ANNALS OF PHYSICS 149, 374–456 (1983)

Quantum Tunnelling in a Dissipative System

A. O. CALDEIRA

Instituto de Fisica "Gleb Wataghin" Universidade Estadual de Campinas, Cidade Universitaria, Barao Geraldo, 13-100 Campinas, Sao Paulo, Brazil

AND

A. J. LEGGETT*

School of Mathematical and Physical Sciences, University of Sussex, Falmer, Brighton BN1 9QH, England

Received July 22, 1982; revised December 20, 1982

In this paper we attempt to motivate, define, and resolve the question "What is the effect of dissipation on quantum tunnelling?" The question is of particular interest in the context of tunnelling of a macroscopic variable such as the trapped flux in a SQUID, where we show that it is crucial to resolve it in the context of tests of the validity of quantum mechanics at the macroscopic level, but it is also relevant to various microscopic tunnelling situations. We define the question as follows: Suppose we have a system, which has a metastable minimum and whose quasiclassical equation of motion in the region near the minimum is given by

$$M\ddot{q} + \eta\dot{q} + \partial V/\partial q = F_{\text{ext}}(t),$$

where the potential $V(q)$ and friction coefficient η are regarded as experimentally determined quantities (and the energy dissipated irreversibly per unit time is simply $\eta\dot{q}^2$). How does the tunnelling behaviour of such a system at $T = 0$ differ from that of one obeying a similar equation, with the same potential $V(q)$ and mass M, but with friction coefficient η equal to zero?

We start by arguing that provided any one degree of freedom of the environment is only weakly perturbed by the motion of the system, then at $T = 0$ it is always possible, *without any loss of generality*, to represent the environment as a bath of harmonic oscillators; moreover, if the damped equation of motion is indeed of the above form, then (barring, possibly, certain apparently pathological cases) it is possible to choose the system-bath coupling to be linear in the (suitably chosen) oscillator coordinates and a function only of the system coordinate (i.e., of the form $\sum_j F_j(q)\, x_j$). In particular this is always possible for the important case of adiabatic

* Current address: Department of Physics, University of Illinois, Urbana, Illinois 61801.

0003-4916/83 $7.50

coupling. In certain cases (which we refer to as cases of "strictly linear" dissipation) it is possible further to restrict $F_j(q)$ to be linear in q. In addition to this linear coupling, the interaction may also introduce an extra term ("counterterm") in the effective potential $V(q)$ felt by the system, which is just such as to cancel unphysical frequency shifts, etc., produced by the linear term. We discuss this point, which we believe to have caused some confusion in the recent literature, in some detail; in particular, we show that such a counterterm arises automatically both when the dissipative coupling is of the electromagnetic type realised, e.g., in a SQUID and when it is the correction to the zeroth-order adiabatic (Born–Oppenheimer) description.

We next apply a variant of the "instanton" technique well known in particle physics to calculate the zero-temperature tunnelling rate out of the metastable minimum. By integrating out the environment variables explicitly, we can represent this rate in the form of a functional integral involving only the system variable $q(\tau)$ but with the effective action containing an interaction term which is nonlocal in (imaginary) "time." (The method is checked by applying it to the exactly soluble problem of a damped harmonic oscillator). As in the nondissipative case, the result for the tunnelling rate Γ can be written, in the WKB limit, in the form

$$\Gamma = A \exp -B/\hbar,$$

where B is the effective action evaluated along the "quasiclassical" path $q(\tau)$ in the inverted potential, and A represents the effect of fluctuations around this path. The effect of dissipation is always to suppress the tunnelling, and in the strictly linear case the values of A and B are unique functions, for a given $V(q)$, of the phenomenological dissipative coefficient η. In the more general case the corresponding function of η gives a lower limit on the tunnelling rate. We carry out a quantitative analysis of the general formula for the physically important case of strictly linear dissipation in a cubically anharmonic potential; in this case we show that the correction to the "undamped" exponent B_0 can be written

$$\Delta B = \Phi(\alpha)\, \eta q_0^2, \qquad \alpha \equiv \eta/2M\omega_0,$$

where η is the phenomenological friction coefficient, ω_0 the frequency of small oscillations around the metastable minimum, and q_0 the distance the (undamped) system would have to travel "under the barrier." The dimensionless factor $\Phi(\alpha)$ is always of order 1; we find its exact forms in the limit of weak and strong damping and put stringent limits on it for the intermediate case. We also obtain the form of the prefactor A in the strong-damping limit, where it is proportional to $\alpha^{9/2}\omega_0$, and give a prescription for calculating it more generally. We generalize our results to the case of nonlinear and frequency-dependent friction. We discuss the limits of validity of our results and some related problems concerning quantum coherence in a dissipative system, and outline possible future lines of development.

1. INTRODUCTION

In discussions of the conceptual foundations of quantum mechanics, a crucial role is played by the so-called "quantum measurement paradox"—formulated in its most spectacular and best-known form by Schrödinger [1] as the Cat Paradox.[1] This paradox arises because it is the general belief that there is no natural limitation to the realm of validity of quantum mechanics, so that it should in principle describe the behaviour of macroscopic bodies such as Geiger counters and cats as well as that of electrons and atoms. Yet it is a curious fact that if one asks what is the experimental evidence that quantum mechanics does apply on the scale of macroscopic bodies, and in particular that linear superpositions of states corresponding to macroscopically different properties can actually exist in nature, then at least until very recently the answer was that there was none. Thus, the most fundamental problem of modern physics, over which philosophers as well as physicists have puzzled for decades, actually rests on an extrapolation which is largely experimentally untested.

