots8$, indicating that the adopted values for the $L = 12$ ladder are in the right ballpark. The resulting agreement between theory and numerics is excellent for both choices of the response profile when λ is small and remains very good also for larger values of λ in case of the second-order continued-fraction approximation (3.166).

In Fig. 3.14b, we show a similar comparison between numerics and theory for the magnetization correlation $\sigma^z_{1,1}\sigma^z_{1,2}$ between the initially aligned spins, adopting the same values for α_v and Δ_v. Since the expectation values approached at long times display some variability, we use the numerical values as indicated in the figure caption for $\langle A\rangle_{\bar\rho_\lambda}$ in the theoretical prediction (3.160). The resulting agreement with the simulation data is again quite pleasing.

This concludes our collection of illustrating examples. We remark that additional examples (involving, among others, spinless fermions and hard-core bosons) can also be found in Refs. [226, 248]. In summary, we demonstrated that the main theoretical result (3.160) of this chapter successfully predicts the perturbed relaxation of several, quite distinct quantum many-body systems, including bosons, fermions, and spins, one-, two-, and infinite-dimensional setups, as well as integrable and nonintegrable models and equilibrating or nonequilibrating reference dynamics.

3.8 Connections, flaws, and prospects

Asking for the response of a given system to a perturbation is a very natural question, which has been investigated in numerous settings and examples in the past. In the context of quantum many-body systems, however, analytical results in this direction are rather scarce. One reason for this may be that the standard approach, perturbation theory à la Rayleigh-Schrödinger, is practically unfeasible or at least questionable due to the unimaginably small energy differences occurring in

the denominators of such a perturbative expansion (see also below Eq. (3.1)). For instance, Fermi's golden rule, which we briefly discussed in Sec. 3.6.3, is traditionally derived within a perturbative scheme, but its applicability to many-body quantum systems can hardly be justified this way. Yet there are a few studies aiming at characterizing the response in such many-body systems in terms of analytical derivations or heuristics, e.g., Refs. [178, 206, 208, 259, 265, 266]. Below we will comment on how our findings here and in Refs. [177, 226, 227, 248] relate to those works which are closest in spirit to our approach, i.e., which adopt tools from typicality and random matrix theory.

The typicality approach itself is also in some sense uncontrolled as discussed in Secs. 2.3 and 3.6.3, especially because it is often hard to determine in advance whether a given system can be emulated faithfully as a typical member of a particular ensemble of systems. Hence it is of utmost importance upon construction of the ensemble to incorporate the key mechanisms governing the situation under study. To conclude this chapter, we will therefore come back to scenarios of perturbed relaxation that are important, but not covered by our present approach (see also Sec. 3.6.3), and speculate about potential extensions to amend the theory accordingly.

Relation to other studies. The conceptual foundations of the approach chosen here were laid in the works [114, 253] by Deutsch, who investigated thermalization (i.e., the equality of time-averaged and thermal expectation values) in the setup (3.1) of randomly perturbed quantum many-body systems (see also Secs. 2.2.2 and 2.3.2). In this regard, he developed, in particular, a method to calculate the second moment (3.30) of eigenvector overlaps $U_{n\mu}$ from (3.5) for perturbations from the Gaussian Orthogonal Ensemble (GOE) and adapted it to estimate also higher-order moments by assuming that the $U_{n\mu}$ are independent Gaussian random variables. While such a simplification is sufficient to establish thermalization of most perturbed systems [114, 257], we observed in Sec. 3.4.4 that it violates the unitarity constraint of the $U_{n\mu}$ and can therefore not be adopted to the problem of calculating (ensemble-averaged) time-dependent expectation values as in (3.6) [248]. Moreover, strictly speaking, a rigorous justification of the original argument from Ref. [114] regarding thermalization can only be given by means of the fourth-order moment as derived here in Eq. (3.35) (see also the final paragraph of Sec. 3.4.3).

Using the same setup as in Ref. [114] with perturbations from the GOE, Nation and Porras extended Deutsch's method to approximate eigenvector overlaps and found an expression for the fourth moment (3.87) [252], which includes some of the corrections, but still violates certain reduction properties, notably Eq. (3.88b) (see Eq. (48) in [252] as well as Fig. 3.7 above). Employing their approximate result, they obtained a relation similar to (3.152) with $g(t) = \mathrm{e}^{-\Gamma|t|/2}$ for the ensemble-averaged time evolution in Ref. [208], in accordance with our result (3.164) for perturbations with an infinite band width (such as the GOE matrices). However, due to the missing reduction property, the class of admissible observables was limited to operators A whose matrix elements $A_{\mu\nu} = {}_0\langle\mu|A|\nu\rangle_0$ are zero unless the index difference $\mu - \nu$ is from a small (nonextensive) set of values. Moreover, concentration of measure (or "self-averaging") of the time evolution as in (3.154) was postulated without proof in Ref. [208].

The relaxation behavior of a small system coupled to a large bath was investigated by Genway et al. in Ref. [206] by means of a random matrix model for the system-bath coupling. The adopted ensemble is generated from a Dyson Brownian motion [267] and leads to a distribution of the matrix elements $V_{\mu\nu}$ that is approximately equal to (3.23) with the step perturbation profile (3.85). The main result of Ref. [206] is an approximation for the ensemble-averaged time-dependent reduced density matrix of the system. When applied to calculate the expectation values of system observables, it leads to a relation akin to (3.152). Notably, Genway et al. also observed a time-domain crossover of the relaxation characteristics from Gaussian to exponential behavior, occurring on a time scale of order Δ_v^{-1} (cf. Sec. 3.6.2 and especially Eq. (3.181)). An argument for concentration of measure in a very closely related setting assuming weak coupling between the system and the bath such that the total perturbation strength $\lambda\sigma_v$ decreases with increasing size of the compound was given in Ref. [205].

Finally, a similar setup with random perturbations exhibiting a step profile (3.85) was also investigated by Richter et al. in Ref. [266] using projection-operator methods. The final expression for the (ensemble-averaged) perturbed expectation values again resembles (3.152), at least if $\langle A\rangle_{\tilde{\rho}_\lambda} = 0$ and

$g_\lambda(t) = \mathrm{e}^{-\Gamma|t|/2}$, even though the derivation adopts some simplifications regarding the correlations of perturbation matrix elements and restrictions on the initial state $\rho(0)$.

To sum up, it is quite remarkable that special cases of (3.152) are recovered by different approaches of varying degree of rigor (see also Ref. [248] for yet another sketchy argument leading to a similar result). This structural stability of the result only corroborates its fundamental nature.

Constrained relaxation. It is not uncommon that realistic systems exhibit certain symmetries or relations between the key operators H_0, V, A, and $\rho(0)$ (see Sec. 3.1), which usually restrict the effective Hilbert space explored by the system. Since the perturbation ensembles introduced in Sec. 3.3 do not explicitly incorporate such symmetries, it is important to ask under which circumstances they may lead to "atypical behavior" in the sense of noticeable deviations of the perturbed relaxation from (3.160). We therefore consider several scenarios.

First, there may be conservation laws Q pertaining to both the unperturbed Hamiltonian H_0 and the perturbation V, i.e., $[H_0, Q] = [V, Q] = 0$ and hence also $[H_\lambda, Q] = 0$ for all λ in (3.1). Such a conservation law Q naturally divides the Hilbert space into sectors of constant Q, i.e., subspaces within which $\langle\psi|Q|\psi\rangle$ takes the same value for all states $|\psi\rangle$. In this case, one should restrict the analysis right from the beginning to the relevant subsectors populated by the initial state $\rho(0)$. For each subsector, provided that it is still of sufficiently high dimension, the corresponding perturbation profile $\sigma_v^2(E)$ from (3.10) should be assessed individually, yielding a possibly distinct response profile $g_\lambda(t)$ for each sector, too. The resulting prediction for the perturbed relaxation then involves a combination of terms similar to (3.160) for the dynamics within each sector as well as interference terms between those sectors.

The simplest (and perhaps most common) case is that $\rho(0)$ lives on just one subsector. An example of this kind is the two-dimensional spin system (3.20), for which both H_0 and V commute with the total magnetization M^z (see below Eq. (3.20b)). Hence we focused on the $M^z = 0$ sector for the analysis of the level distribution (Fig. 3.1) and the perturbation profile (Fig. 3.2) in Sec. 3.2, which was also the sector occupied by the initial states considered in Sec. 3.7 as an example to test the perturbed relaxation prediction (see Fig. 3.10 in particular). The same applies to the example of two coupled spin chains from Fig. 3.14 in Sec. 3.7. Another example for this setting explicitly considered in Sec. 3.7 is the XXZ chain (3.197) (see also Fig. 3.13a), whose Hamiltonian H_λ again commutes with the total magnetization M^z and is, in fact, integrable for all values of λ.

Second, there may be conservation laws Q of the unperturbed system H_0 that are broken by the perturbation. This is an important case as it applies, for example, to the transition from an integrable H_0 to a nonintegrable H_λ ($\lambda > 0$) and is thus the prime example for prethermalization (see Secs. 2.2.3 and 3.5). In this scenario, our typicality approach is generally expected to work as the overwhelming majority of perturbations in any ensemble from Sec. 3.3 will violate the conservation laws similarly to the true V. The opposite situation, where H_λ exhibits special symmetries for some λ whereas H_0 does not, is excluded for obvious reasons since such a scenario would entail a very special perturbation (see also Prerequisite (vi) from Sec. 3.2 as well as the discussion in Sec. 3.6.3).

Third, there may be operators Q that commute with the true perturbation V, $[V, Q] = 0$, but not with the reference Hamiltonian H_0, hence $[H_\lambda, Q] \neq 0$ for all λ. In such a situation, it is not unlikely that the typical behavior of the random perturbations from Sec. 3.3 deviates from the true dynamics because most perturbations from the ensemble will violate $[V, Q] = 0$, contrary to the true perturbation of interest. In particular, examples for the special case $Q = A$ were found to yield a different relaxation behavior in some circumstances in Refs. [238, 259]. However, a setup with $[V, A] = 0$ but $[H_0, A] \neq 0$ does not automatically imply deviations from the relaxation dynamics (3.160) induced by the perturbation ensembles considered here. For instance, the staggered magnetization (3.198) commutes with the perturbation, but not with the reference Hamiltonian in our XXZ-chain example from Eq. (3.197) and Fig. 3.13a, yet the agreement between theory and numerics was found to be very good. Similarly as discussed in Sec. 3.6.3, the considered combination of the initial state and observable must also explicitly scan for the additional symmetry in order for deviations to occur.

Nevertheless, it is certainly desirable in general to include the knowledge about such additional symmetries of the perturbation into the definition of the considered ensembles. One way to do so

could be to introduce a (Gaussian) "penalty weight" proportional to the commutator $[V, Q]$ in the probability distribution (3.24) of the V ensemble. While the averaging procedure can then still be carried out similarly as in Secs. 3.4.2 and 3.4.3, the constraint leads to additional terms that severely complicate the integration over the supersymmetric degrees of freedom, hence we did not manage to arrive at a reasonably clear and simple result.

Local and few-body interactions. The observation that fundamental interactions in physical systems are local and of few-body type is certainly an important characteristic for their dynamics. We already touched upon the shortcomings of our present approach in this regard in Sec. 3.6.3. The conclusion was that the particular setup under study should not probe local or few-body properties explicitly in order to be amenable to the theory, for instance, by restricting to initial states and/or observables which are sufficiently homogeneous on a macroscopic level. Here we will briefly comment on ways to extend the typicality approach so that it respects the locality and few-body character of a given model.

There are a few studies which investigate equilibration and thermalization by means of random matrix ensembles with an additional local or few-body structure [107, 237, 268–271]. Moreover, matrices built from random interactions between a limited number of degrees of freedom—going by the name of *embedded ensembles*—have also been considered to some extent in the random-matrix community, see, for example, Refs. [272–277]. Unfortunately, most of these investigations employed numerical methods by and large, meaning that rather little is known analytically about those ensembles.

An immediate connection to the relaxation theory from this chapter could be established by considering a free (noninteracting) reference system H_0 and model the few-body interactions V by such an embedded ensemble. If the overlap distribution $u(E)$ from (3.32) were available, much of the formalism from Sec. 3.4.4 could be readily adapted to find a first approximation for higher-order moments, even though the additional correlations in the embedded ensemble could invalidate the concomitant, rather coarse reduction of higher-order moments to second order. To include locality as well, one needs to impose a model of the system's geometry, e.g., by means of a lattice. Restricting the interaction to nearby sites or scaling its strength according to the distance between sites can then serve as a proxy for the local character of fundamental interactions.

Apparently, these additional structures introduce quite strong correlations between the matrix elements of the perturbation, and they should manifest in the final result (e.g., a prediction of the perturbed relaxation) in some way. Having less interactions (compared to the perturbation ensembles from Sec. 3.3) thus makes the analytical treatment much more complicated. Moreover, it is not even clear *a priori* whether or not the so-generated random matrix ensembles can be utilized for a typicality argument in the sense that they exhibit concentration of measure for the resulting relaxation dynamics. Pursuing further analytical progress in this direction will thus be a challenging, yet potentially exciting endeavor.

Driven systems. So far, we only considered time-independent perturbations. Then again, it is not uncommon that a given setup depends on some time-dependent control parameter. Prime examples of this kind are (finite-time) quenches [7, 83, 174, 278–281] or periodically driven systems [282, 283]. Hence it is an interesting and quite natural question whether the results of this chapter can be extended to time-dependent couplings $\lambda = \lambda(t)$.

The following two chapters will consider dynamics of quantum many-body systems in such time-dependent settings. In Chapter 4, we investigate so-called echo protocols [284, 285], whose time dependence is still relatively simple as there is only a single parameter change involved. The more general case of largely arbitrary protocols $\lambda(t)$ will be dealt with in Chapter 5.

4 Echo dynamics

As expounded in the introduction (Chapter 1), the ubiquitous irreversibility of macroscopic processes is not reflected in the fundamental laws governing their microscopic constituents, but rather emerges as the practical impossibility to revert the dynamics of a large system resulting from the complicated (or "chaotic" [158, 159, 163]) character of interactions between them. Our understanding of the underlying mechanism, however, is still far from comprehensive.

In this chapter, we will explore aspects of this irreversibility for many-body systems within the realm of nonrelativistic quantum mechanics by means of so-called echo protocols. The goal is to quantify how small deviations in the initial state or the dynamical laws propagate and manifest themselves in macroscopic observables. The fact that macroscopic experiments are generally reproducible despite the practical impossibility to prepare a large system in exactly the same microstate suggests that the usual relaxation behavior of macroscopic observables is somewhat insensitive to such small deviations [181, 182]. On the other hand, it is well known at least for classically chaotic systems that the distance between close-by phase-space points grows with time [158, 159], so it would be generally desirable to extract signatures of this indicator of chaos and irreversibility by means of macroscopic observables.

Echo protocols offer one way to carry out such an investigation. The general idea is to prepare a given system out of equilibrium and to compare its observable dynamics forward in time to an (effectively) time-reversed setup. Small and usually uncontrolled inaccuracies in this time-reversed setting lead to deviations between the two scenarios, which become macroscopically detectable in the nonequilibrium regime. Their scaling with various control parameters can be taken as a measure for how irreversible the considered dynamics is.

In essence, the results presented in this chapter have been published in Refs. [260, 286]. The concept and reasoning behind echo protocols and experiments will be introduced in more detail in Sec. 4.1. Thereafter, we will first consider two examples from classical mechanics in Sec. 4.2 that will help us to work out the distinct behavior of quantum systems in the ensuing sections. Namely, we will examine the effects two different types of inaccuracies or imperfections on the echo dynamics in Secs. 4.3 and 4.4, and conclude with a discussion and comparison between the classical and quantum cases in Sec. 4.5.

4.1 Echo protocols and irreversibility

Protocol and imperfections. As usual, we focus on isolated many-body quantum systems and start from a *reference system* with time-independent Hamiltonian

$$H_0 := \sum_\mu E_\mu |\mu\rangle\!\rangle\langle\!\langle\mu| \,. \tag{4.1}$$

Moreover, we will again monitor the dynamics of the system by means of an experimentally realistic observable A. Denoting the (pure or mixed) state of the system at time t by $\rho(t)$, the principal object of study are thus the time-dependent expectation values $\langle A\rangle_{\rho(t)} := \mathrm{tr}[\rho(t)A]$ as before.

In general, our echo protocols consist of two main phases. For the first phase, the system is prepared in some initial state $\rho(0) := \rho_\mathrm{T}$, henceforth called the *target state*, and subsequently relaxes for a certain *waiting time* τ according to the reference Hamiltonian H_0. During this *forward phase*, the state is thus given by

$$\rho(t) = \rho_\mathrm{f}(t) := \mathrm{e}^{-\mathrm{i}H_0 t}\,\rho_\mathrm{T}\,\mathrm{e}^{\mathrm{i}H_0 t} \tag{4.2}$$

for $0 \leq t \leq \tau$. At time $t = \tau$, the system has reached the *return state* $\rho_\mathrm{R} := \rho_\mathrm{f}(\tau)$. At this point, we perform an effective time reversal by switching to the negative Hamiltonian $-H_0$, thus simulating time running backwards in the subsequent evolution since a true time reversal is unfeasible for obvious reasons. We will comment on the practicality of such a transformation from H_0 to $-H_0$ in a minute and may regard it as a *gedankenexperiment* for the time being. If this effective time

reversal is implemented perfectly, the system will simply trace out the same sequence of states as during the forward phase (4.2), but in reverse order. In particular, we thus end up in the target state ρ_{T} again after time $t = 2\tau$. However, in practice there will be unavoidable errors or imperfections in carrying out this procedure, which will lead to deviations between the forward and backward dynamics. Intriguingly, we can learn something from these imperfections about the investigated system and its irreversibility.

In the following, we will consider two types of imperfections. The first one concerns the return state ρ_{R}, i.e., the initial state for the backward evolution. During the process of changing from the forward Hamiltonian H_0 to the backward Hamiltonian $\tilde{H}_0 := -H_0$, the state of the system may be slightly perturbed, too, resulting in an imperfect return states ρ'_{R}. We model this by acting with a perturbing or *scrambling Hamiltonian* W on ρ_{R} for a short *scrambling time* δ, such that

$$\rho_{\mathrm{s}}(t) := \mathrm{e}^{-\mathrm{i}Wt}\rho_{\mathrm{R}}\mathrm{e}^{\mathrm{i}Wt} \tag{4.3}$$

is the state of the system during the *scrambling phase*, and $\rho'_{\mathrm{R}} := \rho_{\mathrm{s}}(\delta)$. Note that the usage of the word "scrambling" here merely indicates the mixing or perturbing character of the operation and is unrelated to the more specific meaning in the context of out-of-time-ordered correlators. Consequently, the state $\rho(\tau + t) := \rho_{\mathrm{s}}(t)$ for $0 \leq t \leq \delta$. This mechanism of perturbing the return state will be referred to as *imperfect preparation* in the following.

The second type of inaccuracies regards the time-reversed Hamiltonian, which may potentially not be the perfect inverse of the forward Hamiltonian H_0, but could be slightly perturbed such that $\tilde{H} := -H_0 + \epsilon V$ is the Hamiltonian acting during the *backward phase*. Hence

$$\rho_{\mathrm{b}}(t) := \mathrm{e}^{\mathrm{i}(H_0-\epsilon V)t}\,\rho'_{\mathrm{R}}\,\mathrm{e}^{-\mathrm{i}(H_0-\epsilon V)t} \tag{4.4}$$

is the state during that phase and $\rho(\tau + \delta + t) := \rho_{\mathrm{b}}(t)$ for $0 \leq t \leq \tau$. The presence of the perturbation V during the backward evolution will be called *imperfect reversal* in the following. Note that ϵ here quantifies the magnitude of the inaccuracies and should not be confused with the mean level spacing ε of the Hamiltonian introduced in (2.11).

At the end of the entire process, the system will then be in the state $\rho'_{\mathrm{T}} := \rho_{\mathrm{b}}(\tau)$, a perturbed version of the target state whose deviations from the reference state $\rho(0) = \rho_{\mathrm{T}}$ may in principle depend, for fixed perturbations V and W, on the scrambling time δ, the magnitude ϵ of reversal inaccuracies, and the waiting time τ. Altogether, we may summarize our imperfect echo protocol schematically as

$$\rho_{\mathrm{T}} \xrightarrow[H_0]{\tau} \rho_{\mathrm{R}} \xrightarrow[W]{\delta} \rho'_{\mathrm{R}} \xrightarrow[-H_0+\epsilon V]{\tau} \rho'_{\mathrm{T}}\ . \tag{4.5}$$

Echo experiments. To realize an echo protocol such as (4.5) experimentally, we need the possibility to change the sign of the Hamiltonian H_0 of a given system. While such a procedure is unphysical in many situations (e.g., because it would require negative particle masses), it is in fact a well-established technique in spin systems.

The first experiments of this type are the famous *Hahn echoes* [287], where an initially aligned set of noninteracting spins (typically nuclear spins of some liquid) precesses in an external magnetic field. Due to local inhomogeneities of the field strengths, however, the individual spins precess at slightly different frequencies, causing a dephasing such that the magnetization decays. Applying a so-called π pulse after time τ inverts the orientation of all spins, which is equivalent to an inversion of the external field and hence the Hamiltonian. Consequently, the spins start to converge again during the subsequent evolution and will eventually be aligned again after another time period τ, restoring the original magnetization. Due to remnant interactions between the spins as well as with the environment, however, the final alignment will usually not be perfect, i.e., the final magnetization will not quite reach the original level. This is exactly the type of effective time reversal in the presence of imperfections we have in mind with the protocol (4.5).

Imperfect spin echoes are also at the heart of magnetic resonance imaging (MRI) [288]. This well-known medical imaging technique targets the nuclear spin of hydrogen atoms by means of an echo protocol. Different imperfections and concentrations in different tissues lead to different decay

characteristics of the echo signal, which thus allows to identify those tissues. Hence "imperfections" are in fact not a bug, but a feature in this application.

Since the first experiments by Hahn in 1950, the capabilities to perform an effective time reversal have been adapted to many classes of interacting spin systems using so-called magic- or polarization-echo techniques [37–44]. Here the dominant part of the interacting Hamiltonian is effectively reversed by applying sophisticated pulse sequences of radiofrequency external fields during the backward phase. More generally, such techniques can be used experimentally to adjust the coupling parameters in a variety of effective spin systems over a wide range [45–47, 289].

Further suggestions or experimental realizations of an effective time reversal in quantum systems include tuning a cold-atom gas across a Feshbach resonance [290–292] or employing quantum simulators [36, 289] (see also the discussion at the end of Sec. 2.1). In all these setups, there will naturally be inaccuracies in the effective time reversal, both because it is usually only feasible to invert the dominant part of the Hamiltonian and due to limited experimental precision in carrying out the corresponding protocols.

Echo signal. We intend to quantify the influence of imperfections on the echo dynamics in terms of an experimentally realistic observable A. To detect a signature of the imperfections, it is clear that we have to operate in the nonequilibrium regime. Indeed, if we are initially in equilibrium ($\rho \simeq \bar{\rho}$), then the expectation values $\langle A\rangle_{\rho(t)}$ will remain very close to the equilibrium value $\langle A\rangle_{\bar{\rho}}$ for all times during the forward evolution, and the small imperfections during the scrambling and backward phases will not have any noticeable effect either. Therefore, we take it for granted that the systems starts out in a nonequilibrium state ρ_{T}.

Usually (cf. Sec. 2.2), the system will subsequently relax during the forward phase such that the expectation values approach the equilibrium value $\langle A\rangle_{\bar{\rho}}$, which moreover generically coincides with the thermal expectation value $\langle A\rangle_{\rho_{\mathrm{mc}}}$. Of foremost interest in the following are therefore the deviations

$$\mathcal{A}(t) := \langle A\rangle_{\rho(t)} - \langle A\rangle_{\rho_{\mathrm{mc}}} \tag{4.6}$$

of the time-dependent expectation values from the thermal values. For now, this should merely be understood as a constant offset $-\langle A\rangle_{\rho_{\mathrm{mc}}}$ added to the time-dependent expectation value $\langle A\rangle_{\rho(t)}$. We emphasize that the calculations presented below do *not* require that the system should actually equilibrate, let alone thermalize (see also Fig. 4.5b below for an explicit example that does not equilibrate). Nevertheless, the quantity (4.6) will arise naturally as a suitable measure of the echo signal.

In case of a perfect time reversal ($\delta = \epsilon = 0$), we obviously have $\mathcal{A}(\tau + t) = \mathcal{A}(\tau - t)$ for all $0 \leq t \leq \tau$. Due to the delicate calibration of the initial state $\rho(0)$ and the observable A necessary to achieve nonequilibrium conditions at all (see Secs. 2.2 and 2.3), it is reasonable to expect that the essentially uncontrolled inaccuracies will generically push the system closer to equilibrium, i.e.,

$$|\mathcal{A}(\tau + \delta + t)| \lesssim |\mathcal{A}(\tau - t)| \qquad (0 \leq t \leq \tau)\,. \tag{4.7}$$

By quantifying how sensitive the deviations of the perturbed signal $\mathcal{A}(\tau + \delta + t)$ from the perfect signal $\mathcal{A}(\tau - t)$ are, we can thus assess the degree of irreversibility of the considered system: The faster $\mathcal{A}(\tau + \delta + t)$ decays with δ, ϵ, or τ compared to $\mathcal{A}(\tau - t)$, the harder it is to revert the dynamics, and the more extraordinary or special are the initially probed nonequilibrium states.

The central object of study in the following will therefore be the ratio $\mathcal{A}(\tau+\delta+t)/\mathcal{A}(\tau-t)$ between the perturbed and perfect echo signals. The main result of this chapter is an analytical prediction for this *relative echo signal* in isolated many-body quantum systems.

Especially relevant is the ratio of the echo signal at the end of the protocol (4.5) to the signal at the beginning because in the generic case, this is the time point where the system will be farthest away from equilibrium, entailing that the effects of imperfections will be most pronounced. We denote this *relative echo peak height*, which generally depends on the two parameters δ and ϵ characterizing the magnitude of imperfections as well as on the waiting time τ, by

$$\mathcal{F}(\delta, \epsilon, \tau) := \frac{\mathcal{A}(2\tau + \delta)}{\mathcal{A}(0)} = \frac{\langle A\rangle_{\rho'_{\mathrm{T}}} - \langle A\rangle_{\rho_{\mathrm{mc}}}}{\langle A\rangle_{\rho_{\mathrm{T}}} - \langle A\rangle_{\rho_{\mathrm{mc}}}}\,. \tag{4.8}$$

Outline of the derivation. To arrive at an analytical prediction for the relative echo signal and thus also the echo peak (4.8), we will again employ typicality methods. Such an approach is indeed particularly suited to investigate the present problem since we only have limited (if any) knowledge about the imperfections (i.e., the operators V and W in (4.5)) that lead to deviations from the perfect echo. Modeling our partial ignorance by considering suitable ensembles of random operators V and W is thus a very natural strategy, i.e., it is not just convenient, but also inevitable.

From a computational point of view, we follow the same steps as sketched in Sec. 2.3.1. First, we compute the average effect of all imperfections in a chosen ensemble. Second, we prove that the deviations for a single realization of these imperfections from the average effect is practically undetectable for nearly all members of the ensemble. Third, we compare the so-obtained "typical" echo signal to concrete numerical examples.

We will first address the two types of imperfections introduced below Eq. (4.2) individually: the imperfect preparation scenario ($\delta > 0$, $\epsilon = 0$) in Sec. 4.3 and the imperfect reversal scenario ($\delta = 0$, $\epsilon > 0$) in Sec. 4.4. How to combine the two situations will then be explained in Sec. 4.5. Before all that, however, we will consider a classical example in Sec. 4.2 as a reference to highlight certain peculiarities of the quantum echoes.

4.2 Classical examples

In this section, we will briefly explore an example of echo dynamics from the realm of classical mechanics, focusing on the pure "imperfect preparation" scenario ($\epsilon = 0$ in (4.5)) because here the differences to the quantum case are most striking.

Chaos. On the classical level, the chaotic dynamics of nonlinear systems and concepts like sensitive dependence on initial conditions and mixing [158, 159] elucidate how a small uncertainty about the initial state spreads uniformly across the available phase space in time[3]. Notably, initially close-by points in phase space separate exponentially in time at a rate quantified by so-called Lyapunov exponents. Together with a maximum entropy principle, whereby the macroscopic equilibrium state is simply the one with the largest number of compatible microstates (see also Sec. 2.1), this offers an explanation for why chaotic systems with many degrees of freedom approach equilibrium and stay there for most of the time, meaning that deviations, such as Poincaré recurrences, can and will occur but very rarely.

With regard to echo protocols, therefore, a small imprecision in setting up the return state at the point of reversal will usually lead to a final state that differs noticeably from the initial target state. If the dynamics is chaotic in the classical sense, we can generically expect that these differences grow with the waiting time τ. However, as outlined in the introduction to this chapter, differences between two phase-space points are often hardly amenable from a macroscopic point of view, precisely because macroscopic observables cannot distinguish between equivalent microstates. The aim of the following two examples is therefore to demonstrate that macroscopic observables can indeed detect those chaos indicators by means of echo protocols.

Two-dimensional gas in a box. As a first example of echo dynamics in a classical system, we consider a gas of $\mathcal{N}$ two-dimensional spherical particles of radius R and mass m in a box $\mathcal{V} := \{(x, y) : 0 \leq x < L_x, 0 \leq y < L_y\}$ with periodic boundary conditions. The particles essentially interact via hard-core collisions, modeled by a very steep pair potential of the form

$$U(r) := U_0 \left[\left(\tfrac{2R}{r}\right)^{32} - 1 \right]^2 \Theta(2R - r)\,, \tag{4.9}$$

where r denotes the center-of-mass distance of the two interacting particles, U_0 is a constant, and $\Theta(x)$ is the Heaviside step function as usual. Denoting the position and momentum of the ith

[3] Note that due to the conservation of phase space volume under Hamiltonian dynamics, this does not mean that all points in the phase space compatible with the macroscopic configuration can be reached, but there will typically be close-by reachable neighbors for any such point. More precisely, one could say that the diameter of the time-evolved initial volume approaches the diameter of the full compatible phase space.

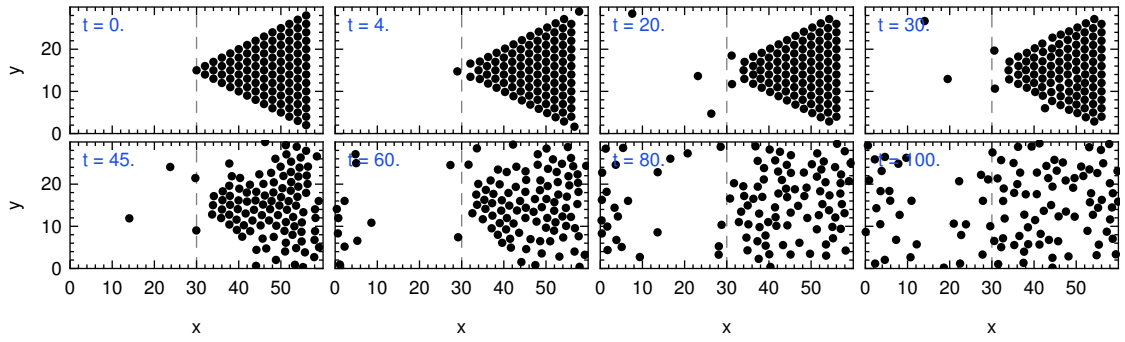


Figure 4.1: Time evolution of a two-dimensional gas of particles in a box of dimensions 60×30 with periodic boundary conditions, described by the classical Hamiltonian function (4.10) with $m = 1$, $R = 1$, $U_0 = 0.2$, prepared in a triangular block with all particles at rest except for the left-most one with initial momentum $(p_{0,x}, p_{0,y}) = (2, 0.1)$. Each panel depicts a snapshot of the system configuration after time t as indicated in the top-left corner.

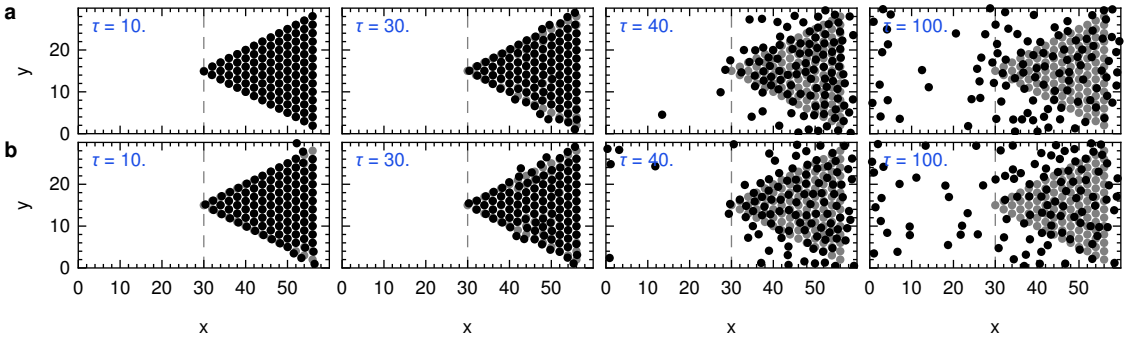


Figure 4.2: Configuration of the system from Fig. 4.1 after completion of the imperfect-preparation echo protocol, i.e., at time $t = 2\tau$ (black disks), for various waiting times τ. Imperfections are introduced by rotations of the reversed momentum vectors by random angles drawn uniformly from $[-\delta, \delta]$ with **a.** $\delta = 0.01$ and **b.** $\delta = 0.04$. For comparison, the perfect echo state is shown as well (gray disks).

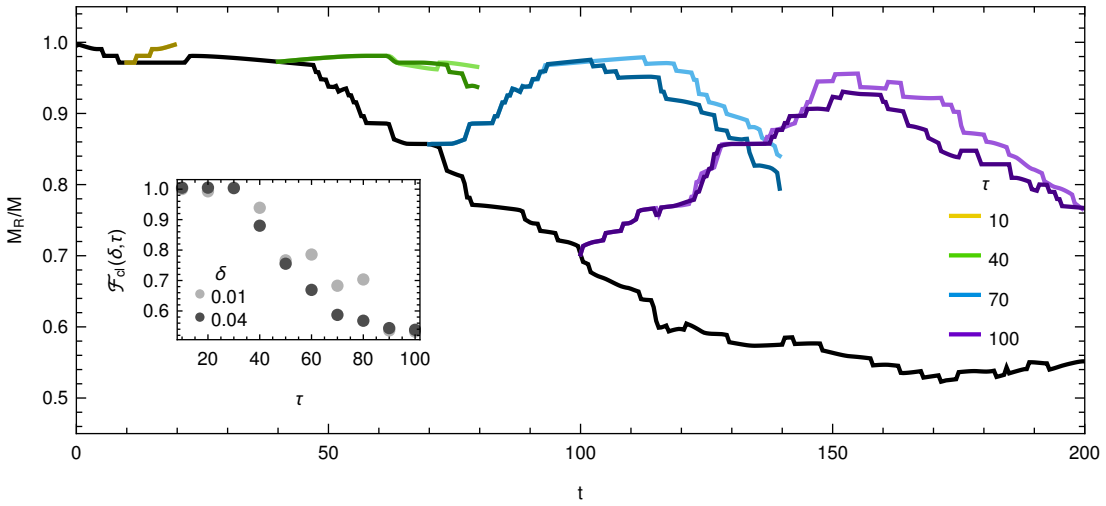


Figure 4.3: Time evolution of the relative mass M_R/M in the right compartment forward in time (black) and under the imperfect-preparation echo protocol (cf. Figs. 4.1 and 4.2) for various waiting times (color-coded as indicated) and imperfections $\delta = 0.01$ (light) and $\delta = 0.04$ (dark). Inset: Echo peak height (4.11) as a function of the waiting time τ for $\delta = 0.01$ (light) and $\delta = 0.04$ (dark).

particle by $\boldsymbol{r}_i$ and $\boldsymbol{p}_i$, respectively, the classical Hamiltonian function of the system thus reads

$$H(\boldsymbol{p},\boldsymbol{r}) = \sum_i \frac{\boldsymbol{p}_i^2}{2m} + \sum_{i<j} U(|\boldsymbol{r}_i - \boldsymbol{r}_j|)\,. \tag{4.10}$$

To study the echo dynamics of macroscopic observables, we divide the box into left and right compartments $\mathcal{V}_\mathrm{L} := \{(x,y) \in \mathcal{V} : x < L_x/2\}$ and $\mathcal{V}_\mathrm{R} := \{(x,y) \in \mathcal{V} : x \geq L_x/2\}$, respectively, and observe the total mass M_R in the right half. In thermal equilibrium, we thus have $M_\mathrm{R,mc} = M/2$ with $M = \mathcal{N}m$ the total mass of all particles.

The system is prepared in a nonequilibrium target state with the particles arranged in a triangle as shown in the top-left panel of Fig. 4.1. The momentum of the left-most particle at $x = L_x/2$, $y = L_y/2$ is set to some finite value $(p_{x,0}, p_{y,0})$, whereas all other particles are at rest initially. We then integrate the Hamiltonian (Newtonian) equations of motion resulting from (4.10) numerically using a velocity-Verlet algorithm.

Several snapshots of the (forward) dynamics for $\mathcal{N} = 105$ particles with $m = 1$, $R = 1$ in a box with $L_x = 60$, $L_y = 30$, interaction strength $U_0 = 0.2$, and initial momentum $(p_{x,0}, p_{y,0}) = (2, 0.1)$ of the left-most particle are shown in Fig. 4.1. The initial impetus of the left particle induces a shock wave propagating through the triangular block and detaching the top and bottom particles on the right ($t \approx 4$). Once the these "corner particles" hit back onto the remaining block ($t \approx 30$), the triangular shape starts to dissolve and the dynamics becomes increasingly complicated and "chaotic," approaching a roughly homogeneous distribution of particles in the box ($t \gtrsim 100$), meaning that the system thermalizes.

To implement the imperfect-preparation protocol, we perform an effective time reversal at $t = \tau$ by flipping all momenta from $\boldsymbol{p}_i(\tau)$ to $-\boldsymbol{p}_i(\tau)$ (yielding the "return state"), such that the system would evolve back towards the triangular target state in the subsequent evolution with the Hamiltonian function from (4.10). Imperfections in the return state are introduced by additionally rotating all $\boldsymbol{p}_i(\tau)$ by random angles δ_i chosen independently from a uniform distribution in the interval $[-\delta, \delta]$.

As a consequence, the particle positions at time $t = 2\tau$ generally deviate from those at $t = 0$. This is illustrated in Fig. 4.2 for two different choices of δ (rows) and four waiting times τ (columns), where we show the resulting final configuration after completion of the entire protocol at $t = 2\tau$ (black disks). Since the chaoticity of the dynamics inevitably entails some sensitivity against numerical inaccuracies as well, we also plot the final state as obtained without perturbations of the return state ($\delta = 0$, gray disks) to ensure that numerical inaccuracies are not (yet) relevant. Qualitatively, the figure suggests that the system is somewhat sensitive to the initial conditions since the deviations between the perfect and perturbed target states (i.e., between the gray and black configurations) clearly increases with increasing waiting time τ (as well as with increasing δ).

Since we eventually intend to quantify this sensitivity by means of macroscopic observables, we also monitor the mass M_R in the right compartment of the box as introduced above. From Fig. 4.2, it is already qualitatively apparent that the associated echo peak

$$\mathcal{F}_\mathrm{cl}(\delta,\tau) := \frac{M_\mathrm{R}(\tau) - M_\mathrm{R,mc}}{M_\mathrm{R}(0) - M_\mathrm{R,mc}} \tag{4.11}$$

will generally decay with both δ and τ. This is corroborated quantitatively in Fig. 4.3, where we depict the time evolution of M_R/M as well as the echo peak height (4.11) as a function of τ and indeed observe that the signal is generally attenuated more the larger δ is and the longer we wait to initiate the reversal.

Classical spins. Another important example, particularly with regard to the realizability of time reversal in the quantum setting, are spin systems. We briefly review the investigations by Fine and co-workers [293–296] concerning echo dynamics and chaoticity of classical spins. The employed classical Hamiltonian function is commonly given by

$$H(\boldsymbol{S}) = \sum_{i \leftrightarrow j, \alpha} J_\alpha\, S_i^\alpha\, S_j^\alpha\,, \tag{4.12}$$

defined on different lattice geometries with $i \leftrightarrow j$ denoting a pair of neighboring sites. The $\boldsymbol{S}_i = (S_i^x, S_i^y, S_i^z)$ represent the classical spin degrees of freedom on the ith site such that $\boldsymbol{S}_i^2 = 1$, and J_x, J_y, J_z are coupling constants.

In Ref. [293, 294] it is demonstrated that, except for the (integrable) Ising model where two of the J_α vanish, the dynamics of these systems is generally chaotic in the sense that the largest Lyapunov exponent is positive, meaning that close-by states in phase space separate exponentially with time. These classical spins systems thus generically exhibit sensitive dependence on initial conditions.

Refs. [295, 296] deal with echo protocols similar to (4.5) in systems of the form (4.12) in particular and monitor the dynamics of a macroscopic observable, namely, the magnetization in the x direction M^x, whose equilibrium value is $M^x_{\mathrm{mc}} = 0$. The imperfections introduced at $t = \tau$ consist of rotations of the spin vectors around randomly chosen axes by random angles between $[-\delta, \delta]$ with $\delta = \pi/100$. For our purposes, the key result of Ref. [295] is that the echo peak height $\mathcal{F}_{\mathrm{cl}}(\delta, \tau) := M^x(2\tau)/M^x(0)$ is found to decay exponentially with the waiting time τ, where the rate is twice the largest Lyapunov exponent. This corroborates the chaotic character of the dynamics and illustrates that echo protocols can indeed detect sensitive dependence on initial conditions. Similarly to the classical two-dimensional gas studied above, the echo signal of classical spins is thus attenuated with the waiting time τ.

Refs. [295, 296] also present a similar analysis for the quantum analog of the system (4.12), where $\boldsymbol{S}_i$ are spin operators of spin quantum number $\frac{1}{2}$ or $\frac{15}{2}$. In the spin-$\frac{1}{2}$ case [295], no exponential decay of the echo peak height $\mathcal{F}(\delta, \tau)$ with τ is found. Instead, the authors report a power-law decay of $\mathcal{F}(\delta, \tau)$ with τ for small times, and the results indicate that $\mathcal{F}(\delta, \tau)$ settles down to a nonzero constant for larger values of τ. In the spin-$\frac{15}{2}$ case [296], the echo signal shows some resemblance to an exponential decay for small τ, yet it again approaches a nonzero plateau value as τ increases. As one might have expected, the behavior of this system of large quantum spins is thus somewhat closer to the classical case than the "more quantum" spin-$\frac{1}{2}$ example, but there are still marked differences especially for large τ. These observations foreshadow the persistence of echo signals in many-body quantum systems to be derived in the subsequent Sec. 4.3 in a much more general setting.

4.3 Imperfect preparation

After our excursion into the realm of classical mechanics, we resume the investigation of echo dynamics in the quantum setting. As sketched in Sec. 4.1, there are two types of imperfections we intend to consider. We begin with the imperfect-preparation scenario here, meaning that $\delta > 0$ and $\epsilon = 0$ in (4.5), leaving us with the protocol

$$\rho_{\mathrm{T}} \xrightarrow[H_0]{\tau} \rho_{\mathrm{R}} \xrightarrow[W]{\delta} \rho'_{\mathrm{R}} \xrightarrow[-H_0]{\tau} \rho'_{\mathrm{T}} \ . \tag{4.13}$$

The inaccuracies thus affect only the short scrambling phase that models potential disruptions when changing from the forward to the backward Hamiltonian, whereas the backward evolution itself is considered to be clean and governed by the perfect inverse $-H_0$. By analogy with the classical examples discussed in Sec. 4.2, this protocol may roughly be understood as a probe for sensitive dependence on initial conditions since it measures how a small difference between the return states ρ_{R} and ρ'_{R} propagates under the influence of the Hamiltonian $\tilde{H}_0 = -H_0$ and manifests itself after the time τ in the target states ρ_{T} and ρ'_{T}.

We remark that, on the one hand, the scrambling phase in this context may be regarded as an actual short time interval during which the system evolves under the perturbation Hamiltonian W. On the other hand, more abstractly, it may simply be viewed as a rotation in the pertinent Hilbert space of dimension N by an "angle" δ around the (properly normalized) "axis" W since the skew-Hermitian operators iW generate the unitary group U(N) or subgroups thereof, depending on the choice of the W ensemble.

4.3.1 Typical echo signal

Prerequisites. We begin by collecting the principal assumptions about the isolated many-body system under study taken for granted in the derivations of this section. The first one is the standard requirement that the system should have a well-defined macroscopic energy $\mathcal{E}$. As explained in Sec. 2.1, this means that there is an energy window $I_\mathcal{E}$ as in (2.9) such that populations $\rho_{\mu\mu}(0) = {}_0\langle\mu|\rho(0)|\mu\rangle_0$ of the initial state with $E_\mu \notin I_\mathcal{E}$ are negligible for the dynamics. Consequently, the state $\rho_\mathrm{f}(t)$ from (4.2) is concentrated on this energy window for all times t, too. We recall that the dimension of this energy window is denoted by N and is exponentially large in the system's degrees of freedom according to (2.10). Contrary to the setting in Sec. 3, however, we do not require the density of states $D(E)$ from (2.11) to be constant, implying that we can also be a little more tolerant regarding the extent of $I_\mathcal{E}$ and allow somewhat larger windows.

The second principal assumption is that the perturbation W acting during the scrambling phase and taking ρ_R to ρ'_R is so small that it does not change the energy of the system macroscopically. As a consequence, the state $\rho(t)$ at any time, i.e., during the entire course of the protocol (4.13), only significantly occupies energy levels with $E_\mu \in I_\mathcal{E}$. Throughout the rest of Sec. 4.3, summations over energy levels are thus implicitly understood to be restricted to the window $I_\mathcal{E}$ of dimension N.

For clarity of the presentation, we furthermore assume that the populations ${}_W\langle n|\rho_\mathrm{R}|n\rangle_W$ of the return state ρ_R in the eigenbasis $\{|n\rangle_W\}$ of the scrambling Hamiltonian W are approximately uniformly distributed within the energy window $I_\mathcal{E}$. On the one hand, this is a natural assumption in the absence of additional knowledge about the type of inaccuracies modeled by W. On the other hand, the following findings can be generalized straightforwardly to take into account the precise distribution of ρ_R in the eigenbasis of W by adapting the results from Ref. [100] accordingly (see also the discussion following Eq. (4.27) below).

Ensemble of scrambling operators. The operator W models largely uncontrolled inaccuracies to which the system is subjected during the scrambling phase. Yet it is clear that the character of the system cannot change completely during that phase because it still entails, for example, the same degrees of freedom. We choose to model this partial knowledge and partial ignorance about the nature of W by an ensemble of random operators similar to the one adopted in Ref. [96]: The eigenvalues E_n^W of W are assumed to be given and fixed, whereas the corresponding eigenvectors $|n\rangle_W$ are drawn at random according to the Haar measure of the appropriate symmetry group associated with the Hilbert space of the energy window $I_\mathcal{E}$. More precisely, the transformation matrices

$$\tilde{U}_{n\mu} := {}_W\langle n|\mu\rangle_0 \tag{4.14}$$

between the eigenbasis $\{|\mu\rangle_0\}$ of the reference Hamiltonian H_0 and the eigenbasis $\{|n\rangle_W\}$ of W will be drawn from that symmetry group, and we explicitly consider the circular unitary ensemble (CUE) associated with the unitary group $\mathrm{U}(N)$ as well as the circular orthogonal ensemble (COE) of symmetric unitary matrices.

This modeling still leaves some flexibility allowing to incorporate additional information about the system via the eigenvalues E_n^W. If no further information is available, one possibility is to choose them related (or equal) to the actual system Hamiltonian H. Alternatively, for instance, the well-known Gaussian Unitary Ensemble (GUE) is obtained by combining the CUE with randomly drawn E_n^W from a semicircle distribution [163]. In the more abstract Lie-group perspective, where the action of $\mathrm{e}^{-\mathrm{i}W\delta}$ is viewed merely as a rotation in Hilbert space, the eigenvalues E_n^W represent weights for the directions represented by the eigenvectors $|n\rangle_W$.

Ensemble-averaged echo signal. To establish a prediction of the echo signal, we first compute the average effect of the scrambling operators W from one of those ensembles on the state $\rho_\mathrm{b}(t)$ during the backward evolution (cf. Eq. (4.4) with $\epsilon = 0$). By means of the transformation matrices $\tilde{U}_{n\mu}$ from (4.14), the matrix elements of $\rho_\mathrm{b}(t)$ in the eigenbasis of H_0 can be written as

$${}_0\langle\mu|\rho_\mathrm{b}(t)|\nu\rangle_0 = \sum_{\alpha,\beta,m,n} \mathrm{e}^{-\mathrm{i}(E_\nu - E_\mu)t}\, \mathrm{e}^{\mathrm{i}(E_n^W - E_m^W)\delta}\, \mathrm{e}^{\mathrm{i}(E_\beta - E_\alpha)\tau}\, {}_0\langle\alpha|\rho_\mathrm{T}|\beta\rangle_0\, \tilde{U}_{m\alpha}\tilde{U}^*_{m\mu}\tilde{U}^*_{n\beta}\tilde{U}_{n\nu}\,. \tag{4.15}$$

The corresponding time-dependent expectation value of the observable A during the backward phase can then be computed as

$$\langle A\rangle_{\rho_\mathrm{b}(t)} = \sum_{\mu,\nu} {}_0\langle\mu|\rho_\mathrm{b}(t)|\nu\rangle_0\, {}_0\langle\nu|A|\mu\rangle_0\,. \tag{4.16}$$

According to the definition of the considered W ensembles, the only random quantities in the relation (4.15) are the $\tilde{U}_{n\mu}$. Similarly as in Chapter 3 (cf. Eq. (3.142)), the ensemble-averaged state is thus essentially determined by the average over four factors of transformation matrix elements. For the CUE and COE considered here, such averages can be evaluated as detailed in Ref. [197] (see also the explicit example from Sec. 2.3.2). Denoting ensemble averages by $\mathbb{E}[\cdots]$ again, we obtain

$$\begin{aligned}\mathbb{E}[\tilde{U}_{n_1\alpha_1}\tilde{U}_{n_2\alpha_2}\tilde{U}^*_{n_1\beta_1}\tilde{U}^*_{n_2\beta_2}] = v_{1,1}\,(\delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2} + \delta_{n_1n_2}\delta_{\alpha_1\beta_2}\delta_{\alpha_2\beta_1})\\ + v_2\,(\delta_{\alpha_1\beta_2}\delta_{\alpha_2\beta_1} + \delta_{n_1n_2}\delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2})\,.\end{aligned} \tag{4.17}$$

Here the so-called symmetry factors $v_{1,1}$ and v_2 take the values $v_{1,1} = \frac{1}{N^2-1}$ and $v_2 = -\frac{1}{N(N^2-1)}$ for the CUE or $v_{1,1} = \frac{N+2}{N(N+1)(N+3)}$ and $v_2 = -\frac{1}{N(N+1)(N+3)}$ for the COE. Since $N \gg 1$ in our setting (see Eq. (2.10)), we can treat both cases simultaneously by settling for the leading-order approximation $v_{1,1} \simeq \frac{1}{N^2}$ and $v_2 \simeq -\frac{1}{N^3}$. Taking the ensemble average in Eq. (4.15), we then find

$$\begin{aligned}\mathbb{E}[{}_0\langle\mu|\rho_\mathrm{b}(t)|\nu\rangle_0] = \mathrm{e}^{\mathrm{i}(E_\nu-E_\mu)(\tau-t)}\, {}_0\langle\mu|\rho_\mathrm{T}|\nu\rangle_0 \left[\frac{1}{N^2}\sum_{m,n}\mathrm{e}^{\mathrm{i}(E_n^W-E_m^W)\delta} - \frac{1}{N^2}\right]\\ + \frac{\delta_{\mu\nu}}{N}\left[1 - \frac{1}{N^2}\sum_{m,n}\mathrm{e}^{\mathrm{i}(E_n^W-E_m^W)\delta}\right].\end{aligned} \tag{4.18}$$

In the first term on the right-hand side of this equation, we identify the matrix elements

$${}_0\langle\mu|\rho_\mathrm{f}(\tau-t)|\nu\rangle_0 = \mathrm{e}^{\mathrm{i}(E_\nu-E_\mu)(\tau-t)}\, {}_0\langle\mu|\rho_\mathrm{T}|\nu\rangle_0 \tag{4.19}$$

of the forward state from (4.2), which would constitute the perfect echo in the absence of inaccuracies. Recalling that we work exclusively within the energy window $I_\mathcal{E}$ from (2.9), we furthermore recognize the matrix elements

$${}_0\langle\mu|\rho_\mathrm{mc}|\nu\rangle_0 = \frac{\delta_{\mu\nu}}{N} \tag{4.20}$$

of the microcanonical density operator (2.5) in the second line of (4.18). To shed light on the remaining terms, we introduce the normalized (or relative) density of states (cf. Eq. (2.11))

$$d_W(E) := \frac{1}{N}\sum_n \delta(E - E_n^W) \tag{4.21}$$

of the scrambling Hamiltonian W as well as its Fourier transform

$$\hat{d}_W(t) := \int \mathrm{d}E\; d_W(E)\,\mathrm{e}^{\mathrm{i}Et}\,. \tag{4.22}$$

Taking all these considerations into account in Eq. (4.18), the ensemble-averaged backward state is thus given by

$$\mathbb{E}[\rho_\mathrm{b}(t)] = \rho_\mathrm{f}(\tau-t)\left[|\hat{d}_W(\delta)|^2 - \frac{1}{N^2}\right] + \rho_\mathrm{mc}\left[1 - |\hat{d}_W(\delta)|^2\right]. \tag{4.23}$$

In view of (2.10), the term proportional to $1/N^2$ here is negligible. With the definition (4.6) and Eq. (4.16), and observing the relation between $\rho(t)$ and $\rho_\mathrm{b}(t)$ as defined below Eq. (4.4), we thus conclude that the ensemble-averaged echo signal relates to the perfect echo as

$$\mathbb{E}[\mathcal{A}(\tau+\delta+t)] = |\hat{d}_W(\delta)|^2\,\mathcal{A}(\tau-t)\,. \tag{4.24}$$

Variance. In the second step of the typicality algorithm, we again consider the fluctuations of the echo signal $\mathcal{A}(\tau+\delta+t)$ for a single realization W of the scrambling Hamiltonian around the average behavior $\mathbb{E}[\mathcal{A}(\tau+\delta+t)]$, i.e., the variance of $\mathcal{A}(\tau+\delta+t)$. In view of Eqs. (4.15) and (4.16), computing this variance amounts to an ensemble average over eight factors of the transformation matrices $\tilde{U}_{n\mu}$ from (4.14). With the aid of Ref. [197], this calculation can be carried out as detailed in Appendix E.6 and leads to

$$\mathbb{E}\Big[\big(\mathcal{A}(\tau+\delta+t)-\mathbb{E}[\mathcal{A}(\tau+\delta+t)]\big)^2\Big] \leq \frac{11(\Delta A)^2}{N}+\mathcal{O}\bigg(\frac{1}{N^2}\bigg)\,. \tag{4.25}$$

As usual, ΔA here denotes the measurement range of A from (2.18). Hence the variance decreases inversely proportionally to the dimension N of the energy window $I_{\mathcal{E}}$ and thus exponentially fast in the degrees of freedom (see Eq. (2.10)).

