Herding Shrödinger’s Supposed Cat. How the Classical World Emerges from the Quantum Universe.

Phil Attard
phil.attard1@gmail.com

The pointer states and preferred basis of the classical world are that of definite positions and momenta. Here it is shown that the reason for the absence of superposition states is the limited resolution with which observations or measurements are made.

I. INTRODUCTION

“Entities are not to be multiplied without necessity”
—Law of Parsimony (William of Occam, 1287–1347).\(^1\)

There are four main attributes that distinguish the quantum world from the classical: first, the quantization of states; second, the uncertainty in simultaneous values of position and momentum; third, occupancy of particle states consequent on wave function symmetrization; and fourth, the superposition of states.

The first two of these are quantified by Planck’s constant, and unless a measurement or phenomenon is contingent on scales of this size or smaller then it will appear to be classical. These two types of quantum phenomena are negligible when the temperature is high enough.\(^2,3\) The third attribute also depends on the value of Planck’s constant, as it turns out that symmetrization effects may be neglected when the relevant particle separations are larger than the thermal wave-length, which is proportional to Planck’s constant. In practice this means that this type of quantum phenomenon is negligible for high enough temperatures and low enough densities.\(^2,3\)

The superposition of states is in a different category to the other three quantum attributes as it does not appear to directly depend upon the value of Planck’s constant (but see the text). Quantum mechanics, specifically Schrödinger’s equation, is a linear theory. Hence the sum of valid wave functions, including macroscopically distinct wave functions, is also a valid wave function. But in the classical universe such linear superpositions of observed states are not themselves observed.

Schrödinger’s cat embodies the divide between the quantum and the classical universe manifest in the superposition of states. Hidden from view, the cat is said to be simultaneously alive and dead, in accordance with the rules of quantum mechanics. But upon observation the cat is exclusively alive or dead, which manifests the prohibition on superposed states in the classical world. Although the common-sense view is that the cat is metaphorical, a significant number of otherwise intelligent scientists appear to take the cat quite literally. How is one to reconcile these qualitatively different pictures? How exactly does classical reality emerge from the underlying quantum equations?

II. ANALYSIS

A. What is Superposition?

Hilbert space is closed under scalar multiplication and addition, \(\Psi = \lambda_1 \Psi_1 + \lambda_2 \Psi_2\), for example. In consequence any wave function can be written as a series of orthonormal basis functions,

\[
|\psi\rangle = \sum_n \psi_n |n\rangle, \quad \psi_n = \langle n | \psi \rangle. \tag{1}
\]

In so far as \(n\) labels a state of the system, this expression suggests that the system in the wave state \(\psi\) is actually in a superposition of states \(n\).

Although the use of the words ‘state’ and ‘superposition’ have an arguable justification in this context, they don’t actually contribute to resolving the present question of the absence of classical superposition. This is because a wave function cannot be measured or observed directly.

Instead it is expectation values that correspond to the physical act of observation or measurement. More precisely, they predict the outcome of a future measurement. The expectation value of an arbitrary operator \(\hat{A}\) when the system is in the normalized wave state \(\psi\) is

\[
\langle \psi | \hat{A} | \psi \rangle = \sum_{n,m} \psi_n^* \psi_m A_{nm} \tag{2}
\]

The second equality separates the diagonal and the off-diagonal contributions.

For the diagonal terms the quantity \(A_{nn}\) has the interpretation of the value of the operator in the state \(n\). Since \(\psi_n^* \psi_n = |\psi_n|^2\) is real and non-negative, it has the interpretation of the probability or weight of the state \(n\). Probability and weight here mean the same as in classical probability theory.\(^4\)

The off-diagonal contributions \(n \neq m\) are more difficult to interpret in terms of classical concepts. One could say that \(A_{nm}\) represents the coupling of states \(n\) and \(m\) for the operator \(\hat{A}\), and that \(\psi_n^* \psi_m\), which is complex, represents the amount of interference between the states \(n\) and \(m\) in the current wave state \(\psi\). This interference reinforces or cancels the coupling.
I argue that the phrase ‘superposition of states’ is appropriate for these off-diagonal contributions. The physical interpretation of pairwise coupling and interference is not inconsistent with the physical picture of the system being simultaneously in all the states $n$. Alternatively, one might interpret it as multiple systems, each in its own state, being superposed on each other. Of course this simultaneous occupancy of the states is hard to visualize classically. Ultimately the physical picture one envisages makes no actual difference since the only thing that is essential is the mathematical manipulation of the wave functions, operators, and expectation value. It would not matter if one simply considers the off-diagonal terms as non-classical probabilities whose meaning is to be garnered from their mathematical manipulation. But in either case it is important that the words used reflect and define the underlying concept, and in a real sense the word ‘superposition’ does convey the idea of the mutual interference of the states in the off-diagonal contribution.