Suppose then that we wish to look for ways of testing this extrapolation. An obvious way to start off is to introduce the idea of a *macroscopic coordinate*. There are many macroscopic systems in nature for which it is possible to separate off one or more coordinates which are recognizably "macroscopic" (cf. below), such as the centre-of-mass coordinate, and which moreover have the property that they are only weakly coupled to the other, microscopic degrees of freedom. For example, the centre-of-mass coordinate of a body falling freely in a uniform gravitational field is completely decoupled from the relative coordinates of the atoms composing it. In other cases, although the original coupling may be quite strong, the microscopic coordinates respond adiabatically to the macroscopic motion and hence the effective coupling which results is weak. In either case, to a first approximation the Lagrangian may be separated into a term referring only to the macroscopic coordinate X and the corresponding velocity and another term, possibly adiabatically dependent on X, which refers to the microscopic coordinates and velocities. If now we make the extrapolation mentioned above, we can quantize the motion of X in the usual way and in fact write down for it a standard (closed) Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t}(X, t) = -\frac{\hbar^2}{2M} \frac{\partial^2 \Psi}{\partial X^2} + V(X, t)\, \Psi(X, t). \tag{1.1}$$

Here the macroscopic variable X need not necessarily have the significance of a geometrical coordinate, nor the parameter M that of a physical mass: cf. below. Now, in general, "appreciably" different values of X will correspond to *macroscopically distinguishable* states of the system; so, if the wave functions $\Psi(X, t)$ which occur

[1] The argument of the first few paragraphs is developed in greater detail in a previous paper by one of us [2]. This paper discusses, inter alia, the widespread misconception that the validity of quantum mechanics on the macroscopic scale (in the sense required to generate the Cat paradox) is demonstrated by the so-called "macroscopic quantum phenomena" (Josephson effect, flux quantization, etc.) seen in superconductors and superfluids.

under physically attainable conditions extend over "appreciable" regions of X, we should in principle expect to be able to observe interference effects between macroscopically distinguishable states and hence to confirm (or refute) our extrapolation of quantum mechanics.

An immediate objection is that interference effects are typically associated with single quantum states and are blurred out by any thermal effects which populate more than one or two energy eigenstates appreciably; but since for a macroscopic system the spacing of energy levels is very tiny, there will be a large amount of thermal disorder even at the lowest temperatures conceivably attainable, and hence we shall never be able to see macroscopic interference effects in the sense envisaged. This objection is not correct. It is of course true that if we consider *all* the energy levels of a macroscopic system, they will usually be very closely spaced;[2] however, to the extent that we can neglect the coupling between the macroscopic and microscopic degrees of freedom, the wave function $\Psi(X, t)$ simply multiplies the wave function of the microscopic coordinates, so that it is only the spacing of the levels described by (1.1) which is relevant. This need not necessarily be small, even for a macroscopic system; for example, if X is not a geometrical coordinate but the flux through a simple LC-circuit, then the level spacing is equal to $\hbar\omega$, where $\omega \equiv (LC)^{-1/2}$ is the resonant frequency of the circuit. With currently attainable circuit parameters it is not difficult to make this spacing large compared to the thermal energy at readily attainable temperature, so that the circuit should certainly be in its quantum-mechanical groundstate. This argument relies crucially on the assumption that we can to a first approximation neglect the coupling between the macroscopic and microscopic degrees of freedom; the validity of this approximation, and the corrections to it, are in a sense just the subject of this paper.

Suppose then that we have our macroscopic system at a low enough temperature to neglect thermal blurring of interference effects, how are we to look for such effects? That is, how can we verify that the system is behaving in a characteristically quantum-mechanical manner? For reasons given elsewhere [2] we believe that the most promising phenomenon to look for is that of quantum tunnelling. Here it is necessary to make an important distinction. The phenomenon we shall be considering in this paper is the macroscopic analog of the tunnelling of an alpha particle out of a nucleus, or of an electron out of an atom in a strong electric field. It is *not* the macroscopic analog of phenomena such as the inversion resonance of the ammonia molecule, the Josephson effect in superconductors or Bloch waves in a metal, which involve the system tunnelling *coherently* between two or more degenerate or nearly degenerate potential minima. From our point of view the crucial difference is that in the former type of phenomenon the phase relationship between the amplitudes for being on different sides of the barrier(s) can be neglected, since once outside the barrier the system never comes back and "interferes with itself;" for the latter type of phenomenon, on the contrary, it is crucial to take it into account. To avoid confusion

[2] Although even this may not be quite as obvious as it seems, for example, for a superconductor very far below its transition temperature [3].

we shall refer to the latter type of phenomenon as quantum *coherence*; the subject of this paper, quantum *tunnelling*, refers to the case where a system decays from a metastable state into a continuum. Although the techniques developed in this paper are quite general and are applicable equally to both types of phenomenon, we believe that in the context of the verification of quantum mechanics on the macroscopic scale the experimental relevance of the two is quite different. In fact, in our opinion the observation of quantum tunnelling on the macroscopic scale is, in principle at least, feasible with existing cryogenic and experimental techniques (and has indeed been reported by several groups—see below); by contrast, the observation of quantum *coherence* on the macroscopic scale seems probably at best a long-term prospect. A discussion of the reasons for this belief, and of the relation between the two types of phenomenon, has been given elsewhere [2]; see also the brief remarks in Section 6 of the present paper.