Typical echo upon imperfect preparation. In this form, the bound in (4.25) combined with Chebyshev's inequality (2.29) implies that, for any given time point t, the probability to observe noticeable deviations between the echo dynamics induced by a single scrambling Hamiltonian W and the average over all W is exponentially suppressed in the degrees of freedom (cf. also Eq. (3.157)). Similarly as in Sec. 3.6.1 (see Eq. (3.159) in particular), we can immediately extend this statement to most time points in any preset time interval. Together with (4.24), we therefore infer that the relative echo signal is given by

$$\frac{\mathcal{A}(\tau+\delta+t)}{\mathcal{A}(\tau-t)} = |\hat{d}_W(\delta)|^2 \tag{4.26}$$

for nearly all imperfections W and times $0 \leq t \leq \tau$. The modifications from the perfect echo arising due to the scrambling phase are thus basically encoded in the Fourier-transformed density of states $\hat{d}_W(t)$. This function $\hat{d}_W(t)$ thus plays the role of a response profile similarly to $g_\lambda(t)$ in Chapter 3 (see Eqs. (3.146) and (3.160)). Evaluating (4.26) at $t=\tau$, the height of the echo peak (4.8), in particular, is typically found to be

$$\mathcal{F}(\delta,0,\tau) = |\hat{d}_W(\delta)|^2\,. \tag{4.27}$$

Hence the relative echo signal and the peak height in particular are given by the squared magnitude of the Fourier transform (4.22) of the scrambling operator's DOS (4.21) evaluated at the scrambling time δ. Remarkably, this relative echo signal is independent of the waiting time τ, in striking contrast to the classical case (see Sec. 4.2 and Refs. [293–295]). This suggests that sensitive dependence on initial conditions, a prime indicator of chaos in classical systems, may not have a direct quantum analog.

As mentioned in the beginning of this subsection, we adopted the working hypothesis that the return state ρ_{R} populates the eigenstates of the scrambling Hamiltonian W approximately homogeneously. If this is not the case and the concrete occupations ${}_W\langle n|\rho_{\mathrm{R}}|n\rangle_W$ are actually known, the procedure from Ref. [100] can be employed to incorporate these imbalances in the echo-signal prediction. The resulting echo signal is still structurally similar to Eq. (4.26), but the integrand of the function $\hat{d}_W(t)$ from (4.22) must be supplemented by an additional weight factor depending on the populations of levels around E, i.e., it becomes the Fourier transform of the "occupation density" $d'_W(E) := \frac{1}{N}\sum_n \delta(E-E_n^W)\,{}_W\langle n|\rho_{\mathrm{R}}|n\rangle_W$. Importantly, the right-hand side of (4.26) will still be independent of t, but it will usually depend on τ now since the occupation density $d'_W(E)$ of $\rho_{\mathrm{R}}=\rho_{\mathrm{f}}(\tau)$ in the eigenbasis of W generally changes as long as the system is out of equilibrium, and even in equilibrium it will keep changing. Nevertheless, if the reference system H_0 equilibrates, then also the modified $d'_W(E)$ will eventually become independent of τ (on a coarse-grained scale) and we will observe a persistent and asymptotically constant echo peak similar to (4.27).

4.3.2 Examples

To verify the somewhat surprising persistence of echo peaks in quantum many-body systems, we consider two explicit examples.

Spin-1/2 XXX chain. The first example is the paradigmatic spin-$\frac{1}{2}$ XXX chain, whose Hamiltonian comprises a one-dimensional lattice of spin degrees of freedom that are coupled via isotropic Heisenberg interactions,

$$H_0 := -\sum_{i=1}^{L-1} \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_{i+1} \,. \tag{4.28}$$

Here $\boldsymbol{\sigma}_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ is a vector of Pauli matrices acting on site i. We prepare the system in the target state $\rho_\mathrm{T} = |\psi\rangle\langle\psi|$ with the Néel state $|\psi\rangle = |{\downarrow}{\uparrow}{\downarrow}{\uparrow}\cdots\rangle$ and monitor the staggered magnetization in the z direction,

$$M_\mathrm{s}^z := \frac{1}{L}\sum_{i=1}^{L} (-1)^i \, \sigma_i^z \,. \tag{4.29}$$

Note that the initial state significantly occupies levels across the entire spectrum of H_0, hence formally the energy window from (2.9) spans the full Hilbert space. (Recall that we do not rely on the—otherwise often employed—assumption of a constant DOS in the present setting.) The scrambling Hamiltonian W is chosen of the form

$$W = \sum_{i<j} \sum_{\alpha,\beta=1}^{3} J_{ij}^{\alpha\beta} \, \sigma_i^\alpha \, \sigma_j^\beta \tag{4.30}$$

with the couplings $J_{ij}^{\alpha\beta}$ drawn independently from a standard normal distribution. On the one hand, this scrambling Hamiltonian thus complies with the general structure of the spin model and its interactions. On the other hand, the erratic, uncontrolled nature of the modeled imperfections is reflected in the random choice of the couplings. Note that in the form (4.30), the operator norm of W scales as L^2. Since we will only consider fixed values of L in the following, this is of no further concern. However, if one intends to compare the effect of the perturbation across different system sizes, the scaling of L can of course make a difference. In particular, in the more abstract perspective of W as a generator of Hilbert space rotations (see below Eqs. (4.13) and (4.14)), the

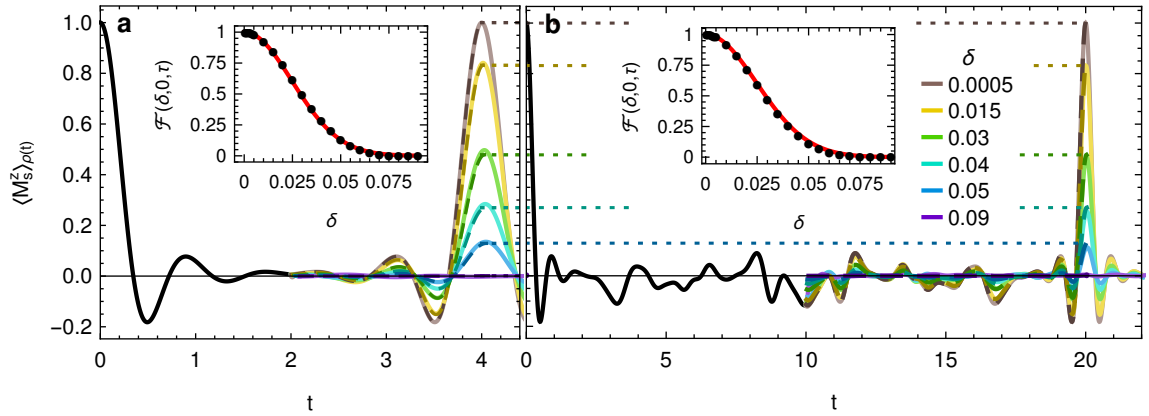


Figure 4.4: Echo dynamics under the imperfect-preparation protocol (4.13) in the spin-$\frac{1}{2}$ XXX chain from (4.28) for various scrambling times δ and waiting times **a.** $\tau = 2$ and **b.** $\tau = 10$. The observable is the staggered magnetization (4.29), the initial (target) state $\rho_\mathrm{T} = |\psi\rangle\langle\psi|$ with $|\psi\rangle = |{\downarrow}{\uparrow}{\downarrow}{\uparrow}\cdots\rangle$ is the Néel state. Solid: Numerical results from exact diagonalization of a chain of length $L = 14$ for the forward (black) and backward (color-coded as indicated in the right panel) dynamics. Dashed: Analytical prediction (4.26) for the perturbed backward dynamics using the analytically known Fourier-transformed DOS $\hat{d}_W(t)$ from (4.32) and $\langle M_\mathrm{s}^z \rangle_{\rho_\mathrm{mc}} = 0$ by symmetry. The dashed lines lie essentially on top of the solid ones and are therefore sometimes hardly distinguishable. Dotted: Levels of the echo peak $\mathcal{F}(\delta, 0, \tau)\,\mathcal{A}(0)$, showing that the height is independent of τ. Insets: Comparison of the numerical (dots) and theoretical (solid line, cf. Eq. (4.27)) echo peak height $\mathcal{F}(\delta, 0, \tau)$ from (4.8) as a function of the scrambling time δ.

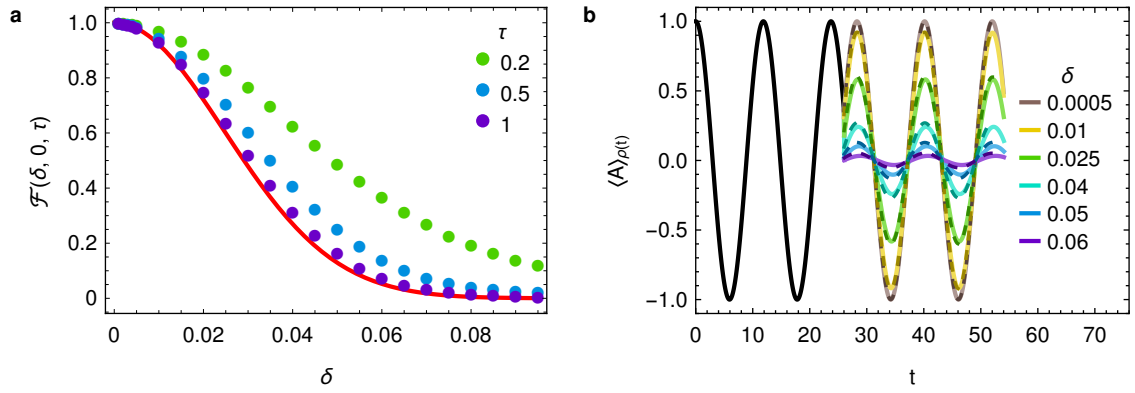


Figure 4.5: Echo dynamics under the imperfect-preparation protocol (4.13) in the spin-$\frac{1}{2}$ XXX chain from (4.28). **a.** Relative echo peak height (4.8) as a function of δ for the same setup as in Fig. 4.4, but very short waiting times τ. Dots: Numerical results from exact diagonalization for various τ as indicated. Solid: Theoretical (τ-independent) prediction (4.27) using $\hat{d}_W(t)$ from (4.32) and $\langle A \rangle_{\rho_{\mathrm{mc}}} = 0$. **b.** Time-dependent expectation values for the observable $A = (|\nu_1\rangle_{00}\langle\nu_2| + |\nu_2\rangle_{00}\langle\nu_1|)/2$ and target state $\rho_{\mathrm{T}} = |\phi\rangle\langle\phi|$ with $|\phi\rangle = (|\nu_1\rangle_0 + |\nu_2\rangle_0)/\sqrt{2}$ for $\tau = 26$ and various scrambling times δ, where $\nu_1 = 7936$ and $\nu_2 = 8448$ (such that $\nu_1 + \nu_2 = 2^{14} = N$).

operator W essentially encodes a rotation axis and should therefore be normalized such that its magnitude is independent of the system size.

Of particular convenience for our present purposes is the fact that the DOS $d_W(E)$ of W is known exactly for $L \gg 1$ [297]. Namely, it assumes the form of a Gaussian of mean zero and variance

$$\sigma^2_{\mathrm{DOS}} = 9L(L-1)/2\,. \tag{4.31}$$

For its Fourier transform from (4.22), we therefore conclude that it is also Gaussian,

$$\hat{d}_W(t) = \mathrm{e}^{-\sigma^2_{\mathrm{DOS}} t^2/2}\,. \tag{4.32}$$

This allows us to compare our theoretical predictions (4.26) and (4.27) to a numerical simulation of the system (4.28) subject to an imperfect-preparation perturbation (4.30) without any sort of fitting. For a chain of length $L = 14$, observing the staggered magnetization (4.29) and starting from the Néel state, such a comparison is carried out in Fig. 4.4. For the various choices of the scrambling and waiting times δ and τ, respectively, the analytical prediction is in excellent agreement with the numerical results and all features of the theory are recovered: The perturbed echo signal resembles the perfect echo, but is attenuated as quantified by the Fourier-transformed DOS of the scrambling operator W: increasingly strongly with increasing δ, but independently of τ.

However, we remark that there is in fact a residual dependence on τ for very short values $\tau \lesssim 1$ below the relaxation time. This is shown in Fig. 4.5a, which displays the relative echo peak height (4.8) as a function of δ for various $\tau \leq 1$. We observe that the mitigation of the echo peak is less than theoretically predicted, with the discrepancies increasing as τ is decreased. As explained below Eq. (4.27), such deviations from the prediction (4.27) for times below the system's relaxation time can occur due to a violation of the assumption that the return state ρ_{R} is uniformly distributed across the levels $|n\rangle_W$ of W (see also third paragraph of this Sec. 4.3.1). Due to its few-body spin structure, the scrambling Hamiltonian (4.30) is still sensitive to polarizations of the individual lattice sites, and the special pattern of the Néel state is only gradually washed out during the forward evolution. More subtlely, and as usual in typicality arguments (cf. Sec. 2.3), deviations may also be caused by dependencies or correlations between H, W, A, and ρ_{T} that are disregarded by the chosen ensemble of scrambling operators. Again, if such correlations play a role, then their effect is expected to be most pronounced when ρ_{R} is still out of equilibrium because this is the regime where correlations are indeed crucial. Similar dependencies for relatively small values of τ have also been observed, for instance, in Refs. [295, 296, 298]; we will come back to this point in the discussion in Sec. 4.5.

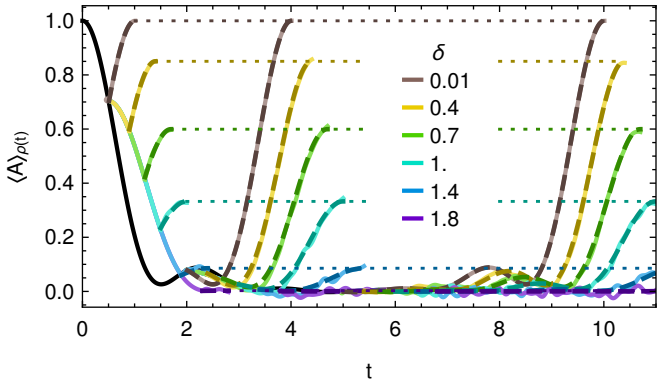


Figure 4.6: Echo dynamics under the imperfect-preparation protocol (4.13) in a random matrix model with H_0 as in (3.27) with level spacing $\varepsilon = 1/512$. The observable A is a random matrix from the GOE, and the target state $\rho_{\mathrm{T}} = |\psi\rangle\langle\psi|$ is constructed according to the dynamical typicality prescription (2.39) with $\kappa = 1$ and Π projecting onto the central 2048 states. Solid: Numerical time evolution from exact diagonalization for a scrambling Hamiltonian W drawn at random from the GUE, waiting times $\tau = 0.5, 2, 5$, and various scrambling times δ as indicated. Dashed: Corresponding theoretical prediction (4.26) with $\hat{d}_W(t)$ from (4.33) and $\langle A\rangle_{\mathrm{mc}} = 0$. Dotted: Levels of the echo peak height (4.27).

Then again, we point out that small-τ deviations of this kind are *not* found universally in the system (4.28) with the perturbation (4.30). An example without such deviations is the setup from Fig. 4.5b, which involves the same H_0 from (4.28) and W from (4.30) as in Figs. 4.4 and 4.5a. However, the initial state is now given by a superposition of two eigenstates $|\nu_1\rangle_0$ and $|\nu_2\rangle_0$ of the reference Hamiltonian H_0 with $\nu_1 = 7936$ and $\nu_2 = 8448$, where the indices run from 1 to $N = 2^L = 16\,384$ and increase with increasing energy E_ν. The observable $A = (|\nu_1\rangle_{00}\langle\nu_2| + |\nu_2\rangle_{00}\langle\nu_1|)/2$ quantifies cross-correlations between these two eigenstates and exhibits perpetual oscillations during the forward evolution since the state $\rho_{\mathrm{f}}(t)$ (cf. Eq. (4.2)) is periodic with frequency $E_{\nu_2} - E_{\nu_1}$. In particular, the system does not equilibrate in this scenario, but the analytical prediction again describes the numerically observed behavior remarkably well.

Random matrix example. As a second illustrational example, we consider an artificial, but clean random matrix system which, by construction, does not exhibit any of the aforementioned correlations. For the reference system, we choose the Hamiltonian H_0 from (3.27) with a level spacing $\varepsilon = 1/512$ and dimension $N = 2^{14} = 16\,384$. The considered observable is drawn randomly from the Gaussian Orthogonal Ensemble (GOE). The initial state is $\rho(0) = |\psi\rangle\langle\psi|$ with $|\psi\rangle$ generated as in (2.39) with $\kappa = 1$ and Π projecting onto the central 2048 states. The scrambling Hamiltonian W is a random matrix from the Gaussian Unitary Ensemble (GUE), so $d_W(E)$ is a semicircle of radius 2 [163] and the corresponding Fourier transform from (4.22) is given by

$$\hat{d}_W(t) = \frac{J_1(2t)}{t} \tag{4.33}$$

with the Bessel function $J_1(x)$ of the first kind of order 1. In essence, this is thus an explicit realization of the ensembles of imperfections introduced around Eq. (4.14).

As in the previous examples from Figs. 4.4 and 4.5, all parameters in (4.26) are known explicitly and the analytical prediction is free of any fit parameters. Theory and numerics are in excellent agreement, even for the shortest waiting time $\tau = 0.5$ where the system has not yet equilibrated upon initiation of the time reversal.

4.4 Imperfect reversal

A second special case of the general echo protocol (4.5) involves only inaccuracies of the imperfect-reversal type (see below Eq. (4.4)), such that $\delta = 0$ and $\epsilon > 0$. Hence the echo protocol for this section reads

$$\rho_{\mathrm{T}} \xrightarrow[H_0]{\tau} \rho_{\mathrm{R}} \xrightarrow[-H_0+\epsilon V]{\tau} \rho'_{\mathrm{T}}\,, \tag{4.34}$$

i.e., we have an imperfect version of the time-reversed Hamiltonian $-H_0$ acting continually during the backward phase, but the state at the point of reversal ρ_R is assumed to be unaffected. As mentioned in Sec. 4.1, this scenario is of particular relevance for echo experiments since in practice only the dominant part of the respective Hamiltonian can be effectively reversed and the precision to carry out the required sophisticated manipulations is ultimately limited.

We remark that (4.34) entails as a special case the so-called Loschmidt echo [284, 299, 300] when choosing $A = \rho_\mathrm{T} = |\psi\rangle\langle\psi|$ with a pure state $|\psi\rangle$. This Loschmidt echo was one of the first popular quantifiers for irreversibility in quantum systems, but it is notoriously difficult to measure in macroscopic systems [285, 298, 301] (see, however, Ref. [300] for an experimental realization).

4.4.1 Typical echo signal

Prerequisites and perturbation ensembles. In the imperfect-reversal protocol (4.34), the structure of the Hamiltonian $\tilde{H}_\epsilon := -H_0 + \epsilon V$ governing the backward evolution is akin to Eq. (3.1), i.e., the setup for perturbed relaxation considered in Chapter 3. Indeed, it is natural to take the inaccuracies present during the backward phase to be "ordinary" perturbations with a structure similar to those introduced in Secs. 3.2 and 3.3. In this spirit, we will consider the same classes of perturbation ensembles from Sec. 3.3 for the operators V in (4.34) again, as defined specifically in Eqs. (3.23) and (3.24). We recall that these comprise, in particular, banded and sparse perturbation matrices $V_{\mu\nu} = {}_0\langle\mu|V|\nu\rangle_0$, both of which are common features in realistic perturbations [130, 134, 209, 219, 230, 232–235].

Consequently, the reference system H_0 should satisfy the same prerequisites as formulated in Sec. 3.2. Notably, and contrary to Sec. 4.3, we require again that the density of states within the energy window (2.9) should be approximately constant as in (2.13).

Ensemble-averaged echo signal and variance. From a mathematical point of view, evaluating the ensemble-averaged echo signal and its variance for the protocol (4.34) and the perturbation ensembles from Sec. 3.3 is merely a corollary of the results from Sec. 3.6.1, even though the physical perspective is quite distinct. By analogy with (3.5), we denote the overlap between an eigenstate $|n\rangle_\epsilon$ of $H_0 - \epsilon V$ and $|\mu\rangle_0$ of H_0 by $U_{n\mu} := {}_\epsilon\langle n|\mu\rangle_0$. The corresponding eigenvalues are E_n^ϵ and E_μ, respectively. Focusing on the expectation values of the observable A during the backward phase, we then obtain

$$\langle A\rangle_{\rho_\mathrm{b}(t)} = \sum_{m,n} \sum_{\substack{\mu_1,\mu_2,\\ \nu_1,\nu_2}} \mathrm{e}^{-\mathrm{i}(E_n^\epsilon - E_m^\epsilon)t}\, \mathrm{e}^{\mathrm{i}(E_{\nu_2} - E_{\mu_1})\tau}\, {}_0\langle\mu_1|\rho_\mathrm{T}|\nu_2\rangle_0\, {}_0\langle\mu_2|A|\nu_1\rangle_0\, U_{m\mu_1} U_{n\mu_2} U^*_{m\nu_1} U^*_{n\nu_2}\,. \tag{4.35}$$

To take the ensemble average, we restrict to the regime of small ϵ since the inaccuracies should not prevail in the resulting dynamics. Employing Prerequisite (v) from Sec. 3.2 to replace $E_n^\epsilon - E_m^\epsilon$ by $E_n - E_m$ upon averaging, we can then substitute the result (3.35) for the fourth moment of eigenvector overlaps and find that

$$\begin{aligned}\mathbb{E}[\langle A\rangle_{\rho_\mathrm{b}(t)}] = {}& \sum_n d^{mn}_{\mu\nu}\, {}_0\langle\mu|\rho_\mathrm{T}|\mu\rangle_0\, {}_0\langle\nu|A|\nu\rangle_0 \\ & + \sum_{m,n} \mathrm{e}^{-\mathrm{i}(E_n - E_m)t} \sum_{\mu,\nu} \Big[\mathrm{e}^{\mathrm{i}(E_\nu - E_\mu)\tau}\, d^{mn}_{\mu\nu}\, {}_0\langle\mu|\rho_\mathrm{T}|\nu\rangle_0 {}_0\langle\nu|A|\mu\rangle_0 + f^{mn}_{\mu\nu}\, {}_0\langle\mu|\rho_\mathrm{T}|\mu\rangle_0\, {}_0\langle\nu|A|\nu\rangle_0\Big]\end{aligned} \tag{4.36}$$

with $d^{mn}_{\mu\nu}$ and $f^{mn}_{\mu\nu}$ as defined in Eqs. (3.35b) and (3.35c), respectively. The remaining sums in this relation can be calculated similarly as in Sec. 3.6.1 (see Eqs. (3.144)–(3.148)), meaning that we eventually arrive at

$$\mathbb{E}[\langle A\rangle_{\rho_\mathrm{b}(t)}] - \langle A\rangle_{\tilde{\rho}_\epsilon} = \mathrm{e}^{-2\pi\alpha_v |t| \epsilon^2} \left[\langle A\rangle_{\rho_\mathrm{f}(\tau - t)} - \langle A\rangle_{\tilde{\rho}_\epsilon}\right] + R(t) \tag{4.37}$$

with the intrinsic perturbation strength $\alpha_v = \sigma_v^2/\varepsilon$ from (3.12), $\rho_\mathrm{f}(t)$ from (4.2), and where $\tilde{\rho}_\epsilon$ and $R(t)$ are defined as in Eqs. (3.150) and (3.151), respectively, observing that $\rho(0) = \rho_\mathrm{T}$. As explained in Sec. 3.6.3, the term $R(t)$ is generally negligible (see the discussion around Eqs. (3.185) and (3.186)) and we can commonly identify $\langle A\rangle_{\tilde{\rho}_\epsilon} = \langle A\rangle_{\rho_\mathrm{mc}}$ (see above Eq. (3.188)). Defining

$\tilde{\alpha}_v := 2\pi\alpha_v$ and employing the definition (4.6), the ensemble-averaged echo signal thus takes the form

$$\mathbb{E}[\mathcal{A}(\tau+t)] = \mathrm{e}^{-\tilde{\alpha}_v t\epsilon^2}\,\mathcal{A}(\tau-t) \qquad (0 \leq t \leq \tau)\,. \tag{4.38}$$

Computing the variance of $\mathcal{A}(\tau+t)$ within the perturbation ensemble proceeds analogously to Sec. 3.6.1 again. Similarly as in Eq. (3.154), we obtain

$$\mathbb{E}\Big[\big(\mathcal{A}(\tau+t) - \mathbb{E}[\mathcal{A}(\tau+t)]\big)^2\Big] \leq \frac{c(\Delta A)^2}{N_v} \tag{4.39}$$

with some constant $c \lesssim \mathcal{O}(10^3)$, the spectral range ΔA of the observable, and the number N_v of unperturbed levels mixed by the perturbation. Considering (3.9), we conclude that the variance decreases again exponentially with the system's degrees of freedom.

Typical echo upon imperfect reversal. Invoking the same machinery as usual, the bound (4.39) on the variance of $\mathcal{A}(\tau+t)$ implies, via Chebyshev's and Markov's inequalities, that the observable echo signal for an individual (sufficiently weak) perturbation drawn at random from any admitted ensemble will be practically indistinguishable from the average (4.38) for nearly all times $t \in [0,\tau]$. Consequently, the relative echo signal in the imperfect-reversal scenario is found to be excellently approximated by

$$\frac{\mathcal{A}(\tau+t)}{\mathcal{A}(\tau-t)} = \mathrm{e}^{-\tilde{\alpha}_v t\epsilon^2} \tag{4.40}$$

for the vast majority of imperfections V and times $0 \leq t \leq \tau$. For the relative height of the echo peak (4.8), this implies

$$\mathcal{F}(0,\epsilon,\tau) = \mathrm{e}^{-\tilde{\alpha}_v \tau\epsilon^2}\,. \tag{4.41}$$

In contrast to the imperfections at the point of reversal discussed in Sec. 4.3, the continuous exposure to perturbations during the backward evolution is reflected in a dependence of the peak height on the waiting time τ. The mitigation of the echo peak is exponential in the waiting time τ (at fixed ϵ) and Gaussian in the magnitude of inaccuracies ϵ (at fixed τ).

We confined ourselves to the limit of weak imperfections here. In view of the considerations from Sec. 3.6.2, the result (4.41) is thus expected to describe the observable echo peak for sufficiently small ϵ or sufficiently large τ. Furthermore, the relations (4.40) and (4.41) are generalized straightforwardly to the regime of larger perturbation strengths and/or smaller times by replacing the right-hand sides with the corresponding response profile (3.146) (i.e., $|g_\epsilon(t)|^2$ and $|g_\epsilon(\tau)|^2$, respectively).

4.4.2 Examples

Spin-1/2 XXX chain. To validate the theoretical predictions (4.40) and (4.41), we consider the spin-$\frac{1}{2}$ XXX chain with Hamiltonian H_0 from (4.28) again. As before, we focus on the staggered magnetization M_s^z from (4.29), and we model the imperfections V acting during the backward evolution by an operator similar to (4.30), i.e., $V = \sum_{i<j}\sum_{\alpha,\beta} J_{ij}^{\alpha\beta}\sigma_i^\alpha\sigma_j^\beta$ with independent, Gaussian distributed $J_{ij}^{\alpha\beta}$ of vanishing mean and unit variance. The target state ρ_T is based on the Néel state $|{\downarrow\uparrow\downarrow\uparrow}\cdots\rangle$, too, but in order to satisfy the requirement of a homogeneous density of states of H_0 (cf. Prerequisite (i)), we apply an additional Gaussian filter $\Pi_{\mathcal{E},\Delta_\mathcal{E}}$ (cf. Eq. (3.193) and Refs. [133, 201, 202]) to obtain $\rho_\mathrm{T} = |\psi\rangle\langle\psi|$ with

$$|\psi\rangle \propto \Pi_{0,\Delta_\mathcal{E}}|{\downarrow\uparrow\downarrow\uparrow}\cdots\rangle \tag{4.42}$$

and $\Delta_\mathcal{E} = 1.3$. For the considered chain of length $L = 14$, this means that roughly 15 % of the total $N = 2^{14}$ levels have energies $E_\mu \in [-\Delta_\mathcal{E}, \Delta_\mathcal{E}]$. The initially observed staggered magnetization $\langle A\rangle_{\rho_\mathrm{T}}$ is thus reduced from its maximal value of 1, but still noticeably out of equilibrium (see Fig. 4.7).

To estimate the relevant intrinsic perturbation strength $\tilde{\alpha}_v$ from (3.12), we focus on the window $I_\mathcal{E} = [-2\Delta_\mathcal{E}, 2\Delta_\mathcal{E}]$, which includes about 95 % of the weight of the Gaussian filter. Upon exact diagonalization of H_0 from (4.28), we then find $\varepsilon^{-1} \approx 962$ for the density of states (inverse mean level spacing). To determine the scale σ_v^2 from (3.11), we inspect the matrix elements $V_{\mu\nu} =$

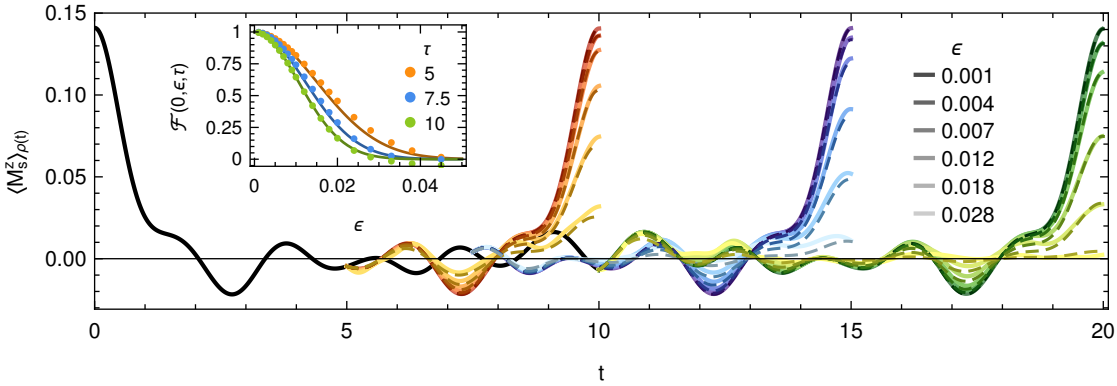


Figure 4.7: Echo dynamics under the imperfect-reversal protocol (4.34) in the spin-$\frac{1}{2}$ XXX chain from (4.28) subject to imperfections V of the form (4.30) for reversal times $\tau = 5, 7.5, 10$ and different perturbation strengths ϵ as indicated. The observable is the staggered magnetization (4.29), the target state is $\rho_{\mathrm{T}} = |\psi\rangle\langle\psi|$ with the filtered Néel state $|\psi\rangle$ from (4.42). Solid: Numerical results from exact diagonalization of a chain of length $L = 14$ for the forward (black) and backward (colored) dynamics for $\tau = 5$ (red-toned curves), $\tau = 7.5$ (blue-toned), and $\tau = 10$ (green-toned). Dashed: Corresponding analytical prediction (4.40) using the numerically determined value $\tilde{\alpha}_v = 440$ and $\langle M_{\mathrm{s}}^z \rangle_{\rho_{\mathrm{mc}}} = 0$. Inset: Numerical (dots) and theoretically predicted (lines) relative echo peak heights (4.8) as a function of the perturbation strength ϵ for the three values of τ.

${}_0\langle\mu|V|\nu\rangle_0$ in the window $I_{\mathcal{E}}$ and average within a band of 1000 states (i.e., for $0 < |E_\mu - E_\nu| \lesssim 1$). This yields $\sigma_v^2 \approx 0.0729$ and hence $\tilde{\alpha}_v = 2\pi\sigma_v^2\varepsilon^{-1} \approx 440$. The thermal expectation value is $\langle M_{\mathrm{s}}^z \rangle_{\rho_{\mathrm{mc}}} = 0$ as before. Consequently, all parameters in the predictions (4.40) and (4.41) are again known explicitly, and we can compare theory and numerics without any fit parameters.

This comparison for the described setup is exemplified in Fig. 4.7, and we observe good agreement of the prediction with the numerical data. Minor deviations occur for small τ and large ϵ. As explained below Eq. (4.41), this is the regime where the approximation $\mathrm{e}^{-\tilde{\alpha}_v t\epsilon^2}$ on the right-hand side of (4.40) must eventually be replaced by the more general response profile (3.146).

Experiments. In standard applications such as magnetic resonance imaging, it is commonly taken for granted that the echo signal decays exponentially with the waiting time τ [288]. The concrete "imperfections" acting in the corresponding setting are remnant interactions among the spins and with the environment (see also the discussion below Eq. (4.5)). The theory developed in this section demonstrates that such an exponential decay can indeed be expected generically as long as the influence of the imperfections is sufficiently weak.

An explicit echo experiment for *interacting* spins was conducted, for example, by Levstein et al. in Ref. [43], investigating nuclear spins in polycristalline samples of cymantrene and ferrocene. The authors observe that the echo peak height indeed decays exponentially with the waiting time τ in the cymantrene sample, whereas their data for the ferrocene sample lead them to suggest a Gaussian decay. As a matter of fact, the imperfections (i.e., the experimentally nonreversible components of the Hamiltonian) are much stronger in the latter case [43], such that deviations from the exponential decay are in line with our general observations below Eq. (4.41). Moreover, in view of the limited size of the data set (only four points), the crossover towards a Gaussian-like shape is compatible with our prediction of an eventually Bessel-like response profile (3.165) since both functional forms could fit the data equally well.

4.5 Combined effect and discussion

Combining imperfect preparation and reversal. As motivated in Sec. 4.1, in any real experiment we will presumably encounter both imperfect-preparation and imperfect-reversal inaccuracies. Hence the generally appropriate echo protocol is of the form (4.5). As long as the underlying mechanisms behind the inaccuracies W (imperfect preparation) and V (imperfect reversal) are

independent in the sense that their effect can be modeled by statistically uncorrelated random operators W and V, it is straightforward to merge the results from Secs. 4.3 and Secs. 4.4. Such an assumption is not unreasonable because the two types of imperfections act during different stages of the protocol and under somewhat distinct circumstances. In particular, the dominant contributions during the backward evolution (imperfect reversal) will usually be due to unreverted parts of the Hamiltonian, whereas the inaccuracies during the scrambling phase (imperfect preparation) are mostly due to experimental limitations such as a finite switching time from H_0 to $-H_0$ and improper isolation during that change.

The relative echo signal under the joint, but independent influence of both imperfections is then simply obtained by multiplying the right-hand sides of Eq. (4.26) and (4.40), i.e.,

$$\boxed{\frac{\mathcal{A}(\tau+\delta+t)}{\mathcal{A}(\tau-t)} = |\hat{d}_W(\delta)|^2 \, \mathrm{e}^{-\tilde{\alpha}_v t \epsilon^2}}. \tag{4.43}$$

This prediction of the backward-phase dynamics for the general echo protocol (4.5) represents the main result of this chapter. Likewise, the relative echo peak height is then given by

$$\mathcal{F}(\delta,\epsilon,\tau) = |\hat{d}_W(\delta)|^2 \, \mathrm{e}^{-\tilde{\alpha}_v \tau \epsilon^2} \, . \tag{4.44}$$

We emphasize once more that both types of inaccuracies typically attenuate the observable echo signal, and for both of them the mitigation becomes stronger with growing magnitude δ or ϵ, but only the continuously acting inaccuracies of the imperfect-reversal type lead to a stronger suppression with increasing waiting time τ.

Relation to previous works. As already mentioned in Sec. 4.1, echo protocols have been explored in a variety of experimental setups; see, for instance, Refs. [37–44, 287]. Then again, there are a few more theoretical studies of echo protocols of the forms (4.13) or (4.34) in the context of quantum many-body systems [285, 295, 296, 298, 301], which are also oriented more strongly towards extracting general principles. Yet these studies assume a somewhat different perspective than our present approach.

Elsayed, Fine, and co-workers [295, 296] concentrated mostly on similarities and differences in the echo peak height $\mathcal{F}(\delta,0,\tau)$ from (4.8) between spin models in classical and quantum mechanics when adopting the imperfect-preparation scenario (cf. Secs. 4.2 and 4.3). Based on predecessor works [293, 294], they reinforce the exponential attenuation of the echo peak height in classical systems. In the quantum setting, by contrast, they find a power-law attenuation with τ for short times τ. In all their numerical simulations of quantum systems, however, $\mathcal{F}(\delta,0,\tau)$ never fully decays as a function of τ and instead eventually settles down to distinctly nonzero values, in accordance with the asymptotic persistence of quantum echoes we identified in Sec. 4.3. (Recall that a residual dependence for short τ below the relaxation time does not contradict our analysis there as explained below Eq. (4.27).)

The studies [285, 298, 301] by Schmitt, Kehrein, and co-workers focus mostly on the thermodynamic limit in quantum systems and explore which of the classical chaos indicators carry over to the quantum setting in formally infinite systems. Our present investigation, which explicitly considers large, but finite systems, therefore assumes a complementary perspective. In particular, Refs. [285, 298] find a persistent contribution for the imperfect-preparation setting in their numerical simulations of finite systems, but employ a somewhat provisional extrapolating projection to argue that this persistent echo vanishes in the thermodynamic limit.

Our results from Sec. 4.3 suggest that the scaling of the scrambling Hamiltonian W with the system size (e.g., the degrees of freedom f) is decisive in this context. If W grows extensively with f, meaning that the relative density of states $d_W(E)$ from (4.21) broadens accordingly, then its Fourier transform $\hat{d}_W(t)$ from (4.22) will decay ever faster as f increases and will eventually become zero instantaneously for all $t > 0$. In contrast, if W does not grow extensively with f and hence the (effective) width of $d_W(E)$ is bounded for all f, then also the echo peaks are predicted to persist for sufficiently small scrambling times δ. A similar observation that the limits $\delta \, \|W\| \to 0$ and $f \to \infty$ do not commute was also made in Ref. [298]. Which type of scaling is appropriate

depends to some extent on the character of the imperfections in a concrete setup and also on the physical motivation for considering imperfect echoes in the first place (see also the discussion below Eq. (4.30)). Recalling that $d_W(E)$ is actually the relative density of states of W in the significantly populated energy window $I_\mathcal{E}$ (cf. Eq. (2.9)), an unbounded broadening of this energy window with f is unphysical because it would entail a macroscopically unrealistic Schrödinger-cat state. This is reinforced by the extension discussed below Eq. (4.27), whereby it is in fact the Fourier transform of the return state's energy distribution $d'_W(E)$ which encodes the effect of inaccuracies. In any case, we point out that chaos theory in classical mechanics does usually not operate in the thermodynamic limit and actually often considers rather small systems. Hence working in the thermodynamic limit is generally not an immediate imperative when comparing chaotic dynamics in classical and quantum systems.

An exemplary investigation of the imperfect-reversal setting in the transverse-field Ising model was conducted in Ref. [301]. The authors suggested a classification of systems into "irreversible" or "reversible" depending on whether the echo peak decays exponentially with τ ("irreversible") or slower ("reversible"). In this sense, the findings from Sec. 4.4 in conjunction with the analysis from Sec. 3.6.2, which showed that the exponential decay of the response profile prevails at large times, indicate that quantum many-body systems are generically "irreversible." However, like all other potential indicators of chaos and irreversibility, such a classification does certainly not cover all practically relevant aspects.

Another such indicator, which gained some popularity recently, are so-called out-of-time-ordered correlators (OTOCs) [302, 303], which are believed to encode a rough analog of Lyapunov exponents from classical chaos theory in quantum systems, even though this correspondence is far from complete [304, 305]. Due to their considerably more involved structure, comprising products of two different observables evaluated at different times, we do not see an immediate relation to our present investigation of echo dynamics and thus refrain from attempting any more detailed comparison. The same also applies to other proposed indicators of quantum chaos such as level statistics [163, 233] or eigenstate thermalization (see Sec. 2.2.2).

Finally, from a methodological point of view, the imperfect-preparation scenario is related to the transportless equilibration theory for isolated many-body quantum systems from Ref. [96] (see also Refs. [97, 100] and Sec. 2.3.2). In this context, the scrambling phase can be understood as a typical relaxation process under the action of the scrambling Hamiltonian, and the system's past and future (i.e., the forward and backward evolution stages) are somewhat trivial time translations of the state that could be gauged away in principle by a redefinition of the observable (see also Ref. [286]). Similarly, and as already mentioned in Sec. 4.4, the imperfect-reversal scenario can essentially be seen as a corollary of the perturbed relaxation theory from Chapter 3 and Ref. [226].

Conclusions. This chapter's study of echo dynamics in isolated many-body quantum systems dealt with two types of imperfections accompanying the concomitant effective time reversal. The general setup stars a quantum system prepared in a nonequilibrium state and evolving for a certain waiting time τ, followed by another period of duration τ during which it evolves under the time-reversed Hamiltonian. The first type of inaccuracies, named imperfect preparation, entails a distortion of the system's state at the point of reversal. The second type, called imperfect reversal, emulates an inaccurate implementation of the time-reversed Hamiltonian. In the spirit of typicality methods, we modeled these imperfections by suitable random operators, calculated their average effect, and showed that the individually observed behavior under most of these inaccuracies is practically indistinguishable from the average. This promotes the average behavior to a prediction for a single run of the suggested protocol.

Both types of imperfections generically lead to an attenuation of the observable echo signal and the peak height after completion of the entire protocol. For sufficiently small inaccuracies of the imperfect-reversal type, the suppression of the peak is Gaussian in the magnitude ϵ of the perturbation and exponential in the waiting time τ, in line with the commonly observed behavior in applications. The inaccuracies of the imperfect-preparation type, in contrast, do not entail such a decay with τ. Whereas a mild dependence of the observable peak height on τ can occur for small τ below the relaxation time, the peaks are predicted to become independent of τ in the long run and to persist *ad infinitum*. This behavior of quantum systems is manifestly different from similar classical examples as demonstrated in Sec. 4.2 and Refs. [293–295].

A detailed test of our analytical predictions against numerical simulations of a spin-$\frac{1}{2}$ XXX chain revealed good agreement, notably without resorting to any kind of fitting procedure. Another noteworthy aspect of these numerical verifications is that the exemplary imperfections V and W were of similar structure and both given by (4.30), but their modeling in terms of random matrix ensembles was quite different. In particular, the crucial characteristic of the W ensembles is the (relative) density of states $d_W(E)$ from (4.21), whereas the crucial characteristic of the V ensembles is the perturbation profile (3.10) or, for the case of sufficiently weak perturbations, the intrinsic strength (3.12). In view of the convincing agreement between the respective theories and numerics in Figs. 4.4, 4.5 and 4.7, the explicitly employed imperfections from (4.30) can thus be embedded into both ensembles, with each of them highlighting or sensing different features of the "true" imperfections.

Changing perspectives, echo protocols such as (4.5) may be seen as an example for a driven system, i.e., a setup in which some external control parameter can be manipulated such that certain components of the Hamiltonian vary in time. The echo protocols (4.13) and (4.34) considered in this chapter involve switching between two distinct (constant) Hamiltonians (or three for the combined protocol (4.5)), so the resulting time dependence is still rather simple. Carrying on along these lines, an inspection of different and more versatile driving protocols will be at the heart of the ensuing Chapter 5.

5 Driven systems

The relaxation theory from Chapter 3 describes how many-body quantum systems approach equilibrium under the influence of reasonably weak perturbations, but still being completely isolated and on their own. In particular, system parameters like interaction strengths or force fields were assumed to be fixed once and for all.

In this chapter, we will soften these restrictions to some extent by studying the response of a given system to perturbations with a time-dependent amplitude. The general procedure is already familiar from Chapter 3: Starting from the known behavior of some reference system, we will derive predictions for the time-dependent expectation values of experimentally realistic observables for a related system subject to time-dependent perturbations. Specific applications include periodically driven many-body quantum systems [282, 283] and finite-time quantum quenches [278–281].

The derivation will build again on a typicality argument, supplemented by a suitable transformation of the time-dependent problem into a family of time-independent problems to which previous methods and results can be applied. The resulting predictions of the dynamics will be most reliable for relatively short times, which can nonetheless extend significantly beyond the characteristic time scale of the driving in certain parameter regimes. Notably, our present approach is thus complementary to the majority of studies for periodically driven systems from the literature, which predominantly focus on the long-time behavior at stroboscopic times, meaning that the dynamics is assessed at integer multiples of the driving period only (see, for example, Refs. [306–309]).

In Sec. 5.1, we will describe the setup more concretely and outline the computational strategy. In Sec. 5.2 we will collect prerequisites of the derivation and assess the expected validity of the eventual theory, which in turn will be established in Sec. 5.3. We will then compare the theory to concrete examples in Sec. 5.4 and finally discuss the result in a broader context in Sec. 5.5.

5.1 Setup and methods

Goal. Similarly as in Chapter 3, we intend to investigate the dynamics of many-body quantum systems under the influence of weak-to-moderate perturbations, meaning that the total Hamiltonian is composed of two operators: a reference Hamiltonian H_0 and a perturbation V, both of which should exhibit basically the same properties as before. Contrary to Chapter 3, however, we now allow the coupling strength $\lambda = \lambda(t)$ to be time dependent, hence the total Hamiltonian

$$H(t) := H_0 + \lambda(t)\, V \tag{5.1}$$

becomes time dependent, too. The function $\lambda(t)$ will be called the *driving protocol*. As before, we study the dynamics in terms of expectation values of experimentally realistic observables A and assume that the system is prepared in some pure or mixed initial state $\rho(0)$ with a well-defined macroscopic energy (see Sec. 2.1). The state

$$\rho(t) = \mathcal{U}(t)\, \rho(0)\, \mathcal{U}(t)^\dagger \tag{5.2}$$

at any later time evolves according to the Liouville-von Neumann equation (2.2), i.e., $\mathcal{U}(t)$ solves Eq. (2.3) with $t_0 = 0$ fixed. Since the Hamiltonian is time dependent, however, this solution is no longer of the simple exponential form $\mathrm{e}^{-\mathrm{i}Ht}$.

Another immediate consequence is that the system is no longer isolated because the adjustment of the coupling parameter $\lambda(t)$ apparently requires some external manipulations. As a result, both the occupations of the (instantaneous) Hamiltonian's energy levels and the energy expectation value $\langle H(t)\rangle_{\rho(t)}$ will generally change with time. Nevertheless, we will assume that these effects are reasonably small on the time scales of interest as will be laid out in more detail in Sec. 5.2 below.

The overall goal of this chapter is again very similar to Chapter 3: Assuming that we know the dynamics $\langle A\rangle_{\rho_0(t)}$ of the undriven reference system with Hamiltonian H_0 and time-evolved state

$\rho_0(t) := \mathrm{e}^{-\mathrm{i}H_0 t}\rho(0)\,\mathrm{e}^{\mathrm{i}H_0 t}$, we aim to predict the time-dependent expectation values $\langle A\rangle_{\rho(t)}$ of the driven system. In the same vein, we also intend to adopt a typicality argument by considering an ensemble of perturbation operators V and showing that nearly all members of the ensemble entail similar dynamics for fixed $\rho(0)$, A, and $\lambda(t)$. However, the time dependence of $H(t)$ poses additional challenges, which we will tackle by means of a Magnus expansion in combination with a set of time-independent auxiliary Hamiltonians as sketched below.

Applications. Time-dependent variations of some parameters of an otherwise isolated system occur in many different situations. Arguably the most intensely studied case involves *periodically driven systems*, meaning that the protocol $\lambda(t)$ satisfies $\lambda(t+\tau) = \lambda(t)$ for some *driving period* $\tau > 0$. The perturbation in this case may consist of, for example, varying electromagnetic fields, the operational cycle of an engine with repeated coupling to different environments, or, more generally, rather arbitrary externally applied forces. In the controlled environment of cold-atom [26–28, 30, 60] or polarization-echo experiments [45–47, 289] (see also Secs. 2.1 and 4.1), such external manipulations can also be exploited to modulate intrinsic properties of the system such as tunneling amplitudes or interaction strengths. Hence the results presented in the following can readily be tested experimentally. Moreover, periodic driving has been suggested and utilized as a means to engineer effective phases with unusual material properties and so-called "Floquet time crystals;" see, for instance, Refs. [310–317]. However, such exotic phases will not be our primary interest in the following since these properties usually emerge in a quasistationary regime at late times, whereas our focus will be on the response at short times when the system is still far from some (quasi)equilibrium state. Hence we mostly have in mind the more "traditional" cases of periodically time-dependent perturbations mentioned first.

Another important example of time-dependent variations that can be modeled by Hamiltonians of the form (5.1) are *quenches* [7, 83, 174]. In this case, which constitutes a popular method to prepare nonequilibrium states (see also Sec. 2.2.3), some parameter of the Hamiltonian (e.g., an interaction strength) is tuned from one value to another one. In fact, we already mentioned this in Sec. 3.1 (see Table 3.1 in particular) as a potential application for the time-independent relaxation theory, too, because indeed such quenches are commonly assumed to happen instantaneously, i.e., the corresponding parameter change is taken to occur suddenly, in a discontinuous way. In any practical application, however, this change will obviously be carried out in a possibly small, but finite time [278–281]. Our present framework is especially suited to address this scenario in a rather general setting because it particularly assesses the short-time response. By comparison with the results from Chapter 3, we can thus discern how the finite quench time affects the ensuing relaxation.

In any case, we point out that the methods adopted in the following are largely independent of the precise time dependence of $\lambda(t)$ and not restricted to periodic functions or functions approaching some constant value. As long as the amplitude does not become overly large, any reasonably well-behaved functional form of $\lambda(t)$ will thus be admitted (see also Sec. 5.2 for more details on the prerequisites), meaning that the theory can in principal be adopted to more general situations than the ones describe above if they happen to be found in a certain setup of interest.

Magnus expansion. As mentioned above, the time-dependent Hamiltonian $H(t)$ entails considerable technical complications because there is no simple solution of the evolution equation (2.3) for the propagator $\mathcal{U}(t)$. A formal solution can be given in terms of a *Dyson series* [1], i.e., a perturbative expansion in powers of $H(t)$ or V if one adopts the interaction picture instead. We will briefly consider such an expansion for a qualitative analysis of the driving effects in Sec. 5.2. However, employing the Dyson series for quantitative estimates has several disadvantages. Most notably, truncating the series at a finite order yields a nonunitary approximation for $\mathcal{U}(t)$, implying that normalization of the state is not guaranteed.

An alternative approximate construction of $\mathcal{U}(t)$ is given by the so-called *Magnus expansion* [318]. In general, the propagator is expressed as

$$\mathcal{U}(t) = \mathrm{e}^{\Omega(t)}\,, \quad \Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t)\,, \tag{5.3}$$

where the individual terms $\Omega_k(t)$ in the exponent consist of integrals over $k-1$ nested commutators of $H(t)$ at different time points. We remark that $\Omega_k(t)$ can be expressed as a function of the corresponding terms of the Dyson series up to order k [318], so there is a direct relation between both approaches, elucidating the perturbative character of the Magnus expansion. Crucially, however, every $\Omega_k(t)$ is skew-Hermitian, $\Omega_k(t)^\dagger = -\Omega_k(t)$, so the truncated series still leads to a unitary operator. The first two terms in the Magnus expansion are

$$\Omega_1(t) = -\mathrm{i}\int_0^t \mathrm{d}t_1\; H(t_1)\,, \qquad \Omega_2(t) = -\frac{1}{2}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\; [H(t_1), H(t_2)]\,, \tag{5.4}$$

and we will generally content ourselves with this low-order truncation. We point out that the question of convergence of the Magnus expansion can generally be a subtle issue. In our present setting, convergence is guaranteed up to times t such that $\int_0^t \mathrm{d}s\;\; \|H(s)\| < \pi$, but can extend to considerably longer times, too [318]. Here $\|H(s)\|$ is the operator norm of $H(s)$. Due to the extensive growth of $H(t)$ with the degrees of freedom, guaranteed convergence is thus generally very limited for typical many-body systems, but the expansion can still remain valuable as an asymptotic series [308, 309]. In general, we therefore cannot expect to obtain predictions for arbitrarily late times from the Magnus expansion, and the restriction to a truncated series will usually diminish the applicability further. Yet the predictions will turn out to be useful for short times or high-frequency driving, and may even capture the entire relaxation process if the system approaches a stationary state sufficiently fast.

For Hamiltonians $H(t)$ of the form (5.1), the first two terms in the Magnus expansion can be written as

$$\Omega_1(t) = -\mathrm{i}\,[H_0\, t + \Lambda_1(t)\, V]\;, \qquad \Omega_2(t) = \left[\Lambda_2(t) - \frac{t}{2}\Lambda_1(t)\right]\,[V, H_0]\,, \tag{5.5}$$

where

$$\Lambda_1(t) := \int_0^t \mathrm{d}s\;\lambda(s)\,, \qquad \Lambda_2(t) := \int_0^t \mathrm{d}s\;\Lambda_1(s) \tag{5.6}$$

are the first and second integrals of the driving protocol $\lambda(t)$.

Auxiliary dynamics. The Magnus expansion provides us with a more explicit expression of the propagator in terms of the operators H_0 and V entering the Hamiltonian (5.1) as well as the time dependence mediated via $\lambda(t)$. However, we still need a way to conveniently express the time-dependent expectation values $\langle A\rangle_{\rho(t)}$ in the eigenbasis $\{|\mu\rangle_0\}$ of the unperturbed Hamiltonian H_0 to adopt a typicality scheme similarly as in Chapter 3.

The idea is to introduce a family of time-independent auxiliary Hamiltonians $\{H^{(t')}\}$ with $t' \geq 0$ being a fixed parameter. The dynamics in any of these auxiliary systems may then be assessed by means of the predictions from Chapter 3 and suitably combined to obtain an approximation of the time-dependent problem. Based on the Magnus expansion (5.3), we therefore define

$$H^{(t')} := \mathrm{i}\Omega(t')/t'\,. \tag{5.7}$$

Starting from the same initial state $\rho(0)$ of interest and keeping t' fixed, any of these Hamiltonians $H^{(t')}$ generates a time evolution given by

$$\rho(t,t') := \mathrm{e}^{-\mathrm{i}H^{(t')}t}\,\rho(0)\,\mathrm{e}^{\mathrm{i}H^{(t')}t}\,. \tag{5.8}$$

Introducing the eigenvalues $E_n^{(t')}$ and eigenvectors $|n(t')\rangle$ of $H^{(t')}$, the observable expectation values under the auxiliary dynamics take the form

$$\langle A\rangle_{\rho(t,t')} = \sum_{m,n} \mathrm{e}^{\mathrm{i}(E_n^{(t')}-E_m^{(t')})t}\,\langle m(t')|\rho(0)|n(t')\rangle\,\langle n(t')|A|m(t')\rangle\,. \tag{5.9}$$

This relation is already structurally similar to Eq. (3.4). Provided that all the $H^{(t')}$ can be written in the form (3.1) and satisfy the prerequisites from Sec. 3.2, we can thus employ the prediction (3.160) for the relaxation under the influence of time-independent perturbations to assess the $\langle A\rangle_{\rho(t,t')}$ Moreover, observing (5.2), (5.3), (5.7), and (5.8), we conclude that the true state $\rho(t)$

under the driven dynamics can be obtained from the auxiliary states $\rho(t,t')$ as $\rho(t)=\rho(t,t)$ and thus

$$\langle A\rangle_{\rho(t)} = \langle A\rangle_{\rho(t,t)}\,. \tag{5.10}$$

We remark that the latter correspondence between the true and the auxiliary dynamics is exact, provided that the Magnus expansion (5.3) converges. The additional approximation necessary here compared to Chapter 3 is that we have to truncate the Magnus expansion after the second order to recast the auxiliary Hamiltonians $H^{(t')}$ in the form (3.1).