I further argue that the word ‘superposition’ ought to be reserved for the off-diagonal terms alone. One should not use it to describe the diagonal contribution because this has the clear interpretation as classical probability. In this theory, giving the probabilities for the possible outcomes of an event is most definitely not a statement that all outcomes occur simultaneously. For example, assigning probabilities to horses in a race does not mean that all the horses will simultaneously win the race. Similarly calculating the diagonal terms in the expectation value is a prediction of the outcome of a future measurement, not a statement that the system will be measured to be simultaneously in all the states. Classical probability can be understood as merely the mathematical representation of physical weight, which are necessarily real and non-negative. Some understand it instead as a measure of information or belief. For our purposes all that matters is that classical probability is restricted to systems in which the occupancy of the states is mutually exclusive. Obviously this mutual exclusivity is the opposite of the normal understanding of the word ‘superposition’. For this reason it is inappropriate to use the latter for the diagonal contribution to the expectation value.

Whether or not one agrees with this definition of the superposition of states, or even if one regards it as overly pedantic, there should be no quibble with the assertion that the diagonal contribution to the expectation value has a classical character, and the non-diagonal contribution is non-classical. And so the question of why there is no superposition of states in the classical universe boils down to the mathematical question of why classically only the diagonal terms contribute to expectation values.

B. Classical Measurement

Perhaps surprisingly, like the other three quantum attributes mentioned above, the key to the answer also lies in the value of Planck’s constant. We shall take it as axiomatic that our observation and measurement of the classical world is limited to a resolution that is large on the scale of Planck’s constant. Of course in reality it is possible to actually measure quantum effects, and so there are some measurement techniques with higher resolution than this. Also, there exist quantum phenomena that are macroscopic in character. But our concern here is the absence of superposition in classical systems, and for our purposes we shall define such systems to be those that are measured or observed with limited resolution. The following analysis of the effects of limited resolution on the superposition of states says nothing about other quantum phenomena.

This definition of a classical measurement as one of limited resolution does not a priori preclude the superposition of states in classical systems. One has to actually show how the laws of quantum mechanics determine that the off-diagonal contribution to an expectation value vanishes in such systems.

It will be shown this definition of a classical system provides a quantitative condition sufficient for the absence of the superposition of states. This establishes a useful criterion to guide thinking and discussion of the physical universe, and it gives the reason for the absence of superposition in this type of classical system. For example, a cat is a macroscopic object the observation of which is large on the scale of Planck’s constant. Its enlightening to give explicitly the reason why Schrödinger’s cat can’t be both dead and alive.

The answer to the question of why and when only the diagonal terms contribute to expectation values begins by defining a set of classical states. This set is not complete in Hilbert space; the subset defined by its span may be called ‘the’ classical space. The classical wave functions that comprise the set will turn out to be locally orthogonal (ie. no overlap) for resolutions not less than about Planck’s constant.

As a motivation for the particular set to be chosen, consider the fact that our brains have evolved to perceive objects, specifically their location and movement. Hence the position and momentum of particles play a key role in the observation and measurement of the classical universe. Indeed since first Newton and then Boltzmann it has been known that the state of a classical system can be described mathematically by the positions and momenta of its constituent particles. Accordingly, the aim here is to show that there can be no superposition of particle states when their positions and momenta are observed or measured on scales larger than Planck’s constant.

Consider the set of wave packets for a sub-system of $N$ particles, $\phi_{pq}(r) = \prod_{j=1}^{N} \phi^{(1)}_{p,q}(r_j)$. Here $r = \{r_1, r_2, \ldots, r_N\}$, with $r_j = \{x_j, y_j, \ldots, d_j\}$ being the representation position in $d$ dimensional space. Suppose that
the wave packets have width $O(\hbar)$ in momentum and position space. Suppose further that the momenta $p$ and positions $q$ are discrete with sufficient separation $\Delta_p$ and $\Delta_q$ between the grid points so that the wave packets do not overlap. The grid points may be called classical phase points.

The wave packet $\phi_{pq}$ makes a low pass filter for changes in position and in momentum. In this sense it provides a complete basis for measurements with no better than classical resolution. We shall assume that any classical phase function is slowly varying on the scale of $\Delta_p$ and $\Delta_q$.