In looking for a suitable system in which to try to observe macroscopic quantum tunnelling we should impose a number of desiderata. First, the system in question must have a metastable state which is separated from a more stable continuum (or near-continuum: see Section 6) of states by a free energy barrier. Moreover, the points at which the system "enters" and "exits from" the barrier should correspond to macroscopically distinguishable[3] states. (There are of course many common situations (e.g., the α-decay of a nucleus within a Geiger counter) where a quantum tunnelling event leads to a macroscopic change in the system, but we should not want to call such cases examples of macroscopic quantum tunnelling). Second, the frequency of small oscillations around the equilibrium position ω_0 should be fairly high, in order that it should be possible to satisfy the condition $\hbar\omega_0 \gg k_B T$ with attainable temperatures T; this is because we wish to be able to identify any transitions which occur as unambiguously due to quantum tunnelling rather than thermal nucleation. Since the thermal escape rate is proportional to $\omega_0 \exp -V_0/k_B T$ (with V_0 the barrier height) while the quantum tunnelling rate can be written (see Section 2) in the form const. $\exp -\alpha V_0/\hbar\omega_0$ with $\alpha \gtrsim 1$, the above criterion follows. Third, the barrier height should not be too great, otherwise the lifetime of the metastable state will be unobservably long; if, for example, we assume $\omega_0 \sim 10^{11}\ \text{sec}^{-1}$, then V_0/k_B should not be more than about 40°K (which is a rather small energy even on the *atomic* scale!) Fourth, it is highly desirable that the essential parameters, in particular the barrier height, should be experimentally variable, that there should be a direct means of registering when a transition has occurred and that noise in the system studied should be low. Finally, it is very important that one should be able to measure the parameters of the system by experiments describable in purely classical terms.

Of the various possible physical systems which might be proposed as candidates, the one which satisfies the above conditions best is probably an (rf) superconducting

[3] We will not enter here into the vexed question of what precisely one means by "macroscopically distinguishable"—a concept which in any case has no sharp borderline; in any specific practical case we feel this is unlikely to cause difficulty.

interference device (SQUID), that is, a superconducting ring interrupted by a Josephson junction (see, e.g., [4]). Indeed, it was suggested many years ago [5] that quantum tunnelling would be the dominant transition mechanism in such a system at sufficiently low temperature. In this case, the macroscopic variable of interest is the flux Φ trapped in the ring. The "potential energy" of such a system is described to a first approximation by the expression (see, e.g., [6])

$$U(\Phi) = \frac{(\Phi - \Phi_x)^2}{2L} - I_c \Phi_0 \cos(2\pi\Phi/\Phi_0). \tag{1.2}$$

Here L is the self-inductance of the ring, I_c the critical current of the Josephson junction, $\Phi_0 \equiv h/2e$ the flux quantum and Φ_x the externally imposed flux through the ring, which in the present context we shall treat as a c-number parameter controllable at will by the experimenter. The first term is the electromagnetic energy arising from the finite self-inductance of the ring, while the second is the phase-locking energy associated with the junction. If the parameters satisfy the condition $2\pi L I_c/\Phi_0 > 1$, then the curve $U(\Phi)$ has at least one metastable minimum for at least some values of Φ_x, and by suitable manipulation of Φ_x it is possible to "trap" the SQUID in this metastable state; by further variation of Φ_x the metastable minimum can be made unstable. Indeed, the normal mode of operation [4] of a standard *rf* SQUID involves repeated trapping and release of the flux by means of variation of Φ_x. In addition to the "potential energy" term (1.2), there is also in the Hamiltonian a "kinetic energy" of the form $\frac{1}{2}C\dot{\Phi}^2$ which is due to the finite capacitance C of the Josephson junction; it is evident that C plays the role of the particle mass in the dynamical problem. Since the problem is now completely isomorphic to that of a simple one-dimensional mechanical system, we expect there to exist a macroscopic wave function $\Psi(\Phi)$ for the trapped flux governed by the Schrödinger equation (1.1) (with the substitutions $X \to \Phi$, $V(X) \to U(\Phi)$, $M \to C$). In particular, we expect quantum tunnelling to occur out of the metastable potential walls. The rate of such tunnelling can be calculated (at first sight at least) by the standard WKB approximation [5, 7] and evidently depends strongly on the parameters L, I_c, and (especially) C; for the values of these parameters typical of a strongly hysteretic *rf* SQUID we expect the rate to be appreciable only when Φ_x is close to the value needed to make the metastable state classically unstable (see [7]). It should be noted that under these conditions the shape of the potential in the region where tunnelling occurs is well represented by a quadratic-plus-cubic form $(\alpha x^2 - \beta x^3)$; this is why, in carrying out quantitative analyses in this paper, we have concentrated on this form of potential. The width of the barrier through which the system has to tunnel is then considerably less than a flux quantum ($0 \cdot 1\Phi_0$ might be typical) but still large enough for the entry and exit points to correspond to macroscopically distinguishable states—at least by our and, we suspect, most people's definition. For less strongly hysteretic SQUIDs such as that used in the experiments reported in [8] the entry and exit points can differ by an amount of order Φ_0. In either case the other desiderata are reasonably well satisfied,

so that such a system seems to be a good place to look for macroscopic quantum tunnelling.