Mapping to perturbed-relaxation setting. Employing the Magnus expansion up to second order from (5.5), the auxiliary Hamiltonians from (5.7) can be approximated as

$$H^{(t')} \approx H_0 + V^{(t')} \tag{5.11}$$

with

$$V^{(t')} := \frac{\Lambda_1(t')}{t'}\,V + \left[\frac{\Lambda_2(t')}{t'} - \frac{\Lambda_1(t')}{2}\right]\,\mathrm{i}[V,H_0]\,. \tag{5.12}$$

Since these auxiliary perturbation operators $V^{(t')}$ are linear in the underlying basic perturbation V from (5.1), this constitutes the desired mapping of the family of auxiliary Hamiltonians $H^{(t')}$ to the class of systems (3.1) treated in Chapter 3. Expressing $V^{(t')}$ in the eigenbasis $\{|\mu\rangle_0\}$ of the unperturbed Hamiltonian H_0, we obtain

$$V^{(t')}_{\mu\nu} = \left\{\frac{\Lambda_1(t')}{t'} - \mathrm{i}(E_\mu - E_\nu)\left[\frac{\Lambda_2(t')}{t'} - \frac{\Lambda_1(t')}{2}\right]\right\} V_{\mu\nu}\,, \tag{5.13}$$

where we wrote $V_{\mu\nu} := {}_0\langle\mu|V|\nu\rangle_0$, $V^{(t')}_{\mu\nu} := {}_0\langle\mu|V^{(t')}|\nu\rangle_0$, etc. for the matrix elements in the unperturbed basis and E_μ for the eigenvalue of H_0 corresponding to the eigenstate $|\mu\rangle_0$, similarly as in Chapter 3. Since the perturbation V is supposed to have similar properties as before, too, it exhibits, in particular, a well-defined perturbation profile $\sigma_v^2(E)$ as specified in Eq. (3.10). In view of (5.13), the operators $V^{(t')}$ therefore have an effective perturbation profile

$$\sigma_v^2(E,t') := \left[\varphi_1(t') + E^2\varphi_2(t')\right]\,\sigma_v^2(E) \tag{5.14}$$

with

$$\varphi_1(t') := \left[\frac{\Lambda_1(t')}{t'}\right]^2\,, \qquad \varphi_2(t') := \left[\frac{\Lambda_2(t')}{t'} - \frac{\Lambda_1(t')}{2}\right]^2\,, \tag{5.15}$$

such that $|V^{(t')}_{\mu\nu}|^2 \simeq \sigma_v^2(E_\mu - E_\nu, t')$. Note that the coupling λ, which was used to control the perturbation strength in Chapter 3, has now been absorbed into the effective perturbation profile $\sigma_v^2(E,t')$ since both $\Lambda_1(t)$ and $\Lambda_2(t)$ are proportional to the amplitude of the driving protocol $\lambda(t)$.

5.2 Prerequisites and expected applicability

Before investigating the consequences of the transformation relating the dynamics generated by $H(t)$ and $H^{(t')}$, we scrutinize under which conditions this strategy can be reasonably expected to work.

Auxiliary Hamiltonians. As explained in Sec. 5.1, we intend to adopt the theoretical predictions from Chapter 3 to the dynamics generated by the auxiliary Hamiltonians $H^{(t')}$ from (5.11). For this to be valid, these $H^{(t')}$ should thus satisfy the prerequisites collected in Sec. 3.2 in particular.

Prerequisite (i), which required a homogeneous density of states $D(E)$ of H_0, does not need any further discussion because the reference Hamiltonian H_0 in (5.11) and (5.1) is assumed to be of similar type as in (3.1), and the same holds for the initial state $\rho(0)$. As before, we will denote the corresponding mean level spacing by ε, i.e., $D(E)\approx\varepsilon^{-1}$.

Prerequisite (ii), which stipulated sufficiently weak perturbations and thereby essentially demanded a property similar to Prerequisite (i) also for the perturbed systems, is more subtle. On the one

hand, the strength of the perturbation $V^{(t')}$ depends on the driving protocol $\lambda(t)$ via the integrals $\Lambda_1(t)$ and $\Lambda_2(t)$ from (5.6). If $\lambda(t)$ is unbiased in the sense that its time average $\overline{\lambda(t)}$ vanishes, then $\Lambda_1(t)$ is bounded and $\Lambda_2(t)$ will grow at most linearly in t. By inspection of the amplitude functions $\varphi_1(t)$ and $\varphi_2(t)$ from (5.15), we understand that the effective perturbation strength as mediated by the driving is thus bounded, too, meaning that the weak-perturbation criterion translates into a corresponding weak-driving prerequisite. On the other hand, for fixed t', the effective perturbation profile $\sigma_v^2(E,t')$ from (5.14) generally decays slower with the energy separation E of the coupled levels than the underlying "pure" profile $\sigma_v^2(E)$ of the operator V. We thus additionally require that $\sigma_v^2(E)$ should decay at least quadratically as $E \to \infty$ such that $\sigma_v^2(E,t')$ remains bounded.

Another subtlety arises from the fact that the energy level occupations in the driven system are not conserved, meaning that the pertinent window of nonnegligibly populated levels generically broadens over the course of time. This spreading of the state in energy is essentially a higher-order effect [309] and therefore does not become apparent in our present setting because of our restriction to second order of the Magnus expansion. Hence the assumption of a homogeneous density of states is unproblematic within our present approximation, but this approximation itself will eventually become questionable at late times, accompanied by a broadening of the energy distribution and a varying density of states. We will come back to this point below.

Since Prerequisite (iii), which called for sufficiently strong perturbations, turned out to be largely dispensable at last in Chapter 3, it should not cause any problems in the present setting either. Likewise, as shown above, Prerequisite (iv) of a well-defined (effective) perturbation profile is automatically satisfied if the underlying operator V from (5.1) exhibits such a well-defined profile.

Prerequisite (v) of sufficiently small level fluctuations relates to similar issues as Prerequisite (ii), but since the analysis around Eq. (3.18) and from Sec. 3.6.3 revealed that there is plenty of room for such fluctuations without a noticeable effect on the dynamics within the relevant relaxation time, we do not expect any complications in the present setting. Finally, Prerequisite (vi), requiring sufficiently small correlations between the perturbation matrix elements $V_{\mu\nu}^{(t')}$, is unaffected by the modifications due to the auxiliary dynamics since the $V^{(t')}$ essentially inherit their matrix structure from V according to (5.13). More precisely, the randomization entailed in the typicality approach still only affects the pure perturbation operator V and the correlations between H_0 and V introduced in $V^{(t')}$ via Eq. (5.12) are explicitly taken into account.

Short times and high frequencies. The above observations suggest that predicting the auxiliary dynamics generated by the approximate $H^{(t')}$ from (5.11) by means of the result (3.160) from Chapter 3 should usually yield satisfactory results without any significant further restrictions. The remaining question regarding the usefulness of the proposed approach is therefore whether or rather under which circumstances the second-order truncation (5.11) is a reliable approximation of the true auxiliary Hamiltonian from (5.7).

We already commented on the general convergence issues of the Magnus expansion below Eq. (5.4). Since the terms $\Omega_k(t)$ are in direct correspondence to the terms up to the same order of the Dyson series, the second-order approximation will generally perform best at short times. In case of periodic driving with a period τ such that $\lambda(t+\tau) = \lambda(t)$, the Floquet theorem states that the propagator $\mathcal{U}(t)$ can be decomposed as $\mathcal{U}_\mathrm{P}(t)\,\mathrm{e}^{-\mathrm{i}H_\mathrm{F}t}$ with the time-independent *Floquet Hamiltonian* H_F and the periodic unitary operator $\mathcal{U}_\mathrm{P}(t) = \mathcal{U}_\mathrm{P}(t+\tau)$ [319]. The Magnus series can then be adapted to this special structure by expanding H_F and $\mathcal{U}_\mathrm{P}(t)$ individually [282, 318]. For sufficiently small periods τ such that $\int_0^\tau \mathrm{d}t\, \|H(t)\| < 0.21$, the series will then converge for all times t. Generally speaking, the Floquet-Magnus series can be regarded as an expansion in the driving period τ and thus constitutes a high-frequency approximation of the propagator. While we will not explicitly exploit the Floquet theorem in our approach since it does not offer any technical advantages for the present problem, remnants of this special suitability for fast driving will become apparent in the final prediction, too (see Sec. 5.3.3 in particular).

We remark that the somewhat trivial limit of extremely slow driving is also reflected correctly: If $\lambda(t)$ can be treated as essentially constant, $\lambda(t) = \lambda_0$, then $\varphi_1(t) = \lambda_0^2$ and $\varphi_2(t) = 0$ according to (5.15), meaning that the perturbation profile $\sigma_v^2(E,t') = \lambda_0^2\,\sigma_v^2(E)$ is mapped back onto the case of time-independent perturbations from Chapter 3.

As with the relaxation theory for time-independent perturbations from Chapter 3, the common locality and few-body properties of possible perturbations V are not explicitly accounted for in the considered ensembles (see Sec. 3.8), hence the initial state should be approximately homogeneous in space on a macroscopic level, or at least the observable should not probe potential inhomogeneities. Put differently, the theory will not cover response behavior involving macroscopic transport and thus considerably longer time scales, but only local, short-time effects of the driving. Since the predictions are not expected to reach into the late-time regime anyway due to the truncated Magnus expansion, this does not pose any additional restrictions, though.

Energy distribution and heating. Finally, we come back to the issue already mentioned in the discussion of Prerequisite (ii) above, namely the fact that the level populations of the (instantaneous) Hamiltonian $H(t)$ are no longer conserved in the time-dependent setting. To get a qualitative understanding of how the energy distribution changes under the driving, we temporarily switch to the *Dirac* or *interaction picture* by defining $A_{\mathrm{int}}(t) := \mathrm{e}^{\mathrm{i}H_0 t}\, A\, \mathrm{e}^{-\mathrm{i}H_0 t}$ and $\rho_{\mathrm{int}}(t) := \mathrm{e}^{\mathrm{i}H_0 t}\, \rho(t)\, \mathrm{e}^{-\mathrm{i}H_0 t}$, meaning that time-dependent expectation values can be computed as $\langle A \rangle_{\rho(t)} = \langle A_{\mathrm{int}}(t) \rangle_{\rho_{\mathrm{int}}(t)}$ as well. In the interaction picture, the time-evolved state is thus obtained from the initial state $\rho(0)$ via the interaction-picture propagator $\mathcal{U}_{\mathrm{int}}(t) := \mathrm{e}^{\mathrm{i}H_0 t}\mathcal{U}(t)$ as $\rho_{\mathrm{int}}(t) = \mathcal{U}_{\mathrm{int}}(t)\rho(0)\mathcal{U}_{\mathrm{int}}(t)^\dagger$. From Eq. (2.3), it follows that $\mathcal{U}_{\mathrm{int}}(t)$ solves

$$\dot{\mathcal{U}}_{\mathrm{int}}(t) = -\mathrm{i}\,\lambda(t)\, V_{\mathrm{int}}(t)\, \mathcal{U}_{\mathrm{int}}(t) \quad \text{with} \quad \mathcal{U}_{\mathrm{int}}(0) = \mathbb{1}\,, \tag{5.16}$$

where $V_{\mathrm{int}}(t) := \mathrm{e}^{\mathrm{i}H_0 t}\, V\, \mathrm{e}^{-\mathrm{i}H_0 t}$. Expanding the solution in a Dyson series [1], we then find

$$\mathcal{U}_{\mathrm{int}}(t) = \mathbb{1} - \mathrm{i}\int_0^t \mathrm{d}t_1\, \lambda(t_1) V_{\mathrm{int}}(t_1) - \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\; \lambda(t_1)\, \lambda(t_2)\, V_{\mathrm{int}}(t_1)\, V_{\mathrm{int}}(t_2) + \ldots \tag{5.17}$$

To first order, the time-dependent expectation values can therefore be calculated as

$$\langle A_{\mathrm{int}}(t) \rangle_{\rho_{\mathrm{int}}(t)} = \langle A_{\mathrm{int}}(t) \rangle_{\rho(0)} + \mathrm{i}\int_0^t \mathrm{d}t_1\, \lambda(t_1)\, \langle [V_{\mathrm{int}}(t_1), A_{\mathrm{int}}(t)] \rangle_{\rho(0)} + \ldots \tag{5.18}$$

This is just the Kubo formula describing the linear response of the observable expectation values to the time-dependent perturbation $\lambda(t)V$. We concentrate on the case of unbiased protocols $\lambda(t)$ such that $\overline{\lambda(t)} = 0$, meaning that the unperturbed Hamiltonian H_0 is also the time-averaged Hamiltonian, $\overline{H(t)} = H_0$. If this was not the case, we could absorb the corresponding offset $\overline{\lambda(t)}V$ into the definition of H_0 for the purpose of the following discussion. Since we are interested in the energy distribution, we now inspect one particular eigenstate $|\mu\rangle_0$ of the unperturbed, time-averaged Hamiltonian H_0 by choosing $A = \Pi_\mu := |\mu\rangle_{00}\langle\mu|$. Substituting into (5.18) and returning to the Schrödinger picture, we obtain

$$\langle \Pi_\mu \rangle_{\rho(t)} = \rho_{\mu\mu}(0) - 2\,\mathrm{Im}\sum_\alpha \rho_{\mu\alpha}(0)\, V_{\alpha\mu} \int_0^t \mathrm{d}t_1\; \lambda(t_1)\, \mathrm{e}^{\mathrm{i}(E_\alpha - E_\mu)t_1} + \ldots \tag{5.19}$$

For the sake of the ensuing argument, we focus on a single-frequency driving $\lambda(t) = \lambda_0 \sin(\omega t)$, observing that similar considerations can be carried out for individual Fourier components of a more general protocol. In other words, if $\hat{\lambda}(\omega)$ is the Fourier transform of $\lambda(t)$, the following reasoning can be applied to any frequency ω with $\hat{\lambda}(\omega) \neq 0$. Note that for periodic driving with period $\tau = 2\pi/\omega$, the frequency spectrum $\hat{\lambda}(\omega)$ is discrete with contributions from integer multiples of the base frequency $2\pi/\tau$ only. Moreover, we may think of $\rho(0) = |\psi\rangle\langle\psi|$ as a pure state with coefficients $\psi_\mu := {}_0\langle\mu|\psi\rangle$ in the unperturbed basis since any mixed state will be a linear combination of such pure states. Hence $\rho_{\mu\alpha}(0) = \psi_\mu \psi_\alpha^*$. Generally speaking, Eq. (5.19) then expresses that the μth unperturbed level can exchange population or occupation "weights" with any other level α. However, significant and persistent changes will only occur if

(i) the matrix element $V_{\mu\alpha}$ is relatively large and

(ii) the energy difference $|E_\alpha - E_\mu|$ roughly matches the driving frequency ω.

Condition (i) directly relates to the perturbation profile $\sigma_v^2(E)$ from (3.10). In particular, for the common case of a banded matrix structure, significant exchanges of population will only occur between states with $|E_\alpha - E_\mu| \lesssim \Delta_v$, where Δ_v is the band width from (3.13), and the farther

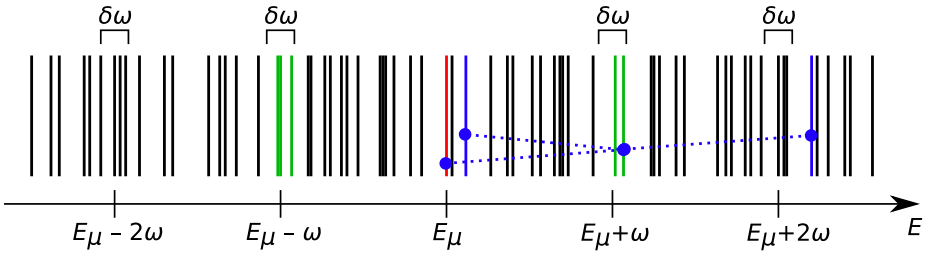


Figure 5.1: Illustration of the resonance condition for the exchange of level populations. The state $|\mu\rangle_0$ at energy E_μ (red) is directly resonant at frequency ω with all states $|\alpha\rangle_0$ such that $|E_\alpha - E_\mu| \approx \omega$ and $|E_\alpha - E_\mu| \lesssim \Delta_v$ (green). The blue states, in turn, are examples of states that are not directly resonant with $|\mu\rangle_0$, but are connected by (second-order) resonance chains via an intermediate state (dotted lines).

apart the levels are in the spectrum, the less they will typically interact. Condition (ii) ensures that the modifications are persistent on the relevant time scales and do not average out over short time intervals. If both conditions are met, we call the states $|\mu\rangle_0$ and $|\alpha\rangle_0$ *directly resonant at frequency* ω. In general, we expect that populations between two resonant states are approximately balanced as time progresses, at least if the life times of both states are of the same order, which should typically be the case for most such pairs in a generic many-body system.

Naturally, the occupation of the μth level will receive further corrections from higher-order terms in the series (5.19). Nevertheless, the general mechanism remains the same. For instance, the second-order correction reads (see also Eq. (5.17))

$$\begin{aligned} &\left| \sum_\alpha V_{\mu\alpha} \psi_\alpha^* \int_0^t \mathrm{d}t_1 \; \lambda(t_1) \, \mathrm{e}^{\mathrm{i}(E_\alpha - E_\mu)t_1} \right|^2 \\ &\quad - 2 \, \mathrm{Re} \sum_{\alpha,\beta} \psi_\alpha^* V_{\alpha\beta} V_{\beta\mu} \psi_\mu \int_0^t \mathrm{d}t_1 \; \lambda(t_1) \, \mathrm{e}^{\mathrm{i}(E_\beta - E_\mu)t_1} \int_0^{t_1} \mathrm{d}t_2 \; \lambda(t_2) \, \mathrm{e}^{\mathrm{i}(E_\alpha - E_\beta)t_2} \, . \end{aligned} \tag{5.20}$$

The terms in the first line can again contribute significantly only if the states $|\mu\rangle_0$ and $|\alpha\rangle_0$ are directly resonant. In the second line, by contrast, population can be exchanged indirectly via an intermediate state $|\beta\rangle_0$. This state $|\beta\rangle_0$, however, must be directly resonant with both $|\mu\rangle_0$ and $|\alpha\rangle_0$.

As sketched in Fig. 5.1, the generally emerging picture is thus as follows: Populations can be exchanged between resonant states. Two states $|\mu\rangle_0$ and $|\alpha\rangle_0$ are resonant if there exists a sequence of states $|\alpha\rangle_0 =: |\beta_0\rangle_0, |\beta_1\rangle_0, \ldots, |\beta_M\rangle_0 := |\mu\rangle_0$ such that, for all $m = 1, \ldots, M$, the states $|\beta_{m-1}\rangle_0$ and $|\beta_m\rangle_0$ are directly resonant at frequency ω and the corresponding Fourier component $\hat{\lambda}(\omega)$ of the driving protocol $\lambda(t)$ is nonvanishing. Such a sequence is called a *resonance chain of order* M. In a sufficiently generic setting, almost all states will be resonant with each other. However, higher-order resonance chains will typically be suppressed: For sufficiently weak driving, the amplitudes $\lambda(t)$ or, more precisely, the product $\lambda(t)\sigma_v(\omega)$ will be small. Since every additional state in a resonance chain is accompanied by an extra factor on the order of $\lambda(t)\sigma_v(\omega)$, longer chains will contribute relatively less. Moreover, the bandedness of the perturbation V entails that the matrix elements $V_{\beta_{m-1}\beta_m}$ decay with $|E_{\beta_{m-1}} - E_{\beta_m}|$, hence contributions from higher frequencies are relatively suppressed, too.

We remark that the ultimate spreading of the populations in energy generically leads to energy absorption and thus heating of the system [306, 320–323]. The reason is that the level density of many-body systems usually increases with increasing energy. A given eigenstate $|\mu\rangle_0$ of the time-averaged Hamiltonian H_0 will therefore be resonant with more states whose energy is higher than E_μ than with states of lower energy, meaning that the occupations are gradually shifted towards higher energies.

Under physically reasonable additional assumptions about the system, such as locality of interactions, the effect of driving frequencies and interaction ranges can be assessed in more rigorous terms than the above resonance-chain argument. For instance, Refs. [307, 309, 324, 325] provide

upper bounds for the heating rate that decay exponentially with the driving frequency, reinforcing the suppression of energy dispersion for fast driving. The phenomenology is thus reminiscent of prethermalization since the system spends a long time near a quasistationary state associated with the initial energy shell [308, 309, 314, 317, 323, 325]. Nevertheless, at very late times, periodically driven many-body systems are generically expected to reach a state of infinite temperature [306, 320–323].

Altogether, populations will be balanced primarily between states that are close-by in energy, even though the spreading will generally increase over time. Mechanisms that are expected to impede and defer this spreading include weak perturbations (small driving amplitude), high driving frequencies, and small perturbation band widths. These observations match favorably with the above-collected criteria for the validity of our typicality approach in combination with the truncated Magnus expansion, supplemented by the additional supporting factor of short-ranged perturbations in energy, i.e., small Δ_v. Whereas the eventual theoretical prediction will thus presumably break down at late times, we anticipate that it should apply the longer the better those three criteria are satisfied.

5.3 Typical response

Having made sure that the results from Chapter 3 should be readily applicable to the auxiliary Hamiltonians $H^{(t')}$, we proceed with the strategy proposed in Sec. 5.1 and establish in Sec. 5.3.1 a prediction for the driven dynamics based on the behavior of the auxiliary systems. As before, there will be a characteristic response-profile function encoding the modifications of the reference dynamics caused by the driving. This response profile will be inspected more closely in Sec. 5.3.2. In Sec. 5.3.3 we will then put our expectations about the regime of applicability as formulated in the previous section to a quantitative test.

5.3.1 Prediction for driven time evolution

Auxiliary dynamics. Provided that the prerequisites from Sec. 5.2 hold, the auxiliary systems with approximate Hamiltonians $H^{(t')}$ of the form (5.11) can be described by the relaxation theory for isolated many-body quantum systems under the influence of time-independent perturbations from Chapter 3. The main result (3.160) then applies to every individual $H^{(t')}$, so the time-dependent expectation values $\langle A\rangle_{\rho(t,t')}$ from (5.9) are predicted to be given by

$$\langle A\rangle_{\rho(t,t')} = \langle A\rangle_{\tilde{\rho}(t')} + |g(t,t')|^2 \left[\langle A\rangle_{\rho_0(t)} - \langle A\rangle_{\tilde{\rho}(t')}\right]. \tag{5.21}$$

Here $\langle A\rangle_{\rho_0(t)}$ denotes the undriven reference dynamics as before, whereas the other quantities are defined similarly as in Chapter 3, but inherit a parametric dependence on the auxiliary time t' via the perturbation profile $\sigma_v^2(E,t')$ from (5.14). To wit, we can introduce for every $H^{(t')}$ an ensemble-averaged resolvent $G(z,t')$ which satisfies $G(z-H_0,t') = \mathbb{E}[(z-H^{(t')})^{-1}]$ (cf. Eq. (3.58)) and solves

$$G(z,t')\left[z - \int \frac{\mathrm{d}E}{\varepsilon}\, G(z-E,t')\,\sigma_v^2(E,t')\right] = 1 \tag{5.22}$$

according to (3.59) (see also Eq. (3.31)). Similarly as in (3.32), we can extract from these ensemble-averaged resolvents an overlap distribution $u(E,t') := \varepsilon \lim_{\eta\to 0+} \operatorname{Im} G(E-\mathrm{i}\eta,t')/\pi$, which describes, in particular, the second moment $\mathbb{E}[|\langle n(t')|\mu\rangle_0|^2] = u(E_n - E_\mu, t')$ of the overlap between the eigenvectors $|n(t')\rangle$ of $H^{(t')}$ (see above Eq. (5.9)) and $|\mu\rangle_0$ of H_0. The function $g(t,t')$ is the corresponding response profile, i.e., the Fourier transform of $u(E,t')$ in E,

$$g(t,t') := \int \frac{\mathrm{d}E}{\varepsilon}\, \mathrm{e}^{\mathrm{i}Et}\, u(E,t'), \tag{5.23}$$

see also Eq. (3.146). Likewise, the state $\tilde{\rho}(t')$ is defined by analogy with $\tilde{\rho}_\lambda$ from (3.150) in terms of its matrix elements in the unperturbed basis as

$${}_0\langle\mu|\tilde{\rho}(t')|\nu\rangle_0 := \delta_{\mu\nu} \sum_\alpha \tilde{u}(E_\mu - E_\alpha, t')\,\rho_{\alpha\alpha}(0), \tag{5.24}$$

where $\tilde{u}(E,t') := \int \mathrm{d}E'\, u(E-E',t')\, u(E',t')/\varepsilon$ (cf. Eq. (3.116)). As discussed above Eq. (3.188), the expectation value $\langle A\rangle_{\tilde{\rho}(t')}$ can commonly be identified with the thermal prediction $\langle A\rangle_{\rho_\mathrm{mc}}$. In practice, we can therefore usually reduce the dependence of Eq. (5.21) on the auxiliary time t' such that

$$\langle A\rangle_{\rho(t,t')} = \langle A\rangle_{\rho_\mathrm{mc}} + |g(t,t')|^2 \left[\langle A\rangle_{\rho_0(t)} - \langle A\rangle_{\rho_\mathrm{mc}}\right]. \tag{5.25}$$

Driven dynamics. As explained above Eq. (5.10), the actual dynamics of the driven system (5.1) can be retrieved from the family of undriven auxiliary dynamics mediated by the Hamiltonians $H^{(t')}$ upon identifying $t'=t$, i.e., the time-dependent expectation values of the auxiliary system $H^{(t)}$ at time t coincide with those of the driven system at time t, cf. Eq. (5.10). Approximating $H^{(t')}$ by the second-order Magnus expansion from (5.7), we can thus employ the result (5.21) to obtain a prediction for the time-dependent expectation values of the driven system,

$$\langle A\rangle_{\rho(t)} = \langle A\rangle_{\tilde{\rho}(t)} + |g(t,t)|^2 \left[\langle A\rangle_{\rho_0(t)} - \langle A\rangle_{\tilde{\rho}(t)}\right]. \tag{5.26}$$

Hence the behavior of the driven system is expected to resemble the undriven dynamics, but experiences modifications according to the generalized response profile $g(t,t')$ from (5.23). More precisely, this response profile modulates the undriven dynamics with respect to a generally also time-dependent base line $\langle A\rangle_{\tilde{\rho}(t)}$. The latter quantity, however, will usually be approximated well by the thermal expectation value associated with the reference system, such that the prediction (5.26) takes the form

$$\boxed{\langle A\rangle_{\rho(t)} = \langle A\rangle_{\rho_\mathrm{mc}} + |g(t,t)|^2 \left[\langle A\rangle_{\rho_0(t)} - \langle A\rangle_{\rho_\mathrm{mc}}\right]}. \tag{5.27}$$

We emphasize that ρ_mc here corresponds to the microcanonical density operator for the initially populated energy window and not to an infinite-temperature state that may possibly be reached at very late times beyond the scope of applicability of our present approach.

The predictions (5.26) and (5.27) are the main results of the present chapter, extending the perturbed relaxation theory from Chapter 3 to explicitly time-dependent perturbations. The general structure is similar to the results (3.160) and (3.188) for time-independent perturbations. Nevertheless, as we will see in the following, the response profile will exhibit considerably more diverse behavior as a consequence of the driving.

5.3.2 Response profile

The key quantity governing the deviations of the driven dynamics from the unperturbed behavior is once again the response profile in its generalized form (5.23). For any fixed t', this function $g(t,t')$ has similar properties as the response profile $g_\lambda(t)$ from (3.146), which describes the effect of time-independent perturbations and was discussed in detail in Sec. 3.6.2. In view of Eqs. (5.26) and (5.27), however, the functional dependence on t for $t'=t$ is of primary interest.

Overview of special cases. For the function $g_\lambda(t)$, explicit solutions could be constructed in three special cases: for weak perturbation or large times, for strong perturbation or small times, and for perturbations with a Breit-Wigner profile (3.74). Unfortunately, for the solutions in the first and third cases, there is no obvious way to generalize them to the present time-dependent setting. The reason is that the perturbation profiles $\sigma_v^2(E)$ pertaining to these two cases were either constant (for weak perturbations) or slowly decaying like $\sigma_v^2(E)\sim E^{-2}$ as $E\to\infty$ (for the Breit-Wigner profile); see also Sec. 3.4.2. The effective perturbation profile $\sigma_v^2(E,t')$ corresponding to the approximate auxiliary Hamiltonian $H^{(t')}$ from (5.11), however, scales like $\sigma_v^2(E,t')\sim E^2\sigma_v^2(E)$ for large E, which renders it infeasible to solve the integral equation (5.22) by means of the methods employed in Sec. 3.4.2.

For the second case of reasonably strong perturbation or small times, by contrast, we can proceed similarly as in Eqs. (3.68) through (3.72) to find an explicit approximation of $g(t,t')$. Fortunately, this is also the most important case in the present setting: On the one hand, the perturbations have to be somewhat stronger than in the setup from Chapter 3 to have a noticeable effect, especially in the case of periodic driving because fast switching of the sign of $\lambda(t)$ will usually revert some of the effects, similarly to the effective partial time reversal protocols studied in Chapter 4. On the other hand, our approach is naturally limited to shorter times as explained in Sec. 5.2, even

though the detailed validity analysis in Sec. 5.3.3 will reveal that the truncated Magnus expansion generally still applies to longer times than the following approximation will.

Narrow profile or strong perturbations. In the same spirit as above Eq. (3.68), we thus turn to the integral equation (5.22) and assume that the effective perturbation profile $\sigma_v^2(E,t')$ decays much faster as a function of E than the ensemble-averaged resolvent $G(z,t')$ does with z. As before, this corresponds to the case of strong mixing of the perturbed and unperturbed eigenvectors due to a relatively strong perturbation $V^{(t')}$ or to the case of a relatively narrow perturbation profile $\sigma_v^2(E,t')$. Approximating $G(z-E,t')$ by $G(z,t')$ in the integrand and substituting (5.14) for the perturbation profile leads again to a purely algebraic equation,

$$[\gamma(t')\,G(z,t')]^2/4 - z\,G(z,t') + 1 = 0 \tag{5.28}$$

with

$$\gamma(t') := \sqrt{4\alpha_v\,[2\Delta_v\,\varphi_1(t') + \Sigma_2\,\varphi_2(t')]}\,, \tag{5.29}$$

where α_v and Δ_v are the intrinsic strength and band width of the true perturbation operator V as introduced in Eqs. (3.12) and (3.13), respectively, and Σ_2 is the second moment (3.175) of the perturbation profile $\sigma_v^2(E)$ (note that Δ_v is half the zeroth moment). Comparing with Eq. (3.69), the occurrence of this second moment Σ_2 indicates that the response profile $g(t,t')$ for driven dynamics is more sensitive to details of the perturbation than its counterpart $g_\lambda(t)$ for the undriven setting.

Solving Eq. (5.28) for $G(z,t')$, we find the same functional form as in (3.70) with γ replaced by $\gamma(t')$. The initial assumption that the decay of $\sigma_v^2(E,t')$ in its first argument is much faster than that of $G(z,t')$ is thus verified self-consistently if $\gamma(t') \gg \Delta_v$. In view of Eq. (5.29) in combination with (5.15) (see also Fig. 5.2), this condition is guaranteed in particular if $|\lambda(t)| \gg \lambda_\mathrm{c}$ for all t, where λ_c is the crossover coupling between the weak- and strong-perturbation regimes introduced in Eq. (3.73). The t'-dependent overlap distribution defined below Eq. (5.22) is then found to be

$$u(E,t') = \frac{2\varepsilon}{\pi\gamma(t')}\sqrt{\gamma(t')^2 - E^2}\,\Theta(\gamma(t')^2 - E^2)\,. \tag{5.30}$$

Substituting into (5.23), we are left with the response profile for sufficiently strong or sufficiently short-ranged (in energy) and generally time-dependent perturbations,

$$g(t,t') = \frac{2J_1(\gamma(t')\,t)}{\gamma(t')\,t}\,, \tag{5.31}$$

where $J_1(x)$ is the first-order Bessel function of the first kind as usual.

Integro-differential equation. General solutions for $g(t,t')$ can only be calculated numerically. In principle, we could solve the integral equation (5.22) and calculate $g(t,t')$ from the Fourier transform of the solution's imaginary part according to (5.23). However, solving (5.22) for every time point $t' = t$ of interest is not very practical. Instead, it is much more convenient to work with an integro-differential representation of $g(t,t')$ directly, i.e., an analog of Eq. (3.173) for $g_\lambda(t)$. Performing similar steps as between Eqs. (3.167) and (3.173), we readily find that $g(t,t')$ solves the integro-differential equation

$$\frac{\partial g(t,t')}{\partial t} = -\int_0^t \mathrm{d}s\; g(t-s,t')\,g(s,t')\,\hat{\sigma}_v^2(s,t')\,, \tag{5.32}$$

where

$$\hat{\sigma}_v^2(t,t') := \int \frac{\mathrm{d}E}{\varepsilon}\,\mathrm{e}^{\mathrm{i}Et}\,\sigma_v^2(E,t') \tag{5.33}$$

is the Fourier transform of the effective perturbation profile $\sigma_v^2(E,t')$ from (5.14) in E. Alternatively, we may substitute (5.14) directly into (5.33) and subsequently into (5.32), leading to

$$\frac{\partial g(t,t')}{\partial t} = -\int_0^t \mathrm{d}s\; g(t-s,t')\,g(s,t')\left[\varphi_1(t') - \varphi_2(t')\frac{\partial^2}{\partial s^2}\right]\hat{\sigma}_v^2(s)\,, \tag{5.34}$$

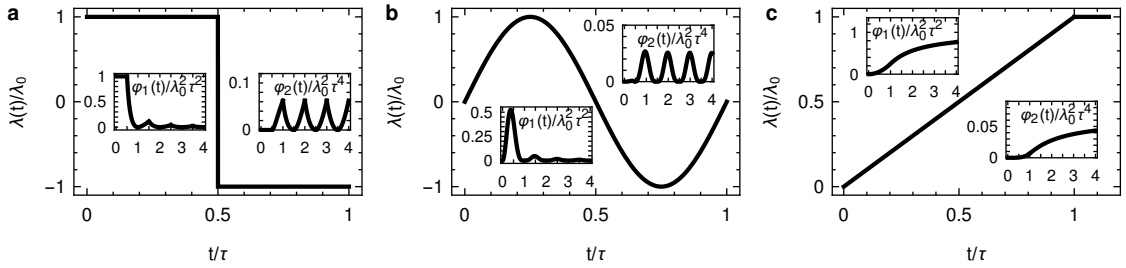


Figure 5.2: Various driving protocols along with the corresponding coefficient functions $\varphi_1(t)$ and $\varphi_2(t)$ from (5.15) in the insets. **a.** Step protocol (5.35); **b.** sinusoidal protocol (5.36); **c.** linear ramp protocol (5.37).

where $\hat{\sigma}_v^2(t)$ is the Fourier transform (3.167) of the perturbation profile $\sigma_v^2(E)$ corresponding to the pure perturbation operator V, and $\varphi_1(t')$ and $\varphi_2(t')$ were defined in (5.15) (see also Fig. 5.2 below for examples). This latter representation in particular reveals explicitly how the driving protocol $\lambda(t)$ and the perturbation operator V from (5.1) determine the response profile and thus the modified dynamics of the driven system compared to the behavior of the reference system with Hamiltonian H_0.

Moreover, for any given driving protocol $\lambda(t)$ and perturbation profile $\sigma_v^2(E)$, the associated response profile $g(t,t')$ and especially its value $g(t,t)$ can be calculated straightforwardly by integrating Eq. (5.32) or Eq. (5.34), respectively, taking account of the initial conditions $g(0,t') = 1$ and $\partial_t g(0,t') = 0$ (cf. Eq. (3.161) and above Eq. (3.176)). Since the integrand on the right-hand side depends only on times $s \leq t$, this can be achieved by simple forward-integration algorithms for any fixed t'. To evaluate the predictions (5.26) or (5.27), the required value $g(t,t)$ is thus obtained by setting $t' = t$ and integrating up to time t. Given the effective double-integral structure, the computing time therefore scales quadratically with the number of time steps.

Finally, for very fast driving and very short times such that $\varphi_1(t) \gg \varphi_2(t)$ (i.e., when the first-order truncation of the Magnus expansion suffices), we observe that Eq. (5.34) reduces to the relation (3.173) for the response profile $g_\lambda(t)$ describing the effect of time-independent perturbations with a parametric dependence $\lambda = \varphi_1(t')$. In this regime, we can thus approximate $g(t,t') \approx g_{\varphi_1(t')}(t)$ and have the additional analytical solutions (3.164) and (3.166) at our disposal again.

5.3.3 Validity analysis

The qualitative considerations from Sec. 5.2 suggest that there are three parameter regimes which favor the applicability of the theory: weak perturbations, small perturbation band widths, and high frequencies. The first two of these criteria pertain to in some sense opposite regimes because the strength of the perturbation is usually assessed relatively to its band width; see, for instance, Eqs. (3.66), (3.71), or (3.73). Moreover, the perturbations need not become too weak because they will not entail any noticeable effect otherwise, particularly in combination with high-frequency driving. Hence we intend to explore the parameter regime in which the theory (5.26) is valid in more quantitative detail.

Testbed. To obviate any additional errors due to an imprecise modeling, we choose a random-matrix model (5.1) which is a direct representative of the perturbation ensembles from Sec. 3.3. The undriven Hamiltonian H_0 is given by (3.27) with a constant level spacing of $\varepsilon = 1/512$. The perturbation matrix $(V_{\mu\nu})$ is drawn according to (3.24) with the marginal distributions $p_{\mu\nu}(v)$ from (3.23) taking a Gaussian form, namely Eq. (3.28) with $s = 0$ (no sparsity). For the perturbation profile $\sigma_v^2(E)$ from (3.10), we focus on the exponential and step shapes from (3.21) and (3.85), respectively.

As for the driving, we consider two periodic protocols with amplitude λ_0 and period τ: a step protocol

$$\lambda(t) = \lambda_0 \operatorname{sgn}\left[\sin\left(\frac{2\pi t}{\tau}\right)\right], \tag{5.35}$$

where $\mathrm{sgn}(x)$ denotes the sign function, and a sinusoidal protocol

$$\lambda(t) = \lambda_0 \, \sin\!\left(\frac{2\pi t}{\tau}\right). \tag{5.36}$$

In addition, we also study a nonperiodic, quench-like protocol in the form of a linear ramp to a target amplitude λ_0 in time τ,

$$\lambda(t) = \lambda_0 \, \Theta(t) \left[\frac{t}{\tau} \, \Theta(\tau - t) + \Theta(t - \tau)\right]. \tag{5.37}$$

These three driving protocols are visualized in Fig. 5.2 together with the functions $\varphi_1(t)$ and $\varphi_2(t)$ defined in (5.15), which encode the effect of the driving protocol on the response profile according to (5.34). For the unbiased periodic protocols, we point out that neither $\varphi_1(t)$ nor $\varphi_2(t)$ are periodic, but they oscillate approximately at the the same frequency as the driving, with $\varphi_1(t)$ being minimal (in fact, vanishing) at multiples of the driving period, whereas $\varphi_2(t)$ becomes maximal there. Moreover, $\varphi_1(t)$ is damped and decays to zero as $t \to \infty$, whereas $\varphi_2(t)$ is asymptotically periodic.

The total Hilbert space dimension of our test system is $N = 2^{14} = 16\,384$, and we choose an unperturbed eigenstate $|\alpha\rangle_0$ from the middle of the spectrum ($\alpha = 2^{13} = 8192$) as the initial state $\rho(0) = |\alpha\rangle_{00}\langle\alpha|$. We then monitor the survival probability or fidelity [284, 326] of this initial state by taking $A = \rho(0)$ as the observable. This entails $\langle A\rangle_{\rho_0(t)} = 1$ and, for sufficiently strong mixing of the eigenvectors such that (3.9) holds, $\langle A\rangle_{\bar{\rho}(t)} \simeq 0$. Consequently, the predicted dynamics according to (5.26) is precisely the squared response profile, $\langle A\rangle_{\rho(t)} = |g(t,t)|^2$.

As demonstrated in Ref. [227] (see Fig. 2 there in particular), such a setup agrees virtually flawlessly with the theory in the case of time-independent perturbations, meaning that any deviations manifesting in the driven setting can be attributed to the additional approximations, notably the truncation of the Magnus series.

Periodic driving. We consider the two periodic protocols (5.35) and (5.36) first. Fig. 5.3 displays a comparison of results for the squared response profile $|g(t,t)|^2$ for a variety of different configurations. All plots contain three types of curves: The solid black ones show the simulated fidelity of the random-matrix model, i.e., the time-dependent expectation values $\langle A\rangle_{\rho(t)}$ of the observable $A = \rho(0)$. If the theory is correct, this fidelity coincides with $|g(t,t)|^2$, hence these simulated curves correspond to the "true" squared response profile. The long-dashed red curves represent the theoretical prediction (5.23) of the response profile within the second-order Magnus approximation, obtained by integrating Eq. (5.34) numerically exactly. The short-dashed blue curves show the strong-perturbation or narrow-band asymptotics (5.31).

Every subfigure contains data for step driving (Eq. (5.35) and Fig. 5.2a) in the left column and data for sinusoidal driving (Eq. (5.36) and Fig. 5.2b) in the right column with periods $\tau = 0.5, 1, 2$ in the three rows, respectively. The driving amplitudes λ_0 and/or the perturbation profiles and band widths are different in the four subfigures, chosen such that the comparison of neighboring ones highlights the effect of one of those parameters.

We inspect how the accuracy of the theory depends on the driving period or frequency first, focusing on the numerically exact predictions (long-dashed, red). As apparent from all subfigures and anticipated in Sec. 5.2, good agreement between theory and numerics is generally found for longer times in case of small periods or high frequencies. We also notice that the effect of the driving is less pronounced in this case: At very high frequencies, the response profile $g(t,t)$ approaches unity, meaning that the system is unable to follow the rapidly oscillating perturbation and is thus expected to essentially reproduce the reference dynamics induced by the time-averaged Hamiltonian H_0.

Next we observe that the shape of the driving profile (step vs. sinusoidal) is seemingly of minor relevance with respect to the accuracy since deviations become noticeable at similar times when all other parameters are fixed, i.e., when comparing left and right panels in every row. In this context, it should also be pointed out that the effective perturbation strength $\overline{\lambda(t)^2}$ is smaller under sinusoidal driving than under step driving.

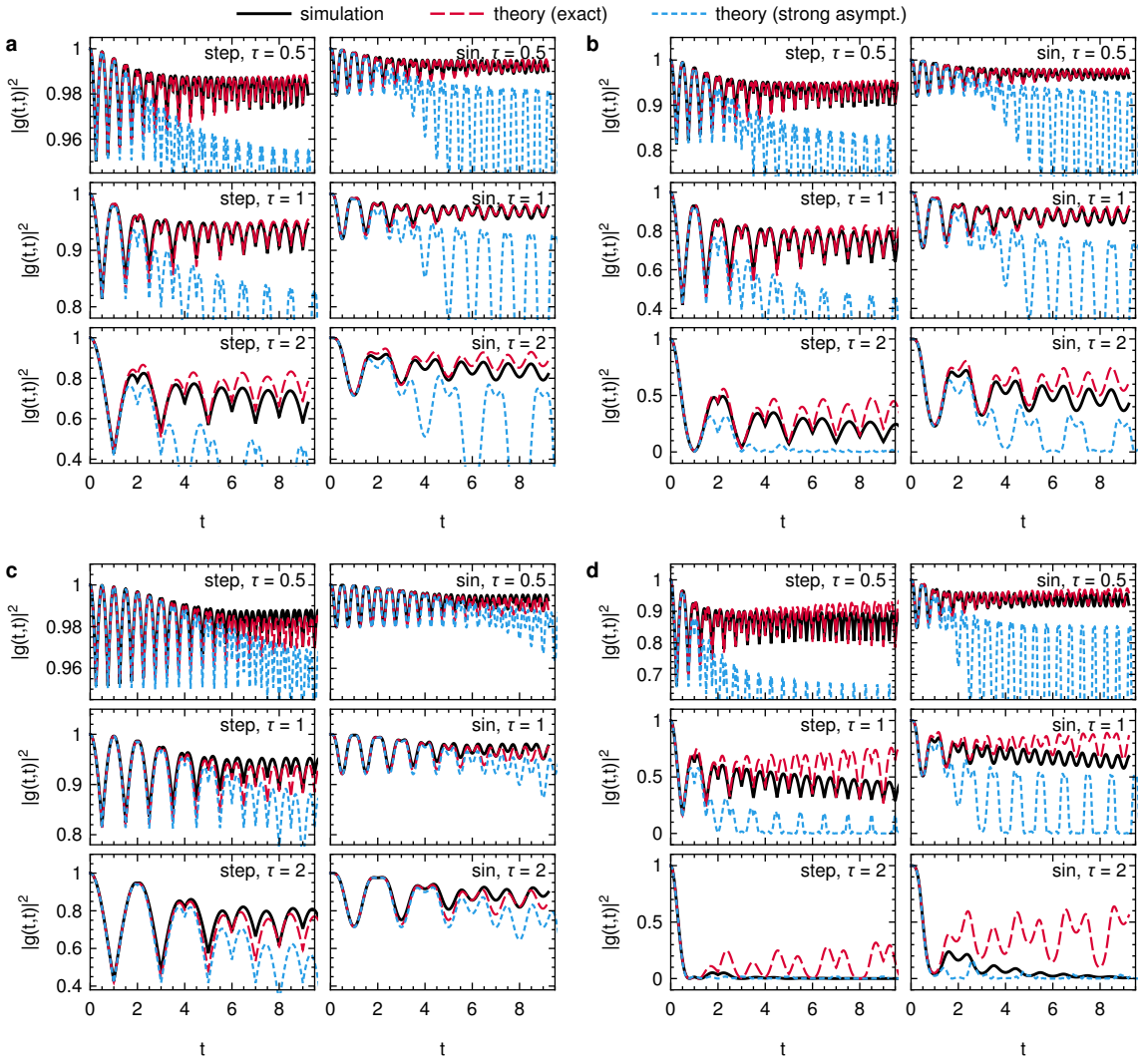


Figure 5.3: Response profiles under driven dynamics for the step and sinusoidal driving protocols $\lambda(t)$ from (5.35) and (5.36), respectively, with periods $\tau = 0.5, 1, 2$ (see top-right corner of each panel) and different amplitudes as well as different perturbation profiles $\sigma_v^2(E)$ (see below). Solid (black): Numerical simulation of a system of dimension $N = 2^{14}$ with Hamiltonian $H(t)$ from (5.1), where H_0 is of the form (3.27) with $\varepsilon^{-1} = 512$ and the perturbation V is sampled randomly according to (3.23) and (3.28) with $s = 0$ and $\sigma_v^2(E)$ as detailed below. The initial state is $\rho(0) = |\alpha\rangle\!\rangle\langle\!\langle\alpha|$ with $\alpha = N/2 = 2^{13}$ and the observable is the fidelity, $A = \rho(0)$. Long-dashed (red): Theoretical prediction (5.27) exploiting $\langle A\rangle_{\rho_0(t)} = 1$ and $\langle A\rangle_{\rho_{\rm mc}} \simeq 0$ such that $\langle A\rangle_{\rho(t)} = |g(t,t)|^2$, where $g(t,t)$ is the numerically exact solution of the integro-differential equation (5.34) for the respective driving protocol and perturbation profile. Short-dashed (blue): Same, but adopting for $g(t,t)$ the respective asymptotic solution (5.31) of (5.34) for strong perturbations and short times. **a.** Exponential perturbation profile (3.21) with $\sigma_v^2 = 1$, $\Delta_v = 250\varepsilon \approx 0.49$, and driving amplitude $\lambda_0 = 0.04$; **b.** exponential perturbation profile (3.21) with $\sigma_v^2 = 1$, $\Delta_v = 250\varepsilon \approx 0.49$, and driving amplitude $\lambda_0 = 0.08$; **c.** step perturbation profile (3.85) with $\sigma_v^2 = 1$, $\Delta_v = 250\varepsilon \approx 0.49$, and driving amplitude $\lambda_0 = 0.04$; **d.** exponential perturbation profile (3.21) with $\sigma_v^2 = 1$, $\Delta_v = 500\varepsilon \approx 0.98$, and driving amplitude $\lambda_0 = 0.08$.

To discuss the role of the driving amplitude (or, more generally, the perturbation strength), it is instructive to compare Figs. 5.3a and b, for which $\lambda_0 = 0.04$ and $\lambda_0 = 0.08$, respectively, whereas the perturbation profile $\sigma_v^2(E)$ is of the exponential form (3.21) with $\Delta_v = 250\varepsilon \approx 0.49$ in both cases. Since the corresponding crossover coupling (3.73) is $\lambda_c \approx 0.014$ for this configuration, both amplitudes correspond to relatively strong perturbations. Similarly as for the driving shape, we conclude that the onset of deviations between theory and numerics is largely independent of the driving amplitude.

On the contrary, the band width Δ_v of the perturbation profile has a significant influence on the accuracy of the prediction. This becomes apparent by comparing Figs. 5.3b and d, both of which have the same driving amplitude $\lambda_0 = 0.08$ and an exponential perturbation profile (3.21), but differ in terms of $\Delta_v = 250\varepsilon \approx 0.49$ in Fig. 5.3b as opposed to $\Delta_v = 500\varepsilon \approx 0.98$ in Fig. 5.3d. As visualized in the figure, the theory reproduces the numerics for considerably longer times when the band width is smaller. This is again in line with the expectations formulated in Sec. 5.2 by means of the arguably somewhat crude perturbation-theoretic picture. There we observed that smaller band widths entail stronger suppression of higher-order resonance chains between the unperturbed energy levels and are thus expected to prevent the dispersal of level populations for longer times. We recall that a similar effect was expected for weaker perturbations, but is apparently much less pronounced (see above), which is highlighted as well by noticing that the crossover coupling (3.73) associated with Fig. 5.3d is $\lambda_c \approx 0.020$, hence $\lambda_0/\lambda_c \approx 4.1$. This should be contrasted with $\lambda_0/\lambda_c \approx 5.7$ in Fig. 5.3b. As far as the distance from the expected crossover is concerned, the setting in Fig. 5.3b thus amounts to a stronger perturbation than the one in Fig. 5.3d, yet the agreement between theory and numerics is markedly better in the former case. We presume that the reason for the greater importance of a small band width for convergence is that it specifically suppresses transitions between levels that are far apart in the unperturbed spectrum, whereas a small perturbation strength attenuates all transitions homogeneously. Hence a small band width prevents the spreading of populations more effectively, especially the dominating contributions due to short resonance chains.

Finally, the relevance of the perturbation profile's shape can be assessed by considering Figs. 5.3a and c, where the exponential form (3.21) is employed in subfigure a and the step form (3.85) in subfigure c. Our first observation by mere comparison of the respective simulation results (solid black curves) is that there are some noticeable differences between the responses to those two profiles. That is to say, the response profiles are certainly similar in both cases, but they are still somewhat more sensitive with respect to the perturbation profile than their counterparts for time-independent perturbations studied in Sec. 3.6.2 (see Fig. 3.9 in particular). Comparing the exact theoretical curves (long-dashed red) to the simulation results, the agreement appears to be slightly better for the exponential than for the step profile, at least for the smaller periods $\tau = 0.5, 1$. Interestingly, the opposite applies when comparing the numerics to the asymptotic response profile (5.31) for strong perturbations (short-dashed blue curves). Here satisfactory agreement is found for considerably longer times in case of the step perturbation profile. The reason is that the crossover between the weak- and strong-perturbation regimes is then significantly sharper, as illustrated, for example, in Fig. 3.9 for the case of time-independent perturbations.

Generally speaking, it should be no surprise that the strong-perturbation asymptotics works better for larger driving amplitudes. Nevertheless, it is somewhat remarkable that it gives equally good (if not better) results for the sinusoidal driving compared to the step driving. This is unexpected in so far as the amplitude is always in the strong-perturbation regime for the step driving, whereas it temporarily becomes arbitrarily small during the sinusoidal protocol.

In summary, the explicit validity analysis for periodic driving confirms our expectations from Sec. 5.2 regarding the favorable effect of high frequencies and small perturbation band widths. By contrast, the third supporting factor identified there, namely weak perturbations, turns out to be of minor importance. Considering the overall scale of $|g(t,t)|^2$ in Fig. 5.3, this is reassuring because the particular combination of small driving amplitudes with high frequencies will not lead to any noticeable effect on the dynamics at all, so moderately large perturbation strengths are certainly of greater interest.

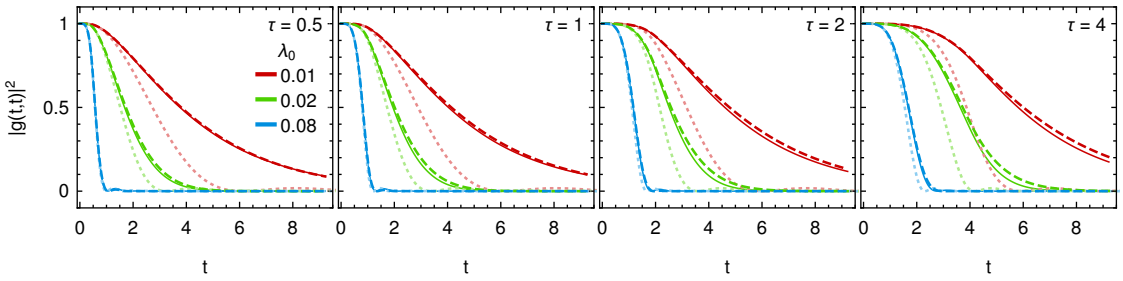


Figure 5.4: Response profiles for a linear quench according to the driving protocol $\lambda(t)$ from (5.35) for four different quench times τ (see top-right corner of each panel) and different target amplitudes λ_0 (color-coded, see legend in left-most panel). Solid: Numerical simulation of the test-system dynamics $\langle A\rangle_{\rho(t)}$ similarly as in Fig. 5.3, utilizing an exponential perturbation profile (3.21) with $\sigma_v^2 = 1$ and $\Delta_v = 500\varepsilon \approx 0.98$. Dashed: Theoretical prediction (5.27) exploiting $\langle A\rangle_{\rho_0(t)} = 1$ and $\langle A\rangle_{\rho_{\mathrm{mc}}} \simeq 0$ such that $\langle A\rangle_{\rho(t)} = |g(t,t)|^2$, where $g(t,t)$ is the numerically exact solution of the integro-differential equation (5.34) for the respective driving protocol and perturbation profile. Dotted: Same, but adopting for $g(t,t)$ the respective asymptotic solution (5.31) of (5.34) for strong perturbations and short times.

Finite-time quenches. To assess the validity in the second principal application of finite-time quenches, we adopt the linear-ramp protocol (5.37) and otherwise consider the same test system as before. In Fig. 5.4, the simulated fidelity for such a linear quench (solid curves) is compared to the numerically exact theoretical prediction of the associated squared response profile $|g(t,t)|^2$ solving (5.34) (dashed curves) as well as to the strong-perturbation asymptotics (5.31) (dotted curves). We content ourselves with an exponential perturbation profile (3.10) with $\sigma_v^2 = 1$ and $\Delta_v = 500\varepsilon \approx 0.98$. Note that this corresponds to the least favorable setup considered in Fig. 5.3 (subfigure d). The general conclusions turned out to be the same for other perturbation profiles and band widths.