As an explicit example, the minimum uncertainty wave packet for $N$ particles in $d$ dimensions is

$$\phi_{pq}(r) = (2\pi \xi^2)^{-dN/4} e^{i(r-q)^2/4\xi^2} e^{-p(r-q)/\hbar}. \quad (3)$$

It is straightforward to show that these wave packets are orthonormal,

$$\langle \phi_{pq} | \phi_{pq'} \rangle = q, \quad \text{and} \quad \langle \phi_{pq} | \hat{p} | \phi_{pq} \rangle = p. \quad (5)$$

The expectation values are

$$\langle \phi_{pq} | \hat{q} | \phi_{pq} \rangle = q, \quad \text{and} \quad \langle \phi_{pq} | \hat{p} | \phi_{pq} \rangle = p.$$  

The grid spacings should be chosen to be very much greater than the root mean square uncertainty in these expectation values,

$$\sqrt{\langle \phi_{pq} | (\hat{q} - q)^2 | \phi_{pq} \rangle} = \xi \ll \Delta_q, \quad (6)$$

and

$$\sqrt{\langle \phi_{pq} | (\hat{p} - p)^2 | \phi_{pq} \rangle} = \frac{\hbar}{2\xi} \ll \Delta_p. \quad (7)$$

Since these fluctuations are negligible on the scale of the classical resolution, the wave packet is a simultaneous position and momentum eigenfunction,

$$\hat{q} | \phi_{pq} \rangle = q | \phi_{pq} \rangle + o(\Delta_q, \Delta_p), \quad (8)$$

and

$$\hat{p} | \phi_{pq} \rangle = p | \phi_{pq} \rangle + o(\Delta_q, \Delta_p), \quad (9)$$

as can be confirmed by direct calculation. The fact that at this level of approximation the wave function is a simultaneous eigenfunction means that the position and momentum operators commute, which is consistent with Planck’s constant being negligible for measurements made with classical resolution.

It is worth mentioning that this result stems from the fact that the wave packet is relatively sharply peaked in both momentum and position space. This means that the grid spacing can be chosen so that distinct wave packets do not overlap,

$$\phi_{p'q'}^*(r) \phi_{p''q''}(r) \approx |\phi_{p'q'}(r)|^2 \delta_{q',q''}, \quad \delta_{p',p''}, \Delta_q \gg \xi, \Delta_p \gg \hbar/\xi. \quad (10)$$

This result holds not only for positions but also for momenta because of the rapid and canceling behavior that occurs if $p' \neq p''$. An analogous conclusion is reached if the argument is made in the momentum representation. One can therefore consider the set of wave packets to be locally orthogonal on the coarse grain grid, which provides a way to picture the absence of the superposition of classical states that is now derived.

An arbitrary normalized wave function can be approximated as the superposition of classical phase points,

$$\psi(r) \approx \sum_{p, q} \psi_{pq} \phi_{pq}(r), \quad \psi_{pq} = \langle \phi_{pq} | \psi \rangle. \quad (11)$$

This effectively filters out the high frequency parts of the wave function by projecting the latter onto the classical space.

It is assumed that all classical phase functions are slowly varying on the scale of the grid spacing. Hence the expectation value of an operator that is an arbitrary function of the momentum and position operators is

$$\langle \psi | A(\hat{p}, \hat{q}) | \psi \rangle = \sum_{p, q, p', q'} \psi_{pq}^* \psi_{p'q'} \langle \phi_{pq} | A(\hat{p}, \hat{q}) | \phi_{p'q'} \rangle$$

$$= \sum_{p, q, p', q'} \psi_{pq}^* \psi_{p'q'} \langle \phi_{pq} | \phi_{p'q'} \rangle \times [A(p', q') + o(\Delta_p, \Delta_q)]$$

$$= \sum_{p, q} \psi_{pq}^* \psi_{pq} A(p, q). \quad (12)$$

One sees that only the diagonal terms contribute to the expectation value, which is to say that there is no interference between the superposed states. Therefore the occupancy of the position and momentum states may be regarded probabilistically rather than physically. The expectation value has the interpretation that the system is in a mixture of pure states, with $|\psi_{p, q}|^2$ being the state probability.

One concludes from this that since the quantum mechanical expectation value is a prediction of the outcome of a measurement or observation, than an expectation value made with classical resolution corresponds to the classical statistical average of possible measured values. Since classical probability is necessarily a prediction about a system in which the occupancy of the possible states is mutually exclusive, this formula for the predicted outcome of a classical measurement implies a system in which superposition or interference effects are prohibited. This probabilistic formulation is obviously consistent with our experience of the classical world, namely one in which the only outcomes ever observed are absent the superposition of states.
III. DISCUSSION

1. Summary

This paper has sought to explain the absence of the superposition of states in the classical world by attributing it to the relatively coarse resolution with which classical measurements or observations are made. Although it is well-known that quantum mechanics applies to the very small and classical mechanics to the very large, this in and of itself does not explain the absence of classical superposition. Arbitrary wave functions can be validly superposed in quantum mechanics, including those that differ macroscopically. Hence there is a puzzle to be resolved here.

The present analysis shows that observations or measurements made with limited precision act like a low-pass filter that suppresses rapid variations in the wave function. In turn this induces a local orthogonality that effectively makes any operator block diagonal. An expectation value thus becomes the equivalent of a classical average, and there is no interference apparent between states that would signify their superposition or simultaneous occupancy.