A closely related system which is also a candidate [9] is a single Josephson junction biassed by a fixed external current I_0. In such a case the relevant "macroscopic" variable is the phase difference φ of the Cooper pair wave function across the junction, and the Hamiltonian conventionally used to describe its behaviour is composed of a "potential energy" (the so-called "washboard potential")

$$U(\varphi) = -\frac{I_0 \Phi_0}{2\pi}\varphi - \frac{I_c \Phi_0}{2\pi}\cos\varphi \tag{1.3}$$

and a "kinetic energy" $\frac{1}{2}C(\Phi_0/2\pi)^2\dot{\varphi}^2$. Comparing (1.3) with (1.2), it is tempting to regard the current-biassed junction as simply the limit of a SQUID with infinite self-inductance, with the correspondence $\Phi \to (\Phi_0/2\pi)\varphi$, $\Phi_x/L \to I_0$. However, there is a subtle point of difference. Two neighbouring local minima of the potential (1.2) are distinguished by the value of the flux Φ threading the ring, which is a perfectly physical and measurable quantity. On the other hand, two neighbouring minima of (1.3) are distinguished only by the fact that φ differs by 2π; since, however, φ is the phase of a wave function, it is only defined modulo 2π anyway. Thus, while it is very reasonable to believe the classical equations of motion derived from (1.3) (which can in fact be derived equally well by alternative methods) the physical meaning of the equation itself is somewhat problematical. (This difficulty is of course not peculiar to Josephson junctions; it occurs equally if we were to try to describe, for example, the motion of a mechanical pendulum driven by a constant external torque by the Hamiltonian (or Lagrangian) technique). We believe that this ambiguity is not a problem in the context of quantum tunnelling provided that this occurs when I_0 is close to the critical current I_c (as is normally the experimentally relevant situation); in this case the values of φ involved in the tunnelling process itself extend over a range much smaller than 2π, so that the variable is unambiguously defined over this region. In this case one should be able to take over the SQUID predictions directly by the above correspondence. However, one should be much more sceptical about using this correspondence in any cases where the phase tunnels through a difference of the order of 2π.

A question may be raised concerning the validity of applying quantum mechanics to "classical" equations which themselves arise from characteristically quantum effects. For example, the potential energy (1.2) involves the flux quantum $\Phi_0 = h/2e$ and is, in the last analysis, a consequence of the quantization of electronic angular momentum. Similarly, in a current-biassed junction the tunnelling variable φ is itself the phase of a quantum-mechanical wave function. We feel that, at least as regards the SQUID, there is no special difficulty here: the fact that we have to start from a classical equation which itself contains quantum effects is in no way peculiar to this case. For example, in discussing a diatomic molecule one is accustomed to write down classical equations of motion for the nuclei; these involve a potential energy which itself arises from characteristically quantum effects involving the electronic

energy levels and, in principle at least, should contain the quantum constant h. Subsequently, one may quantize these equations themselves to get the vibrational levels. Similarly, in the case of a magnetic medium one may start from a classical description in terms of an energy density depending on the local magnetization, and subsequently quantize this, even though one knows that the origin of parts of the energy density is a characteristically quantum-mechanical exchange effect. It seems to us that the only difference in the case of the SQUID is that the dependence on h is particularly simple and hence obvious, whereas in the other cases it often gets buried in more or less phenomenological constants. Turning to the case of a current-biassed junction, we can say that provided we are content to regard this as simply the limit of a SQUID with infinite self-inductance (cf. above) there is no problem. If we do not wish to take this point of view, the matter is a little more delicate. In fact, the question of how far it is legitimate to treat the phase of a wave function as a quantum-mechanical operator has received considerable discussion in the literature, though largely in the context of the difficulties encountered when the number of particles involved is small: see especially [10]. In our case the number of Cooper pairs is very large, so these particular difficulties do not arise. In fact, to the extent that we neglect the current flow into and out of the system (i.e., effectively consider two *isolated* bulk superconductors connected by a Josephson junction) the situation is formally identical to the case of tunnelling between two different bands in a single bulk superconductor. This problem was considered in [11], where arguments were given for treating the relative phase as a quantum-mechanical operator in the limit $N \to \infty$. While these arguments are not wholly rigorous, and should moreover ideally be generalized to the case of practical interest, when the bulk superconductors are not isolated, we believe that they are adequate for the purposes of the present discussion.

There have recently been a number of experiments to look for macroscopic quantum tunnelling both in SQUIDs [8, 12, 13] and in current-biassed junctions [14–17]. The experiments on current-biassed junctions usually identify the tunnelling event by the onset of the finite-voltage state (corresponding to the onset of "rolling" in the washboard potential (1.3)); in the SQUID experiments the Leiden group [8, 12] monitored the trapped flux directly, while the authors of [13] claimed to deduce the occurrence of tunnelling from the behaviour of the tank circuit current-voltage characteristic of a SQUID operated in the standard *rf* mode. One of us has commented briefly elsewhere [18] on some of these experiments.

Apart from SQUIDs and junctions, there are a number of other systems whose behaviour has recently been interpreted as possibly evidence for the quantum tunnelling of an essentially macroscopic variable. These include charge density waves in quasi-one-dimensional conductors [19–21] (cf. also [22]), the vortex–antivortex complex in two-dimensional superconducting films [23] and the photon field in a ring laser [24]. Yet other systems have been studied theoretically [25–27], but as far as we know there are as yet no relevant experiments. For the purposes of the present paper it is most convenient to bear in mind primarily the case of a SQUID; this has the advantages that the tunnelling variable is unarguably macroscopic and can be monitored directly, and, most important, that at least for certain types of junction the

classical dynamics is believed to be well understood and the parameters can be obtained rather directly from purely classical experiments.