The first observation is that the finite quench time delays the relaxation as one might have anticipated: The slower the quench, the longer the system will follow the unperturbed reference dynamics. To compare the effect of finite quench times to instantaneous quenches, we consider the corresponding integro-differential equations (5.34) and (3.173) for the respective response profiles $g(t,t')$ and $g_\lambda(t)$. Inspecting the post-quench dynamics first, we substitute the protocol (5.37) into the definitions of $\varphi_1(t)$ and $\varphi_2(t)$ from (5.15). For late times ($t \gg \tau$), we then find that $\varphi_1(t) \sim \lambda_0^2$ and $\varphi_2(t) \sim \lambda_0\tau^2/16$. If the quench time is very short ($\tau \to 0$), we thus recover the relation (3.173) for instantaneous quenches from the finite-time result (5.34). The slower the quench, the more important does the correction effect due to the $\varphi_2(t')$ term in (5.34) become. For very short times at the beginning of the quench, in turn, we find that $\varphi_1(t) \sim (\lambda_0 t/2\tau)^2$ and $\varphi_2(t) \sim (\lambda_0 t^2/12\tau)^2$. Together with (5.34), we therefore conclude that the rate at which $g(t,t')$ changes with t becomes smaller with increasing τ, leading to the above-observed delayed response for slow quenches.

Coming back to Fig. 5.4, we also notice that deviations between numerics and theory occur at late times and for slow quenches first, conforming with our expectation from Sec. 5.2. Remarkably, the agreement between the numerics (solid lines) and the exact theory (dashed lines) is still very good for all target amplitudes up to the longest quench times considered. Small deviations for slower quenches become noticeable only for weak perturbations since the response profile decays slower in this case, too, such that there is still nontrivial behavior beyond the presumed validity regime of the second-order Magnus approximation. Note that the crossover coupling (3.73) is $\lambda_c \approx 0.020$, meaning that the selected target amplitudes cover the full range from relatively weak to strong perturbations. The strong-perturbation approximation (5.31) for $g(t,t)$ naturally works best for the largest target amplitude $\lambda_0 = 0.08$, but still describes the short-time behavior for all three perturbation strengths, in accordance with the analysis from Sec. 3.6.2.

To conclude, the present validity analysis demonstrates that the theoretical assumptions can be upheld over a significant range of parameters. The dynamics under both periodic and quench-like driving protocols is reproduced excellently for time scales up to several multiples of the characteristic driving time τ (period or quench duration). The interplay of fast driving and small perturbation band widths turns out to be particularly beneficial for the applicability at later times.

5.4 Examples

As in the previous chapters, an important ingredient to establish the usefulness of the obtained relations for the typical dynamics is to test them in concrete physical models. Since it is our perception that the short-time response of many-body systems to time-dependent driving has received only little attention in the literature so far (see also Sec. 5.5 below), we will rely on numerical simulations of our own and thereby re-encounter two spin-$\frac{1}{2}$ systems which already featured as illustrations of the relaxation theory for time-independent perturbations in Sec. 3.7.

Two-dimensional spin-1/2 lattice. In our first example, we look into the 4×4 spin-$\frac{1}{2}$ lattice first introduced in Sec. 3.2 again. The Hamiltonian is of the form (5.1) with H_0 from (3.20a) and V from (3.20b). As before, we focus on an energy window comprising the central 60 % of states with mean level spacing $\varepsilon = 1.90 \times 10^{-3}$. The perturbation profile $\sigma_v^2(E)$ is approximately of the exponential form (3.21) with $\sigma_v^2 = 5.02 \times 10^{-3}$ and $\Delta_v = 7.32$, hence $\alpha_v = 2.64$.

We already investigated the dynamics of this system under time-independent perturbations in Sec. 3.7 (see Fig. 3.10 in particular) and employ the same basic setup here. That is to say, we consider the central magnetization correlation m_c from (3.192) for the initial state $\rho(0) = |\psi\rangle\langle\psi|$ from (3.194) on the one hand and the central spin-flip correlation j_c from (3.195) for an initial state using the dynamical-typicality scheme (2.39) as specified below Eq. (3.195).

Fig. 5.5 shows the resulting undriven behavior together with driven dynamics for three different periods and two amplitudes (solid lines). For the dynamics of m_c in Fig. 5.5a, we adopt the

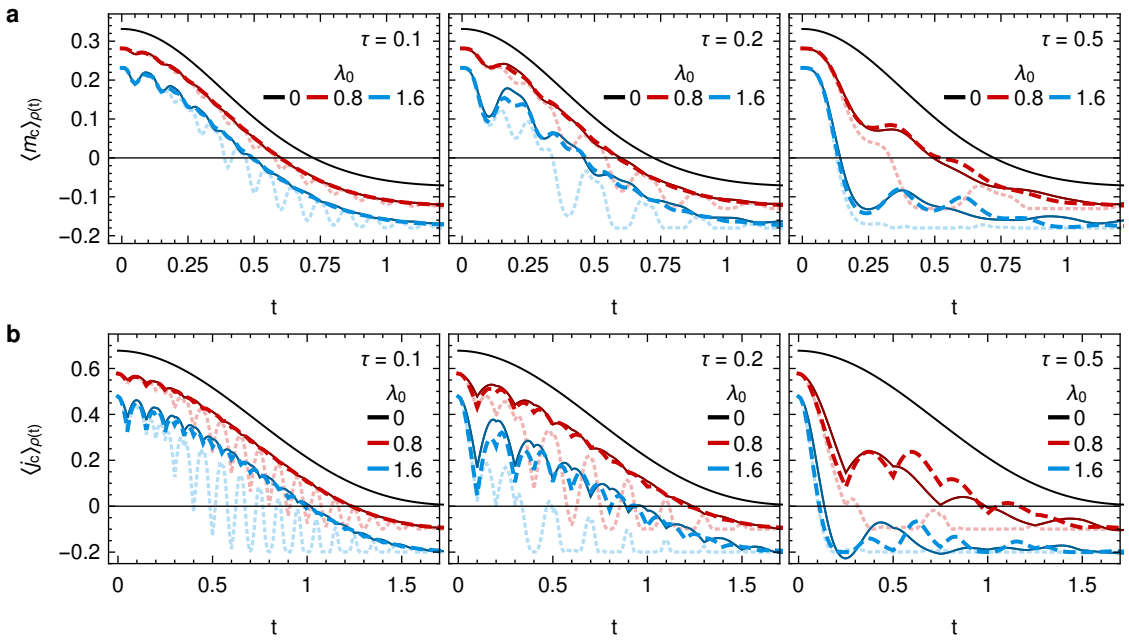


Figure 5.5: Time evolution in the two-dimensional spin-$\frac{1}{2}$ model $H(t) = H_0 + \lambda(t)V$ with H_0 from (3.20a) and V from (3.20b) under periodic driving for periods $\tau = 0.1, 0.2, 0.5$ (see top-right corner of each panel) and amplitudes $\lambda_0 = 0, 0.8, 1.6$ (see legend) for two different observables A and initial states $\rho(0) = |\psi\rangle\langle\psi|$ (see below). Solid: Numerical simulation results obtained using exact diagonalization for the undriven dynamics ($\lambda_0 = 0$) and a fourth-order Runge-Kutta scheme for the driven case ($\lambda_0 > 0$). Dashed: Analytical prediction (5.27), employing the numerical solution for $\lambda_0 = 0$ (solid black curve) for the reference dynamics $\langle A\rangle_{\rho_0(t)}$, the exact numerical solution of (5.34) for the response profile $g(t,t)$ with the empirically determined exponential perturbation profile (3.21) with intrinsic perturbation strength $\alpha_v = 2.64$ and band width $\Delta_v = 7.32$ (cf. Figs. 3.1 and 3.2), and the known $\langle A\rangle_{\rho_\mathrm{mc}}$ (see below). Dotted: Analytical prediction (5.27) adopting the strong-perturbation asymptotics (5.31) for $g(t,t)$ instead. **a.** $A = m_\mathrm{c}$ from (3.192) with $\langle m_\mathrm{c}\rangle_{\rho_\mathrm{mc}} = -0.0805$ (see below Eq. (3.195)), $|\psi\rangle$ from (3.194), and sinusoidal driving protocol (5.36). Curves for $\lambda_0 > 0$ are shifted in steps of -0.05 for better visibility. **b.** $A = j_\mathrm{c}$ from (3.195) with $\langle j_\mathrm{c}\rangle_{\rho_\mathrm{mc}} = 0$ (see below Eq. (3.195)), $|\psi\rangle$ as described below Eq. (3.195), and sinusoidal driving protocol (5.36). Curves for $\lambda_0 > 0$ are shifted in steps of -0.1 for better visibility.

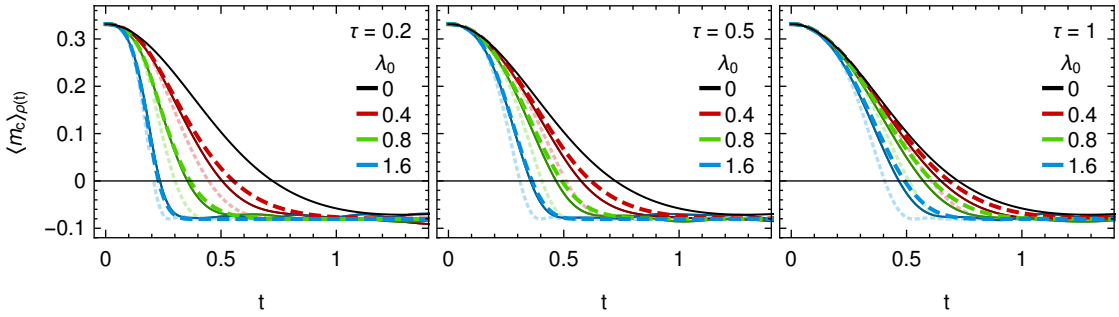


Figure 5.6: Time evolution in the two-dimensional spin-$\frac{1}{2}$ model $H(t) = H_0 + \lambda(t)V$ with H_0 from (3.20a) and V from (3.20b) under the linear-ramp driving protocol (5.37) for quench times $\tau = 0.2, 0.5, 1$ (see top-right corner of each panel) and amplitudes $\lambda_0 = 0, 0.4, 0.8, 1.6$ (see legend). Solid: Numerical simulation results obtained using exact diagonalization for the undriven dynamics ($\lambda_0 = 0$) and a fourth-order Runge-Kutta scheme for the driven case ($\lambda_0 > 0$). Dashed: Analytical prediction (5.27), employing the numerical solution for $\lambda_0 = 0$ (solid black curve) for the reference dynamics $\langle m_\mathrm{c}\rangle_{\rho_0(t)}$, the exact numerical solution of (5.34) for the response profile $g(t,t)$ with the empirically determined exponential perturbation profile (3.21) with intrinsic perturbation strength $\alpha_v = 2.64$ and band width $\Delta_v = 7.32$ (cf. Figs. 3.1 and 3.2), and the known $\langle m_\mathrm{c}\rangle_{\rho_\mathrm{mc}} = -0.0805$ (see below Eq. (3.195)). Dotted: Analytical prediction (5.27) adopting the strong-perturbation asymptotics (5.31) for $g(t,t)$ instead.

sinusoidal driving protocol (5.36), whereas the step protocol (5.35) is used for j_c in Fig. 5.5b to give variety. The theoretical prediction (5.27) is compared to those numerical simulations by exploiting the numerics (solid black curve) for the undriven dynamics $\langle A\rangle_{\rho_0(t)}$ and the known values $\langle m_\mathrm{c}\rangle = -0.0805$ and $\langle j_\mathrm{c}\rangle = 0$ (see below Eq. (3.195)) for the thermal expectation values. The response profile $g(t,t')$ is either evaluated exactly by solving (5.34) numerically (dashed lines) or approximated analytically using the asymptotic solution (5.31) (dotted lines).

The overall conclusions are similar to the observations from Sec. 5.3.3: The agreement between the numerics and the exact theory is very good for high frequencies (small τ) and/or short times. The deviations at late times and lower frequencies appear slightly larger in Fig. 5.5b, but the theory still reflects the dynamics qualitatively. The asymptotic theory for strong perturbations performs less convincingly, but the driving amplitudes do not reach very far into the strong-perturbation regime either since the crossover coupling (3.73) is $\lambda_\mathrm{c} \approx 0.75$.

In Fig. 5.6, we compare numerics and theory for the linear quench protocol (5.37), focusing on the magnetization correlation m_c in an otherwise identical setup. In this more realistic quench setting compared to Fig. 5.4, we again observe the expected delayed response, entailing that the finite-quench dynamics remains closer to the undriven behavior for longer quench times. The theory reproduces the numerical data well in general with small deviations becoming more pronounced at later times.

Two coupled spin-1/2 chains. As a second example system, we consider two isolated spin chains that are coupled by the perturbation, such that $H(t)$ in Eq. (5.1) is composed of H_0 and V as given in (3.202) and (3.203), respectively. For time-independent coupling strengths, a comparison between numerics and the corresponding theory from Chapter 3 had been carried out in Fig. 3.14. We take the same initial state $\rho(0) = |\psi\rangle\langle\psi|$ as specified in Eq. (3.204) and focus on the single-site magnetization $\sigma^z_{1,1}$.

Fig. 5.7a shows the observed dynamics for the sinusoidal driving protocol (5.36) and various amplitudes and periods. As usual, the solid lines correspond to the "true" dynamics as obtained from a numerical integration of Schrödinger's equation using a second-order Suzuki-Trotter decomposition with time step $\Delta t = 0.01$. We also plot the theoretical prediction (5.27) exploiting $\langle\sigma^z_{1,1}\rangle_{\rho_\mathrm{mc}} = 0$ (see dash-dotted lines) and the solution of (5.34) for $g(t,t')$ adopting an exponential perturbation profile (3.21) with $\alpha_v = 0.90$ and $\Delta_v = 5.8$, the same estimates as in Sec. 3.7 (cf. Fig. 3.14). This yields the dashed curves in Fig. 5.7a, which are in excellent agreement with the numerics for small $\tau = 0.25$ (left panel) and likewise match well for the larger value of $\tau = 2$ (right panel). More pronounced deviations become visible for the intermediate value of $\tau = 0.5$ (middle panel) and larger driving amplitudes. Here the theory predicts that the expectation values stabilize near

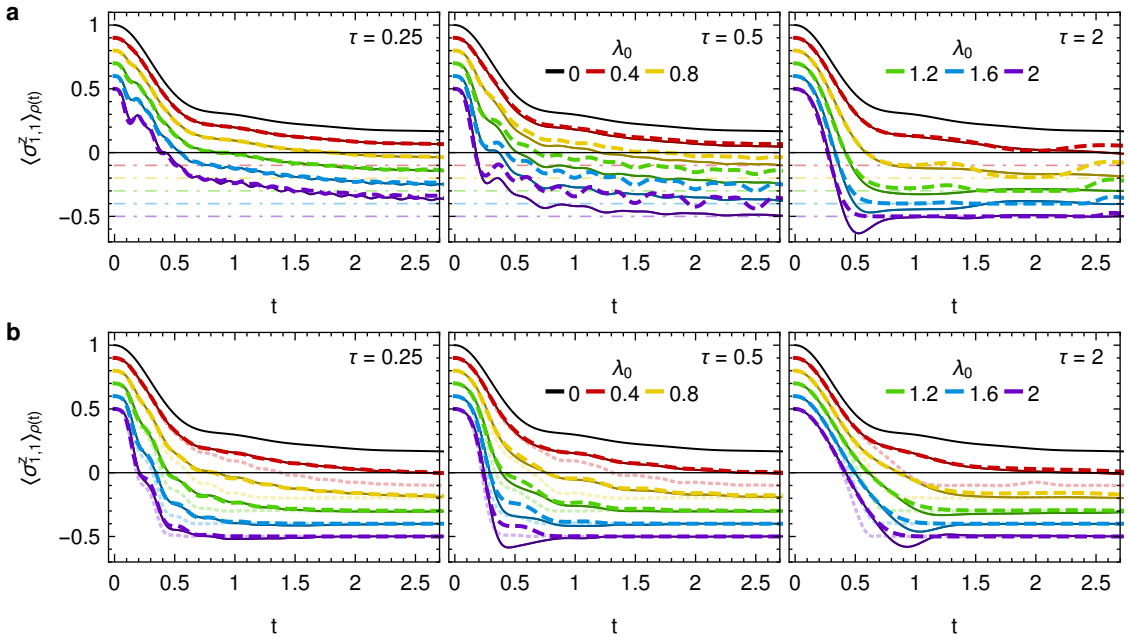


Figure 5.7: Time evolution for two coupled spin-$\frac{1}{2}$ chains with $H(t) = H_0 + \lambda(t)V$, where H_0 and V are given by (3.202) and (3.203), respectively, under periodic driving for periods $\tau = 0.25, 0.5, 2$ (see top-right corner of each panel) and various amplitudes (see legend). The observable $A = \sigma^z_{1,1}$ is a single-site magnetization, the initial state is $\rho(0) = |\psi\rangle\langle\psi|$ with $|\psi\rangle$ from (3.204). Solid: Numerical simulation results obtained using a second-order Suzuki-Trotter decomposition with time step $\Delta t = 0.01$. Dash-dotted (only **a**): Thermal expectation value $\langle\sigma^z_{1,1}\rangle_{\rho_{\mathrm{mc}}} = 0$. Dashed: Analytical prediction (5.27), employing the numerical solution for $\lambda_0 = 0$ (solid black curve) for the reference dynamics $\langle\sigma^z_{1,1}\rangle_{\rho_0(t)}$, the exact numerical solution of (5.34) for the response profile $g(t,t)$ adopting an exponential perturbation profile (3.21) with $\alpha_v = 0.95$ and $\Delta_v = 5.8$ (see Fig. 3.14 and corresponding discussion), and $\langle\sigma^z_{1,1}\rangle_{\rho_{\mathrm{mc}}} = 0$. Dotted (only **b**): Analytical prediction (5.27) adopting the strong-perturbation asymptotics (5.31) for $g(t,t)$ instead. Data for $\lambda_0 > 0$ are shifted in steps of -0.1 for better visibility. **a.** Sinusoidal driving protocol (5.36); **b.** biased-cosine driving protocol (5.38).

a value different from $\langle\sigma^z_{1,1}\rangle_{\rho_{\mathrm{mc}}} = 0$ (dash-dotted lines), whereas the actually observed behavior shows an approach towards $\langle\sigma^z_{1,1}\rangle_{\rho_{\mathrm{mc}}}$. Comparing to the left and right panels, we observe that the numerics indeed stays away from $\langle\sigma^z_{1,1}\rangle_{\rho_{\mathrm{mc}}}$ for higher frequencies (left panel), and the theory indeed predicts an approach of $\langle\sigma^z_{1,1}\rangle_{\rho_{\mathrm{mc}}}$ for smaller frequencies (right panel). Qualitatively, the theory thus entails a similar crossover between the limiting regimes of fast and slow driving as suggested by the numerics, but misses the precise point of the transition.

Note that the crossover coupling $\lambda_{\mathrm{c}} \approx 1.1$, so the figure contains data for rather small amplitudes. Particularly in combination with high frequencies (left panel), it appears that the dynamics barely changes compared to the unperturbed behavior. We emphasize that this is a nontrivial correct prediction of the theory, too. Since $\langle\sigma^z_{1,1}\rangle_{\rho_{\mathrm{mc}}} = 0$ whereas the observed value at long times is noticeably larger than zero, the fact that $\langle\sigma^z_{1,1}\rangle_{\rho(t)}$ as calculated from (5.27) settles down to a nonthermal value in accordance with the numerics is solely encoded in the response profile $g(t,t')$ obtained via (5.34). This should be contrasted with the behavior observed for a constant perturbation shown in Fig. 3.14a: Here the dynamics eventually approaches $\langle\sigma^z_{1,1}\rangle_{\rho_{\mathrm{mc}}} = 0$ for seemingly arbitrarily small $\lambda > 0$. For sufficiently fast driving, the dynamics thus exhibits characteristics reminiscent of prethermalization in the sense that the system avoids to thermalize for an extended period of time (see also the discussion in the subsequent section), and the theory correctly reproduces this behavior.

In Fig. 5.7b, we apply a perhaps experimentally more realistic driving protocol of the form

$$\lambda(t) = \frac{\lambda_0}{2}\left[1 - \cos\left(\frac{2\pi t}{\tau}\right)\right], \tag{5.38}$$

which we imagine to be implemented, for example, by varying the spatial distance between the

two chains. Note that this is a biased protocol since $\overline{\lambda(t)} = \lambda_0/2$, i.e., it can be viewed as a mixture of a pure (unbiased) periodic driving and a finite-time quench (on average). In all other respects, the setup is the same as in Fig. 5.7a. Note that we no longer display the equilibrium value $\langle\sigma^z_{1,1}\rangle_{\rho_{\mathrm{mc}}} = 0$ since all configurations (except for the undriven case $\lambda = 0$) ultimately approach this value. Instead, we now include (similarly as in Figs. 5.5 and 5.6) the theoretical prediction utilizing for $g(t, t')$ the asymptotic form (5.31) for strong perturbations (dotted lines), which has been omitted in Fig. 5.7a for the sake of clarity. We observe that the exact theory (dashed lines) conforms almost flawlessly with the numerics for high frequencies or small amplitudes, whereas minor deviations occur for lower frequencies and larger amplitudes, similar to the ones observed in the undriven scenario from Fig. 3.14a. The strong-perturbation approximation, in turn, performs reasonably well for high frequencies and large amplitudes or short times, i.e., in the regime where it is supposed to do so.

5.5 Discussion

The main result of this chapter is the prediction (5.27) describing the typical short-time response of nearly isolated many-body quantum systems to time-dependent perturbations $\lambda(t)V$ when the operators V are contained in one of the ensembles introduced in Sec. 3.3. As demonstrated for various models and setups in Sec. 3.7, these perturbation ensembles capture the effect of realistic physical perturbations in many different situations, and indeed we found good agreement between the enhanced relaxation theory for driven systems and realistic example systems in Sec. 5.4 here again. Notably, the theory correctly predicts highly nontrivial time-dependent expectation values in those systems remarkably well in the expected regime of validity (see, in particular, Figs. 5.5 and 5.7 as well as Fig. 5.3). Moreover, it applies to largely arbitrary driving protocols as long as the amplitudes do not become overly large, including the important special cases of periodic driving and finite-time quenches.

The connection between the driven dynamics and the undriven behavior is established by the response profile $g(t, t')$, which links the function $\sigma_v^2(E)$ from (3.10) as the decisive perturbation characteristic to the dynamical response via the integro-differential equation (5.34). Integrating this equation numerically is straightforward and thus readily allows to evaluate the theoretical prediction (5.27) quantitatively. Moreover, we also derived an explicit analytical expression (5.31) which solves Eq. (5.34) for reasonably large perturbation strengths or small band widths Δ_v.

Limitations. Even though the theory is applicable to a large variety of different models, it also suffers from some limitations, which are ultimately due to the adopted modeling of perturbations and approximations employed during the derivation. The typicality approach and its embedding of the true perturbation of interest into an ensemble of operators with similar perturbation profiles entails essentially the same restrictions regarding admissible setups as in the time-independent setting. These restrictions (as well as potential strategies to overcome them) have been discussed in detail in Secs. 3.6.2 and 3.8. Notably, we still need to demand that the dynamics should not involve macroscopic transport, i.e., the initial state should be roughly homogeneous in space on a macroscopic level or at least the observable should be insensitive to potential inhomogeneities. Furthermore, regarding the description of finite-time quenches, the often adopted setting where the system is prepared in the ground state of the pre-quench Hamiltonian is unfortunately beyond the scope of our present analysis, too, because the assumption of a homogeneous level density (see Prerequisite (i)) will commonly be violated then. Similarly, quenches between integrable models are presumably not covered if the relevant local integrals of motion change with $\lambda(t)$, whereas quenches between models with identical conserved quantities or from integrable to nonintegrable ones are generally admitted (see also the discussion in Sec. 3.8).

The truncated Magnus expansion utilized during the derivation implies additional restrictions, which were expounded based on theoretical considerations in Sec. 5.2 and verified quantitatively in Sec. 5.3.3. They entail that the theoretical prediction cannot be expected to apply at late times and will be more reliable if the characteristic time scale of the driving protocol is short (or, equivalently, for high driving frequencies). Moreover, relatively narrow perturbation profiles with a small band width (see Eq. (3.13)) have a positive impact on the applicability.

A natural question regards the inclusion of higher-order terms from the Magnus expansion into the auxiliary Hamiltonians (5.7). Generally speaking, such corrections should enhance the applicability further into the regime of longer times and slower driving. Unfortunately, though, such an extension will involve significant technical complications. The virtue of the second-order truncation is that it is still linear in V, which enables a direct mapping onto the setting from Chapter 3. The third-order contribution, by contrast, contains terms of order V^2. In principle, such a contribution could be accounted for by introducing renormalized variances and covariances in the distribution $p(V)$ of V from (3.24). This particularly affects, for example, the step from Eq. (3.47) (or Eq. (3.45)) to Eq. (3.50) in the calculation of the second moment of eigenvector overlaps from Sec. 3.4.2: In lieu of the independent Gaussian distributions (3.49) for the matrix elements $V_{\mu\nu}$, we would have to employ a still Gaussian, but correlated joint distribution $p(V)$ for them. While this does not preclude performing the ensemble average and arriving at an analog of Eq. (3.50), the subsequent steps of the calculation to process the supersymmetric degrees of freedom become considerably more involved. As indicated in Sec. 3.8 already, an efficient way to deal with correlated perturbation ensembles seems to be vital for further progress.

Relation to other studies. Regarding the detailed dynamics of many-body quantum systems subject to arbitrary driving protocols, analytical investigations are generally scarce. When it comes to identifying general principles regarding the behavior of periodically driven systems in particular, the focus in the literature thus far has clearly been on the long-time behavior, usually assessed in terms of stroboscopic dynamics, where states and observable expectation values are considered at integer multiples of the driving period only [306, 309, 323, 325, 327–330]. Since periodically driven systems can be shown to approach a periodic state synchronized with the driving at long times under basically the same prerequisites for which isolated systems equilibrate (cf. Sec. 2.2.1) [328, 331], the asymptotic behavior is indeed essentially characterized by stroboscopic observations. This stroboscopic time evolution can be described in terms of the *Floquet Hamiltonian* H_F (see also Sec. 5.2), an effective, time-independent Hamiltonian related to the propagator for one period via $\mathcal{U}(\tau) = \mathrm{e}^{-\mathrm{i}H_\mathrm{F}\tau}$. Note that, contrary to $\mathcal{U}(\tau)$, the Floquet Hamiltonian H_F is not unique because we can add arbitrary multiples of the driving frequency $\frac{2\pi}{\tau}$ to any of its eigenvalues and still obtain the same time-evolved state at stroboscopic times. Within our approach outlined in Sec. 5.1, we can identify H_F with the auxiliary Hamiltonian $H^{(\tau)}$ from (5.7) due to (5.3). Adopting the second-order truncation (5.11) of the Magnus expansion and assuming an unbiased driving such that $\Lambda_1(\tau) = \overline{\lambda(t)} = 0$, we thus find $H_\mathrm{F} \approx H_0 + \mathrm{i}[V, H_0]\Lambda_2(\tau)/\tau$. Besides the Magnus expansion, also a variety of other approaches have been suggested to calculate H_F approximately; see, for instance, Refs. [330, 332, 333].

To assess the asymptotic behavior of periodically driven systems at late times, one can then adopt well-established tools and concepts from quantum statistical mechanics for time-independent Hamiltonians (cf. Chapter 2) to H_F, leading to effective "periodic thermodynamics" [328]. In the end, many of the concepts presented in Chapter 2 are believed to have a corresponding "Floquet analog," such as a *Floquet ETH* [138, 309, 315, 320], a *periodic Gibbs ensemble* [328, 329], *Floquet many-body localization* [315, 321, 327, 333–335], *Floquet prethermalization* [308, 309, 314, 317, 323, 325], etc.

The generally emerging picture is that periodically driven quantum many-body systems eventually absorb energy indefinitely and heat up to infinite temperature [306, 320–323], meaning that the level populations of the time-averaged Hamiltonian approach a uniform distribution, unless there are mechanism preventing thermalization such as an extensive number of conserved quantities [328, 331] or many-body localization [321, 333, 335]. However, even in the generic case, energy absorption can be suppressed for time scales that are exponentially large in the driving frequency [307, 309, 324, 325], which comes in useful for the applicability of the theory developed here. In fact, the time scale for the onset of heating has been related to the time scale when the Magnus expansion breaks down [306, 322, 327]. In the regime where the Magnus expansion is applicable, which is the maximal time scale up to which the methodological approach from this chapter can possibly work, the energy distribution will thus not spread significantly compared to its initial form, justifying once more our considerations from Sec. 5.2 regarding the homogeneous density of states in the relevant subspace for the dynamics.

An interesting suggestion to extract the matrix structure of the driving operator V in the eigenba-

sis of the reference Hamiltonian H_0 based on the heating rates observed at late times for different driving frequencies was put forward in Ref. [323]. Given the direct relationship between the perturbation profile $\sigma_v^2(E)$ and the response profile $g(t,t')$ as encoded in Eq. (5.34), it might be possible to probe that matrix structure based on the short-time response of the system, too. Note that such a reconstruction should in principle be possible already from the response to time-independent perturbations via Eq. (3.173), but since the corresponding response profile $g_\lambda(t)$ turned out to be rather insensitive to details of $\sigma_v^2(E)$, extremely high measurement precision would be required to extract details beyond the intrinsic strength α_v and the band width Δ_v. In the driven case, by contrast, the response profile $g(t,t')$ exhibited somewhat higher susceptibility to variations of $\sigma_v^2(E)$, and the freedom to choose essentially arbitrary protocols $\lambda(t)$ entails a greater variability of probe dynamics as well. In particular the second moment Σ_2 of $\sigma_v^2(E)$ (cf. Eq. (3.175)) should be accessible with relatively little effort due to Eqs. (5.14) and (5.32) (see also Eq. (5.29)).

Generally speaking, the results of this chapter therefore complement the aforementioned earlier studies of the stroboscopic long-time behavior by a detailed, continuously time-resolved description of the dynamics at short times, which can still cover multiple driving periods provided that the frequency is sufficiently high and the perturbation profile sufficiently narrow. As far as the above collected phenomenology is concerned, the present theory thus seems particularly suited to describe the initial transient of systems exhibiting Floquet prethermalization [308, 309, 314, 317, 325]. Note that this concept of Floquet prethermalization is not restricted to time-averaged or undriven systems H_0 with conserved quantities that are broken by the perturbation V. Instead, it applies likewise to generic (nonintegrable) systems and describes the behavior that a system driven at high frequency spends a long time close to the equilibrium state associated with the time-averaged Hamiltonian H_0 (or, more generally, some truncated approximation of the Floquet Hamiltonian H_F [308, 309]) and the initial energy $\langle H_0 \rangle_{\rho(0)}$, which will commonly be a thermal state, albeit at finite temperature. "True" thermal equilibrium is then reached at later times in the form of an infinite-temperature state. Observing that ρ_mc in Eq. (5.27) refers to the prethermal state corresponding to the initially occupied energy window, the prediction thus precisely describes such Floquet-prethermalization transient dynamics, at least if the undriven system thermalizes. The eventual approach of an infinite-temperature state, however, is not accounted for and cannot be hoped to emerge within a Magnus-expansion scheme as explained above.

Experiments. Much like the theoretical investigations of periodically driven quantum many-body systems, experimental efforts regarding dynamical properties have mostly concentrated on the long-time behavior and the emergence of exotic phases [283]. Unfortunately, we thus could not make out any experimental time-series data in the literature that would allow a direct comparison with the derived theoretical predictions. Nevertheless, obtaining such data should be readily possible in principle with presently available methods and techniques, for instance by means of cold atoms or tunable effective spin systems as employed in polarization-echo experiments (see also Secs. 2.1 and 4.1). The only caveat is to determine the perturbation profile $\sigma_v^2(E)$ in a given experimental setup. We argue that this can be circumvented by considering a similar setup with time-independent perturbations first ($\lambda(t) = \lambda_0$). In this case, the theory from Chapter 3, notably Eq. (3.188) in combination with (3.166) for $g_\lambda(t)$, can be employed to determine the intrinsic strength α_v and band width Δ_v of the relevant perturbation (cf. Eqs. (3.12) and (3.13)). Based on these estimates, the response profile for time-dependent perturbations $g(t,t')$ can then be approximated by solving Eq. (5.34) for a generic perturbation profile such as the exponential form (3.21). As demonstrated in Fig. 5.3 (cf. subfigures a and c), for sufficiently short times and high frequencies, the observable differences of time-dependent expectation values between different perturbation profiles will be relatively small and presumably below the experimental uncertainties. Note that such a method was also effectively adopted in the example from Fig. 5.7, even though the choice of an exponential profile was guided by numerical observations for smaller system sizes (see also the discussion around Fig. 3.14). On the other hand, if the experimental accuracy is higher, this high-resolution information can be exploited to determine the perturbation profile from the experiment as indicated above.

Connection to echo dynamics. The close methodological relationship between the present chapter and Chapter 3 has already been stressed on several occasions. Notwithstanding, the analysis of driven systems here is also similar or perhaps complementary to the echo dynamics studied in Chapter 4. Indeed, the protocol (4.34) for imperfect reversal can be mapped to the model (5.1)

by a suitable definition of H_0, V, and $\lambda(t)$ in terms of the echo operators[4] $H_0^{(\mathrm{e})}$, $V^{(\mathrm{e})}$ and the parameter ϵ. We point out, however, that such a mapping would generally violate Prerequisite (vi) of small correlations between H_0 and V in (5.1). More importantly, though, the resulting perturbation V would not be small because the dominant contribution comes from $H_0^{(\mathrm{e})}$, the reference Hamiltonian of the echo protocol. The echo studies from Chapter 4 are thus not simply a special case of the analysis from this chapter for weak-to-moderate perturbations, but rather describe a different setting. In fact, the roles of H_0 and V are essentially reversed in the echo setup, meaning that the latter could also be viewed as a theory for extremely strong driving, albeit restricted to a single cycle.

The echo protocol (4.13) for the imperfect-preparation scenario, in turn, was meant to model inaccuracies during the process of switching from $H_0^{(\mathrm{e})}$ to $-H_0^{(\mathrm{e})}$. Perhaps more realistically, this switching could be modeled by a finite-time quench from $H(0) = H_0^{(\mathrm{e})}$ to $H(\delta) = -H_0^{(\mathrm{e})}$ mediated by a perturbation $V = -2H_0^{(\mathrm{e})}$, supplemented by additional terms to describe the actual inaccuracies. Hence the pertinent perturbation operator V violates Prerequisite (vi), too. Again, the echo setting thus represents a rather different situation that cannot be simply embedded into the present approach for driven systems.

[4]Note that we introduced the superscript "(e)" for operators pertaining to the echo setting from Chapter 4 since their meaning is generally different from the reference Hamiltonian and perturbation operators in (5.1) for the present chapter.

6 Conclusions

The central theme of this thesis was to find comparatively simple relations for the behavior of many-body systems based on the well-verified theoretical modeling of their individual constituents. For our everyday experience, the relevant microscopic degrees of freedom are atoms and molecules, whose behavior is described by nonrelativistic quantum mechanics. If many degrees of freedom are involved, the microscopic dynamics is usually extremely complicated and hardly analytically tractable. Yet the macroscopically observable behavior is often surprisingly regular. For instance, as expounded in Chapter 2, it is well-established that isolated many-body quantum systems generically equilibrate, meaning that observable expectation values become essentially stationary at long times, and usually even thermalize, i.e., the stationary values coincide with the prediction from the pertinent thermodynamic ensemble (e.g., microcanonical or canonical).

Our primary interest was to describe the dynamics of systems that are initially far from such an equilibrium state under the influence of weak-to-moderate perturbations: How do these systems relax towards equilibrium (Chapter 3)? Can we characterize the irreversibility of the relaxation processes (Chapter 4)? And how does time-dependent driving affect the response (Chapter 5)? In essence, our answers to these questions consisted in analytical predictions for the time evolution of expectation values of experimentally realistic observables in three distinct but related setups.

Typicality. The common procedure to derive these predictions is based on so-called typicality methods [6, 179, 180]. Conceptionally, these exploit the phenomenologically well-established fact that the behavior of a macroscopic system is often characterized by a few parameters (e.g., energy, pressure, magnetization, ...), meaning that it is not necessary to know all complicated microscopic details to describe it. The general idea of the typicality approach is thus to temporarily consider, instead of one particular system, an entire class of similar systems. In our case, this amounts to an ensemble of Hamiltonians H, which are supposed to emulate the key features of the true system of interest as closely as possible. In a first step, we calculated the average dynamics over all H from a given ensemble. Second, we showed that the probability that a single system from the ensemble exhibits noticeable deviations from the average behavior is exponentially suppressed in the degrees of freedom. In the considered many-body setting, the dynamics of nearly all systems in the ensemble is thus practically indistinguishable from the average. In a third step, we argued that the true system of interest is one of those "typical" members, unless there are specific reasons to the contrary. Hence the ensemble-averaged dynamics becomes the aforementioned prediction for the actual behavior of the true system.

From a technical point of view, the pivotal step is to evaluate averages over products of transformation matrix elements between a fixed reference basis of the underlying Hilbert space and the variable eigenbases of the individual H in the ensemble. Depending on the precise problem, we adopted various approaches to do so, most notably so-called supersymmetry methods [163, 220, 222–224] as developed in the context of random matrix theory and extended for our present purposes.

Perturbed relaxation. In Chapter 3, we studied systems whose Hamiltonian $H_\lambda = H_0 + \lambda V$ can be split into two parts: an unperturbed reference system H_0 one the one hand, for which we assume that the dynamics is known, and a perturbation λV on the other hand, which should be sufficiently weak so that the thermodynamical properties of the system do not change noticeably. The dynamical properties can then still be modified significantly, especially if the reference system is in some sense "more regular" or "simpler," for example because it is noninteracting, consists of isolated subsystems, has one or more conservation laws that are broken by the perturbation, etc.

Predicting the response of a given system to a perturbation is a recurrent and important problem in many areas of physics. In the context of many-body systems, however, remarkably little is known analytically. Not least, this is the case because arguably the standard method using ordinary perturbation theory (i.e., expanding the propagator in powers of λ) is usually limited to uninterestingly small time scales below the relevant relaxation time due to the exceedingly dense many-body energy spectrum and the concomitant small denominators in the perturbation series.

We evaded this difficulty by employing a typicality approach for the perturbation operator V. Importantly, our perturbation ensembles admit sparse and banded matrix representations of V

(in the eigenbasis of H_0) as they commonly arise as a result of the local and few-body character of interactions. However, a caveat is that the precise geometry and form of interactions are not considered directly, so relaxation processes involving macroscopic currents may in general show deviations from the theory. Yet our resulting analytical prediction for the time-dependent perturbed expectation values, Eq. (3.160), provides a unified framework for such diverse phenomena as prethermalization, quantum quenches, or the relaxation of system-bath compounds in the absence of macroscopic spatial inhomogeneities. For weak perturbations and late times, in particular, it furthermore demonstrates Fermi's golden rule in the many-body context.

Besides describing the evolution towards equilibrium, the theory also provides a characterization of the eventually reached equilibrium state, which generalizes and concretizes previous results obtained in a similar setup [114, 257]. As in those earlier studies, the overall conclusion is that perturbed many-body systems will generically thermalize in the long run.

Echo dynamics. In Chapter 4, we investigated so-called "echo protocols" [284, 285], which allow us to quantify irreversibility and the stability of nonequilibrium states in isolated many-body quantum systems in terms of macroscopic observables. The general procedure works as follows: A system with Hamiltonian H_0 is prepared in a nonequilibrium *target state* and subsequently relaxes for a certain waiting time τ to reach the *return state*. At this point, an (effective) time reversal is performed by switching to the Hamiltonian $-H_0$, a transformation that can be realized at least approximately, for example, in spin systems or using quantum simulators. If implemented perfectly, this procedure would take the system back to the target state after another period of time τ. However, (a) tiny distortions of the return state as well as (b) imperfections in the time-reversed Hamiltonian, both of which are practically unavoidable in experiments, lead to a different, perturbed target state at the end of the process. This perturbed target state will generically be closer to equilibrium than the initially prepared target state because the imperfections spoil the fine-tuned correlations required for nonequilibrium conditions. The dependence of these differences between the initial and final states on the magnitude of the imperfections and the waiting time τ conveys information about the irreversibility of the dynamics by quantifying how difficult it is to perform an effective time reversal.

We assessed to what extent the difference of the final state from the perfect target state manifests itself in macroscopic observables, modeling the usually uncontrolled and partly unknown imperfections of types (a) and (b) by suitable random operators. By means of a typicality argument, we again extracted analytical predictions for the time-dependent expectation values under the prescribed echo protocols, namely Eq. (4.26) for imperfections of type (a), Eq. (4.40) for type (b), and Eq. (4.43) for their combined effect.

By analogy with classical chaos theory, the imperfect-preparation scenario with imperfections of type (a), where the state at the point of reversal is slightly perturbed, may roughly be regarded as a probe for sensitive dependence on initial conditions. The most remarkable aspect of our findings in this regard is that the strength of the observable echo peak relative to the deflection from equilibrium at the beginning of the protocol is essentially independent of the waiting time τ in the long run. This is in striking contrast to the corresponding scenario in classical chaotic systems, where the echo peak decays with τ as quantified by the concept of Lyapunov exponents.

The imperfections of type (b), pertaining to the imperfect-reversal scenario where the time-reversed Hamiltonian is slightly perturbed as $-H_0 + \epsilon V$, are of particular relevance with respect to experimental realizations of echo dynamics because one can usually only revert the dominant part of the Hamiltonian and the accuracy to carry out the sophisticated manipulations to do so is naturally limited. The obtained result, which is closely related to the investigations from Chapter 3, states that the decay of the relative echo peak is expected to be Gaussian in the magnitude ϵ of the imperfections and exponential in the waiting time τ.

Driven systems. In Chapter 5, we advanced the relaxation theory from Chapter 3 to cover situations with time-dependent couplings of the perturbation, investigating Hamiltonians of the form $H(t) = H_0 + \lambda(t)V$. As before, it is assumed that the dynamics of the time-independent reference system with Hamiltonian H_0 is known. The generalization was achieved by transforming the problem of time evolution under a single time-dependent Hamiltonian into a family of time-evolution problems for time-independent auxiliary Hamiltonians based on a truncated Magnus expansion

[318] of the propagator. These auxiliary systems can then be treated by a suitable extension of the previous theory for the typical relaxation under time-independent perturbations. Combining the typical behavior of all auxiliary systems, we arrived at a theoretical prediction, Eq. (5.27), for the driven dynamics generated by the original, time-dependent Hamiltonian.

The approximations adopted during the aforementioned transformation entail that the resulting response theory works best for short times and fast driving, but still for largely arbitrary protocols $\lambda(t)$ as long as the amplitude remains within reasonable bounds for a perturbative approach to be justified at all. Particularly interesting applications are periodically driven systems or finite-time quenches. Importantly, our theory targets a regime that is largely unexplored in the literature, where the focus regarding periodically driven systems in particular has been on long-time properties of prethermal states and the question of eventual heating [306, 309, 328]. For sufficiently high frequencies, our present theory describes the detailed approach of such prethermal states.

Response profiles. The key element in all of the theoretical predictions was some scalar function that characterizes the principal modifications of the respective reference dynamics induced by the different types of perturbations, which we dubbed the *response profile*. For relaxation under time-independent perturbations and echo inaccuracies of the imperfect-reversal type, the relevant response profile is $g_\lambda(t)$ from (3.146). For echo inaccuracies of the imperfect-preparation type, it is $\hat{d}_W(t)$ from (4.22), and for the response to time-dependent perturbations, it is $g(t,t')$ from (5.23). All of these characteristic functions arise as Fourier transforms of some energy distribution, namely the second moment of eigenvector overlaps between unperturbed and perturbed Hamiltonians in the case of $g_\lambda(t)$ and $g(t,t')$ (see Eq. (3.32) and above Eq. (5.23), respectively), and the scrambling operator's density of states (4.21) in the case of $\hat{d}_W(t)$.

For $g_\lambda(t)$ and $g(t,t')$, we furthermore established the integro-differential equations (3.173) and (5.34), respectively, which relate those response profiles to the perturbation profile (3.10), i.e., the locally averaged magnitude of perturbation matrix elements as a function of the energy difference between the coupled levels of the reference system. Those integro-differential equations thus single out the perturbation profile as the essential characteristic determining the perturbed dynamics within the theories' regimes of applicability.

Interestingly, the equations resemble common relations for response functions, but are distinctly nonlinear with the effective "memory kernel" depending on the respective response-profile function itself. Analytical solutions were obtained for certain special cases, notably for sufficiently weak or sufficiently strong perturbations and for perturbation profiles with a Breit-Wigner shape in the case of $g_\lambda(t)$ as well as for sufficiently strong driving or very short times in the case of $g(t,t')$. For the rest, the response profiles can be calculated numerically by integration of the respective integro-differential equation using standard techniques. Moreover, the function $g_\lambda(t)$ (and to a lesser extent also $g(t,t')$) turned out to be somewhat insensitive with respect to details of the perturbation profile. For $g_\lambda(t)$ in particular, this led to the analytical approximation (3.166), which is presumed to cover essentially all cases of physical relevance and depends only on the perturbation strength and its band width (cf. Eqs. (3.12) and (3.13)).

Verification. We demonstrated the wide applicability of each of the acquired theoretical descriptions by comparison with experimental and/or numerical data from the literature and our own simulations for various distinct model systems. We particularly highlight the entirely parameter-free predictions of nonequilibrium many-body dynamics in realistic models from Figs. 3.10, 4.4, 4.7, 5.5, and 5.6. These demonstrate that the good agreement is not simply a coincidence or the result of excessive fitting. In all other examples, at most two fit parameters sufficed to describe the full time evolution of several different coupling strengths, observables, initial states, or driving protocols at once, which reinforces the explanatory power of the results. Not least, the considered models were quite diverse in their degrees of freedom (including spins, bosons, and fermions), in the interactions between them, as well as in the underlying spatial or lattice structure.

Possible extensions. The adopted typicality approach entails that certain characteristics of the system of interest are disregarded. On the one hand, as mentioned before, this is a desired effect in general since it can be utilized to separate important from unimportant properties with respect to the observable dynamics. On the other hand, a theoretical prediction based on a typicality

argument cannot be expected to apply when primary drivers of those dynamics remain unaccounted for.

Regarding the typicality ensembles employed in this thesis, potentially important but largely disregarded characteristics of physical systems include, in particular, local interactions between only a limited number of constituents (see above) as well as correlations or functional dependencies between Hamiltonians, observables, and initial states. Sometimes it is intuitively clear when such aspects become relevant, for instance if a locally interacting system shows dynamics involving macroscopic transport. Nevertheless, improving our understanding of when those and other properties matter remains an important open issue. In addition, it would naturally be desirable to extend or modify the considered ensembles such that those properties can be taken into account when they are important.

The tiny piece of the puzzle. At last, we come back to the grand question underlying the research of this thesis as phrased in the introduction from Chapter 1: How does the observable macroscopic regularity arise from the fundamental laws describing the motion and interactions of individual atoms and molecules? We believe that the present results indeed contributed a tiny piece of the puzzle as promised in the beginning.

Our starting point was a quantum mechanical description of (nearly) isolated many-body systems, where we focused on three different setups: perturbed relaxation, echo protocols, and short-time dynamics of driven systems. In all three cases, we considered large classes of microscopically distinct systems which shared some essential macroscopic properties, namely the density of states and possibly the perturbation profile (or even just the perturbation strength and band width). We then demonstrated that nearly all systems within any such class exhibit macroscopically indistinguishable dynamics, and we furthermore presented quantitative predictions for this typical behavior. Notably, this incidence of macroscopic regularity among microscopically rather different systems was derived entirely within the framework of quantum mechanics and essentially merely exploited the high number of degrees of freedom. Hence the indifference of large systems to many details at the level of their constituents is indeed somehow built into the fundamental laws, albeit not in an immediately obvious way.

Within all of the three setups, the considered systems were still rather generic and included a large variety of concrete physical models with different degrees of freedom and interactions between them. In fact, one might even argue that the classes were a little too generic because certain aspects common to the microscopic models that make up our reality, especially the locality and few-body character of interactions, were not directly taken into account. Nevertheless, the theory evidently covers a large number of nontrivial and physically important situations in which these properties are not the pivotal mechanism behind the macroscopically perceivable behavior. In other words, the devised predictions successfully characterize the macroscopically observable dynamics based on the microscopic laws in many different scenarios—but there is still a lot to explore.

Acknowledgments

First and foremost, I would like to thank my supervisor Peter Reimann, whose ingenious physical insights laid the foundations for this thesis, for his excellent guidance and support. I am very grateful for his permanent approachability, his modest demeanor, and the constant exchange of ideas, all of which contributed to the pleasant experience that this collaboration felt much more like a collegial interaction than a supervision. Notably, his insistence on physical meaningfulness and his virtue of asking the crucial questions are far from commonplace and have been very inspiring.

I also want to thank Jürgen Schnack for many fascinating insights and discussions about physical and nonphysical topics in formal and informal settings, his readiness to examine this thesis, and his ongoing efforts to stimulate exchange within our E5 community. In the same vein, I am grateful to Ben Balz, who preceded me as a PhD student working on related questions, for numerous stimulating discussions and especially the warm welcome he provided when I joined the team.

Furthermore, I am indebted to Ralf Eichhorn at Nordita in Stockholm, without whom it would probably never have occurred to me to go to Bielefeld nor to Peter to get me there. Besides, I am thankful for our continuing collaboration on stochastic thermodynamics and active matter as well as a two-months stay at Nordita in the summer of 2019 within their visiting PhD program in particular.

Special thanks are due to Christian Eidecker for his very thorough proofreading and numerous helpful comments on the entire manuscript. No less, I appreciate the insightful discussions on topics of and related to this thesis (as well as entertaining activities to distract from it). Likewise, I would like to thank Gregor Sauer and Sönke Ziemer for their careful reading of and several helpful remarks on the nontechnical parts of this thesis (and distraction, naturally). Moreover, I am thankful to Patrick Vorndamme for his genuine curiosity, which led to many critical questions and enlightening debates about those topics and beyond.

In a broader scope, I would like to thank the members of the condensed-matter-theory group at Bielefeld University (aka E5) as well as of the Research Unit FOR 2692 of the Deutsche Forschungsgemeinschaft (DFG) for stimulating input on various occasions. I am also thankful for the, I believe, solid mathematical and theoretical toolkit conveyed during my bachelor and master studies at Friedrich Schiller University Jena, notably in the context of the respective theses supervised by Holger Gies, which facilitated the access to supersymmetry methods, in particular. Furthermore, financial support by the DFG under Grants No. RE 1344/10-1 and 397303734 is also gratefully acknowledged.

Not least, there are quite a few people outside of academia who contributed, mostly indirectly, to the (hopeful) success of this scientific endeavor and to whom I am particularly thankful. I will not intend to list them here, risking embarrassing omissions, since I believe they are well aware of the appreciation, and most of them will probably never read this piece of work either (which is good). Notwithstanding, I do want to explicitly thank my parents and sisters for having been there forever.

A Abbreviations and conventions

Serving as a quick reference, we collect various abbreviations and conventions used throughout the thesis in this appendix, namely a list of abbreviations in Appendix A.1, a list of concepts in Appendix A.2 and a list of symbols in Appendix A.3.

A.1 Abbreviations

The following list of abbreviations is sorted alphabetically.

cf.	compare
DOS	density of states
e.g.	for example
Eq(s).	Equation(s)
ETH	eigenstate thermalization hypothesis
et al.	and others
etc.	et cetera
Fig(s).	Figure(s)
GGE	generalized Gibbs ensemble
GOE	Gaussian Orthogonal Ensemble
GUE	Gaussian Unitary Ensemble
i.e.	that means
LIOM	local integral(s) of motion
MBL	many-body localization
Sec(s).	Section(s)
Tab.	Table

A.2 Concepts

The following list of concepts is sorted alphabetically by key word.

Averages. There are three basic types of averaging procedures employed repeatedly. Averages over typicality ensembles are denoted by

$$\mathbb{E}[\cdots] \,. \tag{A.1}$$

In particular, the ensembles consist of perturbation operators V throughout Chapters 3 and 5 as well as in Sec. 4.4, and of scrambling Hamiltonians W in Sec. 4.3.

Time averages are denoted by an overbar,

$$\overline{\cdots} := \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \mathrm{d}t \, \cdots \,. \tag{A.2}$$

The precise interval depends on the context, most common is the choice $t_1 = 0$ and $t_2 \to \infty$.

Quantum expectation value of observables for a system in the state ρ are denoted by angular brackets,

$$\langle \cdots \rangle_\rho := \mathrm{tr}\,[\rho \cdots] \,. \tag{A.3}$$

Commutator. The commutator of two linear operators A and B is denoted by $[A, B] := AB - BA$ (see below Eq. (2.2)).

Complex conjugation. For a complex number $z \in \mathbb{C}$, its complex conjugate is denoted by $z^* = \mathrm{Re}\, z - \mathrm{i}\, \mathrm{Im}\, z$. For an anticommuting number χ, the formal complex conjugate is denoted by χ^* with the definitions $(\chi^*)^* = -\chi$ and $(\chi_1 \chi_2)^* = \chi_1^* \chi_2^*$, cf. Appendix B.1.

Complex integration measure. Integrals of a complex variable z over the entire plane $\mathbb{C}$ are defined as a double integral of $\mathrm{Re}\, z$ and $\mathrm{Im}\, z$ over the real line. We use the notation

$$\int \mathrm{d}z \mathrm{d}z^* \cdots = 2 \int \mathrm{d}(\mathrm{Re}\, z) \int \mathrm{d}(\mathrm{Im}\, z) \cdots , \tag{A.4}$$

observing that the Jacobian determinant associated with the coordinate transformation

$$\begin{pmatrix} z \\ z^* \end{pmatrix} = \begin{pmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{pmatrix} \begin{pmatrix} \mathrm{Re}\, z \\ \mathrm{Im}\, z \end{pmatrix} \tag{A.5}$$

is $|\partial(z, z^*)/\partial(\mathrm{Re}\, z, \mathrm{Im}\, z)| = 2$.