The analysis was in terms of wave packets situated on a coarse grid in the phase space of the momenta and positions of the particles of the system. The rationale for choosing these as the basis functions lies in the observation that the state of a classical system is completely characterized by the momenta and positions of its constituent particles. The trick in the analysis lies in reconciling this with Heisenberg’s uncertainty principle, which is accomplished by choosing a sufficiently coarse grid for the momentum and position labels of the wave packets. Such a coarse grid serves to define the resolution of the classical observation or measurement. Provided that the classical phase functions are slowly varying on the scale of the grid spacing, then these momenta and positions may be regarded as a continuum for the purposes of classical mechanics.

Classical probability theory played an essential role in the analysis. Superposition was defined to mean the simultaneous occupancy of states, consequent upon which is their mutual interference. In contrast classical probability theory is predicated upon the exclusive occupancy of states: the system must be in one, and only one, state at a time. The distinction between these two cases is reflected in the diagonal and off-diagonal terms in the expectation values. The presence or absence of the latter signifies whether the system is quantum or classical. An expectation value is a prediction about the outcome of an observation or measurement. The case that only diagonal terms contribute to an expectation value is a statement that the possible states of the system are those in which exclusive occupancy holds.

2. Further Questions

Three further questions may be canvassed. First, are these results consistent with the existence of macroscopic quantum phenomena? Second, is the present mechanism the only way for quantum systems to appear classical? And third, are classical states superposed when nobody is looking?

It is evident that macroscopic quantum phenomena exist, for example superfluidity and superconductivity. The present results address the dearth of classically observed superposition states. They do not prohibit macroscopic quantum phenomena per se. Instead the present analysis says that either macroscopic quantum phenomena are not formed from the superposition of classical states, or else classical wave packets do not adequately represent the quantum states relevant for macroscopic quantum phenomena. Both statements are consistent with the present conclusion that when a measurement or observation has limited resolution (ie. projects onto classical space defined by the span of the set of appropriately spaced wave packets), then a superposition of measured or observed states is precluded.

The origin of classicality does not solely reside in the resolution with which observations or measurements are made. The present definition of superposition makes it clear that the absence or cancelation of the off-diagonal terms in the expectation value is essential for a system to be classical. The alternative way that this can occur is if the wave function is decoherent, since, by definition, coherency is required for interference effects. In fact this is the same problem as that of quantum measurement, which related wave function collapse has been attributed to the decoherence induced by interactions with the environment. One can mention the early work of Bohm and the more recent decoherence program initiated by Zeh, Zurek, and Joos the broad progress of which has been reviewed. The decoherence program holds that it is the interactions between a system and its environment that causes the wave function to collapse into the pointer states or the preferred basis of the measuring apparatus. The former are said to be stable over time, and the latter are said to preclude the superposition of states, although different authors use different definitions. As one example, Schlosshauer asks “What singles out the preferred physical quantities in nature — eg. why are physical systems usually observed to be in definite positions rather than in superpositions of positions?” (p. 50) The preferred basis problem also occurs in the Many Worlds Interpretation of quantum mechanics. The precise relationship between decoherence and pointer states in measurement on the one hand, and the emergence of classical behavior and its preferred basis on the other, remains a matter of debate.

In related work, the present author has shown how exchange and entanglement with a thermodynamic reservoir makes the sub-system wave function decoherent, and the density operator diagonal in the entropy basis (see ap-
For an equilibrium system, the Maxwell-Boltzmann weighted distribution of entropy states is sharply peaked and remains stable during the time evolution of the open sub-system (see the author’s quantum and classical analysis of the stochastic evolution of open sub-systems). These two conclusions are a subtle modification of the usual definition that the pointer states are individually stable under time evolution. The author formulates quantum statistical mechanics as a formally exact integral over classical phase space, with commutation function $W$ and symmetrization function $\eta$ formally accounting for the quantum corrections to classical behavior,

$$\langle A(\hat{\mathbf{p}}, \hat{\mathbf{q}}) \rangle^{\pm}_{z,T} = \frac{1}{2\pi} \sum_{N=0}^{\infty} \frac{z^N}{h^{2N} N!} \int d\mathbf{p} dq \ e^{-\beta H(\mathbf{p}, \mathbf{q})} \times A(\mathbf{p}, \mathbf{q}) W_{\pm}(\mathbf{p}, \mathbf{q}).$$ (13)

The derivation of this result, which is summarized in appendix [A], uses exact position and momentum eigenfunctions, not wave packets. The approach shows quantitatively how classical statistical mechanics emerges at high temperatures and low densities, in which limits the quantum corrections vanish, $W_{\pm}(\mathbf{p}, \mathbf{q}) \to 1$ and $\eta^{\pm}(\mathbf{p}, \mathbf{q}) \to 1$. The significance of this result in the present context is that it shows that decoherence is necessary but not sufficient for classical behavior.

Finally, the result of the present analysis—that the predicted outcome of a measurement with classical resolution precludes superposition states—raises the question of whether or not a classical system can exist in a superposition of states when nobody is looking (i.e., prior to measurement). There are two remarks that can be made in response to this.