We now turn to the question which is the central motivation of this paper. As far as we are aware, all work in this area previous to our own [7, 28] has explicitly or implicitly described the tunnelling system by a wave function and used the standard microscopic techniques such as the WKB approximation to calculate the tunnelling rate. Thus, the macroscopic nature of the tunnelling variable does not enter the problem explicitly at all. Yet in fact there is a crucial difference between these macroscopic systems and the well-known examples (field ionization, Stark effect, etc.) to which techniques such as the WKB approximation are routinely applied: namely, *macroscopic systems are inherently dissipative.* To put it differently, a macroscopic system by its very nature always experiences a complex interaction with its environment, one consequence of which is that it continually exchanges energy with it and can in no way be considered as isolated. In particular, insofar as its motion can be described classically, there is always a term in Newton's equation corresponding to the dissipation of energy. To be sure, this characteristic is not exclusive to macroscopic systems: a microscopic system which undergoes tunnelling also interacts with its environment, but in most cases the interaction is sufficiently weak to be ignored or treated as a small perturbation (for example, the coupling to the radiation field in the field-ionization problem). By contrast the dissipation in macroscopic tunnelling systems such as SQUIDs may be very strong: the typical *rf* SQUID used for practical magnetometry is overdamped. Another important difference is that even in the minority of microscopic tunnelling systems where the coupling to the environment is important (such as the case of electron or defect tunnelling in solids, where the coupling to the phonon modes can be very strong), one usually has a good a priori knowledge of the appropriate coupling Hamiltonian, or at least of its principal features. (cf. [29]). In many macroscopic systems, on the other hand, we are ignorant of the detailed mechanism of dissipation and are reduced to describing its effects by phenomenological coefficients of friction, viscosity or similar quantities. For example, it is conventional (and apparently compatible with most of the experimental data) to describe large classes of SQUIDs by the so-called "resistively shunted junction" (RSJ) model [6]. This model assumes that the Josephson junction is shunted by a phenomenological "normal" resistance R_n, whose origin is often unknown in detail. As a result, the classical equation of motion of the trapped flux takes the form

$$C\ddot{\Phi} + R_n^{-1}\dot{\Phi} + \frac{\partial U}{\partial \Phi} = 0, \tag{1.4}$$

where $U(\Phi)$ is given by Eq. (1.2).

It is clear that until we know the effect which the dissipation is likely to have on the quantum tunnelling behaviour of macroscopic systems, we cannot interpret the results of existing or projected experiments as evidence for, still less against, the extrapolation of quantum mechanics to the macroscopic scale [8, 15]. Thus, the

fundamental question to which we address ourselves in this paper is: *What is the effect of dissipation on quantum tunnelling?* While the question is not necessarily exclusive to the macroscopic context, we shall think primarily of this case in our analysis. In the present paper we shall confine ourselves to the limit of zero temperature. Moreover, in the main body of the paper we concentrate on the simplest case, that of frequency-independent linear dissipation (see Eq. (2.8)); generalizations are given at the appropriate points. For pedagogic purposes we work explicitly in terms of a simple mechanical model for which the macroscopic variable has the significance of a geometrical coordinate; the transposition of the results to a SQUID described by Eq. (1.4) is, however, trivial (see Conclusion).

It should be stressed that the question we have just posed, namely, "What is the influence of dissipation on quantum tunnelling?" has no clear meaning, and therefore no unique answer, if posed in isolation from a context such as the one just given. In Section 2 we shall specify the precise meaning given in this paper to the question (which we believe to be the meaning most relevant to all or almost all experimentally realistic cases), and will subsequently conclude that the answer to the question, when taken in this sense, is that dissipation always tends to suppress tunnelling. However, we are emphatically *not* making the claim that the incorporation of any dissipative mechanism anywhere in the neighbourhood of a system will suppress the tunnelling of any variable: such a claim could be refuted by totally trivial counterexamples. (For example, it is intuitively obvious that if in a SQUID part of the bulk superconducting ring is replaced by a thin filament with a critical current considerably less than that of the junction, then the "dissipation" so provided will be accompanied by an *increase* in the tunnelling rate. See also the discussion of the "anomalous" type of case in Appendix C).

We emphasize that we are *not* interested, in the bulk of the present paper, in the effect of dissipation on quantum *coherence*, an effect we shall refer to for brevity as "environmental detuning." (We do make some brief remarks about it in Section 6.) This latter question is closely related to the problem of the effect on quantum coherence of its "observation" by, or interaction with, the environment, which has had considerable discussion in the literature in particular contexts such as NMR ("T_2 relaxation") and impuriton states in solids ("dynamical destruction of the band" [30]) and has been recently treated in general terms in [31, 32]. One of us has given elsewhere [2] a discussion of this topic in the context of macroscopic quantum coherence. It is also important to emphasize that in this paper we calculate the *total* rate of tunnelling out of the metastable groundstate, without making a distinction between "elastic" processes in which the environment remains in its groundstate and "inelastic" ones in which it is excited, (see especially the discussion following Eq. (4.11)).