Fourier transformation. Fourier transforms are mostly computed as characteristic functions $\hat{h}(t)$ for certain energy distributions $h(E)$, following the convention

$$\hat{h}(t) := \int \mathrm{d}E\, \mathrm{e}^{\mathrm{i}Et}\, h(E)\,. \tag{A.6}$$

The overhat notation is generally employed to denote the Fourier transform, with the notable exception of the pair $u(E)$ (overlap distribution, Eq. (3.32)) and $g_\lambda(t)$ (response profile, Eq. (3.146)) as well as their extensions $u(E, t')$ and $g(t, t')$ from Chapter 5 (see Eq. (5.23)). The corresponding inverse transformation is thus

$$h(E) = \int \frac{\mathrm{d}t}{2\pi}\, \mathrm{e}^{-\mathrm{i}Et}\, \hat{h}(t)\,. \tag{A.7}$$

Hilbert space bases and matrix elements. For concrete computations, we almost exclusively adopt eigenbases of relevant Hamiltonians. The following table summarizes the employed Hamiltonians along with the notation used for their eigenvalues, eigenstates, and matrix elements of a self-adjoint operator A:

Hamiltonian	eigenvalues	eigenstates	observable matrix elements
H	E_n	$\|n\rangle$	A_{mn}
H_0	E_μ	$\|\mu\rangle_0$	$A_{\mu\nu}$
H_λ	E_n^λ	$\|n\rangle_\lambda$	A_{mn}^λ
$H^{(t')}$	$E_n^{(t')}$	$\|n(t')\rangle$	$A_{mn}^{(t')}$
W	E_n^W	$\|n\rangle_W$	A_{mn}^W

Particularly note that references to the eigenbasis of the unperturbed Hamiltonian H_0 use Greek indices, whereas Latin indices are used for all other Hamiltonians, notably H_λ.

Intervals. For $a, b \in \mathbb{R}$, open and closed intervals are denoted by $(a, b) := \{x \in \mathbb{R} : a < x < b\}$, $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$, $[a, b) := \{x \in \mathbb{R} : a \leq x < b\}$, $(a, b] := \{x \in \mathbb{R} : a < x \leq b\}$.

Limits of sums and integrals. Sum, products, and integrals without limits for the summation indices or integration variables run over the entire respective domain, typically $\mathbb{N}$ or $\{1, \ldots, N\}$ for sums and products, and $\mathbb{R}$ for integrals.

Pauli matrices. The Pauli matrices are

$$\sigma^x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^y := \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma^z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{A.8}$$

The corresponding raising and lowering operators σ^+ and σ^- are defined as $\sigma^\pm := \sigma^x \pm \mathrm{i}\sigma^y$. If supplemented with subscripts, all these Pauli operators act on a specific subspace of the full Hilbert space, usually associated with a particular lattice site.

Supersymmetry methods. For a general overview of conventions regarding anticommuting numbers, graded algebras, and related supersymmetry concepts, see Appendix B.

Units. The reduced Planck constant is set to unity, $\hbar = 1$. Hence, in particular, time and energy as well as position and momentum are reciprocal to each other.

A.3 Symbols

The following list collects symbolic notation used in different places of the thesis. Definitions that occur only in a single, limited scope are deliberately omitted. Symbols are sorted alphabetically with Greek letters placed according to their English names and little letters before capital letters.

A	generic observable
α_v	intrinsic perturbation strength per level spacing, see Eq. (3.12)
$\mathcal{A}(t)$	deviation of the time-dependent expectation value $\langle A\rangle_{\rho(t)}$ from the thermal value $\langle A\rangle_{\rho_{\mathrm{mc}}}$, see Eq. (4.6)
β	inverse temperature, see above Eq. (2.8)
β_n	Lagrange multipliers of the generalized Gibbs ensemble, see Eq. (2.27)
$B_n(x)$	Chebyshev rational function, see Eq. (3.80)
$\mathbb{C}$	set of complex numbers
$D(E)$	density of states (DOS), see Eq. (2.11)
δ	magnitude of inaccuracies of the imperfect-preparation type in echo protocols, see above Eq. (4.3)
δA	spectral resolution of the observable A, see Eq. (2.19)
ΔA	spectral range of the observable A, see Eq. (2.18)
$\Delta_{\mathcal{E}}$	width of the energy window $I_{\mathcal{E}}$, see Eq. (2.9)
δ_{mn}	Kronecker delta, $\delta_{mn} = 1$ if $m = n$, 0 otherwise
$\delta(x)$	Dirac delta distribution
$\det(\cdots)$	determinant (of a matrix)
Δ_v	perturbation band width, see Eq. (3.13)
$d_W(E)$	relative density of states of the scrambling Hamiltonian W, see Eq. (4.21)
$\hat{d}_W(t)$	Fourier transform of $d_W(E)$, see Eq. (4.22)
$\mathcal{E}$	macroscopic total system energy, see above Eq. (2.5)
$\mathbb{E}[\cdots]$	average over typicality ensemble, see Sec. 2.3.1
E_n	eigenvalue (energy level) of the Hamiltonian H or H_0, see below Eq. (2.1) or below Eq. (3.5), respectively
E_n^λ	eigenvalue (energy level) of the Hamiltonian H_λ, see above Eq. (3.4)
ε	mean level spacing in energy window $I_{\mathcal{E}}$, see Eqs. (2.9) and (2.13)
ϵ	magnitude of inaccuracies of the imperfect-reversal type in echo protocols, see above Eq. (4.4)
f	degrees of freedom, see above Eq. (2.10)
$\mathcal{F}(\delta, \epsilon, \tau)$	relative echo peak height, see Eq. (4.8)

γ	energy scale of $u(E)$ and relaxation constant of $g_\lambda(t)$ for strong perturbations, see Eq. (3.69)
$\gamma(t')$	instantaneous relaxation constant of $g(t,t')$ for strong perturbations, see Eq. (5.29)
Γ	energy scale of $u(E)$ and relaxation constant of $g_\lambda(t)$ for weak perturbations, see Eq. (3.65)
$g_\lambda(t)$	response profile for time-independent perturbations, see Eq. (3.146)
$g(t,t')$	response profile for time-dependent perturbations, see Eq. (5.23)
$G(z)$	scalar ensemble-averaged resolvent of H_λ, see Eqs. (3.57) and (3.58)
$\mathcal{G}(z)$	resolvent of H_λ, see Eq. (3.36)
$G(z,t')$	scalar ensemble-averaged resolvent of $H^{(t')}$, see above Eq. (5.22)
H	generic Hamiltonian
$\mathcal{H}$	generic Hilbert space, see Sec. 2.1
H_0	unperturbed reference Hamiltonian, see Eq. (3.1)
$\mathcal{H}_\mathcal{E}$	Hilbert space associated with the energy window $I_\mathcal{E}$ (and possible further constraints), see below Eq. (2.9)
$H(t)$	time-dependent Hamiltonian, notably of driven systems from Chapter 5, see Eq. (5.1)
$H^{(t')}$	Hamiltonian of auxiliary system for driven dynamics, see Eq. (5.7)
H_λ	Hamiltonian composed of an unperturbed component H_0 and a perturbation λV, see Eq. (3.1)
$I_\mathcal{E}$	macroscopically small, but microscopically large energy window, see Eq. (2.9)
$\operatorname{Im} z$	imaginary part of the complex number z
$J_\nu(x)$	νth Bessel function of the first kind
L	one-dimensional extent (number of sites) of a lattice model
λ	perturbation coupling strength, see Eq. (3.1)
λ_c	crossover coupling between weak- and strong-perturbation regimes, see Eq. (3.73)
$\lambda(t)$	driving protocol, see Eq. (5.1)
$\lvert n\rangle$	eigenstate of the Hamiltonian H, see below Eq. (2.1)
N	number of energy levels within the energy window $I_\mathcal{E}$, dimension of $\mathcal{H}_\mathcal{E}$, see Eqs. (2.9) and (2.10)
$\mathcal{N}$	macroscopic total number of particles, see above Eq. (2.5)
$\mathbb{N}$	set of natural numbers, $\mathbb{N} = \{1, 2, \ldots\}$
$\mathbb{N}_0$	set of natural numbers including zero, $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$
$\lvert n\rangle_\lambda$	eigenstate of the Hamiltonian H_λ, see above Eq. (3.4)
N_v	number of unperturbed energy levels mixed by the perturbation, see Eq. (3.9)
$\mathcal{O}(\cdots)$	Landau symbol, order of magnitude
$p(V)$	probability density function of a perturbation ensemble, see Eq. (3.24)
$\mathbb{P}$	probability measure of a typicality ensemble, see Sec. 2.3.1
Π	generic projection operator
$\Pi_{\mathcal{E},\Delta_\mathcal{E}}$	Gaussian energy filter (projection operator), see Eq. (3.193)
Π_n	projection onto eigenspace of nth eigenvalue, see above Eq. (2.1)

$p_{\max}$	largest occupation probability of an energy level, see Eq. (2.14)
$p_{\mu\nu}(v)$	marginal probability density function of the matrix element $V_{\mu\nu}$, see Eq. (3.23)
$\lvert\psi\rangle$	generic pure state
$\mathrm{PV}\int\cdots$	principal-value integral
Q	generic conserved quantity, see below Eq. (2.4)
$Q^{(n)}$	local integral of motion of an integrable system, see Sec. 2.2.3
$\mathbb{R}$	set of real numbers
$\mathrm{Re}\, z$	real part of the complex number z
ρ	generic density operator of a (pure or mixed) quantum state
$\bar{\rho}$	time-averaged state/diagonal ensemble, see Eqs. (2.16) and (2.17)
$\rho(0)$	density operator of the initial state of a quantum system
ρ_{can}	canonical density operator, see Eq. (2.8)
ρ_{GGE}	density operator of the generalized Gibbs ensemble, see Eq. (2.27)
$\tilde{\rho}_\lambda$	predicted asymptotic long-time state of perturbed relaxation, see Eq. (3.150)
$\rho_\lambda(t)$	time-evolved state with Hamiltonian H_λ
ρ_{mc}	microcanonical density operator, see Eq. (2.5)
ρ_{R}	return state, i.e., state at the point of reversal of an echo protocol, see below Eq. (4.2)
ρ'_{R}	perturbed return state, see below Eq. (4.3)
$\rho(t)$	time-evolved state with Hamiltonian H (possible time-dependent, cf. Eq. (5.1))
$\rho(t,t')$	time-evolved state with Hamiltonian $H^{(t')}$, see Eq. (5.8)
ρ_{T}	target state, i.e., initial state of an echo protocol, see above Eq. (4.2)
ρ'_{T}	perturbed target state, see above Eq. (4.5)
$\mathrm{sdet}(\cdots)$	superdeterminant of a supermatrix, see Eq. (B.18)
$\mathcal{S}(\mathcal{E})$, $\mathcal{S}(\mathcal{E},\mathcal{N},\mathcal{V})$	Boltzmann entropy, see Eq. (2.7)
$\mathrm{sgn}(x)$	sign function, $+1$ if $x>0$, 0 if $x=0$, -1 if $x<0$
σ_0^2	variance of level fluctuations/diagonal perturbation matrix elements, see below Eq. (3.14)
σ_i^α	αth Pauli matrix acting on lattice site i, see below Eq. (3.197)
Σ_n	nth moment of the perturbation profile, see Eq. (3.175)
σ_v^2	intrinsic perturbation strength, see Eq. (3.11)
$\sigma_v^2(E)$	perturbation profile, see Eq. (3.10)
$\sigma_v^2(E,t')$	perturbation profile of $V^{(t')}$, see Eq. (5.14)
$\mathcal{S}(\rho)$	von Neumann entropy, see Eq. (2.6)
$\mathrm{str}(\cdots)$	supertrace of a supermatrix, see Eq. (B.16)
$\mathrm{Sym}(k)$	symmetric group of degree k
T	temperature, see above Eq. (2.8)
$(\cdots)^{\mathsf{T}}$	transpose (of a matrix)

τ	length of a time interval, see Sec. 2.2.1; particularly for echo protocols (Chapter 4): waiting time, i.e., duration of the forward evolution phase, see above Eq. (4.2); particularly for driven systems (Chapter 5): characteristic time scale of the driving protocol (e.g. period, quench time)
t_c	crossover time to exponential/Fermi-golden-rule decay characteristics, see Eq. (3.181)
$\Theta(x)$	Heaviside step function
$T_n(x)$	Chebyshev polynomial of the first kind, see above Eq. (3.80)
$\mathrm{tr}(\cdots)$	trace (of operator/matrix)
$u(E)$	overlap distribution, see Eqs. (3.32) and (3.60)
$\tilde{u}(E)$	self-convolution of $u(E)$, see Eq. (3.116)
$\mathrm{U}(N)$	unitary group of degree N
$U_{n\mu}$	transformation matrix element/eigenstate overlap between the Hamiltonians H_λ and H_0, see Eq. (3.5)
$\tilde{U}_{n\mu}$	transformation matrix element/eigenstate overlap between the Hamiltonians W and H_0, see Eq. (4.14)
$\mathcal{U}(t,t_0)$, $\mathcal{U}(t)$	time-evolution operator/propagator, see Eq. (2.3)
V	perturbation operator, see Eq. (3.1)
$\mathcal{V}$	macroscopic volume of the system, see above Eq. (2.5)
$V^{(t')}$	second-order Magnus approximation for perturbation part of $H^{(t')}$, see Eqs. (5.11) and (5.12)
W	scrambling Hamiltonian, see above Eq. (4.3)
$\xi_V(t)$	deviations of perturbed time-dependent expectation values from ensemble average, see Eq. (3.3)
$\overline{\xi_V^2}(t_1,t_2)$	time-averaged squared deviations of expectation values from ensemble average, see Eq. (3.156)
Z	canonical partition function, see below Eq. (2.8)
$\mathbb{1}$	identity operator of a given vector space or algebra
$\lvert\downarrow\rangle$	"spin down," eigenstate of σ^z with eigenvalue -1
$\lvert\uparrow\rangle$	"spin up," eigenstate of σ^z with eigenvalue $+1$

B Supersymmetry methods

This appendix provides a brief overview of supersymmetry concepts employed explicitly or implicitly in the main text and Sec. 3.4 in particular. For more exhaustive and explanatory introductions to and reviews of the topic, confer Refs. [163, 220–225, 336, 337].

B.1 Anticommuting numbers and graded algebra

Anticommuting numbers. We consider a set $\{\chi_1, \chi_2, \ldots\}$ of *anticommuting* or *Grassmann numbers*. For any two such numbers χ_α and χ_β, we introduce a formal multiplication operation with the defining property that

$$\chi_\alpha \chi_\beta = -\chi_\beta \chi_\alpha \,. \tag{B.1}$$

Hence these anticommuting numbers can be multiplied much like ordinary numbers x and y (e.g., $x, y \in \mathbb{C}$), but the product of two Grassmann numbers is anticommuting in contrast to the product of two ordinary numbers, $xy = yx$, which we will thus also call *commuting* or *c-numbers* in the following. The defining relation (B.1) immediately implies

$$\chi_\alpha^2 = 0 \tag{B.2}$$

for all α. Therefore, in any product of anticommuting numbers, each of the factors can only occur once.

Multiplication of an ordinary c-number a with an anticommuting number χ is commuting by definition,

$$a\chi = \chi a \,. \tag{B.3}$$

Requiring associativity, it follows that in any product of three or more Grassmann numbers, any group of an even number of Grassmann factors similarly behaves like a c-number, e.g.,

$$(\chi_1\chi_2)\chi_3 = \chi_1(\chi_2\chi_3) = -\chi_1(\chi_3\chi_2) = -(\chi_1\chi_3)\chi_2 = (\chi_3\chi_1)\chi_2 = \chi_3(\chi_1\chi_2) \,. \tag{B.4}$$

We point out that the label "c-number" is sometimes assigned in the literature to any type of object that commutes with other factors in a product, thus including the product $\chi_1\chi_2$, for example. Our notion here is more restrictive: c-numbers are elements of the underlying field (usually $\mathbb{C}$).

Grassmann algebra. Upon multiplication with c-numbers and with each other, the basic anticommuting numbers $\chi_1, \chi_2, \ldots$ act as *generators* of the *Grassmann algebra*

$$\mathcal{A} := \operatorname{span}\{\chi_{\alpha_1} \cdots \chi_{\alpha_n} \;:\; n \in \mathbb{N}_0,\ \alpha_1 < \cdots < \alpha_n\} \,. \tag{B.5}$$

where $\operatorname{span} S$ denotes the set of all linear combination of elements in S over the given c-number field. Within this algebra, we can define the subalgebra $\mathcal{A}_\mathrm{B}$ of commuting elements, consisting of all terms that involve products of an even number of Grassmann generators only. This is also called the *bosonic* sector of the Grassmann algebra, and its elements are said to have *even parity*, $\varsigma(a) = 0$ for $a \in \mathcal{A}_\mathrm{B}$. Similarly, we can define the subset $\mathcal{A}_\mathrm{F}$ of anticommuting elements containing only products of an odd number of Grassmann generators. This is also known as the *fermionic* sector of the Grassmann algebra, comprising elements of *odd parity*, $\varsigma(a) = 1$ for $a \in \mathcal{A}_\mathrm{F}$. Note that $\mathcal{A}_\mathrm{F}$ does not form a subalgebra. Obviously, every element $z \in \mathcal{A}$ can be decomposed uniquely into $z_\mathrm{B} \in \mathcal{A}_\mathrm{B}$ and $z_\mathrm{F} \in \mathcal{A}_\mathrm{F}$ such that $z = z_\mathrm{B} + z_\mathrm{F}$.

Complex conjugation. In analogy with ordinary complex numbers, it is often convenient (notably if the underlying field of c-numbers is $\mathbb{C}$) to define a complex structure on the Grassmann algebra, too. In this setting, the Grassmann generators come in pairs, $\chi_1, \chi_1^*, \chi_2, \chi_2^*, \ldots$, but *a priori* there is no connection between χ_α and χ_α^*. In other words, we could just as well label $\chi^* \equiv \xi$, for instance.

However, we define a formal, artificial relationship between the anticommuting numbers χ and χ^*. We call this relation "complex conjugation" and denote it by the star * symbol as well. It is defined as

$$(\chi)^* := \chi^* \quad \text{and} \quad (\chi^*)^* := -\chi\,. \tag{B.6}$$

The choice of sign in the second defining relation is again for convenience. Furthermore, the complex conjugate of a product of several generators is defined as $(\chi_1\chi_2\chi_3\cdots)^* := \chi_1^*\chi_2^*\chi_3^*\cdots$. With these choices, the anticommuting numbers χ_α, χ_α^*, χ_β, and χ_β^* satisfy

$$(\chi_\alpha^*\chi_\beta)^* = (\chi_\alpha^*)^*\chi_\beta^* = \chi_\beta^*\chi_\alpha\,. \tag{B.7}$$

Complex conjugation of arbitrary elements in $\mathcal{A}$ is then defined in terms of ordinary complex conjugation for c-numbers and the above rules for (products of) the Grassmann generators.

B.2 Linear algebra in superspaces

Supervector. As a generalization of ordinary vector spaces, we define a *superspace* consisting of *supervectors*

$$Z = \begin{pmatrix} z \\ \zeta \end{pmatrix}, \quad z = \begin{pmatrix} z_1 \\ \vdots \\ z_{n_\mathrm{B}} \end{pmatrix}, \quad \zeta = \begin{pmatrix} \zeta_1 \\ \vdots \\ \zeta_{n_\mathrm{F}} \end{pmatrix} \tag{B.8}$$

with n_B commuting ("bosonic") components $z_k \in \mathcal{A}_\mathrm{B}$ and n_F anticommuting ("fermionic") components $\zeta_k \in \mathcal{A}_\mathrm{F}$. We will also write $Z_\mathrm{B} \equiv z$ and $Z_\mathrm{F} \equiv \zeta$ to address the bosonic and fermionic sectors, respectively. Naturally, these supervectors can be multiplied by c-numbers as usual, i.e., aZ denotes the supervector Z with all its components multiplied by $a \in \mathbb{C}$ (or whatever the underlying field may be). Similarly, the complex conjugate Z^* of a supervector Z is defined by complex conjugation of all its components, bosonic and fermionic, according to the above-specified rules. Defining the adjoint supervector

$$Z^\dagger := \begin{pmatrix} z^\dagger & \zeta^\dagger \end{pmatrix}, \quad z^\dagger = \begin{pmatrix} z_1^* & \cdots & z_{n_\mathrm{B}}^* \end{pmatrix}, \quad \zeta^\dagger = \begin{pmatrix} \zeta_1^* & \cdots & \zeta_{n_\mathrm{F}}^* \end{pmatrix} \tag{B.9}$$

of the supervector Z, we can also introduce an inner product between supervectors $Z^{(1)}$ and $Z^{(2)}$ as

$$Z^{(1)\dagger}Z^{(2)} = \sum_{k=1}^{n_\mathrm{B}} z_k^{(1)*} z_k^{(2)} + \sum_{k=1}^{n_\mathrm{F}} \zeta_k^{(1)*} \zeta_k^{(2)}\,, \tag{B.10}$$

which satisfies $(Z^{(1)\dagger}Z^{(2)})^* = Z^{(2)\dagger}Z^{(1)}$, in particular, due to the convention (B.6).

Supermatrices. Linear mappings between supervectors are mediated by *supermatrices* such as

$$M = \begin{pmatrix} M_\mathrm{BB} & M_\mathrm{BF} \\ M_\mathrm{FB} & M_\mathrm{FF} \end{pmatrix} \tag{B.11}$$

with $M_\mathrm{BB} \in \mathcal{A}_\mathrm{B}^{n_\mathrm{B}\times n_\mathrm{B}}$, $M_\mathrm{BF} \in \mathcal{A}_\mathrm{F}^{n_\mathrm{B}\times n_\mathrm{F}}$, $M_\mathrm{FB} \in \mathcal{A}_\mathrm{F}^{n_\mathrm{F}\times n_\mathrm{B}}$, $M_\mathrm{FF} \in \mathcal{A}_\mathrm{B}^{n_\mathrm{F}\times n_\mathrm{F}}$. Note that the dimensionalities and domains of the four blocks ensure that

$$MZ = \begin{pmatrix} M_\mathrm{BB} z + M_\mathrm{BF}\zeta \\ M_\mathrm{FB} z + M_\mathrm{FF}\zeta \end{pmatrix} \tag{B.12}$$

is again a supervector in the same superspace, i.e., with the same structure and dimensionality as Z in (B.8).

A standard transformation for ordinary matrices is transposition. If the matrix A has components A_{ij}, then its transpose A^T is defined as the matrix with components $(A^\mathsf{T})_{ij} = A_{ji}$. For supermatrices like M in (B.11), it is convenient to also introduce a *supertranspose*

$$M^{\mathsf{sT}} := \begin{pmatrix} M_\mathrm{BB}^\mathsf{T} & M_\mathrm{FB}^\mathsf{T} \\ -M_\mathrm{BF}^\mathsf{T} & M_\mathrm{FF}^\mathsf{T} \end{pmatrix}. \tag{B.13}$$

The supertranspose thus consists of the ordinary transpose, supplemented by an additional minus sign for the fermionic-bosonic (FB) block. We remark that different conventions exist such that the minus sign is sometimes placed in front of the BF component of the ordinary transpose instead. An important consequence is that $(M_1 M_2)^{\mathsf{sT}} = M_2^{\mathsf{sT}} M_1^{\mathsf{sT}}$ for two supermatrices M_1 and M_2, whereas no such relation holds for the ordinary transpose of a supermatrix due to the anticommuting character of the BF and FB blocks. Note, however, that $(M^{\mathsf{sT}})^{\mathsf{sT}} \neq M$ in general.

Complex conjugation of a supermatrix is again defined by conjugating all components individually. We can then define the Hermitian adjoint as

$$M^\dagger = (M^*)^{\mathsf{sT}} = (M^{\mathsf{sT}})^* \tag{B.14}$$

implying

$$(M_1 M_2)^\dagger = M_2^\dagger M_1^\dagger \quad \text{and} \quad M^{\dagger\dagger} = M\,. \tag{B.15}$$

By analogy with ordinary matrices, the supermatrix M is called *Hermitian* if $M = M^\dagger$. Similarly, a *unitary* supermatrix M preserves the inner product of two supervectors $Z^{(1)}$ and $Z^{(2)}$, i.e., $(MZ^{(1)})^\dagger (MZ^{(2)}) = Z^{(1)\dagger} Z^{(2)}$, implying $M^\dagger M = \mathbb{1}$ in particular. For generalized inner products $Z^{(1)\dagger} L Z^{(2)}$ with a metric tensor L, the preserving transformations M with $M^\dagger L M = L$ are called *pseudounitary*.

Based on the ordinary trace, denoted by the symbol $\mathrm{tr}(\cdots)$, we define the *supertrace* of the supermatrix M as

$$\mathrm{str}\, M = \mathrm{tr}\, M_{\mathrm{BB}} - \mathrm{tr}\, M_{\mathrm{FF}}\,. \tag{B.16}$$

This definition entails linearity, cyclic invariance, and preservation under (super)transposition,

$$\mathrm{str}(M_1 + M_2) = \mathrm{str}\, M_1 + \mathrm{str}\, M_2\,, \quad \mathrm{str}(M_1 M_2 M_3) = \mathrm{str}(M_3 M_1 M_2)\,, \quad \mathrm{str}(M^{\mathsf{sT}}) = \mathrm{str}\, M\,, \tag{B.17}$$

such that the supertrace behaves much like the usual trace operation for ordinary matrices. Another important concept from linear algebra is the determinant $\det(\cdots)$ of a matrix. The appropriate generalization to supermatrices is the *superdeterminant*

$$\mathrm{sdet}\, M = \frac{\det\left(M_{\mathrm{BB}} - M_{\mathrm{BF}} M_{\mathrm{FF}}^{-1} M_{\mathrm{FB}}\right)}{\det M_{\mathrm{FF}}} \tag{B.18}$$

requiring M_{FF} or rather the c-number part of it to be invertible; otherwise the superdeterminant does not exist. If in turn the inverse of M_{BB} exists, the (inverse) superdeterminant can be defined equivalently as

$$(\mathrm{sdet}\, M)^{-1} = \frac{\det\left(M_{\mathrm{FF}} - M_{\mathrm{FB}} M_{\mathrm{BB}}^{-1} M_{\mathrm{BF}}\right)}{\det M_{\mathrm{BB}}}\,. \tag{B.19}$$

If the supermatrix M is block-diagonal such that $M_{\mathrm{BF}} = M_{\mathrm{FB}} = 0$, these definitions entail

$$\mathrm{sdet} \begin{pmatrix} M_{\mathrm{BB}} & 0 \\ 0 & M_{\mathrm{FF}} \end{pmatrix} = \frac{\det M_{\mathrm{BB}}}{\det M_{\mathrm{FF}}} \tag{B.20}$$

in particular. Even more specifically, any supermatrix that is a multiple of the identity has unit superdeterminant, $\mathrm{sdet}(a\mathbb{1}) = 1$ ($a \in \mathcal{A}_{\mathrm{B}}$). Further important consequences are

$$\mathrm{sdet}(M^{\mathsf{sT}}) = \mathrm{sdet}\, M\,, \quad \mathrm{sdet}(M_1 M_2) = \mathrm{sdet}(M_1)\, \mathrm{sdet}(M_2)\,, \quad \text{and } \ln \mathrm{sdet}\, M = \mathrm{str} \ln M\,, \tag{B.21}$$

where the logarithm is defined in terms of the Taylor expansion of $\ln(1 + [M - \mathbb{1}])$ in $M - \mathbb{1}$. Hence the superdeterminant (in combination with the supertrace) behaves similarly to the ordinary determinant for ordinary matrices.

B.3 Differentiation and integration

Functions of anticommuting numbers. Due to (B.2), an arbitrary function f can at most be linear in any of the generators $\chi_1, \chi_2, \ldots$ (or any other anticommuting degree of freedom $\zeta \in \mathcal{A}_{\mathrm{F}}$).

Put differently, the Taylor series of $f(\chi_1, \chi_2, \ldots)$ terminates at first order. Hence the most general function of a single Grassmann variable $f(\chi)$ can be written in the form

$$f(\chi) = a_0 + \chi a_1 \,, \tag{B.22}$$

where $a_0, a_1 \in \mathcal{A}$ may contain commuting and anticommuting terms, but no dependence on χ. Note that in order for f to have definite parity, we need $a_0 \in \mathcal{A}_\mathrm{B} \Leftrightarrow a_1 \in \mathcal{A}_\mathrm{F}$ and vice versa. Similarly, functions of several variables can be written out in a multilinear form. For example, the most general function of three Grassmann generators $f(\chi_1, \chi_2, \chi_3)$ is

$$f(\chi_1, \chi_2, \chi_3) = a_{000} + \chi_1 a_{100} + \chi_2 a_{010} + \chi_3 a_{001} + \chi_1\chi_2 a_{110} + \chi_1\chi_3 a_{101} + \chi_2\chi_3 a_{011} + \chi_1\chi_2\chi_3 a_{111} \,, \tag{B.23}$$

where again all the coefficients a_{ijk} are independent of χ_1, χ_2, and χ_3. Moreover, if f depends on all generators of the Grassmann algebra, the coefficients similar to a_{ijk} can be chosen as c-numbers without loss of generality.

Differentiation. Derivatives of a function $f(\chi)$ with respect to the Grassmann generator χ are defined by analogy with ordinary calculus as the linear part of the corresponding Taylor expansion,

$$\frac{\overrightarrow{\partial}}{\partial\chi}\,(a_0 + \chi a_1) = a_1 \,. \tag{B.24}$$

The so-defined operation is obviously linear. However, care must be taken regarding the order of terms. For instance, if $a_1 \in \mathcal{A}_\mathrm{F}$, then

$$\frac{\overrightarrow{\partial}}{\partial\chi}\,(a_0 + a_1\chi) = \frac{\overrightarrow{\partial}}{\partial\chi}\,(a_0 - \chi a_1) = -a_1 \,. \tag{B.25}$$

Generally speaking, by exploiting linearity, the differential operator $\overrightarrow{\partial}/\partial\chi$ thus behaves like an element of $\mathcal{A}_\mathrm{F}$:

$$\frac{\overrightarrow{\partial}}{\partial\chi}\,(a_0 + a_1\chi) = \frac{\overrightarrow{\partial}}{\partial\chi}\,(a_1\chi) = -a_1\,\frac{\overrightarrow{\partial}}{\partial\chi}\chi = -a_1 \,. \tag{B.26}$$

In other words, $\overrightarrow{\partial}/\partial\chi$ acts on everything to its right and eliminates a Grassmann generator χ if it stands immediately next to it. This is also called the *left-derivative* as the dependent variable χ has to be moved to the left in every product, observing anticommutation, before the derivative can act. Similarly, we can define a *right-derivative*

$$(a_0 + \chi a_1)\,\frac{\overleftarrow{\partial}}{\partial\chi} = (-a_1\chi)\,\frac{\overleftarrow{\partial}}{\partial\chi} = -a_1 \,. \tag{B.27}$$

Since any function is at most linear in the generator χ, higher derivatives always vanish, i.e., $\overrightarrow{\partial}^2/\partial\chi^2 = \overleftarrow{\partial}^2/\partial\chi^2 = 0$. For the product $f(\chi)g(\chi)$ of two functions $f(\chi) = a_0 + \chi a_1$ and $g(\chi) = b_0 + \chi b_1$, the familiar product rule from ordinary calculus holds as well, provided that the respective factors have definite parity and the anticommuting character is respected. Denoting the parity of f by $\varsigma(f)$, we then verify straightforwardly that

$$\frac{\overrightarrow{\partial}}{\partial\chi}\,[f(\chi)g(\chi)] = \frac{\overrightarrow{\partial} f(\chi)}{\partial\chi} g(\chi) + (-1)^{\varsigma(f)} f(\chi)\frac{\overrightarrow{\partial} g(\chi)}{\partial\chi} \,. \tag{B.28}$$

Similarly, derivatives of several variables can be combined by treating every operator like an element of $\mathcal{A}_\mathrm{F}$,

$$\frac{\overrightarrow{\partial}^{\,k}}{\partial\chi_1\cdots\partial\chi_k} := \frac{\overrightarrow{\partial}}{\partial\chi_1}\cdots\frac{\overrightarrow{\partial}}{\partial\chi_k} \,. \tag{B.29}$$

For instance, taking the function $f(\chi_1, \chi_2, \chi_3)$ from (B.23), we find

$$\frac{\overrightarrow{\partial}^{\,2} f(\chi_1, \chi_2, \chi_3)}{\partial\chi_1\partial\chi_3} = -a_{101} + \chi_2 a_{111} \,. \tag{B.30}$$

Grassmann integral. Even though there is no classical geometrical picture such as a number line for Grassmann numbers, we can define an integral over anticommuting degrees of freedom by requiring certain properties of ordinary (Riemann or Lebesgue) integrals. As we will see, demanding *linearity* and *shift invariance of the measure* defines the *Grassmann-Berezin integral* uniquely up to a multiplicative constant. As observed in Eq. (B.22), the most general function of one Grassmann variable is $f(\chi) = a_0 + \chi\, a_1$. By linearity, we mean that

$$\int \mathrm{d}\chi\, (a_0 + \chi\, a_1) \stackrel{!}{=} \left(\int \mathrm{d}\chi\right) a_0 + \left(\int \mathrm{d}\chi\, \chi\right) a_1\,. \tag{B.31}$$

Note that it is meaningless to define a domain of integration because the Grassmann algebra does not have a geometry. In this sense, "$\int \mathrm{d}\chi \cdots$" is to be read as an integral over all possible values of χ, i.e., it is similar to a real integral of the form "$\int_{-\infty}^{\infty} \mathrm{d}x \cdots$". This leads to the second defining property. Since real integrals over an infinite domain are invariant under a shift of the integration variable, we demand that the Grassmann integral be invariant under the transformation $\chi \mapsto \chi' := \chi + \zeta$, where ζ is a second Grassmann variable independent of χ. More precisely,

$$\int \mathrm{d}\chi\, (a_0 + \chi\, a_1) \stackrel{!}{=} \int \mathrm{d}\chi'\, [a_0 + (\chi' - \zeta)\, a_1] = \int \mathrm{d}\chi'\, [(a_0 - \zeta a_1) + \chi' a_1]\,. \tag{B.32}$$

Applying the linearity condition (B.31) to this relation, we immediately find $\int \mathrm{d}\chi = 0$. Furthermore, $\int \mathrm{d}\chi\, \chi$ must evaluate to a constant, which we choose to be 1 by convention (again, different choices can be found in the literature). Hence we define the Grassmann-Berezin integral as

$$\int \mathrm{d}\chi\, (a_0 + \chi a_1) := a_1\,. \tag{B.33}$$

Remarkably, this is the same relation as we obtained for the (left-)derivative in Eq. (B.24): For anticommmuting numbers, differentiation and integration are the same operation. In summary, the defining relations for differentiation and integration of Grassmann variables are, in our convention,

$$\frac{\overrightarrow{\partial}}{\partial\chi}\, (a_0 + \chi a_1) = \int \mathrm{d}\chi\, (a_0 + \chi a_1) = a_1\,, \tag{B.34}$$

where $a_0, a_1 \in \mathcal{A}$ may contain both commuting and anticommuting variables, but no dependence on χ. Similarly to higher-order derivatives, multiple integrals can be evaluated sequentially.

Linear transformations. Besides shifting integration variables, we will also want to scale them occasionally. Therefore, we consider the transformation $\chi \mapsto \chi' := t\chi$ with $t \in \mathcal{A}_{\mathrm{B}}$. Requiring

$$1 = \int \mathrm{d}\chi\, \chi \stackrel{!}{=} \int \mathrm{d}\chi'\, \chi' = \int \mathrm{d}\chi'\, t\chi\,, \tag{B.35}$$

we find that we need to have $\mathrm{d}\chi' = t^{-1}\mathrm{d}\chi$. This should be contrasted with the transformation of differentials for c-numbers, where $\mathrm{d}(tx) = t\,\mathrm{d}x$ ($t = \mathrm{const}$).

For the higher-dimensional generalization, consider a Grassmann vector $\chi = (\chi_1, \ldots, \chi_N)$ and the linear transformation $\chi \mapsto \chi' := T\chi$ with $T \in \mathcal{A}_{\mathrm{B}}^{N\times N}$, i.e., $\chi'_\alpha = \sum_\beta T_{\alpha\beta}\chi_\beta$. The nonvanishing part of the integrand in an integral over $\mathrm{d}\chi' = \mathrm{d}\chi'_1 \cdots \mathrm{d}\chi'_N$ is proportional to

$$\chi'_1 \cdots \chi'_N = \sum_{\alpha_1,\ldots,\alpha_N} T_{1\alpha_1} \cdots T_{N\alpha_N}\, \chi_{\alpha_1} \cdots \chi_{\alpha_N} = \sum_{\sigma\in\mathrm{Sym}(N)} T_{1\sigma(1)} \cdots T_{N\sigma(N)}\, \chi_{\sigma(1)} \cdots \chi_{\sigma(N)}\,, \tag{B.36}$$

where $\mathrm{Sym}(N)$ denotes the symmetric group of degree N, i.e., the set of all permutations of $\{1, \ldots, N\}$. Here the second equality holds because all terms where any two indices α_i and α_j ($i \neq j$) coincide vanish due to the Grassmann character of the χ_α. Rearranging the vector components, we can equivalently write this as

$$\chi'_1 \cdots \chi'_N = \sum_{\sigma\in\mathrm{Sym}(N)} \mathrm{sgn}(\sigma)\, T_{1\sigma(1)} \cdots T_{N\sigma(N)}\, \chi_1 \cdots \chi_N = (\det T)\, \chi_1 \cdots \chi_N\,. \tag{B.37}$$

Consequently, for Grassmann variables the N-dimensional differential transforms as $\mathrm{d}(T\chi) = (\det T)^{-1}\, \mathrm{d}\chi$, in contrast to the c-number case, where $\mathrm{d}(Tx) = (\det T)\, \mathrm{d}x$.

C Gaussian integrals

Gaussian integrals over commuting and anticommuting variables (cf. Appendix B) feature prominently in the derivations of Sec. 3.4 in particular. Here we collect basic properties of such integrals for commuting (Appendix C.1), anticommuting (Appendix C.2), and supersymmetric (Appendix C.3) variables. We also discuss the related supersymmetric Hubbard-Stratonovich transformation in Appendix C.4.

C.1 Commuting variables

One real variable. Since all Gaussian integrals over real- or complex-valued vectors can be reduced to the real, one-dimensional case, this forms the natural starting point of our collection of results. To this end, let $a > 0$. Then

$$I(a) := \int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-ax^2} = \sqrt{\frac{\pi}{a}}\,. \tag{C.1}$$

Indeed, upon substituting $x \mapsto \sqrt{a}x$, we find $I(a) = I(1)/\sqrt{a}$. The square $I(1)^2$ can be computed using polar coordinates,

$$I(1)^2 = \int \mathrm{d}x\,\mathrm{d}y\, \mathrm{e}^{-(x^2+y^2)} = \int_0^{\infty} \mathrm{d}r \int_0^{2\pi} \mathrm{d}\phi\, r\mathrm{e}^{-r^2} = -\pi\, \mathrm{e}^{-r^2}\Big|_{r^2=0}^{\infty} = \pi\,. \tag{C.2}$$

Consequently, we have $I(1) = \sqrt{\pi}$, and (C.1) follows.

One complex variable. Next we consider a Gaussian integral over a single complex variable and find

$$\int \mathrm{d}z\,\mathrm{d}z^*\, \mathrm{e}^{-a|z|^2} = \frac{2\pi}{a}\,. \tag{C.3}$$

To prove this, we first observe that, by definition, the complex integral consists of two real integrals, cf. Eq. (A.4). Therefore, we can exploit the result (C.1) for real-valued Gaussian integrals and obtain

$$\int \mathrm{d}z\,\mathrm{d}z^*\, \mathrm{e}^{-a|z|^2} = 2\int \mathrm{d}(\mathrm{Re}\, z) \int \mathrm{d}(\mathrm{Im}\, z)\, \mathrm{e}^{-a\left[(\mathrm{Re}\, z)^2 + (\mathrm{Im}\, z)^2\right]} = 2\left[\int \mathrm{d}x\, \mathrm{e}^{-ax^2}\right]^2\,, \tag{C.4}$$

leading to (C.3).

Multiple complex variables. We generalize the result to an arbitrary number N of complex variables, collected in a vector $z = (z_1, \ldots, z_N)$. If $A \in \mathbb{C}^{N\times N}$ is Hermitian ($A^\dagger = A$) and positive definite, then

$$\int [\mathrm{d}z\,\mathrm{d}z^*]\, \mathrm{e}^{-z^\dagger A z} = \frac{(2\pi)^N}{\det A}\,. \tag{C.5}$$

Here $[\mathrm{d}z\,\mathrm{d}z^*] := \prod_{\alpha=1}^n \mathrm{d}z_\alpha\,\mathrm{d}z_\alpha^*$. To prove Eq. (C.5), we notice that, since A is Hermitian, there exists a unitary matrix $U \in \mathrm{U}(N)$ such that $UAU^\dagger = \mathrm{diag}(a_1, \ldots, a_N)$, where a_α are the real eigenvalues of A. Moreover, since A is positive definite, $a_\alpha > 0$. Using the coordinate transformation $z \mapsto z' := Uz$, we then obtain

$$\int [\mathrm{d}z\,\mathrm{d}z^*]\, \mathrm{e}^{-z^\dagger A z} = \int [\mathrm{d}z\,\mathrm{d}z^*]\, \mathrm{e}^{-\sum_\alpha z_\alpha'^* a_\alpha z_\alpha'} = \prod_\alpha \int \mathrm{d}z_\alpha'\,\mathrm{d}z_\alpha'^*\, \mathrm{e}^{-a_\alpha |z_\alpha'|^2} = \prod_\alpha \frac{2\pi}{a_\alpha}\,, \tag{C.6}$$

where we used (C.3) in the last step. Hence we have shown (C.5).

According to (C.5), we can express the determinant of a Hermitian, positive definite matrix A in terms of a Gaussian integral. We will now show how individual matrix elements of the inverse A^{-1} arise as the correlators or moments of the Gaussian distribution. To this end, we define the *generating function*

$$\mathcal{Z}(h, h^*) := \int [\mathrm{d}z\,\mathrm{d}z^*]\, \mathrm{e}^{-z^\dagger A z + h^\dagger z + z^\dagger h}\,. \tag{C.7}$$

This is indeed the generating function of moments of a Gaussian distribution (up to normalization) with covariance matrix A^{-1} because

$$\left.\frac{\partial \mathcal{Z}(h,h^*)}{\partial h^*_{\mu_1}\cdots\partial h^*_{\mu_k}\partial h_{\nu_1}\cdots\partial h_{\nu_l}}\right|_{h=h^*=0} = \int [\mathrm{d}z\,\mathrm{d}z^*]\, z_{\mu_1}\cdots z_{\mu_k} z^*_{\nu_1}\cdots z^*_{\nu_l}\,\mathrm{e}^{-z^\dagger A z}\,. \tag{C.8}$$

As usual, h and h^* are treated as independent variables for differentiation, basically in the same way as for the case of integration, in the sense of the coordinate transformation (A.5) above. Since A is positive definite, it is invertible and we can shift the integration variable in (C.7) as $z \mapsto z' := z - A^{-1}h$ to obtain

$$\mathcal{Z}(h,h^*) = \mathrm{e}^{h^\dagger A^{-1} h}\int [\mathrm{d}z'\,\mathrm{d}z'^*]\,\mathrm{e}^{-z'^\dagger A z'} = \frac{(2\pi)^N}{\det A}\,\mathrm{e}^{h^\dagger A^{-1}h}. \tag{C.9}$$

With this explicit expression for the generating function, we can compute arbitrary moments of the form (C.8). For the second-order correlator, we find

$$\int [\mathrm{d}z\,\mathrm{d}z^*]\, z_\mu z^*_\nu\,\mathrm{e}^{-z^\dagger A z} = \frac{(2\pi)^N}{\det A}(A^{-1})_{\mu\nu}\,, \tag{C.10}$$

establishing the aforementioned connection between correlators and matrix elements of A^{-1}. For higher-order correlators, we are left with the *Isserlis-Wick theorem* [239, 240] upon repeated differentiation according to (C.8). It states that the correlator is obtained by summing over all pairs of z and z^* variables and replacing any such pair by the corresponding second-order correlator. In particular, this means that we always need an equal number of z and z^* terms to obtain a nonvanishing contribution. More explicitly, we thus find

$$\int [\mathrm{d}z\,\mathrm{d}z^*]\, z_{\mu_1}\cdots z_{\mu_k} z^*_{\nu_1}\cdots z^*_{\nu_k}\,\mathrm{e}^{-z^\dagger A z} = \frac{(2\pi)^N}{\det A}\sum_{\sigma\in\mathrm{Sym}(k)} (A^{-1})_{\mu_1\nu_{\sigma(1)}}\cdots(A^{-1})_{\mu_k\nu_{\sigma(k)}}\,, \tag{C.11}$$

where $\mathrm{Sym}(k)$ denotes the symmetric group of degree k, i.e., the set of all permutations of $\{1,\ldots,k\}$.

C.2 Anticommuting variables

Two Grassmann variables. The simplest case of a Gaussian integral over anticommuting numbers involves two Grassmannian degrees of freedom as there is no quadratic form of a single Grassmann variable. By analogy with the c-number setting, we choose them as (formal) complex conjugates, χ and χ^*, but emphasize again that complex conjugation is an artificial operation for Grassmann numbers and only introduced for convenience when combining commuting and anticommuting degrees of freedom (cf. Appendix B.1). For arbitrary $a \in \mathcal{A}_\mathrm{B}$, we then find

$$\int \mathrm{d}\chi\,\mathrm{d}\chi^*\,\mathrm{e}^{\chi^* a\chi} = a\,. \tag{C.12}$$

Observing that $\mathrm{e}^{\chi^* a\chi} = 1 + \chi^* a\chi$, this follows directly from the definition of the Grassmann integral (B.34).

Multiple Grassmann variables. For the multidimensional generalization, let $A \in \mathcal{A}_\mathrm{B}^{N\times N}$ be normal so that it is diagonalizable by a unitary transformation $U \in \mathrm{U}(N)$, and denote the eigenvalues by $a_1,\ldots,a_N$. We then find that

$$\int [\mathrm{d}\chi\mathrm{d}\chi^*]\,\mathrm{e}^{\chi^\dagger A\chi} = \det A\,, \tag{C.13}$$

where $[\mathrm{d}\chi\mathrm{d}\chi^*] = \prod_{\alpha=1}^N \mathrm{d}\chi_\alpha \mathrm{d}\chi^*_\alpha$. This can be shown similarly as in the c-number case by transforming $\chi \mapsto \chi\prime := U\chi$ and $\chi^\dagger \mapsto \chi'^\dagger := \chi^\dagger U^\dagger$. This transformation diagonalizes the matrix A, and since $\det U = 1$, we get

$$\int [\mathrm{d}\chi\mathrm{d}\chi^*]\,\mathrm{e}^{\chi^\dagger A\chi} = \int [\mathrm{d}\chi'\mathrm{d}\chi'^*]\,\mathrm{e}^{\sum_\alpha \chi'^\dagger_\alpha a_\alpha \chi'_\alpha} = \prod_\alpha \int [\mathrm{d}\chi'_\alpha \mathrm{d}\chi'^*_\alpha]\,\mathrm{e}^{\chi'^\dagger_\alpha a_\alpha \chi'_\alpha} = \prod_\alpha a_\alpha\,, \tag{C.14}$$

proving the claim.

Defining a generating function

$$\mathcal{Z}(\eta,\eta^*) := \int [\mathrm{d}\chi \mathrm{d}\chi^*]\, \mathrm{e}^{\chi^\dagger A\chi + \eta^\dagger \chi + \chi^\dagger \eta} \tag{C.15}$$

similarly to the c-number case, we can likewise express correlators of anticommuting variables. Assuming that A is invertible, the integral in (C.16) can be evaluated by transforming $\chi \mapsto \chi' := \chi + A^{-1}\eta$, $\chi^\dagger \mapsto \chi'^\dagger := \chi^\dagger + \eta^\dagger A^{-1}$ to give

$$\mathcal{Z}(\eta,\eta^*) = (\det A)\, \mathrm{e}^{-\eta^\dagger A^{-1}\eta}\,. \tag{C.16}$$

Correlation functions can then be computed by differentiation,

$$\frac{\overrightarrow{\partial}^{\,k}}{\partial\eta^*_{\mu_1}\cdots\partial\eta^*_{\mu_k}}\mathcal{Z}(\eta,\eta^*)\frac{\overleftarrow{\partial}^{\,l}}{\partial\eta_{\nu_1}\cdots\partial\eta_{\nu_l}}\bigg|_{\substack{\eta\to 0,\\ \eta^*\to 0}} = \int [\mathrm{d}\chi\mathrm{d}\chi^*]\, \chi_{\mu_1}\cdots\chi_{\mu_k}\chi^*_{\nu_1}\cdots\chi^*_{\nu_l}\, \mathrm{e}^{\chi^\dagger A\chi}\,, \tag{C.17}$$

where, by convention, derivatives are evaluated from inside to outside, i.e., for the left-derivatives, $\overrightarrow{\partial}/\partial\eta^*_{\mu_k}$ is evaluated first and $\overrightarrow{\partial}/\partial\eta^*_{\mu_1}$ last, whereas for the right-derivatives, $\overleftarrow{\partial}/\partial\eta_{\nu_1}$ is evaluated first and $\overleftarrow{\partial}/\partial\eta_{\nu_l}$ last (cf. Appendix B.3). Note that due to anticommutativity, the correlator vanishes if any two of the indices $\mu_1,\ldots,\mu_k$ are the same, and similarly for $\nu_1,\ldots,\nu_l$. Calculating the derivatives leads to the Isserlis-Wick theorem as for complex Gaussian integrals, except that we have to bear in mind anticommutativity and the additional minus sign in the exponent of (C.16). Hence

$$\begin{aligned}&\int [\mathrm{d}\chi\mathrm{d}\chi^*]\, \chi_{\mu_1}\cdots\chi_{\mu_k}\chi^*_{\nu_1}\cdots\chi^*_{\nu_l}\, \mathrm{e}^{\chi^\dagger A\chi}\\ &\quad = -|\varepsilon_{\mu_1\cdots\mu_k}\varepsilon_{\nu_1\cdots\nu_k}|\,(\det A)\sum_{\sigma\in\mathrm{Sym}(k)}\mathrm{sgn}(\sigma)(A^{-1})_{\mu_1\nu_{\sigma(1)}}\cdots(A^{-1})_{\mu_k\nu_{\sigma(k)}}\,,\end{aligned} \tag{C.18}$$

where $\varepsilon_{\mu_1\cdots\mu_k}$ denotes the completely antisymmetric k-dimensional Levi-Civita symbol. (Note the absolute value in Eq. (C.18), so the mere purpose of this prefactor is to ensure that all μ_i and all ν_i are distinct.)

C.3 Supersymmetric variables

Gaussian integrals in superspace can be readily evaluated by combining the results from Appendices C.1 and C.2. Similarly as in Appendix B.2, we consider a supervector $X = (x\ \chi)^{\mathsf{T}}$ consisting of n_{B} commuting variables $x = (x_1,\ldots,x_{n_{\mathrm{B}}}) \in \mathbb{C}^{n_{\mathrm{B}}}$ and n_{F} Grassmann generators $\chi = (\chi_1,\ldots,\chi_{n_{\mathrm{F}}})$, which are accompanied by their n_{F} formal complex conjugates, the generators $\chi^* = (\chi^*_1,\ldots,\chi^*_{n_{\mathrm{F}}})$. Furthermore, let A be a supermatrix as in (B.11). As one might have anticipated, the Gaussian integral associated with the quadratic form $X^\dagger AX$, convergence presumed, yields essentially the superdeterminant, i.e.,

$$\int [\mathrm{d}X\mathrm{d}X^*]\, \mathrm{e}^{-X^\dagger AX} = \frac{(2\pi)^{n_{\mathrm{B}}}\,(-1)^{n_{\mathrm{F}}}}{\mathrm{sdet}\, A}\,. \tag{C.19}$$

Here $[\mathrm{d}X\mathrm{d}X^*] = [\mathrm{d}x\mathrm{d}x^*][\mathrm{d}\chi\mathrm{d}\chi^*]$. To show this, we write out the exponent in terms of the different sectors first,

$$\int [\mathrm{d}X\mathrm{d}X^*]\, \mathrm{e}^{-X^\dagger AX} = \int [\mathrm{d}x\mathrm{d}x^*]\, \mathrm{e}^{-x^\dagger A_{\mathrm{BB}}x}\int [\mathrm{d}\chi\mathrm{d}\chi^*]\, \mathrm{e}^{-\chi^\dagger A_{\mathrm{FF}}\chi - x^\dagger A_{\mathrm{BF}}\chi - \chi^\dagger A_{\mathrm{FB}}x} \tag{C.20}$$

The fermionic integral can be carried out similarly as in Eqs. (C.15) and (C.16), leading to

$$\int [\mathrm{d}X\mathrm{d}X^*]\, \mathrm{e}^{-X^\dagger AX} = (-1)^{n_{\mathrm{F}}}\ \det A_{\mathrm{FF}}\int [\mathrm{d}x\mathrm{d}x^*]\, \mathrm{e}^{-x^\dagger(A_{\mathrm{BB}} - A_{\mathrm{BF}}A_{\mathrm{FF}}^{-1}A_{\mathrm{FB}})x}\,. \tag{C.21}$$

For convergence, we thus need that the c-number component of $A_{\mathrm{BB}} - A_{\mathrm{BF}} A_{\mathrm{FF}}^{-1} A_{\mathrm{FB}}$ is positive definite. The remaining bosonic integral is then of the form (C.5), leaving us with

$$\int [\mathrm{d}X \mathrm{d}X^*] \, \mathrm{e}^{-X^\dagger A X} = (2\pi)^{n_{\mathrm{B}}} \, (-1)^{n_{\mathrm{F}}} \frac{\det A_{\mathrm{FF}}}{\det(A_{\mathrm{BB}} - A_{\mathrm{BF}} A_{\mathrm{FF}}^{-1} A_{\mathrm{FB}})} \,. \tag{C.22}$$

Taking into account the definition (B.18) of the superdeterminant, this is equivalent to Eq. (C.19).

Furthermore, we can again define the generating function

$$\mathcal{Z}(H, H^*) := \int \frac{[\mathrm{d}X \mathrm{d}X^*]}{(2\pi)^{n_{\mathrm{B}}} (-1)^{n_{\mathrm{F}}}} \, \mathrm{e}^{-X^\dagger A X + H^\dagger X + X^\dagger H} = \frac{\mathrm{e}^{H^\dagger A H}}{\operatorname{sdet} A} \tag{C.23}$$

depending on the supervector $H = (h_1 \;\; \cdots \;\; h_{n_{\mathrm{B}}} \;\; \eta_1 \;\; \cdots \;\; \eta_{n_{\mathrm{F}}})^{\mathsf{T}}$. As before, this generating function $\mathcal{Z}(H, H^*)$ encodes the correlation functions and brings them forward upon differentiation,

$$\frac{\overrightarrow{\partial}^{\,k}}{\partial H_{\mu_1}^* \cdots \partial H_{\mu_k}^*} \mathcal{Z}(H, H^*) \frac{\overleftarrow{\partial}^{\,l}}{\partial H_{\nu_1} \cdots \partial H_{\nu_l}} \bigg|_{\substack{H \to 0, \\ H^* \to 0}} = \int \frac{[\mathrm{d}X \mathrm{d}X^*]}{(2\pi)^{n_{\mathrm{B}}} (-1)^{n_{\mathrm{F}}}} \, X_{\mu_1} \cdots X_{\mu_k} X_{\nu_1}^* \cdots X_{\nu_l}^* \, \mathrm{e}^{-X^\dagger A X} \,, \tag{C.24}$$

Here the indices $1, \ldots, n_{\mathrm{B}}$ and $n_{\mathrm{B}} + 1, \ldots, n_{\mathrm{B}} + n_{\mathrm{F}}$, respectively, refer to the bosonic and fermionic components of the supervectors X and H. For bosonic components, the left- and right-derivatives act just like ordinary differential operators on $\mathcal{Z}(H, H^*)$. The Isserlis-Wick theorem generalizes in the obvious way, i.e.,

$$\begin{aligned} &\int \frac{[\mathrm{d}X \mathrm{d}X^*]}{(2\pi)^{n_{\mathrm{B}}} (-1)^{n_{\mathrm{F}}}} \, X_{\mu_1} \cdots X_{\mu_k} X_{\nu_1}^* \cdots X_{\nu_k}^* \, \mathrm{e}^{-X^\dagger A X} \\ &\quad = \frac{1}{\operatorname{sdet} A} \sum_{\sigma \in \mathrm{Sym}(k)} \varepsilon_\sigma \, (A^{-1})_{\mu_1 \nu_{\sigma(1)}} \cdots (A^{-1})_{\mu_k \nu_{\sigma(k)}} \,, \end{aligned} \tag{C.25}$$

where ε_σ is 1 if the permutation σ is of even parity in the elements corresponding to *fermionic* indices ν_i and -1 otherwise, and provided that none of the fermionic indices appear twice among the μ_i or the ν_i, respectively.