First, the classical evolution equations do not permit the bifurcation of an initially non-superposed classical system, and such equations are time reversible. Assuming that the act of measurement or observation is a negligible perturbation, if the system is only ever observed in one classical state at a time, then it must always be in one classical state at a time, even when it is not observed.

Second, since the only information we have about the world comes from measurement, and since one cannot take seriously a scientific theory that is not falsifiable, it is pointless and meaningless to speculate about that which cannot be measured or observed.

Alternative to this second remark, it is permissible to develop theories about regimes that are beyond measurement, but such theories should be judged by aesthetics, including simplicity, and by their consequences when extrapolated to the measurable world. Evidently the simplest view consistent with the non-superposition of classical states upon measurement or observation, is that classical states are not superposed prior to measurement or when nobody is looking.

Simplicity as a criterion for acceptable scientific theories is attributed to William of Occam (1287–1347). The so-called Law of Parsimony is often rendered as “Entities are not to be multiplied without necessity.” This could well be a slogan for the present paper. In this spirit the present author is of the opinion that it would be of some benefit to apply Occam’s razor to Schrödinger’s cat.

1 “Occam’s razor” (Wikipedia, accessed 17 September, 2018).
2 P. Attard, Quantum Statistical Mechanics: Equilibrium and Non-Equilibrium Theory from First Principles, (IOP Publishing, Bristol, 2015).
3 P. Attard, Entropy Beyond the Second Law, Thermodynamics and Statistical Mechanics for Equilibrium, Non-Equilibrium, Classical, and Quantum Systems, (IOP Publishing, Bristol, 2018).
4 P. Attard, Thermodynamics and Statistical Mechanics: Equilibrium by Entropy Maximisation (Academic Press, London, 2002).
5 A. Messiah, Quantum Mechanics, (North-Holland, Amsterdam, Vols I and II, 1961).
6 E. Merzbacher, Quantum Mechanics, (Wiley, New York, 2nd ed., 1970).
7 D. Bohm, Phys. Rev. 85, 180 (1952).
8 H. D. Zeh, “On the Interpretation of Measurement in Quantum Theory”, Fnd. Phys. 1, 69 (1970).
9 W. H. Zurek, “Environment-Induced Superselection Rules”, Phys. Rev. D 26, 1862 (1982).
10 E. Joos and H. D. Zeh, “The Emergence of Classical Properties Through Interaction with the Environment” Z. Phys. B 59, 223 (1985).
11 W. H. Zurek, “Decoherence, Einselection, and the Quantum Origins of the Classical”, Rev. Mod. Phys. 75, 715 (2003). [arXiv:quant-ph/0105127]
12 M. Schlosshauer, “Decoherence, the Measurement Problem, and Interpretations of Quantum Mechanics”, Rev. Mod. Phys. 76 (2004), 1267. [arXiv:quant-ph/0312059v4].
13 M. Schlosshauer, Decoherence and the Quantum-to-Classical Transition, (Springer, 2007).
14 B. Galvan, “On the Preferred-Basis Problem and its Possible Solutions”, arXiv:quant-ph/1008.3708v1.
15 R. E. Kastner, “Einselection of Pointer Observables: the New H-Theorem?”, Studies in History and Philosophy of Modern Physics. 48, 56 (2014). [arXiv:1406.4126].
16 W. H. Zurek, “Pointer Basis of Quantum Apparatus: Into what mixture does the Wave Packet Collapse?”, Phys. Rev. D 24, 1516 (1981).
17 W. H. Zurek, “Preferred States, Predictability, Classicality and the Environment-Induced Decoherence”, Progr. Theor. Phys. 89, 281 (1993).
18 M. Tegmark and H. S. Shapiro, “Decoherence Produces Coherent States: An Explicit Proof for Harmonic Chains”, Phys. Rev. E 50 2538 (1994). [arXiv:gr-qc/9402026].
19 H. M. Wiseman and J. A. Vaccaro, “Maximally Robust Unraveling”, Phys. Lett. A 250, 241 (1998). [arXiv:quant-ph/9709014].
20 J. P. Paz and W. H. Zurek, “Quantum Limit of Dec-
coherence: Environment-Induced Superselection of Energy Eigenstates”, Phys. Rev. Lett. 82 5181 (1999). \url{arXiv:quant-ph/9811026}

21 M. Busse and K. Hornberger, “Emergence of Pointer States in a Non-Perturbative Environment”, J. Phys. A 42 362001 (2009). \url{arXiv:0905.4609}

22 P. Attard, “Quantum Statistical Mechanics Results for Argon, Neon, and Helium Using Classical Monte Carlo”, \url{arXiv:1702.00096} (2017).

23 P. Attard, Non-Equilibrium Thermodynamics and Statistical Mechanics: Foundations and Applications (Oxford University Press, Oxford, 2012).