The plan of the paper is as follows: In Section 2 we set up the problem, briefly review the standard results for tunnelling in an isolated system and discuss the precise meaning of the question "What is the influence of dissipation on tunnelling?" In Section 3 we formulate a model for the dissipation mechanism which is sufficiently general to cover all cases likely to be of practical interest, whether the original

coupling of the system to the environment was weak or not (in the latter case, we have adiabatic coupling as discussed above). We also give a discussion of the "frequency renormalization" effect (see below). Section 4 is the heart of the paper: using a technique borrowed from particle physics, with appropriate developments, we derive a general expression (Eq. (4.27)) for the effective action in a damped system and hence an expression (4.30) for the tunnelling rate of the system out of a metastable minimum. Section 5 presents a quantitative analysis of the expression for the tunnelling rate for the physically important case of cubic anharmonicity: we obtain exact expressions for the effective bounce exponent in the limit of weak and strong damping, and upper and lower limits on it for the general case. We also find the dependence of the prefactor on damping in the strong-damping limit, and estimate the constant. In Section 6 we examine the physical interpretation and limits of validity of our results and mention some outstanding problems. Section 7 is a brief conclusion. There are four appendices[4]: A discusses in detail the correct form of the Lagrangian for systems with electromagnetic (or adiabatic) coupling, B tests our general formalism on the exactly soluble problem of the damped harmonic oscillator, C justifies the form of system-environment interaction postulated in Section 3, and D proves some purely mathematical results needed in Section 5.

There are two points, neither of them directly connected with quantum tunnelling, on which the reader may feel that we have gone into excessive and tedious detail. One is the question of justification (Appendix C) of the harmonic-oscillator bath representation used for the environment (in Eq. (5) of our previous work [28] and Eq. (3.5) of the present paper). Our aim here is to show that this is not in fact merely a "model," that is, a simplified version of the real situation which is expected to show the main qualitative features, but (at least at zero temperature) is actually a quite general and exact description (barring possible pathological counterexamples) of any situation in which the environment is only weakly perturbed by the motion of the system. We have little doubt that once the ground is firmly established, it will be possible to short-circuit the rather laborious procedure of Appendix C and Section 4 by appropriate formal analytic continuation techniques (cf. [29]); however, for this first venture into more or less unknown territory, it seems wise to attempt a derivation as explicit and complete as possible.[5] The second point on which we have laboured (Section 2 and Appendix A) is the question of the "frequency renormalization" effect described in Section 2, and in particular why it does not occur for systems such as SQUIDs. The reason is as follows. Although we ourselves, possibly with the benefit of hindsight, would regard the qualitative result that dissipation of the kind considered in this paper tends to *suppress* quantum tunnelling as almost obvious intuitively, there are by now at least three papers [33–35] in the literature which claim to have demonstrated the precise opposite, and similar claims have also been expressed in conference discussions, etc. We believe that most, if not all, of the

[4] The reason for the present order of the Appendices is purely historical. The logical order is C, A, B, D.

[5] Those who regard Eq. (3.2) as sufficiently plausible to need no detailed justification might well wish to skip the somewhat turgid arguments of Appendix C.

arguments advanced in favour of such claims have been based on a misunderstanding of the frequency-renormalization phenomenon, and in particular of when it does and does not occur, and have therefore devoted a considerable amount of space to a discussion of it, in the hope of clearing up this point once and for all. It may save the reader time if we remark that the basic question boils down to this: Suppose we take a simple LC-circuit and connect a piece of (normally resistive) wire with zero inductance in parallel with the capacitor. Then (a) does the resistor shift the natural resonance frequency of the circuit (other than by the usual substitution $\omega_0 \to \sqrt{\omega_0^2 - \gamma^2}$ etc.)? (b) if it does not, how do we guarantee that our microscopic model of the resistor and its coupling to the rest of the circuit reflects this fact? Those who find the answer to question (a) sufficiently obvious that they regard a discussion of (b) as superfluous and tedious (a view with which we have considerable sympathy) might wish to skip the relevant parts of Section 2 and the whole of Appendix A.

The core results of this paper (that is, the essential content of Section 2, most of Sections 3 and 4, and Appendix A) were given in our earlier letter [28], for want of space in some cases without derivation.[6] (See also [7].) Parts of the present paper which go substantially beyond the results stated in [28] include the discussion of the adiabatic case (Appendix C), the whole of the quantitative analysis of Section 5, the discussion of the limits of validity of our theory (Section 6) and the generalization to nonlinear dissipation (Section 4 and Appendix C). The application to SQUIDs has been discussed by one of us in a conference contribution [18] and has been used in the analysis of some of the recent experiments [12, 16, 17]; it has also been applied by Kurkijärvi to a discussion of the ultimate sensitivity of a SQUID magnetometer [36]. Finally, a recent criticism [34] of our work has been refuted elsewhere [37].

2. Formulation of the Problem

We shall be interested in this paper in the following problem. At zero temperature we have a system which is characterized by some principal coordinate q and is subject, inter alia, to a c-number external potential $V(q)$ which has a single metastable minimum at a point which we arbitrarily choose as the origin of q; the zero of potential is chosen to lie at the bottom of this metastable minimum, i.e., $V(0) = 0$. The system in question may be (but need not necessarily be) macroscopic, and the coordinate q need not have the significance of geometrical position; for example, in the case of a SQUID q would represent the magnetic flux trapped in the ring (see Introduction). We assume that the potential $V(q)$ is fairly smooth and has the general form shown in Fig. 1; note in particular that $V(q)$ is taken to be negative for all points $q > q_0$, where q_0 is the "exit point" of the system from the barrier (i.e., the nonzero value of q for which $V(q) = 0$ (Fig. 1)). The object of this assumption is

[6] However, the distinction between "quasi-linear" and "strictly linear" dissipation (see Section 2) was not made in [28], and the statements made there about the generality of our result should be read with this in mind.