C.4 Hubbard-Stratonovich transformation

As a particular application of Gaussian integrals in superspace, we demonstrate explicitly the supersymmetric Hubbard-Stratonovich transformation, which allows to reduce integrals with exponents of fourth order in a supervector to second-order Gaussian ones; see also Eqs. (3.51) and (3.103) from the main text.

One auxiliary supermatrix. The basic form of the supersymmetric Hubbard-Stratonovich transformation [220, 243, 244] to decouple a homogeneous quartic term in the exponent of a superintegral is defined by the identity

$$\exp\biggl[-\frac{\lambda^2}{2} \sum\nolimits_{\alpha,\beta} \operatorname{str}\Bigl(X_\alpha X_\alpha^\dagger L X_\beta X_\beta^\dagger L\Bigr)\biggr] = \int \frac{\mathrm{d}R}{2\pi} \, \exp\biggl[-\operatorname{str}\Bigl(\frac{R^2}{2\lambda^2} + \mathrm{i}R \sum\nolimits_\alpha X_\alpha X_\alpha^\dagger L\Bigr)\biggr] \tag{C.26}$$

with N two-dimensional supervectors $X_\alpha = (x_\alpha \;\; \chi_\alpha)$ and an ordinary diagonal matrix $L = \mathrm{diag}(\ell_{\mathrm{B}}, \ell_{\mathrm{F}})$, where $\ell_{\mathrm{B}}, \ell_{\mathrm{F}} \in \mathbb{R}$. Furthermore, the integration variable is a supermatrix

$$R = \begin{pmatrix} r_1 & \rho \\ \rho^* & \mathrm{i}r_2 \end{pmatrix}, \quad r_1, r_2 \in \mathbb{R}, \; \rho, \rho^* \in \mathcal{A}_{\mathrm{F}} \tag{C.27}$$

and $[\mathrm{d}R] \equiv \mathrm{d}r_1 \, \mathrm{d}r_2 \, \mathrm{d}\rho \, \mathrm{d}\rho^*$. To prove this relation, we observe that the terms in the exponent on the left-hand side of Eq. (C.26) expand to

$$\operatorname{str}\Bigl(X_\alpha X_\alpha^\dagger L X_\beta X_\beta^\dagger L\Bigr) = \ell_{\mathrm{B}}^2 \, x_\alpha^* x_\alpha x_\beta^* x_\beta - \ell_{\mathrm{F}}^2 \, \chi_\alpha^* \chi_\alpha \chi_\beta^* \chi_\beta - \ell_{\mathrm{B}} \ell_{\mathrm{F}} \, x_\alpha^* \chi_\alpha \chi_\beta^* x_\beta + \ell_{\mathrm{B}} \ell_{\mathrm{F}} \chi_\alpha^* x_\alpha x_\beta^* \chi_\beta \,. \tag{C.28}$$

We now inspect the right-hand side of (C.26). In the exponent there, we find the terms

$$\operatorname{str}(R^2) = r_1^2 + r_2^2 + 2\rho\rho^* \tag{C.29}$$

and

$$\operatorname{str}(RX_\alpha X_\alpha^\dagger L) = \ell_{\mathrm{B}} x_\alpha^* r_1 x_\alpha + \ell_{\mathrm{B}} x_\alpha^* \rho \chi_\alpha + \ell_{\mathrm{F}} \chi_\alpha^* \rho^* x_\alpha + \mathrm{i}\ell_{\mathrm{F}} \chi_\alpha^* r_2 \chi_\alpha \,. \tag{C.30}$$

The total integral over R thus factorizes into a bosonic and a fermionic contribution,

$$\begin{aligned} &\int \frac{\mathrm{d}R}{2\pi} \, \exp\biggl[-\operatorname{str}\Bigl(\frac{R^2}{2\lambda^2} + \mathrm{i}R \sum\nolimits_\alpha X_\alpha X_\alpha^\dagger L \Bigr) \biggr] \\ &\quad = \int \frac{\mathrm{d}r_1 \, \mathrm{d}r_2}{2\pi} \, \exp\left[-\frac{r_1^2 + r_2^2}{2\lambda^2} - \mathrm{i}r_1 \ell_{\mathrm{B}} \sum_\alpha x_\alpha^* x_\alpha + r_2 \ell_{\mathrm{F}} \sum_\alpha \chi_\alpha^* \chi_\alpha \right] \\ &\qquad \times \int \mathrm{d}\rho \, \mathrm{d}\rho^* \, \exp\left[\frac{\rho^* \rho}{\lambda^2} - \mathrm{i}\rho \ell_{\mathrm{B}} \sum_\alpha x_\alpha^* \chi_\alpha + \mathrm{i}\rho^* \ell_{\mathrm{F}} \sum_\alpha \chi_\alpha^* x_\alpha \right] . \end{aligned} \tag{C.31}$$

Evaluating these Gaussian integrals along the lines of Appendices C.1 and C.2, we find

$$\begin{aligned} &\int \frac{\mathrm{d}R}{2\pi} \, \exp\biggl[-\operatorname{str}\Bigl(\frac{R^2}{2\lambda^2} + \mathrm{i}R \sum\nolimits_\alpha X_\alpha X_\alpha^\dagger L \Bigr) \biggr] \\ &\quad = \exp\left[-\frac{\lambda^2}{2} \sum_{\alpha,\beta} \bigl(\ell_{\mathrm{B}}^2 x_\alpha^* x_\alpha x_\beta^* x_\beta - \ell_{\mathrm{F}}^2 \chi_\alpha^* \chi_\alpha \chi_\beta^* \chi_\beta - \ell_{\mathrm{B}} \ell_{\mathrm{F}} x_\alpha^* \chi_\alpha \chi_\beta^* x_\beta + \ell_{\mathrm{B}} \ell_{\mathrm{F}} \chi_\alpha^* x_\alpha x_\beta^* \chi_\beta \bigr) \right] . \end{aligned} \tag{C.32}$$

In view of (C.28), this is precisely the left-hand side of Eq. (C.26), and thus the identity has been proved.

Several auxiliary supermatrices. If the summands on the left-hand side of Eq. (C.26) are weighted by an additional factor depending on α and β (like in Eq. (3.51)), we need to invoke N auxiliary supermatrices R_α of the form (C.27). The Hubbard-Stratonovich transformation is then mediated by the identity

$$\begin{aligned} &\exp\Bigl[-\tfrac{1}{2} \sum\nolimits_{\alpha,\beta} s_{\alpha\beta} \, \operatorname{str}(X_\alpha X_\alpha^\dagger L X_\beta X_\beta^\dagger L) \Bigr] \\ &\quad = \int \frac{[\mathrm{d}R]}{(2\pi)^N} \exp\Bigl[-\tfrac{1}{2} \sum\nolimits_{\alpha,\beta} (s^{-1})_{\alpha\beta} \operatorname{str}(R_\alpha R_\beta) + \mathrm{i} \sum\nolimits_\alpha \operatorname{str}(R_\alpha X_\alpha X_\alpha^\dagger L) \Bigr] . \end{aligned} \tag{C.33}$$

As before, the X_α are N two-dimensional supervectors. Furthermore, $s = (s_{\alpha\beta})$ is a real symmetric matrix with inverse s^{-1} and $[\mathrm{d}R] = \prod_\alpha \mathrm{d}R_\alpha$. The proof of this relation works essentially analogously to the case from Eq. (C.26) with only one Hubbard-Stratonovich supermatrix R, except that the resulting Gaussian integrals over the auxiliary degrees of freedom $r_{1\alpha}$, $r_{2\alpha}$, ρ_α, and ρ_α^* are now N-dimensional, too. Namely, after substitution of the parametrization (C.27) for every R_α, the right-hand side of (C.33) becomes

$$\begin{aligned} &\int \frac{[\mathrm{d}R]}{(2\pi)^N} \exp\Bigl[-\tfrac{1}{2} \sum\nolimits_{\alpha,\beta} (s^{-1})_{\alpha\beta} \operatorname{str}(R_\alpha R_\beta) + \mathrm{i} \sum\nolimits_\alpha \operatorname{str}(R_\alpha X_\alpha X_\alpha^\dagger L) \Bigr] \\ &\quad = \int \frac{[\mathrm{d}R]}{(2\pi)^N} \exp\left\{ -\frac{1}{2} \sum_{\alpha,\beta} \bigl[r_{1\alpha} (s^{-1})_{\alpha\beta} r_{1\beta} + r_{2\alpha} (s^{-1})_{\alpha\beta} r_{2\beta} \bigr] - \mathrm{i}\ell_{\mathrm{B}} \sum_\alpha r_{1\alpha} x_\alpha^* x_\alpha + \ell_{\mathrm{F}} \sum_\alpha r_{2\alpha} \chi_\alpha^* \chi_\alpha \right\} \\ &\qquad \times \exp\left[\sum_{\alpha,\beta} \rho_\alpha^* (s^{-1})_{\alpha\beta} \rho_\beta - \mathrm{i}\ell_{\mathrm{B}} \sum_\alpha \rho_\alpha x_\alpha^* \chi_\alpha + \mathrm{i}\ell_{\mathrm{F}} \sum_\alpha \rho_\alpha^* \chi_\alpha^* x_\alpha \right] \end{aligned} \tag{C.34}$$

$$= \exp\left[-\frac{1}{2} \sum_{\alpha,\beta} s_{\alpha\beta} \bigl(\ell_{\mathrm{B}}^2 x_\alpha^* x_\alpha x_\beta^* x_\beta - \ell_{\mathrm{F}}^2 \chi_\alpha^* \chi_\alpha \chi_\beta^* \chi_\beta - \ell_{\mathrm{B}} \ell_{\mathrm{F}} x_\alpha^* \chi_\alpha \chi_\beta^* x_\beta + \ell_{\mathrm{B}} \ell_{\mathrm{F}} \chi_\alpha^* x_\alpha x_\beta^* \chi_\beta \bigr) \right] . \tag{C.35}$$

Comparison with the left-hand side of Eq. (C.33) establishes the proof.

D Saddle-point approximation

The saddle-point method is a powerful approach to evaluate integrals depending exponentially on a large parameter N approximately or even asymptotically exactly as $N \to \infty$. We will only sketch the key ideas and results here and refer to Refs. [221, 245, 246], for example, for a more thorough discussion of convergence, higher-order corrections, and other technicalities.

The foundation is Laplace's method for purely real integrals, which will be described in Appendix D.1. Appendix D.2 presents the generalization to the case of complex contour integrals, i.e., what is commonly known as the "saddle-point method" in the literature. Appendix D.3 discusses the extension to supersymmetric integrals.

D.1 Laplace's method

Base case. The basic case is an integral over a real interval $[a, b]$ of the form

$$I_N := \int_a^b \mathrm{d}x \, \mathrm{e}^{Nf(x)} \, g(x) \,, \tag{D.1}$$

where f and g are real-valued functions independent of N. For large values of N, the integrand is dominated by the region where $f(x)$ assumes its maximal value because the larger $f(x)$, the stronger $\mathrm{e}^{Nf(x)}$ grows with N. In other words, the relative influence of points in the vicinity of the maximizing $\hat{x} \in (a, b)$ (i.e., $f(\hat{x}) \geq f(x)$ for all $x \in [a, b]$) increases with increasing N. Hence the idea is to approximate $f(x)$ in the exponent in (D.1) by a Taylor series around $\hat{x}$,

$$I_N = \int_a^b \mathrm{d}x \, \mathrm{e}^{Nf(x_0) + \frac{N}{2} f''(\hat{x}) \, (x-\hat{x})^2 + \mathcal{O}((x-\hat{x})^3)} \, g(x) \,. \tag{D.2}$$

Note that the first derivative vanishes and the second derivative is negative because $\hat{x}$ describes a maximum by assumption. Furthermore, we take for granted that the maximum does not lie at the boundaries of the interval $[a, b]$. Cases with $\hat{x}$ at the boundary or degenerate maxima (such that $f''(\hat{x}) = 0$) can nonetheless be treated by similar methods [246]. Since the integral is dominated by a small region around $\hat{x}$, we can truncate the Taylor expansion at second order. Indeed, the exponential in (D.2) now has a Gaussian shape with width ("standard deviation") $\sigma_N := 1/\sqrt{N|f''(\hat{x})|}$ and hence becomes ever narrower (relatively speaking) as N is increased. For sufficiently large N, we can thus furthermore approximate $g(x)$ by $g(\hat{x})$, such that

$$I_N \simeq \mathrm{e}^{Nf(\hat{x})} \, g(\hat{x}) \int_a^b \mathrm{d}x \, \mathrm{e}^{\frac{N}{2} f''(\hat{x}) \, (x-\hat{x})^2} \,. \tag{D.3}$$

Moreover, since the integrand quickly becomes negligibly small if x is more than a few multiples of σ_N away from $\hat{x}$, we can extend the domain of integration to the entire real line. The resulting Gaussian integral was evaluated in Appendix C.1 (see Eq. (C.1)), yielding

$$I_N = \int_a^b \mathrm{d}x \, \mathrm{e}^{Nf(x)} \, g(x) \simeq \mathrm{e}^{Nf(\hat{x})} \, g(\hat{x}) \sqrt{\frac{2\pi}{-Nf''(\hat{x})}} . \tag{D.4}$$

This is the leading-order asymptotics for the integral I_N as $N \to \infty$. If desired, higher-order corrections could be calculated in principle by including higher-order terms in the Taylor series of $f(x)$ and $g(x)$ [246]. Furthermore, if the function $f(x)$ exhibits several maxima of equal importance (equal height) within the domain of integration, we have to sum in (D.4) over the contributions of all such $\hat{x}$ where $f(x)$ assumes the global maximum. Even though we did not encounter a saddle point (it is hiding in the complex plane), the result (D.4) is a first special case of a saddle-point approximation.

Multivariate extension. It is straightforward to extend the method to functions $f, g : \Omega \to \mathbb{R}$, defined on a d-dimensional, simply connected domain $\Omega \subseteq \mathbb{R}^d$. To this end, we consider the integral

$$I_N := \int_\Omega \mathrm{d}x\, \mathrm{e}^{Nf(x)}\, g(x)\,. \tag{D.5}$$

Expanding $f(x)$ to second order around a global maximum $\hat{x} \in \Omega$ from the interior of Ω yields

$$f(x) = f(\hat{x}) + \frac{1}{2}\sum_{i,j}(x_i - \hat{x}_i) f''(\hat{x})_{ij}(x_j - \hat{x}_j) + \mathcal{O}\big((x-\hat{x})^3\big)\,, \tag{D.6}$$

where $f''(\hat{x})$ denotes the Hessian matrix of $f(x)$ evaluated at $\hat{x}$, i.e., the matrix of second derivatives $f''(x)_{ij} = \partial^2 f(x)/\partial x_i \partial x_j$. We assume that $\hat{x}$ is a simple maximum such that $f''(\hat{x})$ is negative definite. The contribution from such a maximum can then be approximated by a d-dimensional Gaussian integral by analogy with the one-dimensional case, leading to

$$I_N = \int_\Omega \mathrm{d}x\, \mathrm{e}^{Nf(x)}\, g(x) \simeq \sum\nolimits_{\hat{x}} \mathrm{e}^{Nf(\hat{x})}\, g(\hat{x}) \left(\frac{2\pi}{N}\right)^{d/2} [\det(-f''(\hat{x}))]^{-1/2} \tag{D.7}$$

for the total integral, where, in case there are several relevant maxima of $f(x)$ in Ω, we sum over all their locations $\hat{x}$ on the right-hand side .

D.2 Saddle-point method for complex integrals

Constant phase and steepest descent. The integrand in (D.1) was purely real, i.e., $f(x)$, $g(x)$, N, and x were all real-valued. We would like to generalize the asymptotic relation (D.4) to integrals of the form

$$I_N := \int_C \mathrm{d}z\, \mathrm{e}^{Nf(z)}\, g(z)\,, \tag{D.8}$$

where C is a contour in the complex plane, N is a large, real parameter as before, and $f(z)$ and $g(z)$ are complex-valued, analytic functions on a domain $\Omega \subseteq \mathbb{C}$ such that $C \subseteq \Omega$.

We first observe that taking $g(z)$ to be complex-valued does not entail any serious complications compared to the case from (D.1) because the integral can be split into two contributions involving the real and imaginary parts of $g(z)$, respectively. Allowing for complex-valued $f(z)$, by contrast, can change the phenomenology distinctly. A first guess towards an approximation could be that the integral I_N is dominated again by regions around the maximum of $\mathrm{Re}\, f(z)$ for $z \in C$. However, $\mathrm{Im}\, f(z)$ does not necessarily vanish at such a maximum, meaning that the Taylor expansion has a purely imaginary contribution at linear order. While the Gaussian integral obtained by truncating the Taylor expression at second order will usually still be convergent, a nonvanishing imaginary part can entail rapid oscillations as $N \to \infty$ which in turn diminish the effect of the narrow real part and implicate that the integral is not dominated by the region around that maximum of $\mathrm{Re}\, f(z)$ in general. Moreover, Cauchy's theorem implies that the value of I_N is independent of the precise contour C so long as the end points are fixed and a possible deformation can be dragged back continuously onto the original contour without crossing any singularities of the integrand. Hence a maximum of $\mathrm{Re}\, f(z)$ along one contour may not be a maximum anymore along an equivalent second contour.

This freedom to choose the precise path of integration in fact provides the key to the saddle-point approximation in the complex case. Indeed, the problem is essentially traced back to the basic real case from (D.1) if we find a valid deformed contour $C' \subseteq \Omega$ whose end points agree with those of C and for which $f(z)$ has a constant imaginary part so that the phase of the exponential in (D.8) is constant. To wit, substituting $f(z) = u(z) + \mathrm{i}v$ with real-valued $u(z)$ and constant $v \in \mathbb{R}$ into Eq. (D.8) yields

$$I_N = \mathrm{e}^{\mathrm{i}Nv} \int_{C'} \mathrm{d}z\, \mathrm{e}^{Nu(z)}\, g(z) = \mathrm{e}^{\mathrm{i}Nv} \int_a^b \mathrm{d}t\, \mathrm{e}^{Nu(\zeta(t))}\, g(\zeta(t))\, \dot{\zeta}(t)\,, \tag{D.9}$$

where $z = \zeta(t)$ with $t \in [a, b]$ parametrizes the path C'. Separating the real and imaginary parts of $g(\zeta(t))\dot{\zeta}(t)$, the remaining integrals are then indeed of the form (D.1) and can be evaluated using (D.4) or its pertinent extensions [246]. This approach is also called the *method of steepest descent* because starting from the maximum of $u(z)$, the contour C' of constant $v = \operatorname{Im} f(z)$ is also the contour along which $u(z) = \operatorname{Re} f(z)$ decays most rapidly as a consequence of the Cauchy-Riemann equations. Moreover, if $\hat{z}$ denotes the location of the maximum such that $u'(\hat{z}) = 0$, then also $f'(\hat{z}) = 0$ since the derivative vanishes along the curve with constant $\operatorname{Im} f(z) = v$ and $f(z)$ is analytic. Hence $\hat{z}$ is a stationary point of $f(z)$ and, as another consequence of the Cauchy-Riemann equations, a *saddle point* of $u(z)$ because $\partial^2 u/\partial x^2 = -\partial^2 u/\partial y^2$ with $x = \operatorname{Re} z$, $y = \operatorname{Im} z$.

Employing a deformed contour C' that describes a path of steepest descent cutting through the saddle points of $f(z)$, Laplace's method can thus be used to find the leading-order behavior of the integral (D.8) as $N \to \infty$ or even a full asymptotic expansion by means of the appropriate extensions to calculate higher-order corrections [245, 246]. In practice, a suitable contour C' may consist of several segments. Some of them involve a constant phase of $f(z)$, so their contribution can be determined by Laplace's method, and others connect those segments of constant phase such that the integrand ideally vanishes or is at least subleading [246].

Leading-order approximation from vicinity of saddle points. Finding a contour C' consisting of contributing constant-phase and negligible joining segments can generally become cumbersome or even impossible. Fortunately, it suffices to inspect the neighborhood of the dominant saddle points if we are only interested in the leading-order approximation. For this purpose, we must still deform the contour such that it passes through the relevant saddle points along the direction of steepest descent, but it need not be of constant phase away from those saddles as long as it stays in the corresponding valley, i.e., the real part of $f(z)$ does not exceed the value at the saddle point.

If $\hat{z}$ is a relevant saddle point, we can parametrize the integration contour in the vicinity of $\hat{z}$ as $z = \zeta(t) := \hat{z} + t\,\mathrm{e}^{-\mathrm{i}\theta}$ with a fixed phase $\theta \in (-\pi, \pi]$ and t running from $-\eta$ to η ($\eta > 0$ fixed). Moreover, we denote the second derivative of $f(z)$ at $\hat{z}$ by $f''(\hat{z}) =: f_2\,\mathrm{e}^{\mathrm{i}\phi}$ with $f_2 > 0$ and $\phi \in (-\pi, \pi]$. The Taylor expansion of $f(z)$ around $\hat{z}$ along the path $z = \zeta(t)$ then reads

$$f(\zeta(t)) = f(\hat{z}) + \frac{1}{2} f_2\, t^2\, \mathrm{e}^{\mathrm{i}(\phi+2\theta)} + \mathcal{O}(t^3)\,. \tag{D.10}$$

For $\zeta(t)$ to pass through $\hat{z}$ along the direction of steepest descent, we thus need to choose θ such that $\phi + 2\theta = (2n+1)\pi$ ($n \in \mathbb{Z}$), e.g., $\theta = (\pi - \phi)/2$. The contribution from the immediate neighborhood of $\hat{z}$ to the integral (D.8) can thus be written as

$$\mathrm{e}^{Nf(\hat{z})+\mathrm{i}(\pi-\phi)/2}\, g(\hat{z}) \int_{-\eta}^{\eta} \mathrm{d}t\; \mathrm{e}^{-\frac{N}{2} f_2\, t^2}\,. \tag{D.11}$$

For sufficiently large N, we can extend the domain of the remaining integral to the entire real line similarly as in Appendix D.1 and are left with a standard real Gaussian integral. If $\hat{z}$ is the only dominant saddle, the integral (D.8) can thus be approximated as

$$I_N = \int_C \mathrm{d}z\; \mathrm{e}^{Nf(z)}\, g(z) \simeq \mathrm{e}^{Nf(\hat{z})+\mathrm{i}(\pi-\phi)/2}\, g(\hat{z}) \sqrt{\frac{2\pi}{N f_2}} = \mathrm{e}^{Nf(\hat{z})}\, g(\hat{z}) \sqrt{\frac{2\pi}{-N f''(\hat{z})}}\,. \tag{D.12}$$

If there are several relevant saddle points, their contributions have to be summed accordingly. Furthermore, this result can again be generalized to analytic functions $f, g : \Omega \to \mathbb{C}$ on a d-dimensional complex and simply connected domain $\Omega \subseteq \mathbb{C}^d$. Given a d-dimensional contour $C \subseteq \Omega$, it takes essentially the same form as in the real-valued case (cf. Eq. (D.7)), provided that the original contour C can be deformed such that it passes through the dominant saddle points $\{\hat{z}\}$ from the interior of Ω in the direction of steepest descent and stays in the corresponding valleys in between. In this case,

$$I_N = \int_C \mathrm{d}z\; \mathrm{e}^{Nf(z)}\, g(z) \simeq \sum\nolimits_{\hat{z}} \mathrm{e}^{Nf(\hat{z})}\, g(\hat{z}) \left(\frac{2\pi}{N}\right)^{d/2} \left[\det(-f''(\hat{z}))\right]^{-1/2}\,. \tag{D.13}$$

D.3 Supersymmetric extension

Anticommuting variables. To adapt the saddle-point method to superspace, we should verify that the result (D.13) has a formal analog for anticommuting variables. We content ourselves with the simplest case. Noting that the Taylor expansion of a function of a single anticommmuting variable terminates at first order, we thus consider two anticommuting numbers χ, χ^* and an integral of the form

$$I_N := \int \mathrm{d}\chi\, \mathrm{d}\chi^*\, \mathrm{e}^{Nf(\chi,\chi^*)}\, g(\chi,\chi^*)\,, \tag{D.14}$$

where $f(\chi,\chi^*)$ and $g(\chi,\chi^*)$ are functions of even Grassmann parity (see also Appendix B). Their general forms are

$$f(\chi,\chi^*) = f_{00} + \chi\, f_{10} + \chi^*\, f_{01} + \chi\chi^*\, f_{11}\,, \tag{D.15a}$$
$$g(\chi,\chi^*) = g_{00} + \chi\, g_{10} + \chi^*\, g_{01} + \chi\chi^*\, g_{11} \tag{D.15b}$$

with $f_{00}, f_{11}, g_{00}, g_{11} \in \mathcal{A}_\mathrm{B}$ and $f_{01}, f_{10}, g_{01}, g_{10} \in \mathcal{A}_\mathrm{F}$, respectively. Splitting off the c-number component of $f(\chi,\chi^*)$ and expanding the exponential for the remaining terms, we can evaluate the integral (D.14) explicitly and find

$$I_N = \mathrm{e}^{Nf_{00}} \left[N^2 f_{01}\, f_{10}\, g_{00} + N(f_{01}\, g_{10} - f_{10}\, g_{01} + f_{11}\, g_{00}) + g_{11}\right]. \tag{D.16}$$

Next we try to relate this result to an expansion around the "stationary points" of $f(\chi,\chi^*)$. We first observe that $f(\chi,\chi^*)$ from (D.15a) can be expanded as

$$\begin{aligned} f(\chi,\chi^*) = f(\hat\chi,\hat\chi^*) &+ (\chi-\hat\chi)\frac{\vec\partial}{\partial\chi} f(\chi,\chi^*)\Big|_{\hat\chi,\hat\chi^*} + (\chi^*-\hat\chi^*)\frac{\vec\partial}{\partial\chi^*} f(\chi,\chi^*)\Big|_{\hat\chi,\hat\chi^*} \\ &+ \frac{1}{2}(\chi-\hat\chi)\frac{\vec\partial}{\partial\chi} f(\chi,\chi^*)\frac{\overleftarrow\partial}{\partial\chi^*}\Big|_{\hat\chi,\hat\chi^*}(\chi^*-\hat\chi^*) + \frac{1}{2}(\chi^*-\hat\chi^*)\frac{\vec\partial}{\partial\chi^*} f(\chi,\chi^*)\frac{\overleftarrow\partial}{\partial\chi}\Big|_{\hat\chi,\hat\chi^*}(\chi-\hat\chi) \end{aligned} \tag{D.17}$$

for arbitrary $\hat\chi, \hat\chi^* \in \mathcal{A}_\mathrm{F}$. Choosing $\hat\chi = f_{01}\, f_{11}^{-1}$ and $\hat\chi^* = -f_{10}\, f_{11}^{-1}$ (the "saddle point"), the first order terms of this Taylor expansion vanish. The integral (D.14) can thus be written as

$$I_N = \mathrm{e}^{Nf(\hat\chi,\hat\chi^*)} \int \mathrm{d}\chi\, \mathrm{d}\chi^*\, \exp\left[\frac{N}{2}\begin{pmatrix}\chi-\hat\chi\\ \chi^*-\hat\chi^*\end{pmatrix}^\mathsf{T} f''(\hat\chi,\hat\chi^*)\begin{pmatrix}\chi-\hat\chi\\ \chi^*-\hat\chi^*\end{pmatrix}\right] g(\chi,\chi^*)\,, \tag{D.18}$$

where we introduced the notation

$$f''(\chi,\chi^*) = \begin{pmatrix} \frac{\vec\partial}{\partial\chi} f(\chi,\chi^*)\frac{\overleftarrow\partial}{\partial\chi} & \frac{\vec\partial}{\partial\chi} f(\chi,\chi^*)\frac{\overleftarrow\partial}{\partial\chi^*} \\ \frac{\vec\partial}{\partial\chi^*} f(\chi,\chi^*)\frac{\overleftarrow\partial}{\partial\chi} & \frac{\vec\partial}{\partial\chi^*} f(\chi,\chi^*)\frac{\overleftarrow\partial}{\partial\chi^*} \end{pmatrix} = \begin{pmatrix} 0 & f_{11} \\ -f_{11} & 0 \end{pmatrix} \tag{D.19}$$

for the matrix of second derivatives of $f(\chi,\chi^*)$. In the spirit of the saddle-point approximation for ordinary integrals, we now replace $g(\chi,\chi^*)$ by $g(\hat\chi,\hat\chi^*)$ in the integrand in (D.18) as well and shift the integration variables in the remaining Gaussian integral. Employing Eqs. (D.15), we obtain

$$\mathrm{e}^{Nf(\hat\chi,\hat\chi^*)}\, g(\hat\chi,\hat\chi^*) = \frac{\mathrm{e}^{Nf_{00}}}{f_{11}}\left(N\, f_{01}\, f_{10}\, g_{00} + f_{01}\, g_{10} - f_{10}\, g_{01} + f_{11}\, g_{00} - f_{01}\, f_{10}\, f_{11}^{-1}\, g_{11}\right) \tag{D.20}$$

for the prefactor, and

$$\int \mathrm{d}\chi\, \mathrm{d}\chi^*\, \exp\left[\frac{N}{2}\begin{pmatrix}\chi\\ \chi^*\end{pmatrix}^\mathsf{T} f''(\hat\chi,\hat\chi^*)\begin{pmatrix}\chi\\ \chi^*\end{pmatrix}\right] = \sqrt{N^2 \det[f''(\hat\chi,\hat\chi^*)]} = N\, f_{11} \tag{D.21}$$

for the remaining Gaussian integral. Comparing the exact result (D.16) to the product of Eqs. (D.20) and (D.21), we conclude that we can write

$$I_N = \int \mathrm{d}\chi\, \mathrm{d}\chi^*\, \mathrm{e}^{Nf(\chi,\chi^*)}\, g(\chi,\chi^*) \simeq \mathrm{e}^{Nf(\hat\chi,\hat\chi^*)}\, g(\hat\chi,\hat\chi^*)\sqrt{N^2 \det[f''(\hat\chi,\hat\chi^*)]} \tag{D.22}$$

to leading order in N. Hence the approximation is of a similar structure as in the complex case (cf. Eq. (D.13)) with the decisive difference that the product of $N^{d/2}$ and the square root of the Hessian determinant occurs in the numerator rather than in the denominator. The difference is thus similar to the difference between Gaussian integrals over commuting or anticommuting variables (see Appendix C), precisely because such Gaussian integrals are the key element leading to this contribution. In view of the unification of commuting and anticommuting Gaussian integrals by means of the superdeterminant (see Appendix C.3), it should not be too surprising that a similar unification can be achieved regarding the saddle-point approximation.

Saddle-point approximation in superspace. Combining the insights about complex contour integrals and the anticommuting saddle-point approximation, the results generalize essentially in the expected way to integrals in superspace. To sketch the pertinent extension, we consider integrals of the form

$$I_N := \int \frac{[\mathrm{d}X]}{(2\pi)^d} \, \mathrm{e}^{Nf(X)} \, g(X) \,, \tag{D.23}$$

where $f(X)$ and $g(X)$ are functions of even Grassmann parity with a generally complex-valued c-number component and X is a $4d$-dimensional supervector of the form

$$X = (\, x_1 \;\; \cdots \;\; x_{2d} \;\; \chi_1 \;\; \cdots \;\; \chi_{2d} \,)^{\mathsf{T}} \tag{D.24}$$

with $x_i \in \mathbb{R}$ (or a subset thereof) and Grassmann generators χ_i. Note that the relevant superintegrals from the main text over Hubbard-Stratonovich auxiliary matrices R of the form (3.52) or (3.104) can be cast into this form by collecting the matrix elements of R in a suitable vector X. For instance, we may identify $x_\alpha \equiv r_{1\alpha}$, $x_{N+\alpha} \equiv r_{2\alpha}$, $\chi_\alpha \equiv \rho_\alpha$, and $\chi_{N+\alpha} \equiv \rho_\alpha^*$ in (3.52). In particular, the x_i may or may not consist of pairs of real and imaginary parts of a complex variable, and the χ_i may or may not consist of pairs of formally conjugated Grassmann generators.

Moreover, we assume that the contour for the bosonic variables has already been chosen such that it passes through the relevant saddle points of the function f along the direction of steepest descent and is routed through the valleys in between (see, for example, Ref. [221] for a particularly thorough discussion of parametrization and convergence issues).

For notational convenience, we will address the components of X as $X_i = x_i$ and $X_{2d+i} = \chi_i$ for $i = 1, \ldots, 2d$ in the following. As before, we can expand the function $f(X)$ in the exponent into a Taylor series around $X = \hat{X}$ and obtain

$$f(X) = f(\hat{X}) + \sum_i (X_i - \hat{X}_i) \frac{\overrightarrow{\partial}}{\partial X_i} f(X) \bigg|_{X=\hat{X}} + \frac{1}{2} \sum_{i,j} (X_i - \hat{X}_i) \frac{\overrightarrow{\partial}}{\partial X_i} f(X) \frac{\overleftarrow{\partial}}{\partial X_j} \bigg|_{X=\hat{X}} (X_j - \hat{X}_j) + \ldots \tag{D.25}$$

up to second order. Note that the left- and right-derivatives act just like ordinary derivatives on $f(X)$ for the bosonic variables. In the following, we will denote

$$f''(\hat{X})_{ij} := \frac{\overrightarrow{\partial}}{\partial X_i} f(X) \frac{\overleftarrow{\partial}}{\partial X_j} \bigg|_{X=\hat{X}} \tag{D.26}$$

for short. Next we choose for $\hat{X}$ the dominant saddle points, such that the first derivative of f vanishes, and approximate the integrand in (D.23) by Gaussian forms in their vicinity,

$$I_N \simeq \sum_{\hat{X}} \mathrm{e}^{Nf(\hat{X})} \, g(\hat{X}) \int \frac{[\mathrm{d}X]}{(2\pi)^d} \, \mathrm{e}^{\frac{N}{2} \sum_{i,j} (X_i - \hat{X}_i) f''(\hat{X})_{ij} (X_j - \hat{X}_j)} \tag{D.27}$$

Observing that $f''(\hat{X})$ is a regular $(4d \times 4d)$ supermatrix, the remaining integral can be evaluated by means of the methods from Appendix C.3 and yields

$$I_N = \int \frac{[\mathrm{d}X]}{(2\pi)^d} \, \mathrm{e}^{Nf(X)} \, g(X) \simeq \sum\nolimits_{\hat{X}} \mathrm{e}^{Nf(\hat{X})} \, g(\hat{X}) \, \mathrm{sdet} \left[f''(\hat{X}) \right]^{-1/2} . \tag{D.28}$$

This final approximation is thus indeed structurally similar to the saddle-point approximations (D.7) and (D.13) for multidimensional real or complex integrals, except that the determinant is replaced by a superdeterminant and all constant scalar contributions (prefactor involving N, sign in front of the second derivative) cancel between the fermionic and bosonic contributions.

E Details on derivations

Here we elaborate in more detail on various steps of the derivations from Chapters 3 and 4.

E.1 Ensemble variance of the dynamics under diagonal perturbations (Sec. 3.2)

In this appendix, we show Eq. (3.16). Squaring (3.6) with $E_n^\lambda = E_n + \epsilon_n^\lambda$, the ensemble average $\mathbb{E}[(\langle A\rangle_{\rho_\lambda(t)})^2]$ can be evaluated similarly as in Eq. (3.15) by a careful analysis of all possible cases of matching indices between the four level-fluctuation variables ϵ_μ^λ occurring in the corresponding exponential. Subtracting the square of (3.15), a lengthy but straightforward rearrangement of terms yields

$$\mathbb{E}\big[\xi_V(t)^2\big] = \mathbb{E}\Big[\big(\langle A\rangle_{\rho_\lambda(t)}\big)^2\Big] - \big(\mathbb{E}\big[\langle A\rangle_{\rho_\lambda(t)}\big]\big)^2 = \mathcal{V}_1 + \mathcal{V}_2 + \mathcal{V}_3 + \mathcal{V}_4 + \mathcal{V}_5 + \mathcal{V}_6 \tag{E.1}$$

with

$$\mathcal{V}_1 := -\left(1 - \mathrm{e}^{-\lambda^2\sigma_0^2 t^2}\right) \sum_\mu [\rho_{\mu\mu}(0)]^2 (A_{\mu\mu})^2 \,, \tag{E.2a}$$

$$\mathcal{V}_2 := 4\mathrm{e}^{-2\lambda^2\sigma_0^2 t^2} \left(2 - \mathrm{e}^{-\lambda^2\sigma_0^2 t^2}\right) \mathrm{Re} \sum_{\mu,\nu} \rho_{\mu\mu}(0) A_{\mu\mu} [\rho_0(t)]_{\mu\nu} A_{\nu\mu} \,, \tag{E.2b}$$

$$\mathcal{V}_3 := \left(1 + \mathrm{e}^{-2\lambda^2\sigma_0^2 t^2} - 2\mathrm{e}^{-3\lambda^2\sigma_0^2 t^2} + \mathrm{e}^{-4\lambda^2\sigma_0^2 t^2}\right) \sum_{\mu,\nu} \left([\rho_0(t)]_{\mu\nu}\right)^2 (A_{\nu\mu})^2 \,, \tag{E.2c}$$

$$\mathcal{V}_4 := -\mathrm{e}^{-\lambda^2\sigma_0^2 t^2} \left(2 - \mathrm{e}^{-\lambda^2\sigma_0^2 t^2}\right) \sum_{\mu,\nu} |[\rho_0(t)]_{\mu\nu} A_{\nu\mu}|^2 \,, \tag{E.2d}$$

$$\mathcal{V}_5 := -2\mathrm{e}^{-2\lambda^2\sigma_0^2 t^2} \left(1 - \mathrm{e}^{-\lambda^2\sigma_0^2 t^2}\right) \mathrm{Re} \sum_\mu \left([\rho_0(t) A]_{\mu\mu}\right)^2 \,, \tag{E.2e}$$

$$\mathcal{V}_6 := 2\left(1 - \mathrm{e}^{-\lambda^2\sigma_0^2 t^2}\right) \sum_\mu |[\rho_0(t) A]_{\mu\mu}|^2 \,. \tag{E.2f}$$

With the exception of $\mathcal{V}_2$, the sums appearing in these terms can be upper-bounded in modulus by $p_{\max} \|A\|^2$, where $\|A\|$ denotes the operator norm of A and $p_{\max}$ is the largest population of a single energy level (see Eq. (2.14)). Furthermore, the various time-dependent prefactors can be bounded by constants of order unity, such that $|\mathcal{V}_1| \leq \|A\|^2 p_{\max}$, $|\mathcal{V}_3| \leq \frac{17}{16} \|A\|^2 p_{\max}$, $|\mathcal{V}_4| \leq \|A\|^2 p_{\max}$, $|\mathcal{V}_5| \leq \frac{8}{27} \|A\|^2 p_{\max}$, and $|\mathcal{V}_6| \leq 2 \|A\|^2 p_{\max}$. The sum in $\mathcal{V}_2$ is less than or equal to $\sqrt{p_{\max}} \|A\|^2$ in modulus, hence $|V_2| \leq 4 \|A\|^2 \sqrt{p_{\max}}$ Altogether, we thus find

$$\mathbb{E}\big[\xi_V(t)^2\big] \leq \|A\|^2 \left(4\sqrt{p_{\max}} + \tfrac{11}{2} p_{\max}\right) . \tag{E.3}$$

Since $\xi_V(t)$ from (3.3) is invariant upon adding an arbitrary constant to A, we can assume $\|A\| = \Delta A/2$ without loss of generality, where ΔA is the spectral range of A from (2.18). Observing that $\sqrt{p_{\max}} \geq p_{\max}$ since $p_{\max} \leq 1$, Eq. (E.3) then implies Eq. (3.16).

E.2 Saddle-point integral for the fourth overlap moment (Sec. 3.4.3)

In this appendix, we expound the evaluation of the integral (3.109) over the saddle-point manifold, eventually resulting in (3.111).

Parametrization of the saddle-point manifold. In a first step, we specify a suitable parametrization of the supermatrix $Q = T\Lambda T^{-1}$ serving as the integration variable in Eq. (3.109). To this end, we introduce (2×2)-dimensional supermatrices

$$A := \exp \begin{pmatrix} 0 & -\alpha^* \\ \alpha & 0 \end{pmatrix} \quad \text{and} \quad B := \exp \begin{pmatrix} 0 & -\mathrm{i}\beta^* \\ \mathrm{i}\beta & 0 \end{pmatrix} \tag{E.4}$$

with anticommuting $\alpha, \alpha^*, \beta, \beta^*$. The supermatrix A is unitary, $A^\dagger A = \mathbb{1}$, whereas B is pseudounitary, $B^\dagger k B = k$ with $k = \mathrm{diag}(1, -1)$. Defining $\tilde{\tau} := \mathrm{diag}(\tau_\mathrm{B}, \tau_\mathrm{F})$ with $\tau_\mathrm{B}, \tau_\mathrm{F} \in \mathbb{C}$, $|\tau_\mathrm{F}| \leq 1$, the pseudounitary transformation matrices T, which span the saddle-point manifold and satisfy $T^\dagger L T = L$ (see below Eq. (3.108)), can be parametrized as [163, 223, 338]

$$T = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} \sqrt{1 + k|\tilde{\tau}|^2} & k\tilde{\tau}^* \\ \tilde{\tau} & \sqrt{1 + k|\tilde{\tau}|^2} \end{pmatrix} \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix}. \tag{E.5}$$

Note that the left and right matrices are block-diagonal in the boson-fermion decomposition, while the middle matrix is block-diagonal in the retarded-advanced decomposition. Expressing the BB and FF eigenvalues as $\tau_{\mathrm{B,F}} = r_{\mathrm{B,F}} \mathrm{e}^{\mathrm{i}\phi_{\mathrm{B,F}}}$ with $r_\mathrm{B} \in [0, \infty)$, $r_\mathrm{F} \in [0, 1]$, and $\phi_\mathrm{B}, \phi_\mathrm{F} \in [0, 2\pi)$, the integration measure associated with the parametrization (E.5) reads [163]

$$\mathrm{d}\mu(T) = \frac{\mathrm{d}r_\mathrm{B}\, r_\mathrm{B}\, \mathrm{d}\phi_\mathrm{B}\, \mathrm{d}r_\mathrm{F}\, r_\mathrm{F}\, \mathrm{d}\phi_\mathrm{F}}{\pi^2\, (r_\mathrm{B}^2 + r_\mathrm{F}^2)^2} \mathrm{d}\alpha\, \mathrm{d}\alpha^*\, \mathrm{d}(\mathrm{i}\beta)\, \mathrm{d}(\mathrm{i}\beta^*)\,. \tag{E.6}$$

Adopting the parametrization (E.5) for T in the definition of $Q = T\Lambda T^{-1}$ with Λ from (3.100), we obtain

$$Q = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \tilde{Q} \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix} \tag{E.7}$$

with

$$\tilde{Q} := \begin{pmatrix} \ell_\mathrm{B} & 0 & -\sqrt{\ell_\mathrm{B}^2 - 1}\, \mathrm{e}^{-\mathrm{i}\phi_\mathrm{B}} & 0 \\ 0 & \ell_\mathrm{F} & 0 & \sqrt{1 - \ell_\mathrm{F}^2}\, \mathrm{e}^{-\mathrm{i}\phi_\mathrm{F}} \\ \sqrt{\ell_\mathrm{B}^2 - 1}\, \mathrm{e}^{\mathrm{i}\phi_\mathrm{B}} & 0 & -\ell_\mathrm{B} & 0 \\ 0 & \sqrt{1 - \ell_\mathrm{F}^2}\, \mathrm{e}^{\mathrm{i}\phi_\mathrm{F}} & 0 & -\ell_\mathrm{F} \end{pmatrix}. \tag{E.8}$$

Here $\ell_\mathrm{B} := 1 + 2r_\mathrm{B}^2 \in [1, \infty)$ and $\ell_\mathrm{F} := 1 - 2r_\mathrm{F}^2 \in [-1, 1]$, so the corresponding integration measure is given by

$$\mathrm{d}\mu(Q) = -\frac{\mathrm{d}\ell_\mathrm{B}\, \mathrm{d}\phi_\mathrm{B}\, \mathrm{d}\ell_\mathrm{F}\, \mathrm{d}\phi_\mathrm{F}}{(2\pi)^2 (\ell_\mathrm{B} - \ell_\mathrm{F})^2}\, \mathrm{d}\alpha\, \mathrm{d}\alpha^*\, \mathrm{d}\beta\, \mathrm{d}\beta^*\,. \tag{E.9}$$

Thus we fixed an explicit parametrization for the integration variable Q in (3.109) and specified the measure $\mathrm{d}\mu(Q)$.

Explicit form of the integrand. Next, we determine the precise form of the integrand in (3.109) upon substitution of the parametrization (E.7) for Q. To deal with the supermatrix elements in the last two lines of Eq. (3.109), we define the functions

$$K_\nu(\ell) := \left[(z_1^+ - E_\nu)(z_2^- - E_\nu) - \mathrm{i}\Gamma\Delta_z \ell/2 + \Gamma^2/4\right]^{-1}. \tag{E.10}$$

The four supermatrix elements occurring in (3.109) can then be written as

$$\begin{aligned}
&(\mathrm{i}\Gamma Q/2 + \bar{z} + \Delta_z \Lambda/2 - E_{\nu_i})^{-1}_{1\mathrm{B},1\mathrm{B}} \\
&= K_{\nu_i}(\ell_\mathrm{B})(z_2^- - E_{\mu_i} - \mathrm{i}\Gamma\ell_\mathrm{B}/2)(1 + \alpha\alpha^*) - K_{\nu_i}(\ell_\mathrm{F})(z_2^- - E_{\nu_i} - \mathrm{i}\Gamma\ell_\mathrm{F}/2)\, \alpha\alpha^*\,, \qquad \text{(E.11a)} \\
&(\mathrm{i}\Gamma Q/2 + \bar{z} + \Delta_z \Lambda/2 - E_{\nu_i})^{-1}_{1\mathrm{B},2\mathrm{B}} \\
&= \mathrm{i}\Gamma K_{\nu_i}(\ell_\mathrm{B}) \sqrt{\ell_\mathrm{B}^2 - 1}\, \mathrm{e}^{-\mathrm{i}\phi_\mathrm{B}} \left(\tfrac{1}{2} + \tfrac{\alpha\alpha^* - \beta\beta^*}{4} - \tfrac{\alpha\alpha^*\beta\beta^*}{8}\right) + \Gamma K_{\nu_i}(\ell_\mathrm{F}) \sqrt{1 - \ell_\mathrm{F}^2}\, \mathrm{e}^{-\mathrm{i}\phi_\mathrm{F}}\, \alpha^*\beta/2\,, \qquad \text{(E.11b)} \\
&(\mathrm{i}\Gamma Q/2 + \bar{z} + \Delta_z \Lambda/2 - E_{\nu_i})^{-1}_{2\mathrm{B},1\mathrm{B}} \\
&= -\mathrm{i}\Gamma K_{\nu_i}(\ell_\mathrm{B}) \sqrt{\ell_\mathrm{B}^2 - 1}\, \mathrm{e}^{\mathrm{i}\phi_\mathrm{B}} \left(\tfrac{1}{2} + \tfrac{\alpha\alpha^* - \beta\beta^*}{4} - \tfrac{\alpha\alpha^*\beta\beta^*}{8}\right) - \Gamma K_{\nu_i}(\ell_\mathrm{F}) \sqrt{1 - \ell_\mathrm{F}^2}\, \mathrm{e}^{\mathrm{i}\phi_\mathrm{F}}\, \alpha\beta^*/2\,, \qquad \text{(E.11c)}
\end{aligned}$$

$$
\begin{aligned}
&(\mathrm{i}\Gamma Q/2 + \bar{z} + \Delta_z \Lambda/2 - E_{\nu_i})^{-1}_{2\mathrm{B},2\mathrm{B}} \\
&= K_{\nu_i}(\ell_\mathrm{B})(z_1^+ - E_{\nu_i} + \mathrm{i}\Gamma\ell_\mathrm{B}/2)(1-\beta\beta^*) + K_{\nu_i}(\ell_\mathrm{F})(z_1^+ - E_{\mu_i} + \mathrm{i}\Gamma\ell_\mathrm{F}/2)\,\beta\beta^* \,. \qquad \text{(E.11d)}
\end{aligned}
$$

Turning to the exponent in (3.109), we expand it to first order in Δ_z since $\Delta_z \sim \varepsilon$ in the regime where the pseudounitary symmetry holds. Exploiting the saddle-point equation (3.107), this leads to

$$
-\operatorname{str}\sum\nolimits_\alpha \ln(\mathrm{i}\Gamma Q/2 + \bar{z} + \Delta_z\Lambda/2 - E_\alpha) = \mathrm{i}\pi\Delta_z(\ell_\mathrm{B} - \ell_\mathrm{F})/\varepsilon + \mathcal{O}(\Delta_z^2)\,. \qquad \text{(E.12)}
$$

Substituting (E.9), (E.11), and (E.12) into the integrand, the integral (3.109) becomes

$$
\begin{aligned}
&\mathbb{E}\big[\mathcal{G}_{\nu_1\mu_1}(z_1^+)\,\mathcal{G}_{\nu_2\mu_2}(z_2^-)\big] \\
&\quad= \int_1^\infty \mathrm{d}\ell_\mathrm{B} \int_{-1}^1 \mathrm{d}\ell_\mathrm{F} \frac{1}{(\ell_\mathrm{B}-\ell_\mathrm{F})^2} \int_0^{2\pi} \frac{\mathrm{d}\phi_\mathrm{B}}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\phi_\mathrm{F}}{2\pi} \int \mathrm{d}\alpha\,\mathrm{d}\alpha^*\,\mathrm{d}\beta\,\mathrm{d}\beta^*\, \exp\big[\tfrac{\mathrm{i}\pi\Delta_z}{\varepsilon}(\ell_\mathrm{B}-\ell_\mathrm{F})\big] \qquad \text{(E.13)} \\
&\qquad\qquad \times \big[\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\,(D_{00}+D_{11}+D_{10}+D_{01}) + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\,(F_{00}+F_{11}+F_\mathrm{S}+F_\mathrm{C})\big]
\end{aligned}
$$

with

$$
\begin{aligned}
D_{00} &:= K_{\nu_1}(\ell_\mathrm{B})K_{\nu_2}(\ell_\mathrm{B})(z_1^+ - E_{\nu_2} + \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{B})(z_2^- - E_{\nu_1} - \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{B})\,, && \text{(E.14a)} \\
D_{11} &:= -\alpha\alpha^*\beta\beta^*\big[K_{\nu_1}(\ell_\mathrm{B})(z_2^- - E_{\mu_1} - \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{B}) - K_{\nu_1}(\ell_\mathrm{F})(z_2^- - E_{\nu_1} - \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{F})\big] \\
&\qquad\times\big[K_{\nu_2}(\ell_\mathrm{B})(z_1^+ - E_{\nu_2} + \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{B}) - K_{\nu_2}(\ell_\mathrm{F})(z_1^+ - E_{\nu_2} + \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{F})\big]\,, && \text{(E.14b)} \\
D_{10} &:= \alpha\alpha^*\big[K_{\nu_1}(\ell_\mathrm{B})(z_2^- - E_{\nu_1} - \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{B}) - K_{\nu_1}(\ell_\mathrm{F})(z_2^- - E_{\nu_1} - \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{F})\big]\,K_{\nu_2}(\ell_\mathrm{B})(z_1^+ - E_{\nu_2} + \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{B})\,, && \text{(E.14c)} \\
D_{01} &:= -\beta\beta^* K_{\nu_1}(\ell_\mathrm{B})(z_2^- - E_{\nu_2} + \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{B})\big[K_{\nu_2}(\ell_\mathrm{B})(z_1^+ - E_{\nu_2} + \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{B}) - K_{\nu_2}(\ell_\mathrm{F})(z_1^+ - E_{\mu_2} + \tfrac{\mathrm{i}\Gamma}{2}\ell_\mathrm{F})\big]\,, && \text{(E.14d)} \\
F_{00} &:= \tfrac{\Gamma^2}{4}K_{\nu_1}(\ell_\mathrm{B})K_{\nu_2}(\ell_\mathrm{B})(\ell_\mathrm{B}^2 - 1), && \text{(E.14e)} \\
F_{11} &:= -\alpha\alpha^*\beta\beta^*\frac{\Gamma^2}{4}\big[K_{\nu_1}(\ell_\mathrm{B})K_{\nu_2}(\ell_\mathrm{B})(\ell_\mathrm{B}^2-1) + K_{\nu_1}(\ell_\mathrm{F})K_{\nu_2}(\ell_\mathrm{F})(1-\ell_\mathrm{F}^2)\big]\,, && \text{(E.14f)} \\
F_\mathrm{S} &:= (\alpha\alpha^* - \beta\beta^*)\frac{\Gamma^2}{4}K_{\nu_1}(\ell_\mathrm{B})K_{\nu_2}(\ell_\mathrm{B})\,, && \text{(E.14g)} \\
F_\mathrm{C} &:= -\tfrac{\mathrm{i}\Gamma^2}{4}\sqrt{\ell_\mathrm{B}^2-1}\sqrt{1-\ell_\mathrm{F}^2}\,\Big[\alpha\beta^*\,K_{\nu_1}(\ell_\mathrm{B})K_{\nu_2}(\ell_\mathrm{F})\,\mathrm{e}^{-\mathrm{i}(\phi_\mathrm{B}-\phi_\mathrm{F})} + \alpha^*\beta\,K_{\nu_1}(\ell_\mathrm{F})K_{\nu_2}(\ell_\mathrm{B})\,\mathrm{e}^{\mathrm{i}(\phi_\mathrm{B}-\phi_\mathrm{F})}\Big]\,. && \text{(E.14h)}
\end{aligned}
$$

Note that the terms have thus been classified according to the branches of δ contractions and their dependence on the anticommuting integration variables.