24 M. Brune, E. Hagley, J. Dreyer, X. Maitre, A. Maali, C. Wunderlich, J. M. Raimond, and S. Haroche, “Observing the Progressive Decoherence of the “Meter” in a Quantum Measurement”, Phys. Rev. Lett. 77, 4887 (1996).

25 L. Hackermüller, K. Hornberger, B. Brezger, A. Zeilinger, and M. Arndt, “Decoherence of Matter Waves by Thermal Emission of Radiation”, Nature 427, 711 (2004).

26 E. Wigner, “On the Quantum Correction for Thermodynamic Equilibrium”, Phys. Rev. 40, 749, (1932).

27 J. G. Kirkwood, “Quantum Statistics of Almost Classical Particles”, Phys. Rev. 44, 31, (1933).

Appendix A: Decoherence

This analysis summarizes that in §7.2 and §7.3 of Ref. [3].

We seek the average of an operator on an open sub-system. To this end we begin with the weight of states of an isolated sub-system

\[ \omega_{nn} = 1, \quad \omega = I. \]  \hspace{1cm} (A1)

This follows from Schrödinger’s equation and the axiom that time is uniform.

Label the non-degenerate energy macrostates by a Greek letter, and the degenerate microstates by a Roman letter,

\[ \hat{H} | \alpha g, h \rangle = E_{\alpha g, h} | \alpha g, h \rangle. \]  \hspace{1cm} (A2)

In view of the preceding result that microstates of an isolated sub-system have unit weight, the entropy of an energy state is the logarithm of the number of degenerate states that it contains, \( S^E_{\alpha} = k_B \ln N^E_{\alpha} \), where \( N^E_{\alpha} = \sum_{g \in \alpha} \).

For an open sub-system that can exchange energy with a reservoir, the total energy is \( E^{\text{tot}} = E^s + E^R \). The reservoir entropy for a particular energy is just

\[ S^s(E^R_a) = k_B \ln N^R_{a}, \quad \text{with} \quad N^R_{a} = \sum_{g \in a} \langle \text{Er} \rangle. \]  \hspace{1cm} (A3)

One also has the usual thermodynamic formula

\[ S^s(E^R_a) = S^s(E^{\text{tot}} - E^n) = \text{const.} - \frac{E^n}{T_r}. \]  \hspace{1cm} (A4)

The constant can be discarded, as can the superscript on the reservoir temperature.

With \( \{ | \alpha g, h \rangle \} \) an orthonormal basis for the sub-system and \( \{ | \beta_{gh} \rangle \} \) an orthonormal basis for the reservoir, the most general wave function of the total system is

\[ \psi^{\alpha, \beta} = \sum_{\alpha, \beta} c_{\alpha g, \beta h} | \alpha g, h \rangle | \beta_{gh} \rangle. \]  \hspace{1cm} (A5)

Due to energy conservation, the wave function is entangled, which means that the expansion coefficient cannot be dyadic. In the second equality, the reservoir macrostate is defined implicitly, \( E^R_a + E^R_{\beta h} = E^{\text{tot}} \).

With this, the expectation value of a sub-system operator is

\[ \langle \psi^{\alpha, \beta} | \hat{A}^s | \psi^{\alpha, \beta} \rangle = \sum_{\alpha, \beta} \sum_{\alpha', \beta'} c^*_{\alpha' g', h'} c_{\alpha g, h} \langle \alpha' g', h' | \hat{A}^s | \alpha g, h \rangle \]  \hspace{1cm} (A6)

The third equality follows because the relationship between \( \alpha \) and \( \beta_{\alpha} \) is bijective (the macrostates are non-degenerate). \( \langle \alpha \beta_{\alpha} | \beta_{\alpha} \rangle = \delta_{\alpha', \alpha} \delta_{\beta', \beta} \). At this stage the density matrix in the energy (ie. entropy) representation consists of principal blocks on the diagonal.

Since the microstates of the total system have equal weight, apart from overall normalization the coefficients of the total wave function all have unit magnitude,

\[ c_{\alpha g, h} = e^{i \theta_{\alpha g, h}}. \]  \hspace{1cm} (A7)

The phases are uniformly and randomly distributed, and so averaging over them gives

\[ \langle \alpha g, h | c_{\alpha g, h} \rangle = \langle e^{i \theta_{\alpha g, h}} \rangle_{\text{stat}} = \delta_{g, g'}. \]  \hspace{1cm} (A8)