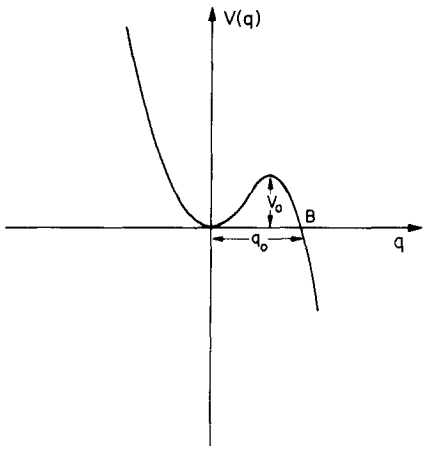

FIG. 1. The form of potential $V(q)$ considered in our calculation.

to guarantee that once the system has left the metastable well, it will have no probability amplitude for returning there in any finite time, so that we can neglect "quantum coherence" effects in the sense defined in the Introduction. (Whether this assumption is justifiable for any specific physical system of interest is a different question and is discussed in Section 6).

Throughout this paper we shall use language appropriate to the case where q is a real geometrical coordinate and has associated with it a mass M; in that case the Lagrangian of the system as described so far is simply

$$L(q, \dot{q}) = \tfrac{1}{2}M\dot{q}^2 - V(q). \tag{2.1}$$

In most cases where q is not a geometrical coordinate, the Lagrangian can still be written in the form (2.1) provided that M is understood as the appropriate parameter (e.g., in the case of a SQUID ring it is the capacitance of the Josephson junction). We denote the frequency of small oscillations around the metastable minimum by ω_0:

$$\omega_0 \equiv [M^{-1}(d^2V/dq^2)_{q=0}]^{1/2} \tag{2.2}$$

and assume for the moment that the height V_0 of the barrier (Fig. 1) is large compared to $\hbar\omega_0$, so that the WKB approximation is applicable to the tunnelling behaviour of the system as described so far. For future reference we now summarize the results of using this approximation for the isolated system.

If the system is known to be initially localized in the metastable potential well, then the probability per unit time that it escapes from the well is given by the standard formula

$$P_0 = A_0 \exp -B_0/\hbar, \qquad A_0 \equiv C_0\omega_0(B_0/2\pi\hbar)^{1/2}, \tag{2.3}$$

where B_0 is the WKB integral

$$B_0 \equiv 2\int_0^{q_0} (2MV(q))^{1/2}\, dq \tag{2.4}$$

and C_0 is a dimensionless constant of order unity which depends on the shape of the potential $V(q)$ and can be calculated by any of a number of standard methods. Corrections to Eq. (2.3) are at most of relative order $\hbar\omega_0/V_0$, and in almost all situations of practical interest in the present context it is the exponential factor which is likely to dominate the tunnelling behaviour, particularly when we are interested in the dependence of tunnelling rate on some external control parameter (such as the externally imposed flux in a SQUID). A case of particular interest is that of a "quadratic-plus-cubic" potential, which is likely to describe many of the systems of interest to us in the regime where tunnelling is appreciable (see Introduction)

$$V(q) = \tfrac{1}{2}M\omega_0^2 q^2 - \beta q^3 \equiv \frac{27}{4} V_0 \left\{ \left(\frac{q}{q_0}\right)^2 - \left(\frac{q}{q_0}\right)^3 \right\}, \tag{2.5}$$

where $q_0 \equiv \frac{1}{2}M\omega_0^2/\beta$ is the coordinate of the "exit point." For such a potential it turns out [7] that

$$B_0 = \frac{36}{5}\frac{V_0}{\omega_0}, \qquad C_0 = (60)^{1/2}. \tag{2.6}$$

We now wish to enquire how the presence of dissipation affects the probability of tunnelling. It is necessary to formulate the precise question we wish to answer rather carefully. If we were treating the motion of the system classically using Eq. (2.1), we would of course produce an equation of motion of the form

$$M\ddot{q} + \frac{dV}{dq} = F_{\text{ext}}(t), \tag{2.7}$$

where for generality we have added a time-dependent "external" potential $-qF_{\text{ext}}(t)$ to the Lagrangian (2.1). Equation (2.7) is, of course, by Ehrenfest's theorem also true as an operator equation in quantum mechanics; however, in general it is not very informative, since when we take its expectation value (indicated by pointed brackets) we cannot identify $\langle \partial V/\partial q \rangle$ with $(\partial V/\partial q)_{q=\langle q \rangle}$. Nevertheless, because of our assumptions that the potential $V(q)$ is reasonably smooth and that the inequality $V_0 \gg \hbar\omega_0$ is satisfied, we may reasonably infer that this identification is indeed approximately valid; therefore a *quantum-mechanical* system, whether trapped in the metastable well or with sufficient energy to surmount the barrier, will have an expectation value $\langle q(t) \rangle$ of its coordinate which approximately satisfies the classical equation of motion (2.7). (Naturally, effects which vanish in the limit $\hbar \to 0$, such as tunnelling itself, are neglected in this approximation.) The (possibly hypothetical) system which behaves

in this way will form, as it were, the reference system with which our real-life (damped) system will be compared.