Evaluation of the integral. To evaluate the integral (E.13), we analyze the eight different contributions from (E.14) individually. At first sight, only the terms D_{11} and F_{11} involving the full set $\alpha\alpha^*\beta\beta^*$ of anticommuting variables would contribute, because if any of them is missing, the corresponding Grassmann integral yields zero. Upon closer inspection, however, we notice that the integration measure is singular in the commuting variables at the boundary of the associated domain where $\ell_\mathrm{B} = \ell_\mathrm{F} = 1$, i.e., $Q = \Lambda$ according to (E.8) or, equivalently, $T = \mathbb{1}$. If this singularity is not lifted by the remaining terms in the integrand, we have to invoke the Parisi-Sourlas-Efetov-Wegner (PSEW) theorem [163, 220, 338–341] to assess the total integral of "$0\cdot\infty$" type. According to the PSEW theorem, the integral then assumes the value of the integrand at the origin, i.e., for $T = \mathbb{1}$ or $\ell_\mathrm{B} = \ell_\mathrm{F} = 1$ in our present parametrization. The conditions of the theorem are fulfilled for the contributions from D_{00} and F_{00}. Substituting $\ell_\mathrm{B} = \ell_\mathrm{F} = 1$, we find

$$
\int D_{00} = D_{00}|_{\ell_\mathrm{B}=\ell_\mathrm{F}=1} = G_\mathrm{w}^-(z_1^+ - E_{\nu_1})\,G_\mathrm{w}^+(z_2^- - E_{\nu_2}) \qquad \text{(E.15)}
$$

and $\int F_{00} = 0$, where $G_\mathrm{w}^\pm(E) = (E \mp \mathrm{i}\Gamma/2)^{-1}$ was defined in (3.110). The bosonic integrals of D_{10} and D_{01} do not diverge due to an additional factor of $(\ell_\mathrm{B} - \ell_\mathrm{F})$ in the integrand, hence the total integrals $\int D_{10} = \int D_{01} = 0$ due to the vanishing Grassmann contribution. Similarly, the bosonic integral of F_C is convergent, but the Grassmann factors are incomplete, hence $\int F_\mathrm{C} = 0$. In case of F_S, in turn, the contributions proportional to $\alpha\alpha^*$ and $\beta\beta^*$ are identical up to a sign and thus cancel, implying that $\int F_\mathrm{S} = 0$, too.

Finally, there remain the integrals of D_{11} and F_{11}. The respective Grassmann integrals readily give $\int \mathrm{d}\alpha\mathrm{d}\alpha^*\mathrm{d}\beta\mathrm{d}\beta^*\,\alpha\alpha^*\beta\beta^* = 1$, and the integrals over ϕ_B and ϕ_F both yield factors of 2π canceling against the corresponding factors in the measure. The remaining integrals over ℓ_B and ℓ_F can be expressed in terms of the exponential integral function [342]

$$\mathrm{Ei}(z) := -\,\mathrm{PV}\int_{-z}^{\infty}\mathrm{d}t\,\frac{\mathrm{e}^{-t}}{t}\,, \tag{E.16}$$

where "PV" indicates that the integral should be evaluated in the principal-value sense. More precisely, it is convenient to define the two auxiliary functions

$$\mathcal{I}(v,c) := \int_1^\infty \mathrm{d}\ell\,\frac{\mathrm{e}^{\mathrm{i}v\ell}}{\ell+c} = -\mathrm{e}^{-\mathrm{i}vc}\,\mathrm{Ei}(\mathrm{i}v[c+1])\,, \tag{E.17}$$

where the last equality holds for $\mathrm{Im}\,v > 0$ and $|\arg(c+1)| < \pi$ [342, 343], and

$$\mathcal{J}(v,c) := \int_{-1}^1 \mathrm{d}\ell\,\frac{\mathrm{e}^{-\mathrm{i}v\ell}}{\ell+c} = \mathrm{e}^{\mathrm{i}vc}\left[\mathrm{Ei}(-\mathrm{i}v[c+1]) - \mathrm{Ei}(-\mathrm{i}v[c-1])\right], \tag{E.18}$$

which holds for $c \notin [-1,1]$ [342, 343]. To compute the integrals of D_{11} and F_{11}, we decompose the integrand into partial fractions in ℓ_B and ℓ_F and employ the definitions (E.17) and (E.18) of $\mathcal{I}(v,c)$ and $\mathcal{J}(v,c)$, respectively. Abbreviating $a_{ij} := z_i - E_{\nu_j}$ and $c_k := 2\mathrm{i}a_{1k}a_{2k}/\varGamma\varDelta_z + \mathrm{i}\varGamma/2\varDelta_z$, we find

$$\begin{aligned}\int D_{11} = {}& \frac{(a_{11}a_{21} - a_{21}\varDelta_z + \frac{\varGamma^2}{4})(a_{12}a_{22} + a_{12}\varDelta_z + \frac{\varGamma^2}{4})}{(a_{11}a_{21} - a_{12}a_{22})^2\varDelta_z^2} \\ & \times \left[\mathcal{I}(\tfrac{\pi\varDelta_z}{\varepsilon}, c_1) - \mathcal{I}(\tfrac{\pi\varDelta_z}{\varepsilon}, c_2)\right]\left[\mathcal{J}(\tfrac{\pi\varDelta_z}{\varepsilon}, c_1) - \mathcal{J}(\tfrac{\pi\varDelta_z}{\varepsilon}, c_2)\right]\end{aligned} \tag{E.19}$$

and

$$\begin{aligned}\int F_{11} = {}& \frac{1}{(E_{\nu_1} - E_{\nu_2})\varDelta_z(E_{\nu_1} + E_{\nu_2} - z_1^+ - z_2^-)} \\ & \times \left\{ \frac{(a_{11} - \frac{\mathrm{i}\varGamma}{2})(a_{21} + \frac{\mathrm{i}\varGamma}{2})\,\mathcal{J}(\frac{\pi\varDelta_z}{\varepsilon}, c_1)}{\varDelta_z}\left[\frac{2\pi(a_{11} + \frac{\mathrm{i}\varGamma}{2})(a_{21} - \frac{\mathrm{i}\varGamma}{2})\,\mathcal{I}(\frac{\pi\varDelta_z}{\varepsilon}, c_1)}{\varGamma\varepsilon} - \mathrm{e}^{\mathrm{i}\pi\varDelta_z/\varepsilon}\right]\right. \\ & - \frac{(a_{12} - \frac{\mathrm{i}\varGamma}{2})(a_{22} + \frac{\mathrm{i}\varGamma}{2})\,\mathcal{J}(\frac{\pi\varDelta_z}{\varepsilon}, c_2)}{\varDelta_z}\left[\frac{2\pi(a_{12} + \frac{\mathrm{i}\varGamma}{2})(a_{22} - \frac{\mathrm{i}\varGamma}{2})\,\mathcal{I}(\frac{\pi\varDelta_z}{\varepsilon}, c_2)}{\varGamma\varepsilon} - \mathrm{e}^{\mathrm{i}\pi\varDelta_z/\varepsilon}\right] \\ & \left. -\mathrm{i}\varGamma\cos(\tfrac{\pi\varDelta_z}{\varepsilon})\big[\mathcal{I}(\tfrac{\pi\varDelta_z}{\varepsilon}, c_1) - \mathcal{I}(\tfrac{\pi\varDelta_z}{\varepsilon}, c_2)\big] + \varGamma\sin(\tfrac{\pi\varDelta_z}{\varepsilon})\big[c_1\,\mathcal{I}(\tfrac{\pi\varDelta_z}{\varepsilon}, c_1) - c_2\mathcal{I}(\tfrac{\pi\varDelta_z}{\varepsilon}, c_2)\big]\right\}.\end{aligned} \tag{E.20}$$

To simplify these expressions, we notice that $v = \pi\varDelta_z/\varepsilon$ and $c = c_k$ in all arguments for $\mathcal{I}(v,c)$ and $\mathcal{J}(v,c)$. Hence the real part entering the exponential integral functions $\mathrm{Ei}(z)$ in (E.17) and (E.18) is given by $\pm\mathrm{i}vc_k = \mp\pi(2a_{1k}a_{2k}/\varGamma\varepsilon + \varGamma/2\varepsilon)$, at least if $\eta \to 0$ in z_1^+ and z_2^-. Since $\varGamma/\varepsilon = N_v \gg 1$ (see Eq. (3.9)), we can approximate $\mathrm{Ei}(z)$ in these expressions by the asymptotic relation $\mathrm{Ei}(z) \sim \mathrm{e}^z/z$ as $\mathrm{Re}\,z \to \pm\infty$ (observing a branch-cut discontinuity of 2π in the imaginary part along the negative real line). This leads to

$$\int D_{11} \approx \left(\frac{\varGamma\varepsilon}{2\pi\varDelta_z}\right)^2\left[\frac{(a_{12} - \frac{\mathrm{i}\varGamma}{2})(a_{21} + \frac{\mathrm{i}\varGamma}{2})}{(a_{11} + \frac{\mathrm{i}\varGamma}{2})^2(a_{12} + \frac{\mathrm{i}\varGamma}{2})(a_{21} - \frac{\mathrm{i}\varGamma}{2})(a_{22} - \frac{\mathrm{i}\varGamma}{2})^2} - \frac{\exp\left(\frac{2\pi\mathrm{i}\varDelta_z}{\varepsilon}\right)}{(a_{11}^2 + \varGamma^2/4)(a_{22}^2 + \varGamma^2/4)}\right] \tag{E.21}$$

and

$$\int F_{11} \approx \frac{\mathrm{i}\varGamma\varepsilon/\pi\varDelta_z}{(a_{11} + \frac{\mathrm{i}\varGamma}{2})(a_{12} + \frac{\mathrm{i}\varGamma}{2})(a_{21} - \frac{\mathrm{i}\varGamma}{2})(a_{22} - \frac{\mathrm{i}\varGamma}{2})}. \tag{E.22}$$

The total integral (3.109) is then obtained by combining the three nonvanishing contributions (E.15), (E.21), and (E.22) according to (E.13), i.e.,

$$\mathbb{E}\big[\mathcal{G}_{\nu_1\mu_1}(z_1^+)\,\mathcal{G}_{\nu_2\mu_2}(z_2^-)\big] = \delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\left[\int D_{00} + \int D_{11}\right] + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\int F_{11}\,. \tag{E.23}$$

The second relevant ensemble average of retarded and advanced resolvents, $\mathbb{E}\big[\mathcal{G}_{\nu_1\mu_1}(z_1^-)\,\mathcal{G}_{\nu_2\mu_2}(z_2^+)\big]$, is obtained from (E.23) by exchanging all labels 1 and 2, i.e., $n_1 \leftrightarrow n_2$, $\mu_1 \leftrightarrow \mu_2$, and $\nu_1 \leftrightarrow \nu_2$. Substituting the explicit expressions and taking the limit $\eta \to 0$, we eventually recover Eq. (3.111).

E.3 Example for the alternative overlap-moment approximation (Sec. 3.4.4)

As an example for the approximate evaluation of eigenvector overlap moments (3.29) by the computational scheme sketched in Sec. 3.4.4 and explained in more detail in the Supplemental Material of Ref. [226], we present how to calculate the fourth moment (3.128) in this approach. In the first step, we express the orthogonalized $\hat{v}_\mu^{n_k}$ in terms of the independent $v_\mu^{n_k}$ via (3.126), adopting $a_k = 1$ (see below Eq. (3.127)). For $n_1 \neq n_2$, we thus obtain

$$\begin{aligned}
&\mathrm{E}\big[\hat{v}_{\mu_1}^{n_1}\hat{v}_{\mu_2}^{n_2}\hat{v}_{\nu_1}^{n_1*}\hat{v}_{\nu_2}^{n_2*}\big]\big|_{n_1\neq n_2} \\
&= \mathrm{E}\big[v_{\mu_1}^{n_1}v_{\mu_2}^{n_2}v_{\nu_1}^{n_1*}v_{\nu_2}^{n_2*}\big] + \mathrm{E}\Big[v_{\mu_1}^{n_1}\Big(v_{\mu_2}^{n_1}\sum\nolimits_\alpha v_\alpha^{n_1*}v_\alpha^{n_2}\Big)v_{\nu_1}^{n_1*}\Big(v_{\nu_2}^{n_1*}\sum\nolimits_\beta v_\beta^{n_1}v_\beta^{n_2*}\Big)\Big] \\
&\quad - \mathrm{E}\Big[v_{\mu_1}^{n_1}\Big(v_{\mu_2}^{n_1}\sum\nolimits_\alpha v_\alpha^{n_1*}v_\alpha^{n_2}\Big)v_{\nu_1}^{n_1*}v_{\nu_2}^{n_2*}\Big] - \mathrm{E}\Big[v_{\mu_1}^{n_1}v_{\mu_2}^{n_2}v_{\nu_1}^{n_1*}\Big(v_{\nu_2}^{n_1*}\sum\nolimits_\alpha v_\alpha^{n_1}v_\alpha^{n_2*}\Big)\Big]\,.
\end{aligned} \tag{E.24}$$

As observed in Sec. 3.4.4, we thus need to calculate the moments of four, six, and eight factors of independent, complex Gaussian vector components v_μ^n. According to the Isserlis-Wick theorem [239, 240], such averages are given by the sum of all possible combinations of pairing up factors of v_μ^m and v_ν^{n*} (such that $m = n$ and $\mu = \nu$) and multiplying their second moments. For the first term in the second line of (E.24) specifically, this yields

$$\begin{aligned}
\mathrm{E}\big[v_{\mu_1}^{n_1}v_{\mu_2}^{n_2}v_{\nu_1}^{n_1*}v_{\nu_2}^{n_2*}\big] &= (\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2} + \delta_{n_1n_2}\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1})\,\mathrm{E}\big[|v_{\mu_1}^{n_1}|^2\big]\,\mathrm{E}\big[|v_{\mu_2}^{n_2}|^2\big] \\
&= (\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2} + \delta_{n_1n_2}\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1})\,u(E_{n_1}-E_{\mu_1})\,u(E_{n_2}-E_{\mu_2})\,.
\end{aligned} \tag{E.25}$$

Similarly, after a careful inspection of the different possibilities of matching indices, the two terms in the third line involving six factors of v_μ^n are found to be given by

$$\begin{aligned}
\mathrm{E}\Big[v_{\mu_1}^{n_1}\Big(v_{\mu_2}^{n_1}\sum\nolimits_\alpha v_\alpha^{n_1*}v_\alpha^{n_2}\Big)v_{\nu_1}^{n_1*}v_{\nu_2}^{n_2*}\Big] &= (\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1})\,u(E_{n_1}-E_{\mu_1})\,u(E_{n_1}-E_{\mu_2}) \\
&\quad\times\{u(E_{n_2}-E_{\nu_2}) + \delta_{n_1n_2}\,[1+u(E_{n_1}-E_{\nu_1})]\}\,,
\end{aligned} \tag{E.26}$$

$$\begin{aligned}
\mathrm{E}\Big[v_{\mu_1}^{n_1}v_{\mu_2}^{n_2}v_{\nu_1}^{n_1*}\Big(v_{\nu_2}^{n_1*}\sum\nolimits_\alpha v_\alpha^{n_1}v_\alpha^{n_2*}\Big)\Big] &= (\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1})\,u(E_{n_1}-E_{\mu_1})\,u(E_{n_1}-E_{\mu_2}) \\
&\quad\times\{u(E_{n_2}-E_{\mu_2}) + \delta_{n_1n_2}\,[1+u(E_{n_1}-E_{\mu_1})]\}\,,
\end{aligned} \tag{E.27}$$

where we exploited that $\sum_\alpha u(E_n - E_\alpha) = 1$ by definition (see Eq. (3.61)). Additionally employing $\sum_\alpha u(E_{n_1}-E_\alpha)\,u(E_{n_2}-E_\alpha) = \tilde{u}(E_{n_1}-E_{n_2})$ (cf. Eqs. (3.62) and (3.116)), the second term in the second line, which comprises eight factors of v_μ^n, reduces to

$$\begin{aligned}
&\mathrm{E}\Big[v_{\mu_1}^{n_1}\Big(v_{\mu_2}^{n_1}\sum\nolimits_\alpha v_\alpha^{n_1*}v_\alpha^{n_2}\Big)v_{\nu_1}^{n_1*}\Big(v_{\nu_2}^{n_1*}\sum\nolimits_\beta v_\beta^{n_1}v_\beta^{n_2*}\Big)\Big] \\
&= (\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2} + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1})\,u(E_{n_1}-E_{\mu_1})\,u(E_{n_1}-E_{\mu_2}) \\
&\quad\times\Big\{u(E_{n_1}-E_{\mu_1})\,u(E_{n_2}-E_{\mu_1}) + u(E_{n_1}-E_{\mu_2})\,u(E_{n_2}-E_{\mu_2}) + \tilde{u}(E_{n_1}-E_{n_2}) \\
&\qquad + \delta_{n_1n_2}\Big[1 + 2u(E_{n_1}-E_{\mu_1}) + 2u(E_{n_1}-E_{\mu_2}) + \big(u(E_{n_1}-E_{\mu_1}) + u(E_{n_2}-E_{\mu_2})\big)^2\Big]\Big\}\,.
\end{aligned} \tag{E.28}$$

Observing that every factor of $u(E)$ is of order N_v^{-1} and restricting to the leading order in any of the $\delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}$ or $\delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}$ branches, we thus conclude that (E.24) becomes

$$\begin{aligned}
&\mathrm{E}\big[\hat{v}_{\mu_1}^{n_1}\hat{v}_{\mu_2}^{n_2}\hat{v}_{\nu_1}^{n_1*}\hat{v}_{\nu_2}^{n_2*}\big]\big|_{n_1\neq n_2} \\
&= \delta_{\mu_1\nu_1}\delta_{\mu_2\nu_2}\,u(E_{n_1}-E_{\mu_1})\,u(E_{n_2}-E_{\mu_2}) \\
&\quad + \delta_{\mu_1\nu_2}\delta_{\mu_2\nu_1}\,u(E_{n_1}-E_{\mu_1})\,u(E_{n_1}-E_{\mu_2})\,[\tilde{u}(E_{n_1}-E_{n_2}) - u(E_{n_2}-E_{\mu_1}) - u(E_{n_2}-E_{\mu_2})]\,.
\end{aligned} \tag{E.29}$$

For $n_1 = n_2$, in turn, we only have one vector $\hat{v}^{n_1}$ to consider, which is approximated by the corresponding independent v^{n_1}. Hence the fourth moment is given by (E.25) with $n_1 = n_2$. Combined with (E.29) for $n_1 \neq n_2$, we then recover Eqs. (3.128) and (3.129) from the main text.

E.4 Exploiting ensemble properties for the prethermalization bound (Sec. 3.5)

In this appendix, we prove Eqs. (3.135), (3.136), and (3.137), which establish certain properties of the ensemble average of the operator $W(t)$ from Eq. (3.134) (see also above Eq. (3.131)) based on characteristics of the underlying perturbation ensemble.

Series expansion of $W(t)$. Recalling the definition $W(t) := \mathrm{e}^{-\mathrm{i}H_\lambda t}\mathrm{e}^{\mathrm{i}H_0 t}$, we can represent the matrix elements $W_{\mu\nu}(t) := {}_0\langle\mu|W(t)|\nu\rangle_0$ of this operator in the unperturbed eigenbasis $\{|\mu\rangle_0\}$ by exploiting the following result from Ref. [344]: Given a complex analytic function $\phi(z)$ and linear operators A and B, the operator $\phi(A+B)$, which is defined in terms of the power-series representation of $\phi(z)$, can be expanded as

$$\phi(A+B) = \sum_{n=0}^{\infty} \frac{1}{n!} C_n(A,B)\, \phi^{(n)}(A)\,. \tag{E.30}$$

Here $\phi^{(n)}(z)$ denotes the nth derivative of $\phi(z)$, and the operators $C_n(A,B)$ satisfy the recurrence relation

$$C_0 := \mathbb{1}\,, \quad C_n := [A, C_{n-1}] + BC_{n-1}\,, \tag{E.31}$$

where $\mathbb{1}$ is the identity operator and $[A,B] = AB - BA$ the commutator as usual. Letting $\phi(z) := \mathrm{e}^{-\mathrm{i}zt}$, $A := H_0$, and $B := \lambda V$ so that $W(t) = \phi(A+B)\,\mathrm{e}^{\mathrm{i}H_0 t}$, it follows straightforwardly that

$$W_{\mu\nu}(t) = \sum_{n=0}^{\infty} \frac{(-\mathrm{i}t)^n}{n!} {}_0\langle\mu|C_n|\nu\rangle_0 \tag{E.32}$$

with

$$C_0 := \mathbb{1}\,, \quad C_n := [H_0, C_{n-1}] + \lambda V C_{n-1}\,. \tag{E.33}$$

Proof of Eq. (3.135). First, we demonstrate Eq. (3.135), i.e., we show that $\mathbb{E}[W(t)]$ is diagonal in the unperturbed eigenbasis $\{|\mu\rangle_0\}$ if the V ensemble is invariant under phase changes of the basis vectors $\{|\mu\rangle_0\}$ (see below Eq. (3.22)). More precisely, consider a second set of unperturbed basis states $\{|\mu\rangle_0'\}$ obtained from the original one by multiplying all $|\mu\rangle_0$ by arbitrarily chosen sign factors $s_\mu \in \{-1,1\}$, such that $|\mu\rangle_0' = s_\mu|\mu\rangle_0$. If the matrix elements ${}_0\langle\mu|V|\nu\rangle_0$ and ${}_0'\langle\mu|V|\nu\rangle_0'$ have identical statistical properties for all μ and ν, meaning that all their moments agree, it follows immediately from the definition (E.33) that also

$$\mathbb{E}[{}_0\langle\mu|C_n|\nu\rangle_0] = \mathbb{E}[{}_0'\langle\mu|C_n|\nu\rangle_0'] = s_\mu s_\nu\, \mathbb{E}[{}_0\langle\mu|C_n|\nu\rangle_0] \tag{E.34}$$

for all n. If $\mu = \nu$, this relation is trivially fulfilled due to $s_\mu^2 = 1$. If $\mu \neq \nu$, however, the product $s_\mu s_\nu$ can be either $+1$ or -1 depending on the particular choice of the s_μ. Since (E.34) must hold regardless of this choice, it follows that $\mathbb{E}[{}_0\langle\mu|C_n|\nu\rangle_0] = 0$ if $\mu \neq \nu$. Together with (E.32), we therefore find that

$$\mathbb{E}[W_{\mu\nu}(t)] = \delta_{\mu\nu}\, \mathbb{E}[w_\mu(t)] \quad \text{with} \quad w_\mu(t) := \sum_{n=0}^{\infty} \frac{(-\mathrm{i}t)^n}{n!} {}_0\langle\mu|C_n|\mu\rangle_0\,. \tag{E.35}$$

Proof of Eq. (3.136). Second, we prove Eq. (3.136), i.e., we establish that the ensemble average $\mathbb{E}[w_\mu(t)]$ of the just-defined function $w_\mu(t)$ from (E.35) is independent of μ under Prerequisites (i) and (iv) from Sec. 3.2.

Prerequisite (i) demanded a constant density of states ε^{-1} of H_0 within the energy shell $I_\mathcal{E}$. Hence we can approximate $E_\nu - E_\mu$ by $(\nu-\mu)\varepsilon$ and therefore freely shift indices in such energy differences, i.e., $E_{\nu+\alpha} - E_{\mu+\alpha} \approx E_\nu - E_\mu$. Note that the induced error can be estimated using (3.17) and strictly bounded from above using (3.19).

Prerequisite (iv) stipulated that the variance of $V_{\mu\nu}$ should only depend on the difference $E_\mu - E_\nu$ (cf. Eq. (3.10)), and since the relevant properties of the pertinent V ensembles are largely determined by the first two moments (see the discussion below Eq. (3.48)), this property can be extended to arbitrary moments. In particular, we thus take for granted that

$$\mathbb{E}[V_{\nu_0\nu_1} V_{\nu_1\nu_2} \cdots V_{\nu_{k-1}\nu_k}] = v_k\big(\{E_{\nu_i} - E_{\nu_{i-1}}\}_{i=1}^k\big) \tag{E.36}$$

for arbitrary indices $\nu_0, \ldots, \nu_k$ and all $k \in \mathbb{N}$, where v_k is a function which only depends on the differences $E_{\nu_1} - E_{\nu_0}, \ldots, E_{\nu_k} - E_{\nu_{k-1}}$.

According to (E.35), it suffices to show that $\mathbb{E}[{}_0\langle\mu|C_n|\mu\rangle_0]$ is independent of μ to conclude the same property for $\mathbb{E}[w_\mu(t)]$. By definition (see Eq. (E.33)), each operator C_n consists of a sum of terms involving powers of V and commutators of these powers with the unperturbed Hamiltonian H_0. By inserting complete sets of states $\mathbb{1} = \sum_{\nu_i} |\nu_i\rangle_{00}\langle\nu_i|$ between factors of V, every ${}_0\langle\mu|C_n|\mu\rangle_0$ is found to be a sum of terms of the general form

$$F(\mu; k; s_1, \ldots, s_k) = \sum_{\nu_1,\ldots,\nu_{k-1}} V_{\mu\nu_1}(E_{\nu_1} - E_\mu)^{s_1} V_{\nu_1\nu_2}(E_{\nu_2} - E_{\nu_1})^{s_2} \cdots V_{\nu_{k-1}\mu}(E_\mu - E_{\nu_{k-1}})^{s_k} \quad \text{(E.37)}$$

with $k \leq n$, $s_i \in \mathbb{N}_0$, and

$$k + \sum_{i=1}^{k} s_i = n\,. \quad \text{(E.38)}$$

We remark that the precise combination of $F(\mu; k; \{s_i\})$ terms contributing to any ${}_0\langle\mu|C_n|\mu\rangle_0$ can be computed from the recurrence relation (E.33) (see also Ref. [344]), but is irrelevant for the ensuing argument. Adopting (E.36), the ensemble average of Eq. (E.37) reduces to

$$\begin{aligned}\mathbb{E}[F(\mu; k; s_1, \ldots, s_k)] = \sum_{\nu_1,\ldots,\nu_{k-1}} &(E_{\nu_1} - E_\mu)^{s_1}(E_{\nu_2} - E_{\nu_1})^{s_2} \cdots (E_\mu - E_{\nu_{k-1}})^{s_k} \\ &\times v_k\big(E_{\nu_1} - E_\mu, \{E_{\nu_i} - E_{\nu_{i-1}}\}_{i=2}^{k-1}, E_\mu - E_{\nu_{k-1}}\big)\,.\end{aligned} \quad \text{(E.39)}$$

Next we shift the indices ν_i to $\nu_i + \mu$ for all $i = 1, \ldots, k-1$ and exploit Prerequisite (i) as specified above Eq. (E.36). This leads to

$$\begin{aligned}\mathbb{E}[F(\mu; k; s_1, \ldots, s_k)] = \sum_{\nu_1,\ldots,\nu_{k-1}} &(E_{\nu_1})^{s_1}(E_{\nu_2} - E_{\nu_1})^{s_2} \cdots (-E_{\nu_{k-1}})^{s_k} \\ &\times v_k\big(E_{\nu_1}, \{E_{\nu_i} - E_{\nu_{i-1}}\}_{i=2}^{k-1}, -E_{\nu_{k-1}}\big)\,.\end{aligned} \quad \text{(E.40)}$$

Since the right-hand side of this relation is manifestly independent of μ, the same must hold for $\mathbb{E}[F(\mu; k; s_1, \ldots, s_k)]$. Consequently, also $\mathbb{E}[{}_0\langle\mu|C_n|\mu\rangle_0]$ and thus $\mathbb{E}[w_\mu(t)]$ from (E.35) are independent of μ, which implies Eq. (3.136) from the main text with

$$\hat{w}(t) := \sum_{n=0}^{\infty} \frac{(-\mathrm{i}t)^n}{n!} \mathbb{E}[{}_0\langle\mu|C_n|\mu\rangle_0]\,. \quad \text{(E.41)}$$

Proof of Eq. (3.137). Third and last, we show that Eq. (3.132) reduces to (3.137) if the perturbation ensemble exhibits the previously exploited properties and the distribution of the diagonal matrix elements $V_{\mu\mu}$ is invariant under inversion of the sign, meaning that $V_{\mu\mu}$ and $-V_{\mu\mu}$ have identical statistical properties. To this end, as observed above Eq. (3.137), it suffices to show that $\hat{w}(t)$ is real-valued. In view of the definition (E.41), we thus need to show that $\mathbb{E}[{}_0\langle\mu|C_n|\mu\rangle_0]$ is purely real for even n and purely imaginary for odd n. Since all ${}_0\langle\mu|C_n|\mu\rangle_0$ are sums of terms of the form (E.37), it is sufficient to verify that

$$\mathbb{E}[F(\mu; k; s_1, \ldots, s_k)] = (-1)^n\, \mathbb{E}[F(\mu; k; s_1, \ldots, s_k)]^*\,, \quad \text{(E.42)}$$

where n, k, and s_i are related via (E.38). Without loss of generality, we can shift the overall energy scale such that $E_0 = 0$ and hence $E_{-\mu} = -E_\mu$ according to Prerequisite (i) (see above Eq. (E.36)). Observing (E.38), we thus find that

$$(E_{-\nu_1})^{s_1}(E_{-\nu_2} - E_{-\nu_1})^{s_2} \cdots (-E_{-\nu_{k-1}})^{s_k} = (-1)^{n-k}\,(E_{\nu_1})^{s_1}(E_{\nu_2} - E_{\nu_1})^{s_2} \cdots (-E_{\nu_{k-1}})^{s_k}\,. \quad \text{(E.43)}$$

Next we recall that $V_{\mu\nu}$ and $-V_{\mu\nu}$ are assumed to have identical statistical properties for all μ and ν in the present setting. Replacing all $V_{\nu_{i-1}\nu_i}$ by $-V_{\nu_{i-1}\nu_i}$ in (E.36), we therefore conclude that the functions v_k characterizing the moments of the $V_{\mu\nu}$ satisfy

$$v_k\big(\{E_{\nu_i} - E_{\nu_{i-1}}\}_{i=1}^{k}\big) = (-1)^k\, v_k\big(\{E_{\nu_i} - E_{\nu_{i-1}}\}_{i=1}^{k}\big)\,. \quad \text{(E.44)}$$

Furthermore, Hermiticity of V implies $V_{\mu\nu} = V_{\nu\mu}^*$ and thus

$$v_k\big(\{E_{\nu_i} - E_{\nu_{i-1}}\}_{i=1}^k\big)^* = v_k\big(\{E_{\nu_{i-1}} - E_{\nu_i}\}_{i=1}^k\big)\,. \tag{E.45}$$

Combining the latter two equations, we conclude that

$$v_k\big(\{E_{\nu_i} - E_{\nu_{i-1}}\}_{i=1}^k\big) = (-1)^k\, v_k\big(\{E_{\nu_{i-1}} - E_{\nu_i}\}_{i=1}^k\big)^*\,. \tag{E.46}$$

Finally, we consider Eq. (E.40) again. Replacing all summation indices ν_i by $-\nu_i$, this relation can be written as

$$\begin{aligned}\mathbb{E}[F(\mu;k;s_1,\ldots,s_k)] = \sum_{\nu_1,\ldots,\nu_{k-1}} &(E_{-\nu_1})^{s_1}(E_{-\nu_2}-E_{-\nu_1})^{s_2}\cdots(-E_{-\nu_{k-1}})^{s_k}\\ &\times v_k\big(E_{-\nu_1},\{E_{\nu_{i-1}}-E_{\nu_i}\}_{i=2}^{k-1},-E_{-\nu_{k-1}}\big)\,.\end{aligned} \tag{E.47}$$

Substituting (E.43) in the first line and utilizing (E.46) in conjunction with $E_{-\nu} = -E_\nu$ (see above Eq. (E.43)) in the second line, we obtain

$$\begin{aligned}\mathbb{E}[F(\mu;k;s_1,\ldots,s_k)] = (-1)^n \sum_{\nu_1,\ldots,\nu_{k-1}} &(E_{\nu_1})^{s_1}(E_{\nu_2}-E_{\nu_1})^{s_2}\cdots(-E_{\nu_{k-1}})^{s_k}\\ &\times v_k\big(E_{\nu_1},\{E_{\nu_i}-E_{\nu_{i-1}}\}_{i=2}^{k-1},-E_{\nu_{k-1}}\big)^*\,.\end{aligned} \tag{E.48}$$

Comparing (E.40) and (E.48), it follows that Eq. (E.42) holds, which in turn implies Eq. (3.137).

E.5 Bound for the remnant term in the ensemble-averaged dynamics (Sec. 3.6.3)

In this appendix, we prove the bound (3.185) for the term $\tilde{R}(t)$ from (3.184) appearing in the ensemble-averaged time evolution (3.183). To this end, we first observe that $\int \mathrm{d}\omega\; \tilde{r}(\omega,t) = 0$ for all t (see also above Eq.(3.182)). Hence we can subtract a constant from the observable A without changing the value of $\tilde{R}(t)$. Defining $\tilde{A} := A - \langle A\rangle_{\rho_{\mathrm{mc}}}\mathbb{1}$, we can thus rewrite (3.184) as

$$\tilde{R}(t) = \sum_{\mu,\nu} \rho_{\mu\mu}(0)\,\tilde{A}_{\nu\nu}\,\tilde{r}(E_\mu - E_\nu, t)\,, \tag{E.49}$$

where the matrix elements $\tilde{A}_{\nu\nu} = {}_0\langle\nu|\tilde{A}|\nu\rangle_0$ quantify how strongly the observable A violates the (diagonal) ETH (see also Eq. (3.182)).

In a first step, we inspect the function $\tilde{r}(\omega,t)$ from (3.184). We recall that $u(E)$ is given by the Breit-Wigner distribution (3.33) and thus $\tilde{u}(E)$ from (3.116) also assumes the Breit-Wigner form with Γ replaced by 2Γ. Hence

$$\tilde{r}(\omega,t) = \frac{1}{\pi N_v}\left[1 - \cos(\omega t) - \frac{\Gamma\sin(\omega|t|)}{\omega}\right]\frac{1}{1+(\omega/\Gamma)^2}\,, \tag{E.50}$$

where we used $\Gamma/\varepsilon = N_v$ as well (see Eq. (3.9) and the explanation above it). For fixed $\omega \in \mathbb{R}\setminus\{0\}$, this function is bounded and periodic in t. The maximal value is attained at times

$$t_k^{\max} := \frac{1}{|\omega|}\left[\arctan\left(\frac{\Gamma}{|\omega|}\right) + (2k+1)\pi\right] \quad (k\in\mathbb{Z})\,. \tag{E.51}$$

To evaluate $\tilde{r}(\omega,t)$ at one such maximum, we utilize $\sin(\arctan(x)+\pi) = -x/\sqrt{1+x^2}$ and $\cos(\arctan(x)+\pi) = -1/\sqrt{1+x^2}$ and find

$$\tilde{r}(\omega,t) \le \frac{2\varepsilon}{\pi|\omega|}\,. \tag{E.52}$$

Similarly, $\tilde{r}(\omega,t)$ assumes its minimum for fixed $\omega \neq 0$ at times

$$t_k^{\min} := \frac{1}{|\omega|}\left[\arctan\left(\frac{\Gamma}{|\omega|}\right) + 2k\pi\right] \quad (k\in\mathbb{Z})\,. \tag{E.53}$$

Once again evaluating at such a minimum with $\sin(\arctan x) = x/\sqrt{1+x^2}$ and $\cos(\arctan x) = 1/\sqrt{1+x^2}$, we obtain

$$\tilde{r}(\omega, t) \geq -\frac{\varepsilon}{\pi|\omega|}\,. \tag{E.54}$$

Combining (E.52) and (E.54), we thus find

$$|\tilde{r}(\omega, t)| \leq \frac{2\varepsilon}{\pi|\omega|}\,. \tag{E.55}$$

To continue, we observe that Eq. (3.183) for the ensemble-averaged time evolution asserts that deviations from the long-time limit $\langle A\rangle_{\tilde{\rho}_\lambda}$ decay with $|g_\lambda(t)|^2 = \mathrm{e}^{-\Gamma|t|}$ in time. Assuming that we can measure such deviations with a resolution κ as detailed below Eq. (3.185), the term $\tilde{R}(t)$ can only become relevant at times $t \leq t_{\kappa *}$ with the "resolvable time" $t_{\kappa *} := -\frac{\ln \kappa}{\Gamma}$.

To obtain a useful bound on $|\tilde{R}(t)|$ within this regime, we split the double sum in (E.49) into two parts with $|E_\mu - E_\nu| < \Gamma/b$ and $|E_\mu - E_\nu| \geq \Gamma/b$, respectively, where $b > 0$ is chosen such that

$$b \arctan b = -\ln \kappa\,. \tag{E.56}$$

Explicitly,

$$\tilde{R}_1(t) := \sum_{\substack{\mu,\nu \\ |E_\mu - E_\nu| < \Gamma/b}} \rho_{\mu\mu}(0)\, \tilde{A}_{\nu\nu}\, \tilde{r}(E_\mu - E_\nu, t)\,, \tag{E.57a}$$

$$\tilde{R}_2(t) := \sum_{\substack{\mu,\nu \\ |E_\mu - E_\nu| \geq \Gamma/b}} \rho_{\mu\mu}(0)\, \tilde{A}_{\nu\nu}\, \tilde{r}(E_\mu - E_\nu, t)\,. \tag{E.57b}$$

Coming back to the function $\tilde{r}(\omega, t)$, we begin with the first case, $|\omega| = |E_\mu - E_\nu| < \Gamma/b$. As observed above, the first local maximum of the absolute value $|\tilde{r}(\omega, t)|$, corresponding to the first minimum of $\tilde{r}(\omega, t)$, lies at $t_0^{\min} = \arctan(\Gamma/|\omega|)/|\omega|$. Since $x \mapsto x \arctan x$ is a monotonically increasing function and $|\omega| < \Gamma/b$, we conclude that $t_0^{\min} > b \arctan(b)/\Gamma = t_{\kappa *}$. Consequently, the first maximum of $|\tilde{r}(\omega, t)|$ is assumed beyond the resolvable time $t_{\kappa *}$. Moreover, since $|\tilde{r}(\omega, 0)| = 0$, the function $|\tilde{r}(\omega, t)|$ is monotonically increasing on the interval $[0, t_{\kappa *}]$ and can therefore be bounded from above by

$$|\tilde{r}(\omega, t)| \leq |\tilde{r}(\omega, t_{\kappa *})| \leq \frac{1 + b \arctan b}{\pi N_v} \tag{E.58}$$

for all $t \in [0, t_{\kappa *}]$. Note that we assumed $\kappa < 1$ here. For completeness, we remark that (E.58) remains valid for $\omega = 0$ (which was previously excluded from the discussion) because

$$|\tilde{r}(0, t)| = \frac{\varepsilon t}{\pi} \leq \frac{\varepsilon t_{\kappa *}}{\pi} = \frac{b \arctan b}{\pi N_v}\,. \tag{E.59}$$

Using the triangle inequality in (E.57a) together with the bound (E.58), we can conclude that

$$|\tilde{R}_1(t)| \leq \frac{1 + b \arctan b}{\pi N_v} \sum_{\substack{\mu,\nu \\ |E_\mu - E_\nu| < \Gamma/b}} \rho_{\mu\mu}(0)\, |\tilde{A}_{\nu\nu}|\,. \tag{E.60}$$

For the second part where $|\omega| = |E_\mu - E_\nu| \geq \Gamma/b$, we can directly use the inequality (E.55) and obtain

$$|\tilde{r}(\omega, t)| \leq \frac{2b}{\pi N_v}\,. \tag{E.61}$$

Substituting into (E.57b), we are left with

$$|\tilde{R}_2(t)| \leq \frac{2b}{\pi N_v} \sum_{\substack{\mu,\nu \\ |E_\mu - E_\nu| \geq \Gamma/b}} \rho_{\mu\mu}(0)\, |\tilde{A}_{\nu\nu}|\,. \tag{E.62}$$

Combining (E.60) and (E.62), we finally obtain

$$|\tilde{R}(t)| \leq |\tilde{R}_1(t)| + |\tilde{R}_2(t)| \leq \frac{C_\kappa}{N_v} \sum_\mu \rho_{\mu\mu}(0) \sum_\nu |\tilde{A}_{\nu\nu}| = \frac{C_\kappa}{N_v} \Delta_{\mathrm{mc}}(A)\,, \tag{E.63}$$

with $C_\kappa := \max\{2b, 1 + b \arctan b\}/\pi$, b defined implicitly as a function of κ according to (E.56), and $\Delta_{\mathrm{mc}}(A)$ as defined in Eq. (3.182). Hence we derived Eq. (3.185) from the main text.

E.6 Ensemble variance of the echo signal under imperfect preparation (Sec. 4.3.1)

In this appendix, we show Eq. (4.25). Exploiting (4.4) and (4.6), we first rewrite the variance of $\mathcal{A}(\tau+\delta+t)$ as

$$\mathbb{E}\Big[\big(\mathcal{A}(\tau+\delta+t)-\mathbb{E}[\mathcal{A}(\tau+\delta+t)]\big)^2\Big] = \mathbb{E}[\big(\langle A\rangle_{\rho_\mathrm{b}(t)}\big)^2] - \big(\mathbb{E}[\langle A\rangle_{\rho_\mathrm{b}(t)}]\big)^2 \tag{E.64}$$

Upon substitution of (4.15) into (4.16) and observing that $\rho(0)=\rho_\mathrm{T}$, we obtain

$$\begin{aligned}\langle A\rangle_{\rho_\mathrm{b}(t)} = \sum_{n_1,n_2}\sum_{\mu_1,\mu_2}\sum_{\nu_1,\nu_2} \mathrm{e}^{\mathrm{i}(E_{\nu_1}-E_{\mu_2})t}\,\mathrm{e}^{\mathrm{i}(E^W_{n_2}-E^W_{n_1})\delta}\,\mathrm{e}^{\mathrm{i}(E_{\nu_2}-E_{\mu_1})\tau}\,\rho_{\mu_1\nu_2}(0)\,A_{\mu_2\nu_1} \\ \times \tilde{U}_{n_1\mu_1}\tilde{U}_{n_2\mu_2}\tilde{U}^*_{n_1\nu_1}\tilde{U}^*_{n_2\nu_2}\,.\end{aligned} \tag{E.65}$$

The ensemble average of these time-dependent expectation values during the backward phase was found to be given by

$$\mathbb{E}\big[\langle A\rangle_{\rho_\mathrm{b}(t)}\big] = \big(\langle A\rangle_{\rho_\mathrm{f}(\tau-t)} - \langle A\rangle_{\rho_\mathrm{mc}}\big)\,|\hat{d}_W(\delta)|^2 + \langle A\rangle_{\rho_\mathrm{mc}} - \frac{1}{N^2}\langle A\rangle_{\rho_\mathrm{f}(\tau-t)} \tag{E.66}$$

in the main text, cf. Eqs. (4.16) and (4.23). Squaring Eq. (E.65) gives

$$\begin{aligned}\big(\langle A\rangle_{\rho_\mathrm{b}(t)}\big)^2 = \sum_{n_1\ldots n_4}\sum_{\mu_1\cdots\mu_4}\sum_{\nu_1\cdots\nu_4} \mathrm{e}^{\mathrm{i}(E_{\nu_1}-E_{\mu_2}+E_{\nu_3}-E_{\mu_4})t}\,\mathrm{e}^{\mathrm{i}(E^W_{n_2}-E^W_{n_1}+E^W_{n_4}-E^W_{n_3})\delta}\,\mathrm{e}^{\mathrm{i}(E_{\nu_2}-E_{\mu_1}+E_{\nu_4}-E_{\mu_3})\tau} \\ \times\ \rho_{\mu_1\nu_2}(0)\,\rho_{\mu_3\nu_4}(0)\,A_{\mu_2\nu_1}\,A_{\mu_4\nu_3}\,\tilde{U}_{n_1\mu_1}\tilde{U}_{n_2\mu_2}\tilde{U}_{n_3\mu_3}\tilde{U}_{n_4\mu_4}\tilde{U}^*_{n_1\nu_1}\tilde{U}^*_{n_2\nu_2}\tilde{U}^*_{n_3\nu_3}\tilde{U}^*_{n_4\nu_4}\,.\end{aligned} \tag{E.67}$$

To compute the average of this expression over the ensemble of scrambling operators W, we thus need the average over eight factors of transformation matrix elements $\tilde{U}_{n\mu}$ (cf. Eq. (4.14)). Similarly to the fourth moment from Eq. (4.17), this average can be evaluated with the aid of Ref. [197] by summing over all possible ways of pairing up the first and second indices of $\tilde{U}$ and $\tilde{U}^*$ factors. This leads to

$$\mathbb{E}\big[\tilde{U}_{n_1\mu_1}\tilde{U}_{n_2\mu_2}\tilde{U}_{n_3\mu_3}\tilde{U}_{n_4\mu_4}\tilde{U}^*_{n_1\nu_1}\tilde{U}^*_{n_2\nu_2}\tilde{U}^*_{n_3\nu_3}\tilde{U}^*_{n_4\nu_4}\big] = \sum_{P,P'\in\mathrm{Sym}(4)} v_{P,P'}\prod_{j=1}^{4}\delta_{n_j n_{P(j)}}\,\delta_{\mu_j\nu_{P'(j)}}\,, \tag{E.68}$$

where Sym(4) denotes the symmetric group of degree 4, i.e., the set of all permutations of $\{1,2,3,4\}$. Hence the sum in (E.68) comprises $(4!)^2 = 576$ terms. The symmetry factors $v_{P,P'}$ depend only on the cyclic structure of the composed permutation $P^{-1}P' \in \mathrm{Sym}(4)$. As in the main text, it is sufficient to consider the leading order in $N \gg 1$, to which the $v_{P,P'}$ coincide for the CUE and COE ensembles. Consulting Ref. [197], we find that

$$v_{1,1,1,1} \simeq N^{-4}\,,\quad v_{2,1,1} \simeq -N^{-5}\,,\quad v_{2,2} \simeq N^{-6}\,,\quad v_{3,1} \simeq 2N^{-6}\,,\quad v_4 \simeq -5N^{-7}\,, \tag{E.69}$$

where the subscripts label the five different combinations of cycle lengths in Sym(4). Plugging (E.68) into (E.67), we observe that the sums over the n_j and μ_j,ν_j factorize, so we can analyze them separately. For each $P \in \mathrm{Sym}(4)$, the corresponding sum over the n_j is of the form

$$F_P := \sum_{n_1\cdots n_4} \mathrm{e}^{\mathrm{i}(E^W_{n_2}-E^W_{n_1}+E^W_{n_4}-E^W_{n_3})\delta}\prod_{j=1}^{4}\delta_{n_j n_{P(j)}}\,. \tag{E.70}$$

The $4! = 24$ different permutations P yield eight different terms as listed in Tab. E.1. Note that the order in N depends again on the cyclic structure of P only.

For each $P' \in \mathrm{Sym}(4)$, in turn, the sums over μ_j,ν_j are of the form

$$\begin{aligned}G_{P'} := \sum_{\mu_1\cdots\mu_4}\sum_{\nu_1\cdots\nu_4} \mathrm{e}^{\mathrm{i}(E_{\nu_1}-E_{\mu_2}+E_{\nu_3}-E_{\mu_4})t}\,\mathrm{e}^{\mathrm{i}(E_{\nu_2}-E_{\mu_1}+E_{\nu_4}-E_{\mu_3})\tau} \\ \times\ \rho_{\mu_1\nu_2}(0)\,\rho_{\mu_3\nu_4}(0)\,A_{\mu_2\nu_1}\,A_{\mu_4\nu_3}\prod_{j=1}^{4}\delta_{\mu_j\nu_{P'(j)}}\,.\end{aligned} \tag{E.71}$$

Table E.1: Contributions to the average of Eq. (E.67) from the sums over n_j for the different permutations $P \in \mathrm{Sym}(4)$, cf. Eq. (E.70).

cycles	P (cycle notation)	F_P
$1,1,1,1$	(1)(2)(3)(4)	$N^4\,\lvert\hat{d}_W(\delta)\rvert^4$
$2,1,1$	(1 2)(3)(4), (1 4)(2)(3), (1)(2 3)(4), (1)(2)(3 4)	$N^3\,\lvert\hat{d}_W(\delta)\rvert^2$
	(1 3)(2)(4)	$N^3\,\hat{d}_W(\delta)^2\,\hat{d}^*_W(2\delta)$
	(1)(2 4)(3)	$N^3\,\hat{d}_W(2\delta)\,\hat{d}^*_W(\delta)^2$
$2,2$	(1 2)(3 4), (1 4)(2 3)	N^2
	(1 3)(2 4)	$N^2\,\lvert\hat{d}_W(2\delta)\rvert^2$
$3,1$	(1 2 3)(4), (1 3 2)(4), (1 2 4)(3), (1 4 2)(3), (1 3 4)(2), (1 4 3)(2), (1)(2 3 4), (1)(2 4 3)	$N^2\,\lvert\hat{d}_W(\delta)\rvert^2$
4	(1 2 3 4), (1 2 4 3), (1 3 2 4), (1 3 4 2), (1 4 2 3), (1 4 3 2)	N

Table E.2: Contributions to the average of Eq. (E.67) from the sums over μ_j, ν_j for the different permutations $P' \in \mathrm{Sym}(4)$, cf. Eq. (E.71).

P' (cycle notation)	$G_{P'}$
(1)(2)(3)(4), (1 3)(2 4)	$(\langle A\rangle_{\rho_\mathrm{f}(\tau-t)})^2$
(1 2)(3)(4), (1)(2)(3 4), (1 2 4 3), (1 3 4 2)	$N\langle A\rangle_{\rho_\mathrm{mc}}\langle A\rangle_{\rho_\mathrm{f}(\tau-t)}$
(1 2)(3 4)	$N^2(\langle A\rangle_{\rho_\mathrm{mc}})^2$
(1)(2 3)(4), (1 4)(2)(3), (1 3 2 4), (1 4 2 3)	$\mathrm{tr}[A^2\rho_\mathrm{f}(\tau-t)^2]$
(1 2 3)(4), (1)(2 3 4), (1 2 4)(3), (1 3 4)(2)	$N\langle A\rangle_{\rho_\mathrm{mc}}\mathrm{tr}[A\rho_\mathrm{f}(\tau-t)^2]$
(1 3 2)(4), (1 4 2)(3), (1 4 3)(2), (1)(2 4 3)	$\langle A^2\rangle_{\rho_\mathrm{f}(\tau-t)}$
(1 3)(2)(4), (1)(2 4)(3)	$\mathrm{tr}\{[\rho_\mathrm{f}(\tau-t)A]^2\}$
(1 2 3 4)	$N^2\langle A\rangle_{\rho_\mathrm{mc}}\mathrm{tr}[\rho_\mathrm{T}^2]$
(1 4 3 2)	$N\langle A^2\rangle_{\rho_\mathrm{mc}}$
(1 4)(2 3)	$N\langle A^2\rangle_{\rho_\mathrm{mc}}\mathrm{tr}[\rho_\mathrm{T}^2]$

The various permutations P' lead to the total of ten different expressions for $G_{P'}$ collected in Tab. E.2.

To find the ensemble average of (E.67), we then have to combine the symmetry factors $v_{P,P'}$ from (E.69), the F_P from Eq. (E.70) and Tab. E.1, and the $G_{P'}$ from Eq. (E.71) and Tab. E.2 for each $P, P' \in \mathrm{Sym}(4)$. Assessing the order of each of these contributions in N is most conveniently achieved by means of computer algebra software. We find a total of nine terms of order 1, which cancel exactly in (E.64) against the nine terms of order 1 resulting from $(\mathbb{E}[\langle A\rangle_{\rho_\mathrm{b}(t)}])^2$, i.e., by squaring Eq. (E.66). Consequently, the variance (E.64) vanishes to order 1. At order N^{-1}, the average of (E.67) contributes 42 terms and the square of (E.66) adds another six terms, leading to

$$\begin{aligned}
\mathbb{E}\Big[\big(\langle A\rangle_{\rho_\mathrm{b}(t)}\big)^2\Big] - \mathbb{E}\big[\langle A\rangle_{\rho_\mathrm{b}(t)}\big]^2 = \frac{1}{N}\Big\{&2\langle A\rangle_{\rho_\mathrm{f}(\tau-t)}\langle A\rangle_{\rho_\mathrm{mc}} - (\langle A\rangle_{\rho_\mathrm{mc}})^2\,\mathrm{tr}(\rho_\mathrm{T}^2)\\
&+2|\hat{d}_W(\delta)|^2\left[(\langle A\rangle_{\rho_\mathrm{f}(\tau-t)})^2 - \langle A\rangle_{\rho_\mathrm{f}(\tau-t)}\langle A\rangle_{\rho_\mathrm{mc}} + \mathrm{tr}[A^2\rho_\mathrm{f}(\tau-t)^2]\right.\\
&\qquad\left.-2\,\mathrm{tr}[A\rho_\mathrm{f}(\tau-t)^2]\langle A\rangle_{\rho_\mathrm{mc}} + 2(\langle A\rangle_{\rho_\mathrm{mc}})^2\,\mathrm{tr}(\rho_\mathrm{T}^2)\right]\\
&+\left[\hat{d}_W(2\delta)\hat{d}_W(\delta)^{*2} + \hat{d}_W(\delta)^2\hat{d}_W(2\delta)^*\right]\left[(\langle A\rangle_{\rho_\mathrm{mc}})^2\,\mathrm{tr}(\rho_\mathrm{T}^2) - 2\,\mathrm{tr}[A\rho_\mathrm{f}(\tau-t)^2] + \mathrm{tr}\{[\rho_\mathrm{f}(\tau-t)A]^2\}\right]\\
&-|\hat{d}_W(\delta)|^4\left[2\,\mathrm{tr}[A^2\rho_\mathrm{f}(\tau-t)^2] - 8\,\mathrm{tr}[A\rho_\mathrm{f}(\tau-t)^2]\langle A\rangle_{\rho_\mathrm{mc}}\right.\\
&\qquad\left.+2\,\mathrm{tr}\{[\rho_\mathrm{f}(\tau-t)A]^2\} + 5(\langle A\rangle_{\rho_\mathrm{mc}})^2\,\mathrm{tr}(\rho_\mathrm{T}^2)\right]\Big\} + \mathcal{O}\left(\frac{1}{N^2}\right).
\end{aligned} \tag{E.72}$$

We content ourselves with a simple upper bound of this quantity by exploiting the triangle inequality on the right-hand side. The various combinations of traces over the observable and states can all be bounded by the squared operator norm $\|A\|^2$ of A. Moreover, we can exploit $|\hat{d}_W(\delta)| \leq 1$. Altogether, we then find

$$\mathbb{E}\Big[\big(\langle A\rangle_{\rho_{\mathrm{b}}(t)}\big)^2\Big] - \mathbb{E}\big[\langle A\rangle_{\rho_{\mathrm{b}}(t)}\big]^2 \leq \frac{42\,\|A\|^2}{N} + \mathcal{O}\left(\frac{1}{N^2}\right). \tag{E.73}$$

Together with (E.64) and taking $\|A\| = \Delta A/2$ without loss of generality as usual because the variance is unchanged when adding an arbitrary constant, we finally obtain Eq. (4.25).

References

[1] J. J. Sakurai, *Modern Quantum Mechanics*, revised ed. (Addison-Wesley, Reading, MA, 1993).

[2] D. Griffiths, *Introduction to quantum mechanics* (Prentice Hall, Englewood Cliffs, NJ, 1995).

[3] J. Gibbs, *Elementary principles in statistical mechanics: developed with special reference to the rational foundation of thermodynamics* (Dover, New York, 1960).

[4] L. Landau and E. Lifshitz, *Statistical Physics* (Pergamon, Oxford, 1970).

[5] C. Kittel and H. Kroemer, *Thermal Physics*, 2nd ed. (W. H. Freeman, US, 1980).

[6] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics* (Springer, Berlin, Heidelberg, 2010).

[7] C. Gogolin and J. Eisert, "Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems," Rep. Prog. Phys. **79**, 056001 (2016).

[8] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, "From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics," Adv. Phys. **65**, 239 (2016).