Hence

\[ \sum_{h \in \beta_{\alpha}} \langle \text{Er} \rangle e_{\alpha' g', h}^{*} c_{\alpha g, h} = \sum_{h \in \beta_{\alpha}} \langle \text{Er} \rangle e_{\alpha' g', h}^{*} c_{\alpha g, h} = \delta_{g, g'} \sum_{h \in \beta_{\alpha}} \langle \text{Er} \rangle e_{\alpha' g', h}^{*}. \] \hspace{1cm} (A9)

since \( k_B \ln N^{\text{Er}}_{\beta_{\alpha}} = S_{\beta_{\alpha}} = -E^R_{\alpha'}/T_r \). With this the expectation value of the sub-system operator in the total system is

\[ \langle \hat{A}^s \rangle_{\text{stat}} = \frac{1}{Z} \sum_{\alpha, g} e^{-\beta E_{\alpha}^s} a^{*}_{\alpha g} a_{\alpha g} = \text{Tr} \hat{\varphi} \hat{A}^s. \] \hspace{1cm} (A10)

where the probability operator is \( \hat{\varphi} = Z^{-1} e^{\beta \hat{R}} \) and \( \beta = 1/k_B T_r \). One sees that there is no superposition of
states in this averaged expectation value; the wave function of the sub-system has collapsed into energy (entropy) microstates. Since the variables that are not exchanged with the reservoir (eg. number in this case) belong to the degenerate states, the states corresponding to these have also collapsed.

It's worthwhile showing that the random phase approach gives the same result as an integral over the total wave space, namely

$$\langle \hat{A}^\dagger \rangle_{\text{stat}} = \frac{1}{Z'} \int d\psi_{\text{tot}} \hat{A}'(\psi_{\text{tot}})$$

$$= \frac{1}{Z'} \sum_{\alpha, g} \sum_{h} \langle \text{Er} \rangle A^{s, E}_{\alpha g', \alpha g} \int d\psi \frac{c^{*}_{\alpha g', h} c_{\alpha g, h}}{N(\psi_{\text{tot}})}$$

$$= \frac{1}{Z'} \sum_{\alpha, g} \sum_{h} \langle \text{Er} \rangle A^{s, E}_{\alpha g, \alpha g} \int d\psi \frac{|c_{\alpha g, h}|^2}{N(\psi_{\text{tot}})}$$

$$= \frac{1}{Z} \sum_{\alpha, g} e^{-E^{\alpha}_{s, g}/k_B} A^{s, E}_{\alpha g, \alpha g} \times \text{const.}$$

$$= \frac{1}{Z} \sum_{\alpha, g} e^{-E^{\alpha}_{s, g}/k_B} A^{s, E}_{\alpha g, \alpha g}. \quad (A11)$$

The third equality follows since the terms in which the integrand is odd vanish; the only non-vanishing terms have $g = g'$. The fourth equality follows because all the integrations give the same constant independent of $\alpha$, $g$, and $h$.

This analysis shows that that the statistical average of a sub-system quantum operator is just the von Neumann trace, with the density operator being the Maxwell-Boltzmann (or Gibbs more generally) operator, which is diagonal in the entropy basis. For the present, the important point is that this mixture formulation results directly from the decoherence induced in the wave function by exchange, entanglement, and subsequent averaging over the reservoir states. The time scale for entanglement and decoherence is likely much shorter than that for equilibration.

**Appendix B: Phase Space Representation**

This analysis summarizes and improves that given in §7.4 of Ref. [3].

The grand partition function for a sub-system that is entangled with a reservoir by the exchange of energy and number is\(^3\)

$$\Xi^\pm = \sum_{N=0}^{\infty} \frac{z^N}{N!} \sum_{\hat{p}} \langle \hat{P}_n | e^{-\beta \hat{H}} | n \rangle$$

$$= \text{TR}^\prime e^{-\beta \hat{H}}. \quad (B1)$$

Here $z$ is the fugacity, $N$ is the number of particles, $\beta$ is the inverse temperature, and $|n\rangle$ is a complete set of energy eigenfunctions. Also, $\hat{P}$ is the permutation operator, $p$ is its parity, the upper sign is for bosons and the lower sign is for fermions. This has the form of a trace over distinct states. Similarly the statistical average of an operator is

$$\langle \hat{A} \rangle_{z, \beta}^{\pm} = \frac{1}{\Xi^\pm} \text{TR}^\prime e^{-\beta \hat{H}} \hat{A}. \quad (B2)$$

As these are in the form of a trace, they can be written in any basis.

The position representation for $N$ particles in $d$ dimensions is $r = \{r_1, r_2, \ldots, r_N\}$, with $r_j = \{r_{jx}, r_{jy}, \ldots, r_{jd}\}$. The orthonormal momentum eigenfunctions in the position representation are\(^5\)

$$|p\rangle \equiv \frac{e^{-p r/\hbar}}{V N^{d/2}}, \quad (B3)$$

where $V = L^d$ is the volume of the sub-system. The momentum label $p$ is discrete with the spacing between momentum states being $\Delta_p = 2\pi\hbar/L$. The position eigenfunctions in the position representation are just the Dirac-\($\delta$\) functions

$$|q\rangle \equiv \delta(r - q). \quad (B4)$$

The momentum and position eigenfunctions each form a complete set. Hence the partition function can be re-written

$$\Xi^\pm = \sum_{N=0}^{\infty} \frac{z^N}{N!} \sum_{\hat{p}} \langle \hat{P}_p | e^{-\beta \hat{H}} | p \rangle$$

$$= \sum_{N=0}^{\infty} \frac{z^N}{N!} \sum_{\hat{p}} \langle \hat{P}_p | e^{-\beta \hat{H}} | p \rangle$$

$$\times \langle q | e^{-\beta \hat{H}} | p \rangle$$

$$= \sum_{N=0}^{\infty} \frac{z^N}{N!} \sum_{\hat{p}} \langle \hat{P}_p | q \rangle$$

$$\times \langle q | e^{-\beta \hat{H}} \hat{P}_p | p \rangle$$

$$\equiv \sum_{N=0}^{\infty} \frac{z^N}{N!} \int d\Gamma e^{-\beta \hat{H}(\Gamma)} W(\Gamma) \eta^\pm(\Gamma). \quad (B5)$$

Here a point in classical phase has been denoted $\Gamma \equiv \{p, q\}$. The commutation function $W$ is defined via\(^3\)

$$e^{-\beta \hat{H}(\Gamma)} W(\Gamma) \equiv \langle q | e^{-\beta \hat{H}} | p \rangle / \langle p | q \rangle. \quad (B6)$$

and the symmetrization function is\(^3\)

$$\eta^\pm(\Gamma) \equiv \sum_{\hat{p}} \langle \hat{P}_p | q \rangle / \langle p | q \rangle. \quad (B7)$$
An infinite resummation of the symmetrization function casts it as a series of loop permutations, which is quite tractable computationally. The commutation function and the symmetrization function are complex. But the imaginary parts of each are odd in momentum, and so the grand partition function is real.

Analogous analysis for the statistical average yields the same result with $e^{-\beta H} \Rightarrow e^{-\beta H} \hat{A}$,

$$\langle \hat{A} \rangle _{z,\beta} ^\pm = \frac{1}{2^\pm} \sum _{N=0} ^\infty \frac{z^N}{N!} \int d\Gamma \times e^{-\beta H(\Gamma)} A(\Gamma) W_A(\Gamma) \eta ^\pm (\Gamma). \quad (B8)$$

Here $\hat{A} = A(\hat{p}, \hat{q})$, and

$$e^{-\beta H(\Gamma)} A(\Gamma) W_A(\Gamma) \equiv \frac{\langle q | e^{-\beta H} A(\hat{p}, \hat{q}) | p \rangle}{\langle q | p \rangle}. \quad (B9)$$

For a Hermitian operator $A(\hat{p}, \hat{q})$ is real, as must be the statistical average. Since $\eta ^\pm (\hat{p}, \hat{q})^* = \eta ^\pm (-\hat{p}, -\hat{q})$, it is sufficient that $A(\hat{p}, \hat{q}) W_A(\hat{p}, \hat{q})^* = A(-\hat{p}, -\hat{q}) W_A(-\hat{p}, -\hat{q})$, in order for the imaginary part to integrate to zero. If $A(\hat{p}, \hat{q}) = A_p(\hat{p}) + A_q(\hat{q})$, then it may be shown that one can make the replacement $W_A(\Gamma) \Rightarrow W(\Gamma)$.

It is worth mentioning that the loop factorization of the symmetrization function facilitates the evaluation of averages that can be written as a derivative of the grand potential.

Since the average must be real, the imaginary part must always integrate to zero no matter how one formulates the integrand. Nevertheless, it may be more efficient computationally to write it in the most symmetric fashion using the fact that the trace operation is insensitive to the order of the operators or of the position and momentum eigenfunctions. Hence one can use the symmetric integrand

$$e^{-\beta H(\Gamma)} A(\Gamma) W_A(\Gamma) \eta ^\pm (\Gamma)$$

Finally, at high temperatures,

$$W(\Gamma) \rightarrow 1, \quad \beta \rightarrow 0,$$

and at high temperatures or low densities,

$$\eta ^\pm (\Gamma) \rightarrow 1, \quad \beta \rightarrow 0 \text{ or } z \rightarrow 0. \quad (B12)$$

In these limits the partition function and statistical average become purely classical.

Because $W_A$ and $\eta ^\pm$ are complex, the phase space integral does not generally take on the form of a classical average, since the weights in the latter must be real and non-negative. As just mentioned, the integrand can be written in a symmetric fashion that is real, but there is no guarantee that it is non-negative. Also, in the general case the commutation function $A$ depends upon the phase function being averaged, whereas a classical weight density should depend only on the state independent of the function of the state being averaged. The functions $W_A$ and $\eta$ only become fully classical in the high temperature, low density regime. So although decoherence from the entanglement with the reservoir leads directly to the suppression of the superposition of both energy states and phase (i.e. momentum and position) states, decoherence alone is not sufficient to ensure full classical behavior.