[9] T. Mori, T. N. Ikeda, E. Kaminishi, and M. Ueda, "Thermalization and prethermalization in isolated quantum systems: a theoretical overview," J. Phys. B **51**, 112001 (2018).

[10] J. M. Deutsch, "Eigenstate thermalization hypothesis," Rep. Prog. Phys. **81**, 082001 (2018).

[11] D. Griffiths, *Introduction to elementary particles* (Wiley, New York, 1987).

[12] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, 1st ed. (Perseus, Reading, MA, 1995).

[13] F. Wilczek, "The universe is a strange place," Nucl. Phys. B Proc. Suppl. **134**, 3 (2004).

[14] T. Padmanabhan, "Obtaining the non-relativistic quantum mechanics from quantum field theory: issues, folklores and facts," Eur. Phys. J. C **78**, 563 (2018).

[15] M. Schlosshauer, ed., *Elegance and Enigma* (Springer, Berlin, Heidelberg, 2011).

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley, Hoboken, NJ, 2006).

[17] P. Reimann, "Foundation of statistical mechanics under experimentally realistic conditions," Phys. Rev. Lett. **101**, 190403 (2008).

[18] P. Reimann, "Canonical thermalization," New J. Phys. **12**, 055027 (2010).

[19] N. Linden, S. Popescu, A. J. Short, and A. Winter, "Quantum mechanical evolution towards thermal equilibrium," Phys. Rev. E **79**, 061103 (2009).

[20] P. Reimann and M. Kastner, "Equilibration of isolated macroscopic quantum systems," New J. Phys. **14**, 043020 (2012).

[21] T. Farrelly, F. G. S. L. Brandão, and M. Cramer, "Thermalization and return to equilibrium on finite quantum lattice systems," Phys. Rev. Lett. **118**, 140601 (2017).

[22] P. Reimann, B. N. Balz, J. Richter, and R. Steinigeweg, "Temporal relaxation of gapped many-body quantum systems," Phys. Rev. B **101**, 094302 (2020).

[23] A. J. Short and T. C. Farrelly, "Quantum equilibration in finite time," New J. Phys. **14**, 013063 (2012).

[24] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, UK, 2007).

[25] A. J. Daley, "Quantum trajectories and open many-body quantum systems," Adv. Phys. **63**, 77 (2014).

[26] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch, and I. Bloch, "Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms," Nature **415**, 39 (2002).

[27] O. Mandel, M. Greiner, A. Widera, T. Rom, T. W. Hänsch, and I. Bloch, "Controlled collisions for multi-particle entanglement of optically trapped atoms," Nature **425**, 937 (2003).

[28] T. Kinoshita, T. Wenger, and D. S. Weiss, "A quantum Newton's cradle," Nature **440**, 900 (2006).

[29] S. Hofferberth, I. Lesanovsky, B. Fischer, T. Schumm, and J. Schmiedmayer, "Non-equilibrium coherence dynamics in one-dimensional Bose gases," Nature **449**, 324 (2007).

[30] I. Bloch, J. Dalibard, and W. Zwerger, "Many-body physics with ultracold gases," Rev. Mod. Phys. **80**, 885 (2008).

[31] M. Gring, M. Kuhnert, T. Langen, T. Kitagawa, B. Rauer, M. Schreitl, I. Mazets, D. A. Smith, E. Demler, and J. Schmiedmayer, "Relaxation and prethermalization in an isolated quantum system," Science **337**, 1318 (2012).

[32] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, "Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional bose gas," Nat. Phys. **8**, 325 (2012).

[33] T. Langen, T. Gasenzer, and J. Schmiedmayer, "Prethermalization and universal dynamics in near-integrable quantum systems," J. Stat. Mech: Theory Exp. **2016**, 064009 (2016).

[34] T. Langen, R. Geiger, and J. Schmiedmayer, "Ultracold atoms out of equilibrium," Annu. Rev. Cond. Mat. Phys. **6**, 201 (2015).

[35] R. P. Feynman, "Simulating physics with computers," Int. J. Theor. Phys. **21**, 467 (1982).

[36] I. M. Georgescu, S. Ashhab, and F. Nori, "Quantum simulation," Rev. Mod. Phys. **86**, 153 (2014).

[37] H. Schneider and H. Schmiedel, "Negative time development of a nuclear spin system," Phys. Lett. A **30**, 298 (1969).

[38] W.-K. Rhim, A. Pines, and J. S. Waugh, "Violation of the spin-temperature hypothesis," Phys. Rev. Lett. **25**, 218 (1970).

[39] W.-K. Rhim, A. Pines, and J. S. Waugh, "Time-reversal experiments in dipolar-coupled spin systems," Phys. Rev. B **3**, 684 (1971).

[40] S. Zhang, B. H. Meier, and R. R. Ernst, "Polarization echoes in NMR," Phys. Rev. Lett. **69**, 2149 (1992).

[41] R. Kimmich, J. Niess, and S. Hafner, "Quadrupolar magic echoes," Chem. Phys. Lett. **190**, 503 (1992).

[42] S. Hafner, D. Demco, and R. Kimmich, "Magic echoes and NMR imaging of solids," Solid State Nucl. Magn. Reson. **6**, 275 (1996).

[43] P. R. Levstein, G. Usaj, and H. M. Pastawski, "Attenuation of polarization echoes in nuclear magnetic resonance: A study of the emergence of dynamical irreversibility in many-body quantum systems," J. Chem. Phys. **108**, 2718 (1998).

[44] G. Usaj, H. M. Pastawski, and P. R. Levstein, "Gaussian to exponential crossover in the attenuation of polarization echoes in NMR," Molecular Physics **95**, 1229 (1998).

[45] K. X. Wei, C. Ramanathan, and P. Cappellaro, "Exploring localization in nuclear spin chains," Phys. Rev. Lett. **120**, 070501 (2018).

[46] K. X. Wei, P. Peng, O. Shtanko, I. Marvian, S. Lloyd, C. Ramanathan, and P. Cappellaro, "Emergent prethermalization signatures in out-of-time ordered correlations," Phys. Rev. Lett. **123**, 090605 (2019).

[47] M. Niknam, L. F. Santos, and D. G. Cory, "Sensitivity of quantum information to environment perturbations measured with a nonlocal out-of-time-order correlation function," Phys. Rev. Research **2**, 013200 (2020).

[48] D. P. DiVincenzo, "Quantum computation," Science **270**, 255 (1995).

[49] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010).

[50] J. Clarke and F. K. Wilhelm, "Superconducting quantum bits," Nature **453**, 1031 (2008).

[51] IBM Quantum Experience (IBM Q), `https://www.ibm.com/quantum-computing/`, access: 12/09/2020.

[52] F. Arute *et al.*, "Quantum supremacy using a programmable superconducting processor," Nature **574**, 505 (2019).

[53] R. Blatt and C. F. Roos, "Quantum simulations with trapped ions," Nat. Phys. **8**, 277 (2012).

[54] E. A. Martinez, C. A. Muschik, P. Schindler, D. Nigg, A. Erhard, M. Heyl, P. Hauke, M. Dalmonte, T. Monz, P. Zoller, and R. Blatt, "Real-time dynamics of lattice gauge theories with a few-qubit quantum computer," Nature **534**, 516 (2016).

[55] A. A. Zhukov, S. V. Remizov, W. V. Pogosov, and Y. E. Lozovik, "Algorithmic simulation

of far-from-equilibrium dynamics using quantum computer," Quantum Inf. Process. **17**, 223 (2018).

[56] A. Chiesa, F. Tacchino, M. Grossi, P. Santini, I. Tavernelli, D. Gerace, and S. Carretta, "Quantum hardware simulating four-dimensional inelastic neutron scattering," Nat. Phys. **15**, 455 (2019).

[57] A. Smith, M. S. Kim, F. Pollmann, and J. Knolle, "Simulating quantum many-body dynamics on a current digital quantum computer," npj Quantum Information **5**, 106 (2019).

[58] S. Trotzky, L. Pollet, F. Gerbier, U. Schnorrberger, I. Bloch, N. V. Prokof'ev, B. Svistunov, and M. Troyer, "Suppression of the critical temperature for superfluidity near the Mott transition," Nat. Phys. **6**, 998 (2010).

[59] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, "Many-body localization in a quantum simulator with programmable random disorder," Nat. Phys. **12**, 907 (2016).

[60] P. Bordia, H. Lüschen, U. Schneider, M. Knap, and I. Bloch, "Periodically driving a many-body localized quantum system," Nat. Phys. **13**, 460 (2017).

[61] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe, "Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator," Nature **551**, 601 (2017).

[62] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, "Probing many-body dynamics on a 51-atom quantum simulator," Nature **551**, 579 (2017).

[63] G. Vidal, "Efficient classical simulation of slightly entangled quantum computations," Phys. Rev. Lett. **91**, 147902 (2003).

[64] G. Vidal, "Efficient simulation of one-dimensional quantum many-body systems," Phys. Rev. Lett. **93**, 040502 (2004).

[65] G. Vidal, "Classical simulation of infinite-size quantum lattice systems in one spatial dimension," Phys. Rev. Lett. **98**, 070201 (2007).

[66] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, "Time-dependent density-matrix renormalization-group using adaptive effective Hilbert spaces," J. Stat. Mech: Theory Exp. **2004**, P04005 (2004).

[67] S. R. White and A. E. Feiguin, "Real-time evolution using the density matrix renormalization group," Phys. Rev. Lett. **93**, 076401 (2004).

[68] U. Schollwöck, "The density-matrix renormalization group," Rev. Mod. Phys. **77**, 259 (2005).

[69] B. Tang, E. Khatami, and M. Rigol, "A short introduction to numerical linked-cluster expansions," Comput. Phys. Commun. **184**, 557 (2013).

[70] K. Mallayya and M. Rigol, "Quantum quenches and relaxation dynamics in the thermodynamic limit," Phys. Rev. Lett. **120**, 070603 (2018).

[71] J. Richter and R. Steinigeweg, "Combining dynamical quantum typicality and numerical linked cluster expansions," Phys. Rev. B **99**, 094419 (2019).

[72] H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner, "Nonequilibrium dynamical mean-field theory and its applications," Rev. Mod. Phys. **86**, 779 (2014).

[73] C. Bartsch and J. Gemmer, "Dynamical typicality of quantum expectation values," Phys. Rev. Lett. **102**, 110403 (2009).

[74] T. A. Elsayed and B. V. Fine, "Regression relation for pure quantum states and its implications for efficient computing," Phys. Rev. Lett. **110**, 070404 (2013).

[75] R. Steinigeweg, J. Gemmer, and W. Brenig, "Spin-current autocorrelations from single pure-state propagation," Phys. Rev. Lett. **112**, 120601 (2014).

[76] P. Bocchieri and A. Loinger, "Quantum recurrence theorem," Phys. Rev. **107**, 337 (1957).

[77] L. S. Schulman, "Note on the quantum recurrence theorem," Phys. Rev. A **18**, 2379 (1978).

[78] N. Linden, S. Popescu, A. J. Short, and A. Winter, "On the speed of fluctuations around thermodynamic equilibrium," New J. Phys. **12**, 055021 (2010).

[79] A. J. Short, "Equilibration of quantum systems and subsystems," New J. Phys. **13**, 053009 (2011).

[80] B. N. Balz and P. Reimann, "Equilibration of isolated many-body quantum systems with respect to general distinguishability measures," Phys. Rev. E **93**, 062107 (2016).

[81] M. Cramer and J. Eisert, "A quantum central limit theorem for non-equilibrium systems: exact local relaxation of correlated states," New J. Phys. **12**, 055020 (2010).

[82] J. Sirker, N. P. Konstantinidis, F. Andraschko, and N. Sedlmayr, "Locality and thermalization in closed quantum systems," Phys. Rev. A **89**, 042104 (2014).

[83] F. H. L. Essler and M. Fagotti, "Quench dynamics and relaxation in isolated integrable quantum spin chains," J. Stat. Mech: Theory Exp. **2016**, 064002 (2016).

[84] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, "Long-time behavior of macroscopic quantum systems," Eur. Phys. J. H **35**, 173 (2010).

[85] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, "Normal typicality and von Neumann's quantum ergodic theorem," Proc. Roy. Soc. A **466**, 3203 (2010).

[86] S. Goldstein, T. Hara, and H. Tasaki, "Time scales in the approach to equilibrium of macroscopic quantum systems," Phys. Rev. Lett. **111**, 140401 (2013).

[87] H. Tasaki, "Typicality of thermal equilibrium and thermalization in isolated macroscopic quantum systems," J. Stat. Phys. **163**, 937 (2016).

[88] P. Reimann, "Typicality of pure states randomly sampled according to the Gaussian adjusted projected measure," J. Stat. Phys. **132**, 921 (2008).

[89] H. Tasaki, "From quantum dynamics to the canonical distribution: General picture and a rigorous example," Phys. Rev. Lett. **80**, 1373 (1998).

[90] R. Gallego, H. Wilming, J. Eisert, and C. Gogolin, "What it takes to avoid equilibration," Phys. Rev. A **98**, 022135 (2018).

[91] A. S. L. Malabarba, L. P. García-Pintos, N. Linden, T. C. Farrelly, and A. J. Short, "Quantum systems equilibrate rapidly for most observables," Phys. Rev. E **90**, 012121 (2014).

[92] M. Žnidarič, C. Pineda, and I. García-Mata, "Non-Markovian behavior of small and large complex quantum systems," Phys. Rev. Lett. **107**, 080404 (2011).

[93] F. G. S. L. Brandão, P. Ćwikliński, M. Horodecki, P. Horodecki, J. K. Korbicz, and M. Mozrzymas, "Convergence to equilibrium under a random Hamiltonian," Phys. Rev. E **86**, 031101 (2012).

[94] M. Cramer, "Thermalization under randomized local Hamiltonians," New J. Phys. **14**, 053051 (2012).

[95] L. Masanes, A. J. Roncaglia, and A. Acín, "Complexity of energy eigenstates as a mechanism for equilibration," Phys. Rev. E **87**, 032137 (2013).

[96] P. Reimann, "Typical fast thermalization processes in closed many-body systems," Nat. Commun. **7**, 10821 (2016).

[97] B. N. Balz and P. Reimann, "Typical relaxation of isolated many-body systems which do not thermalize," Phys. Rev. Lett. **118**, 190601 (2017).

[98] M. Schiulaz, E. J. Torres-Herrera, and L. F. Santos, "Thouless and relaxation time scales in many-body quantum systems," Phys. Rev. B **99**, 174313 (2019).

[99] S. Goldstein, T. Hara, and H. Tasaki, "Extremely quick thermalization in a macroscopic quantum system for a typical nonequilibrium subspace," New J. Phys. **17**, 045002 (2015).

[100] P. Reimann, "Transportless equilibration in isolated many-body quantum systems," New J. Phys. **21**, 053014 (2019).

[101] T. Farrelly, "Equilibration of quantum gases," New J. Phys. **18**, 073014 (2016).

[102] T. R. de Oliveira, C. Charalambous, D. Jonathan, M. Lewenstein, and A. Riera, "Equilibration time scales in closed many-body quantum systems," New J. Phys. **20**, 033032 (2018).

[103] L. P. García-Pintos, N. Linden, A. S. L. Malabarba, A. J. Short, and A. Winter, "Equilibration time scales of physically relevant observables," Phys. Rev. X **7**, 031027 (2017).

[104] R. Heveling, L. Knipschild, and J. Gemmer, "Comment on 'Equilibration time scales of physically relevant observables'," Phys. Rev. X **10**, 028001 (2020).

[105] E. H. Lieb and D. W. Robinson, "The finite group velocity of quantum spin systems," Comm. Math. Phys. **28**, 251 (1972).

[106] D. Nickelsen and M. Kastner, "Classical Lieb-Robinson bound for estimating equilibration

timescales of isolated quantum systems," Phys. Rev. Lett. **122**, 180602 (2019).

[107] D. Nickelsen and M. Kastner, "Modelling equilibration of local many-body quantum systems by random graph ensembles," Quantum **4**, 273 (2020).

[108] P. Gaspard, *Chaos, Scattering and Statistical Mechanics* (Cambridge University Press, Cambridge, UK, 1998).

[109] Y. G. Sinai, "Dynamical systems with elastic reflections," Russ. Math. Surv. **25**, 137 (1970).

[110] L. A. Bunimovich, "On the ergodic properties of nowhere dispersing billiards," Commun. Math. Phys. **65**, 295 (1979).

[111] N. Simányi, "Proof of the ergodic hypothesis for typical hard ball systems," Ann. Henri Poincaré **5**, 203 (2004).

[112] J. Bricmont, "Science of chaos or chaos in science?" Ann. N.Y. Acad. Sci. **775**, 131 (1995).

[113] P. Reimann and M. Evstigneev, "Quantum versus classical foundation of statistical mechanics under experimentally realistic conditions," Phys. Rev. E **88**, 052114 (2013).

[114] J. M. Deutsch, "Quantum statistical mechanics in a closed system," Phys. Rev. A **43**, 2046 (1991).

[115] M. Srednicki, "Chaos and quantum thermalization," Phys. Rev. E **50**, 888 (1994).

[116] M. Rigol, V. Dunjko, and M. Olshanii, "Thermalization and its mechanism for generic isolated quantum systems," Nature **452**, 854 (2008).

[117] J. v. Neumann, "Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik," Z. Phys. **57**, 30 (1929), english translation by R. Tumulka: "Proof of the ergodic theorem and the H-theorem in quantum mechanics," Eur. Phys. J. H **35**, 201 (2010).

[118] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, "Canonical typicality," Phys. Rev. Lett. **96**, 050403 (2006).

[119] S. Popescu, A. J. Short, and A. Winter, "Entanglement and the foundations of statistical mechanics," Nat. Phys. **2**, 754 (2006).

[120] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghi, "Approach to thermal equilibrium of macroscopic quantum systems," Phys. Rev. E **81**, 011109 (2010).

[121] S. Goldstein, D. A. Huse, J. L. Lebowitz, and R. Tumulka, "Thermal equilibrium of a macroscopic quantum system in a pure state," Phys. Rev. Lett. **115**, 100402 (2015).

[122] M. D. Choi, "Almost commuting matrices need not be nearly commuting," Proc. Amer. Math. Soc. **102**, 529 (1988).

[123] W. De Roeck, C. Maes, and K. Netočný, "Quantum macrostates, equivalence of ensembles, and an H-theorem," J. Math. Phys. **47**, 073303 (2006).

[124] P. Reimann, "Generalization of von Neumann's approach to thermalization," Phys. Rev. Lett. **115**, 010403 (2015).

[125] R. V. Jensen and R. Shankar, "Statistical behavior in deterministic quantum systems with few degrees of freedom," Phys. Rev. Lett. **54**, 1879 (1985).

[126] M. V. Berry, "Regular and irregular semiclassical wavefunctions," J. Phys. A: Math. Gen. **10**, 2083 (1977).

[127] M. Srednicki, "The approach to thermal equilibrium in quantized chaotic systems," J. Phys. A: Math. Gen. **32**, 1163 (1999).

[128] M. Rigol, "Breakdown of thermalization in finite one-dimensional systems," Phys. Rev. Lett. **103**, 100403 (2009).

[129] G. Biroli, C. Kollath, and A. M. Läuchli, "Effect of rare fluctuations on the thermalization of isolated quantum systems," Phys. Rev. Lett. **105**, 250401 (2010).

[130] S. Genway, A. F. Ho, and D. K. K. Lee, "Thermalization of local observables in small Hubbard lattices," Phys. Rev. A **86**, 023609 (2012).

[131] R. Steinigeweg, J. Herbrych, and P. Prelovšek, "Eigenstate thermalization within isolated spin-chain systems," Phys. Rev. E **87**, 012118 (2013).

[132] E. Khatami, G. Pupillo, M. Srednicki, and M. Rigol, "Fluctuation-dissipation theorem in an isolated system of quantum dipolar bosons after a quench," Phys. Rev. Lett. **111**, 050403 (2013).

[133] R. Steinigeweg, A. Khodja, H. Niemeyer, C. Gogolin, and J. Gemmer, "Pushing the limits of

the eigenstate thermalization hypothesis towards mesoscopic quantum systems," Phys. Rev. Lett. **112**, 130403 (2014).

[134] W. Beugeling, R. Moessner, and M. Haque, "Off-diagonal matrix elements of local operators in many-body quantum systems," Phys. Rev. E **91**, 012144 (2015).

[135] R. Mondaini, K. R. Fratus, M. Srednicki, and M. Rigol, "Eigenstate thermalization in the two-dimensional transverse field Ising model," Phys. Rev. E **93**, 032104 (2016).

[136] R. Mondaini and M. Rigol, "Eigenstate thermalization in the two-dimensional transverse field Ising model. II. Off-diagonal matrix elements of observables," Phys. Rev. E **96**, 012157 (2017).

[137] L. F. Santos and M. Rigol, "Localization and the effects of symmetries in the thermalization properties of one-dimensional quantum systems," Phys. Rev. E **82**, 031130 (2010).

[138] H. Kim, T. N. Ikeda, and D. A. Huse, "Testing whether all eigenstates obey the eigenstate thermalization hypothesis," Phys. Rev. E **90**, 052105 (2014).

[139] R. Mondaini, K. Mallayya, L. F. Santos, and M. Rigol, "Comment on 'Systematic construction of counterexamples to the eigenstate thermalization hypothesis'," Phys. Rev. Lett. **121**, 038901 (2018).

[140] N. Shiraishi and T. Mori, "Systematic construction of counterexamples to the eigenstate thermalization hypothesis," Phys. Rev. Lett. **119**, 030601 (2017).

[141] N. Shiraishi and T. Mori, "Shiraishi and Mori reply," Phys. Rev. Lett. **121**, 038902 (2018).

[142] A. C. Cassidy, C. W. Clark, and M. Rigol, "Generalized thermalization in an integrable lattice system," Phys. Rev. Lett. **106**, 140405 (2011).

[143] J.-S. Caux and F. H. L. Essler, "Time evolution of local observables after quenching to an integrable model," Phys. Rev. Lett. **110**, 257203 (2013).

[144] J.-S. Caux, "The quench action," J. Stat. Mech: Theory Exp. **2016**, 064006 (2016).

[145] T. N. Ikeda, Y. Watanabe, and M. Ueda, "Finite-size scaling analysis of the eigenstate thermalization hypothesis in a one-dimensional interacting Bose gas," Phys. Rev. E **87**, 012125 (2013).

[146] V. Alba, "Eigenstate thermalization hypothesis and integrability in quantum spin chains," Phys. Rev. B **91**, 155123 (2015).

[147] T. Mori, "Weak eigenstate thermalization with large deviation bound," arXiv:1609.09776 (2016).

[148] E. Iyoda, K. Kaneko, and T. Sagawa, "Fluctuation theorem for many-body pure quantum states," Phys. Rev. Lett. **119**, 100601 (2017).

[149] T. Kuwahara and K. Saito, "Eigenstate thermalization from the clustering property of correlation," Phys. Rev. Lett. **124**, 200604 (2020).

[150] R. Nandkishore and D. A. Huse, "Many-body localization and thermalization in quantum statistical mechanics," Annu. Rev. Cond. Mat. Phys. **6**, 15 (2015).

[151] D. A. Abanin and Z. Papić, "Recent progress in many-body localization," Ann. Phys. **529**, 1700169 (2017).

[152] P. W. Anderson, "Absence of diffusion in certain random lattices," Phys. Rev. **109**, 1492 (1958).

[153] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, "Weak ergodicity breaking from quantum many-body scars," Nat. Phys. **14**, 745 (2018).

[154] Z. Lan, M. van Horssen, S. Powell, and J. P. Garrahan, "Quantum slow relaxation and metastability due to dynamical constraints," Phys. Rev. Lett. **121**, 040603 (2018).

[155] S. Choi, C. J. Turner, H. Pichler, W. W. Ho, A. A. Michailidis, Z. Papić, M. Serbyn, M. D. Lukin, and D. A. Abanin, "Emergent SU(2) dynamics and perfect quantum many-body scars," Phys. Rev. Lett. **122**, 220603 (2019).

[156] W. W. Ho, S. Choi, H. Pichler, and M. D. Lukin, "Periodic orbits, entanglement, and quantum many-body scars in constrained models: Matrix product state approach," Phys. Rev. Lett. **122**, 040603 (2019).

[157] J.-S. Caux and J. Mossel, "Remarks on the notion of quantum integrability," J. Stat. Mech: Theory Exp. **2011**, P02023 (2011).

[158] P. Gaspard, "Dynamical chaos and nonequilibrium statistical mechanics," Int. J. Mod. Phys. B **15**, 209 (2001).

[159] E. Ott, *Chaos in Dynamical Systems*, 2nd ed. (Cambridge University Press, Cambridge, UK, 2002).

[160] G. Orwell, *Animal farm: a fairy story* (Signet Classics, New York, 1996).

[161] F. Franchini, *An Introduction to Integrable Techniques for One-Dimensional Quantum Systems* (Springer, Cham, Switzerland, 2017).

[162] E. Ilievski, M. Medenjak, T. Prosen, and L. Zadnik, "Quasilocal charges in integrable lattice systems," J. Stat. Mech: Theory Exp. **2016**, 064008 (2016).

[163] F. Haake, *Quantum Signatures of Chaos*, 3rd ed. (Springer, Berlin, Heidelberg, 2010).

[164] L. Vidmar and M. Rigol, "Generalized Gibbs ensemble in integrable lattice models," J. Stat. Mech: Theory Exp. **2016**, 064007 (2016).

[165] M. Rigol, V. Dunjko, V. Yurovsky, and M. Olshanii, "Relaxation in a completely integrable many-body quantum system: An ab initio study of the dynamics of the highly excited states of 1d lattice hard-core bosons," Phys. Rev. Lett. **98**, 050405 (2007).

[166] B. Pozsgay, E. Vernier, and M. A. Werner, "On generalized Gibbs ensembles with an infinite set of conserved charges," J. Stat. Mech: Theory Exp. **2017**, 093103 (2017).

[167] J.-S. Caux and R. M. Konik, "Constructing the generalized Gibbs ensemble after a quantum quench," Phys. Rev. Lett. **109**, 175301 (2012).

[168] M. Kollar and M. Eckstein, "Relaxation of a one-dimensional Mott insulator after an interaction quench," Phys. Rev. A **78**, 013626 (2008).

[169] A. Iucci and M. A. Cazalilla, "Quantum quench dynamics of the Luttinger model," Phys. Rev. A **80**, 063619 (2009).

[170] P. Calabrese, F. H. L. Essler, and M. Fagotti, "Quantum quench in the transverse-field Ising chain," Phys. Rev. Lett. **106**, 227203 (2011).

[171] E. Ilievski, J. De Nardis, B. Wouters, J.-S. Caux, F. H. L. Essler, and T. Prosen, "Complete generalized Gibbs ensembles in an interacting theory," Phys. Rev. Lett. **115**, 157201 (2015).

[172] T. Langen, S. Erne, R. Geiger, B. Rauer, T. Schweigler, M. Kuhnert, W. Rohringer, I. E. Mazets, T. Gasenzer, and J. Schmiedmayer, "Experimental observation of a generalized Gibbs ensemble," Science **348**, 207 (2015).

[173] P. Reimann, "Dynamical typicality approach to eigenstate thermalization," Phys. Rev. Lett. **120**, 230601 (2018).

[174] A. Mitra, "Quantum quench dynamics," Annu. Rev. Cond. Mat. Phys. **9**, 245 (2018).

[175] J. Berges, S. Borsányi, and C. Wetterich, "Prethermalization," Phys. Rev. Lett. **93**, 142002 (2004).

[176] M. Moeckel and S. Kehrein, "Interaction quench in the Hubbard model," Phys. Rev. Lett. **100**, 175702 (2008).

[177] P. Reimann and L. Dabelow, "Typicality of prethermalization," Phys. Rev. Lett. **122**, 080603 (2019).

[178] K. Mallayya, M. Rigol, and W. De Roeck, "Prethermalization and thermalization in isolated quantum systems," Phys. Rev. X **9**, 021027 (2019).

[179] J. L. Lebowitz, "Boltzmann's entropy and time's arrow," Phys. Today **46**, 32 (1993).

[180] S. Goldstein, "Typicality and notions of probability in physics," in *Probability in Physics*, edited by Y. Ben-Menahem and M. Hemmo (Springer, Berlin, Heidelberg, 2012) pp. 59–71.

[181] B. N. Balz, J. Richter, J. Gemmer, R. Steinigeweg, and P. Reimann, "Dynamical typicality for initial states with a preset measurement statistics of several commuting observables," in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*, edited by F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso (Springer, Cham, Switzerland, 2018) pp. 413–433.

[182] P. Reimann and J. Gemmer, "Why are macroscopic experiments reproducible? Imitating the behavior of an ensemble by single pure states," Physica A **552**, 121840 (2020).

[183] M. Talagrand, "A New Look at Independence," Ann. Prob. **24**, 1 (1996).

[184] A. Giannopoulos and V. Milman, "Concentration property on probability spaces," Adv.

Math. **156**, 77 (2000).

[185] M. Ledoux, *The concentration of measure phenomenon* (American Mathematical Society, Providence, RI, 2001).

[186] H. Touchette, "The large deviation approach to statistical mechanics," Phys. Rep. **478**, 1 (2009).

[187] M. Hutchinson, "A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines," J. Commun. Stat. Simul. **19**, 433 (1990).

[188] J. Schnack, J. Richter, T. Heitmann, J. Richter, and R. Steinigeweg, "Finite-size scaling of typicality-based estimates," Z. Naturforsch. A **75**, 465 (2020).

[189] J. Schnack, J. Richter, and R. Steinigeweg, "Accuracy of the finite-temperature Lanczos method compared to simple typicality-based estimates," Phys. Rev. Research **2**, 013186 (2020).

[190] J. Jaklič and P. Prelovšek, "Lanczos method for the calculation of finite-temperature quantities in correlated systems," Phys. Rev. B **49**, 5065 (1994).

[191] J. Skilling, "The eigenvalues of mega-dimensional matrices," in *Maximum Entropy and Bayesian Methods: Cambridge, England, 1988*, edited by J. Skilling (Springer, Dordrecht, Netherlands, 1989) pp. 455–466.

[192] A. Hams and H. De Raedt, "Fast algorithm for finding the eigenvalue distribution of very large matrices," Phys. Rev. E **62**, 4365 (2000).

[193] S. Sugiura and A. Shimizu, "Thermal pure quantum states at finite temperature," Phys. Rev. Lett. **108**, 240401 (2012).

[194] S. Lloyd, *Black Holes, Demons and the Loss of Coherence: How complex systems get information and what they do with it*, Ph.D. thesis, Rockefeller University (1988).

[195] J. Gemmer and G. Mahler, "Distribution of local entropy in the Hilbert space of bi-partite quantum systems: origin of Jaynes' principle," Eur. Phys. J. B **31**, 249 (2003).

[196] E. Schrödinger, "Energieaustausch nach der Wellenmechanik," Ann. Phys. **388**, 956 (1927).

[197] P. W. Brouwer and C. W. J. Beenakker, "Diagrammatic method of integration over the unitary group, with applications to quantum transport in mesoscopic systems," J. Math. Phys. **37**, 4904 (1996).

[198] P. Reimann, "Typicality for generalized microcanonical ensembles," Phys. Rev. Lett. **99**, 160404 (2007).

[199] S. Sugiura and A. Shimizu, "Canonical thermal pure quantum state," Phys. Rev. Lett. **111**, 010401 (2013).

[200] P. Reimann, "Dynamical typicality of isolated many-body quantum systems," Phys. Rev. E **97**, 062129 (2018).

[201] C. Presilla and U. Tambini, "Selective relaxation method for numerical solution of Schrödinger problems," Phys. Rev. E **52**, 4495 (1995).

[202] S. Garnerone and T. R. de Oliveira, "Generalized quantum microcanonical ensemble from random matrix product states," Phys. Rev. B **87**, 214426 (2013).

[203] B. V. Fine, "Typical state of an isolated quantum system with fixed energy and unrestricted participation of eigenstates," Phys. Rev. E **80**, 051130 (2009).

[204] M. P. Müller, D. Gross, and J. Eisert, "Concentration of measure for quantum states with a fixed expectation value," Commun. Math. Phys. **303**, 785 (2011).

[205] G. Ithier and F. Benaych-Georges, "Dynamical typicality of embedded quantum systems," Phys. Rev. A **96**, 012108 (2017).

[206] S. Genway, A. F. Ho, and D. K. K. Lee, "Dynamics of thermalization and decoherence of a nanoscale system," Phys. Rev. Lett. **111**, 130408 (2013).

[207] G. Ithier, S. Ascroft, and F. Benaych-Georges, "Typical equilibrium state of an embedded quantum system," Phys. Rev. E **96**, 060102 (2017).

[208] C. Nation and D. Porras, "Quantum chaotic fluctuation-dissipation theorem: Effective Brownian motion in closed quantum systems," Phys. Rev. E **99**, 052139 (2019).

[209] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong, "Random-matrix physics: spectrum and strength fluctuations," Rev. Mod. Phys. **53**, 385 (1981).

[210] T. Guhr, A. Müller–Groeling, and H. A. Weidenmüller, "Random-matrix theories in quantum physics: common concepts," Phys. Rep. **299**, 189 (1998).

[211] M. L. Mehta, *Random matrices* (Academic Press, Amsterdam, San Diego, CA, 2004).

[212] E. P. Wigner, "On the statistical distribution of the widths and spacings of nuclear resonance levels," Math. Proc. Cambridge Philos. Soc. **47**, 790 (1951).

[213] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," Ann. Math. **62**, 548 (1955).

[214] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions II," Ann. Math. **65**, 203 (1957).

[215] E. P. Wigner, "On the distribution of the roots of certain symmetric matrices," Ann. Math. **67**, 325 (1958).

[216] E. P. Wigner, "Random matrices in physics," SIAM Rev. **9**, 1 (1967).

[217] A. D. Mirlin and Y. V. Fyodorov, "Universality of level correlation function of sparse random matrices," J. Phys. A: Math. Gen. **24**, 2273 (1991).

[218] Y. V. Fyodorov and A. D. Mirlin, "Statistical properties of random banded matrices with strongly fluctuating diagonal elements," Phys. Rev. B **52**, R11580 (1995).

[219] Y. V. Fyodorov, O. A. Chubykalo, F. M. Izrailev, and G. Casati, "Wigner random banded matrices with sparse structure: Local spectral density of states," Phys. Rev. Lett. **76**, 1603 (1996).

[220] K. Efetov, "Supersymmetry and theory of disordered metals," Adv. Phys. **32**, 53 (1983).

[221] J. Verbaarschot, H. Weidenmüller, and M. Zirnbauer, "Grassmann integration in stochastic quantum physics: The case of compound-nucleus scattering," Phys. Rep. **129**, 367 (1985).

[222] J. A. Zuk, "Introduction to the supersymmetry method for the Gaussian random matrix ensembles," arXiv:cond-mat/9412060 (1996).

[223] A. D. Mirlin, "Statistics of energy levels and eigenfunctions in disordered and chaotic systems: Supersymmetry approach," arXiv:cond-mat/0006421 (2000).

[224] F. A. Berezin, *Introduction to Superanalysis* (Springer, Dordrecht, Netherlands, 2010).

[225] T. Guhr, "Supersymmetry in random matrix theory," arXiv:1005.0979 (2010).

[226] L. Dabelow and P. Reimann, "Relaxation theory for perturbed many-body quantum systems versus numerics and experiment," Phys. Rev. Lett. **124**, 120602 (2020).

[227] L. Dabelow, P. Vorndamme, and P. Reimann, "Modification of quantum many-body relaxation by perturbations exhibiting a banded matrix structure," Phys. Rev. Research **2**, 033210 (2020).

[228] P. Pechukas, "Distribution of energy eigenvalues in the irregular spectrum," Phys. Rev. Lett. **51**, 943 (1983).

[229] T. Yukawa, "New approach to the statistical properties of energy levels," Phys. Rev. Lett. **54**, 1883 (1985).

[230] M. Feingold, D. M. Leitner, and O. Piro, "Semiclassical structure of Hamiltonians," Phys. Rev. A **39**, 6507 (1989).

[231] I. Arad, T. Kuwahara, and Z. Landau, "Connecting global and local energy distributions in quantum spin models on a lattice," J. Stat. Mech: Theory Exp. **2016**, 033301 (2016).

[232] N. P. Konstantinidis, "Thermalization away from integrability and the role of operator off-diagonal elements," Phys. Rev. E **91**, 052111 (2015).

[233] F. Borgonovi, F. Izrailev, L. Santos, and V. Zelevinsky, "Quantum chaos and thermalization in isolated systems of interacting particles," Phys. Rep. **626**, 1 (2016).

[234] D. Jansen, J. Stolpp, L. Vidmar, and F. Heidrich-Meisner, "Eigenstate thermalization and quantum chaos in the Holstein polaron model," Phys. Rev. B **99**, 155130 (2019).

[235] V. V. Flambaum and F. M. Izrailev, "Statistical theory of finite Fermi systems based on the structure of chaotic eigenstates," Phys. Rev. E **56**, 5144 (1997).

[236] G. Guennebaud, B. Jacob, *et al.*, "Eigen v3.3.4," `http://eigen.tuxfamily.org` (2017), access: 21/02/2017.

[237] R. Hamazaki and M. Ueda, "Atypicality of most few-body observables," Phys. Rev. Lett. **120**, 080603 (2018).

[238] R. Heveling, L. Knipschild, and J. Gemmer, "Modeling the impact of Hamiltonian perturbations on expectation value dynamics," Z. Naturforsch. A **75**, 475 (2020).

[239] L. Isserlis, "On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables," Biometrika **12**, 134 (1918).

[240] G. C. Wick, "The evaluation of the collision matrix," Phys. Rev. **80**, 268 (1950).

[241] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1998).

[242] Z. Pluhař, H. Weidenmüller, J. Zuk, C. Lewenkopf, and F. Wegner, "Crossover from orthogonal to unitary symmetry for ballistic electron transport in chaotic microstructures," Ann. Phys. **243**, 1 (1995).

[243] R. L. Stratonovich, "On a method of calculating quantum distribution functions," Sov. Phys. Doklady **2**, 416 (1957).

[244] J. Hubbard, "Calculation of partition functions," Phys. Rev. Lett. **3**, 77 (1959).

[245] E. T. Copson, *Asymptotic Expansions*, Cambridge Tracts in Mathematics (Cambridge University Press, Cambridge, UK, 1965).

[246] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, 1st ed. (Springer, New York, 1999).

[247] Y. V. Fyodorov, private communication.

[248] L. Dabelow and P. Reimann, "Typical relaxation of perturbed many-body quantum systems," unpublished.

[249] J. P. Boyd, *Chebyshev & Fourier Spectral Methods* (Dover, Mineola, NY, 1989).

[250] B. Fornberg, *A Practical Guide to Pseudospectral Methods* (Cambridge University Press, Cambridge, UK, 1996).

[251] E. Süli and D. F. Mayers, *An Introduction to Numerical Analysis* (Cambridge University Press, Cambridge, UK, 2003).

[252] C. Nation and D. Porras, "Off-diagonal observable elements from random matrix theory: distributions, fluctuations, and eigenstate thermalization," New J. Phys. **20**, 103003 (2018).

[253] J. M. Deutsch, "A closed quantum system giving ergodicity," unpublished, `https://deutsch.physics.ucsc.edu/pdf/quantumstat.pdf`.

[254] G. Ithier and S. Ascroft, "Statistical diagonalization of a random biased Hamiltonian: the case of the eigenvectors," J. Phys. A: Math. Theor. **51**, 48LT01 (2018).

[255] P.-O. Löwdin, "On the nonorthogonality problem," in *Advances in Quantum Chemistry*, Vol. 5, edited by P.-O. Löwdin (Academic Press, 1970) pp. 185 – 199.

[256] L. Piela, *Ideas of quantum chemistry* (Elsevier, Amsterdam, 2014).

[257] P. Reimann, "Eigenstate thermalization: Deutsch's approach and beyond," New J. Phys. **17**, 055025 (2015).

[258] T. Yoshizawa, E. Iyoda, and T. Sagawa, "Numerical large deviation analysis of the eigenstate thermalization hypothesis," Phys. Rev. Lett. **120**, 200604 (2018).

[259] L. Knipschild and J. Gemmer, "Stability of quantum dynamics under constant Hamiltonian perturbations," Phys. Rev. E **98**, 062103 (2018).

[260] L. Dabelow and P. Reimann, "Predicting imperfect echo dynamics in many-body quantum systems," Z. Naturforsch. A , 20190383 (2020).

[261] A. Flesch, M. Cramer, I. P. McCulloch, U. Schollwöck, and J. Eisert, "Probing local relaxation of cold atoms in optical superlattices," Phys. Rev. A **78**, 033608 (2008).

[262] D. Giuliano, D. Rossini, P. Sodano, and A. Trombettoni, "XXZ spin-$\frac{1}{2}$ representation of a finite-U Bose-Hubbard chain at half-integer filling," Phys. Rev. B **87**, 035104 (2013).

[263] P. Barmettler, M. Punk, V. Gritsev, E. Demler, and E. Altman, "Relaxation of antiferromagnetic order in spin-1/2 chains following a quantum quench," Phys. Rev. Lett. **102**, 130603 (2009).

[264] K. Balzer, F. A. Wolf, I. P. McCulloch, P. Werner, and M. Eckstein, "Nonthermal melting of Néel order in the Hubbard model," Phys. Rev. X **5**, 031039 (2015).

[265] L. van Hove, "The approach to equilibrium in quantum statistics: A perturbation treatment to general order," Physica **23**, 441 (1957).

[266] J. Richter, F. Jin, L. Knipschild, H. De Raedt, K. Michielsen, J. Gemmer, and R. Steinigeweg, "Exponential damping induced by random and realistic perturbations," Phys. Rev. E **101**, 062133 (2020).

[267] F. J. Dyson, "A Brownian-motion model for the eigenvalues of a random matrix," J. Math. Phys. **3**, 1191 (1962).

[268] V. K. B. Kota, A. Relaño, J. Retamosa, and M. Vyas, "Thermalization in the two-body random ensemble," J. Stat. Mech: Theory Exp. **2011**, P10028 (2011).

[269] S. K. Haldar, N. D. Chavda, M. Vyas, and V. K. B. Kota, "Fidelity decay and entropy production in many-particle systems after random interaction quench," J. Stat. Mech: Theory Exp. **2016**, 043101 (2016).

[270] J. M. Magán, "Random free fermions: An analytical example of eigenstate thermalization," Phys. Rev. Lett. **116**, 030401 (2016).

[271] S. Sugimoto, R. Hamazaki, and M. Ueda, "Test of eigenstate thermalization hypothesis based on local random matrix theory," arXiv:2005.06379 (2020).

[272] O. Bohigas and J. Flores, "Two-body random Hamiltonian and level density," Phys. Lett. B **34**, 261 (1971).

[273] J. French and S. Wong, "Validity of random matrix theories for many-particle systems," Phys. Lett. B **33**, 449 (1970).

[274] V. V. Flambaum and F. M. Izrailev, "Excited eigenstates and strength functions for isolated systems of interacting particles," Phys. Rev. E **61**, 2539 (2000).

[275] V. V. Flambaum and F. M. Izrailev, "Entropy production and wave packet dynamics in the Fock space of closed chaotic many-body systems," Phys. Rev. E **64**, 036220 (2001).

[276] M. Vyas and T. H. Seligman, "Random matrix ensembles for many-body quantum systems," AIP Conf. Proc. **1950**, 030009 (2018).

[277] V. K. B. Kota and N. D. Chavda, "Random k-body ensembles for chaos and thermalization in isolated systems," Entropy **20** (2018), 10.3390/e20070541.

[278] M. Moeckel and S. Kehrein, "Crossover from adiabatic to sudden interaction quenches in the Hubbard model: prethermalization and non-equilibrium dynamics," New J. Phys. **12**, 055016 (2010).

[279] R. Sachdeva, T. Nag, A. Agarwal, and A. Dutta, "Finite-time interaction quench in a Luttinger liquid," Phys. Rev. B **90**, 045421 (2014).

[280] T. Puškarov and D. Schuricht, "Time evolution during and after finite-time quantum quenches in the transverse-field Ising chain," SciPost Phys. **1**, 003 (2016).

[281] B. Schoenauer and D. Schuricht, "Finite-time quantum quenches in the XXZ Heisenberg chain," Phys. Rev. B **100**, 115418 (2019).

[282] M. Bukov, L. D'Alessio, and A. Polkovnikov, "Universal high-frequency behavior of periodically driven systems: from dynamical stabilization to Floquet engineering," Adv. Phys. **64**, 139 (2015).

[283] A. Eckardt, "Colloquium: Atomic quantum gases in periodically driven optical lattices," Rev. Mod. Phys. **89**, 011004 (2017).

[284] T. Gorin, T. Prosen, T. H. Seligman, and M. Žnidarič, "Dynamics of Loschmidt echoes and fidelity decay," Phys. Rep. **435**, 33 (2006).

[285] M. Schmitt and S. Kehrein, "Irreversible dynamics in quantum many-body systems," Phys. Rev. B **98**, 180301 (2018).

[286] L. Dabelow and P. Reimann, "Persistent many-body quantum echoes," Phys. Rev. Research **2**, 023216 (2020).

[287] E. L. Hahn, "Spin echoes," Phys. Rev. **80**, 580 (1950).

[288] J. A. B. Marinus T. Vlaardingerbroek, *Magnetic Resonance Imaging* (Springer, Berlin, Heidelberg, 2010).

[289] M. Gärttner, J. G. Bohnet, A. Safavi-Naini, M. L. Wall, J. J. Bollinger, and A. M. Rey, "Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet," Nat. Phys. **13**, 781 (2017).

[290] A. Widera, S. Trotzky, P. Cheinet, S. Fölling, F. Gerbier, I. Bloch, V. Gritsev, M. D.

Lukin, and E. Demler, "Quantum spin dynamics of mode-squeezed Luttinger liquids in two-component atomic gases," Phys. Rev. Lett. **100**, 140401 (2008).

[291] F. M. Cucchietti, "Time reversal in an optical lattice," J. Opt. Soc. Am. B **27**, A30 (2010).

[292] C. Weiss, "Effective time-reversal via periodic shaking," J. Phys. Conf. Ser. **414**, 012032 (2013).

[293] A. S. de Wijn, B. Hess, and B. V. Fine, "Largest Lyapunov exponents for lattices of interacting classical spins," Phys. Rev. Lett. **109**, 034101 (2012).

[294] A. S. de Wijn, B. Hess, and B. V. Fine, "Lyapunov instabilities in lattices of interacting classical spins at infinite temperature," J. Phys. A: Math. Theor. **46**, 254012 (2013).

[295] B. V. Fine, T. A. Elsayed, C. M. Kropf, and A. S. de Wijn, "Absence of exponential sensitivity to small perturbations in nonintegrable systems of spins 1/2," Phys. Rev. E **89**, 012923 (2014).

[296] T. A. Elsayed and B. V. Fine, "Sensitivity to small perturbations in systems of large quantum spins," Phys. Scr. **T165**, 014011 (2015).

[297] L. Erdős and D. Schröder, "Phase transition in the density of states of quantum spin glasses," Math. Phys. Anal. Geom. **17**, 441 (2014).

[298] M. Schmitt, D. Sels, S. Kehrein, and A. Polkovnikov, "Semiclassical echo dynamics in the Sachdev-Ye-Kitaev model," Phys. Rev. B **99**, 134301 (2019).

[299] A. Peres, "Stability of quantum motion in chaotic and regular systems," Phys. Rev. A **30**, 1610 (1984).

[300] B. Dóra, F. Pollmann, J. Fortágh, and G. Zaránd, "Loschmidt echo and the many-body orthogonality catastrophe in a qubit-coupled Luttinger liquid," Phys. Rev. Lett. **111**, 046402 (2013).

[301] M. Schmitt and S. Kehrein, "Effective time reversal and echo dynamics in the transverse field Ising model," EPL (Europhys. Lett.) **115**, 50001 (2016).

[302] J. Maldacena, S. H. Shenker, and D. Stanford, "A bound on chaos," J. High Energy Phys. **2016**, 106 (2016).

[303] B. Swingle, "Unscrambling the physics of out-of-time-order correlators," Nat. Phys. **14**, 988 (2018).

[304] Q. Hummel, B. Geiger, J. D. Urbina, and K. Richter, "Reversible quantum information spreading in many-body systems near criticality," Phys. Rev. Lett. **123**, 160401 (2019).

[305] S. Pilatowsky-Cameo, J. Chávez-Carlos, M. A. Bastarrachea-Magnani, P. Stránský, S. Lerma-Hernández, L. F. Santos, and J. G. Hirsch, "Positive quantum Lyapunov exponents in experimental systems with a regular classical limit," Phys. Rev. E **101**, 010202 (2020).

[306] L. D'Alessio and M. Rigol, "Long-time behavior of isolated periodically driven interacting lattice systems," Phys. Rev. X **4**, 041048 (2014).

[307] D. A. Abanin, W. De Roeck, and F. Huveneers, "Exponentially slow heating in periodically driven many-body systems," Phys. Rev. Lett. **115**, 256803 (2015).

[308] T. Kuwahara, T. Mori, and K. Saito, "Floquet–Magnus theory and generic transient dynamics in periodically driven many-body quantum systems," Ann. Phys. **367**, 96 (2016).

[309] T. Mori, T. Kuwahara, and K. Saito, "Rigorous bound on energy absorption and generic relaxation in periodically driven quantum systems," Phys. Rev. Lett. **116**, 120401 (2016).

[310] N. H. Lindner, G. Refael, and V. Galitski, "Floquet topological insulator in semiconductor quantum wells," Nat. Phys. **7**, 490 (2011).

[311] M. Aidelsburger, M. Atala, M. Lohse, J. T. Barreiro, B. Paredes, and I. Bloch, "Realization of the Hofstadter Hamiltonian with ultracold atoms in optical lattices," Phys. Rev. Lett. **111**, 185301 (2013).

[312] H. Miyake, G. A. Siviloglou, C. J. Kennedy, W. C. Burton, and W. Ketterle, "Realizing the Harper Hamiltonian with laser-assisted tunneling in optical lattices," Phys. Rev. Lett. **111**, 185302 (2013).

[313] A. G. Grushin, A. Gómez-León, and T. Neupert, "Floquet fractional Chern insulators," Phys. Rev. Lett. **112**, 156801 (2014).

[314] D. V. Else, B. Bauer, and C. Nayak, "Prethermal phases of matter protected by time-

translation symmetry," Phys. Rev. X **7**, 011026 (2017).

[315] R. Moessner and S. L. Sondhi, "Equilibration and order in quantum Floquet matter," Nat. Phys. **13**, 424 (2017).

[316] H.-J. Schmidt, J. Schnack, and M. Holthaus, "Periodic thermodynamics of the Rabi model with circular polarization for arbitrary spin quantum numbers," Phys. Rev. E **100**, 042141 (2019).

[317] F. Machado, D. V. Else, G. D. Kahanamoku-Meyer, C. Nayak, and N. Y. Yao, "Long-range prethermal phases of nonequilibrium matter," Phys. Rev. X **10**, 011043 (2020).

[318] S. Blanes, F. Casas, J. Oteo, and J. Ros, "The Magnus expansion and some of its applications," Phys. Rep. **470**, 151 (2009).

[319] G. Teschl, *Ordinary Differential Equations and Dynamical Systems (Graduate Studies in Mathematics)* (American Mathematical Society, Providence, RI, 2012).

[320] A. Lazarides, A. Das, and R. Moessner, "Equilibrium states of generic quantum systems subject to periodic driving," Phys. Rev. E **90**, 012110 (2014).

[321] P. Ponte, A. Chandran, Z. Papić, and D. A. Abanin, "Periodically driven ergodic and many-body localized quantum systems," Ann. Phys. **353**, 196 (2015).

[322] T. Ishii, T. Kuwahara, T. Mori, and N. Hatano, "Heating in integrable time-periodic systems," Phys. Rev. Lett. **120**, 220602 (2018).

[323] K. Mallayya and M. Rigol, "Heating rates in periodically driven strongly interacting quantum many-body systems," Phys. Rev. Lett. **123**, 240603 (2019).

[324] D. Abanin, W. De Roeck, W. W. Ho, and F. Huveneers, "A rigorous theory of many-body prethermalization for periodically driven and closed quantum systems," Commun. Math. Phys. **354**, 809 (2017).

[325] D. A. Abanin, W. De Roeck, W. W. Ho, and F. Huveneers, "Effective Hamiltonians, prethermalization, and slow energy absorption in periodically driven many-body systems," Phys. Rev. B **95**, 014112 (2017).

[326] E. J. Torres-Herrera and L. F. Santos, "Quench dynamics of isolated many-body quantum systems," Phys. Rev. A **89**, 043620 (2014).

[327] L. D'Alessio and A. Polkovnikov, "Many-body energy localization transition in periodically driven systems," Ann. Phys. **333**, 19 (2013).

[328] A. Lazarides, A. Das, and R. Moessner, "Periodic thermodynamics of isolated quantum systems," Phys. Rev. Lett. **112**, 150401 (2014).

[329] Z. Lenarčič, F. Lange, and A. Rosch, "Perturbative approach to weakly driven many-particle systems in the presence of approximate conservation laws," Phys. Rev. B **97**, 024302 (2018).

[330] M. Vogl, P. Laurell, A. D. Barr, and G. A. Fiete, "Flow equation approach to periodically driven quantum systems," Phys. Rev. X **9**, 021037 (2019).

[331] A. Russomanno, A. Silva, and G. E. Santoro, "Periodic steady regime and interference in a periodically driven quantum system," Phys. Rev. Lett. **109**, 257201 (2012).

[332] A. Eckardt and E. Anisimovas, "High-frequency approximation for periodically driven quantum systems from a Floquet-space perspective," New J. Phys. **17**, 093039 (2015).

[333] D. A. Abanin, W. D. Roeck, and F. Huveneers, "Theory of many-body localization in periodically driven systems," Ann. Phys. **372**, 1 (2016).

[334] A. Lazarides, A. Das, and R. Moessner, "Fate of many-body localization under periodic driving," Phys. Rev. Lett. **115**, 030402 (2015).

[335] P. Ponte, Z. Papić, F. Huveneers, and D. A. Abanin, "Many-body localization in periodically driven systems," Phys. Rev. Lett. **114**, 140401 (2015).

[336] K. Efetov, *Supersymmetry in Disorder and Chaos* (Cambridge University Press, Cambridge, UK, 1996).

[337] J. Verbaarschot, "The supersymmetric method in random matrix theory and applications to QCD," AIP Conf. Proc. **744**, 277 (2004).

[338] M. R. Zirnbauer, "Anderson localization and non-linear sigma model with graded symmetry," Nucl. Phys. B **265**, 375 (1986).

[339] G. Parisi and N. Sourlas, "Random magnetic fields, supersymmetry, and negative dimen-

sions," Phys. Rev. Lett. **43**, 744 (1979).

[340] A. McKane, "Reformulation of $n \to 0$ models using anticommuting scalar fields," Phys. Lett. A **76**, 22 (1980).

[341] F. Wegner, "Algebraic derivation of symmetry relations for disordered electronic systems," Z. Phys. B Cond. Mat. **49**, 297 (1983).

[342] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed., edited by A. Jeffrey and D. Zwillinger (Elsevier, Burlington, MA, 2007).

[343] Erdelyi, A., ed., *Tables of Integral Transforms*, Vol. 1 (McGraw-Hill, New York, 1954).

[344] K. Kumar, "Expansion of a function of noncommuting operators," J. Math. Phys. **6**, 1923 (1965).