The real-life system, then, is assumed to obey in the region in question not Eq. (2.7) but a *damped* quasiclassical equation of motion of the form

$$M\ddot{q} + \eta\dot{q} + (dV/dq) = F_{\text{ext}}(t). \tag{2.8}$$

(Note that $F_{\text{ext}}(t)$ is a "true" external force, i.e., such that the work done per unit time is simply $\dot{q}F_{\text{ext}}(t)$ and the power dissipated is correspondingly $\eta\dot{q}^2$. In the more general case $\eta = \eta(q)$ this condition needs to be explicitly specified; see Appendix C.) More precisely, the expectation value of the coordinate operator is assumed to satisfy (2.8) (with $\partial V/\partial q$ interpreted as $(\partial V/\partial q)_{q=\langle q\rangle}$, as above). That the system does indeed obey Eq. (2.8), and the value of the (q-independent) friction coefficient η, is something which in any particular case must be inferred from experiment; for most cases of practical interest there are standard ways of doing this. In writing (2.8) we have made the simplest possible assumption, namely, that the frictional force is simply proportional to velocity, and we shall carry this assumption through the bulk of this paper; however, the generalization to nonlinear dissipation (given by a damping term of the form $\eta(q)\dot{q}$) is very straightforward and is given at appropriate points in the text. The generalization to the case where the dissipation involves higher time derivatives is also mentioned at appropriate points.

Thus, the object of this paper is to compare the tunnelling rate of a system described (under appropriate conditions) by the damped quasi-classical equation of motion (2.8) with that of the reference system described by the undamped Eq. (2.7). It cannot be too strongly stressed that in making the comparison the potential $V(q)$ is assumed to be *the same* in the two cases (cf. below). This is our definition of the meaning of the question "What is the effect of dissipation on quantum tunnelling?" We suspect that the recent controversy about the sign of the effect (see Introduction) may have been at least partly due to the lack of such a generally agreed definition. It is appropriate, therefore, to digress for a moment at this stage to see why there is a problem.

In the next section we are going to introduce dissipation into the behaviour of our system by coupling it to a sufficiently complex environment. Now, if we start with an isolated system and introduce terms in the Lagrangian (or Hamiltonian) which couple it to its environment, then in general we will produce a mechanism of dissipation *and also other effects*. For example, in the simple case in which both system and environment are simple harmonic oscillators, most forms of coupling (and in particular a coordinate–coordinate coupling) will shift the natural oscillation frequency of the system; in quantum-mechanical terms this is just the familiar level-repulsion effect. More generally, we shall find (cf. Section 4) that such coupling tends among other things to renormalize the original potential $V(q)$; in particular, if the coupling is linear in the system coordinate (the case of primary interest in this paper) $V(q)$ acquires an extra (negative) term $\frac{1}{2}M\Delta\omega^2 q^2$, where the negative quantity $\Delta\omega^2$ plays the role of a correction to the squared small-oscillation frequency ω_0^2. Such

renormalization effects can be very large, and if $|\Delta\omega^2| > \omega_0^2$ can even render the original metastable potential minimum unstable; in some microscopic tunnelling systems they can play a very important role (cf. [29]). We refer to these effects as "frequency-renormalization" effects (in a more general context "potential renormalisation" might be a better term).

Now, whether the frequency-renormalization phenomenon does indeed occur as a real physical effect depends crucially on the nature of the physical system considered. This question has of course really nothing to do with tunnelling as such, and to emphasize this we shall illustrate the point with two systems which both perform small oscillations around a *stable* equilibrium position. Consider first an example which is not strictly macroscopic but is well suited to our purpose, the interaction of a collective degree of freedom in a nucleus with the single-particle modes. The simplest description of such an interaction is by a term of the general form $\sum_j F_j(q)\, x_j$, where x_j is some coordinate associated with the single-particle mode j. Clearly this type of coupling will not only provide a dissipation mechanism for the "system" but will also shift its natural frequency; in particular the effect of those parts of the "environment" (most of it) which have natural frequencies large compared to ω_0 may be taken into account by supposing that they adjust adiabatically to the system and thereby lower the effective restoring force, i.e., the quadratic term in the potential energy. Note that in general there is no simple relationship between this "frequency renormalisation" and the damping constant, although of course there are the usual Kramers–Kronig relations between the real and imaginary parts of the complete frequency-dependent system response. In this case the frequency-renormalization effect has a real physical significance.

Now consider a simple LC-circuit; to avoid irrelevant problems we assume that it is constructed entirely of superconducting materials and is at zero temperature. It will execute small undamped harmonic oscillations of the current (or flux) with frequency $\omega_0 = (LC)^{-1/2}$. Suppose now that we connect in parallel with the capacitor a piece of wire made of normal metal of finite conductivity. In elementary electrical engineering terms, we would naturally describe the effect of this element by assigning to it a phenomenological resistance R; we might, depending on the details of the geometry, also have to assign to it its own inductance and/or capacitance, but these are irrelevant to the point at issue and we shall assume they are zero. (More generally, we might have to describe the wire by a complex frequency-dependent impedance $Z(\omega)$; our assumption is that $Z(\omega)$ is real and equal to a constant R for a range of frequencies ω_c large compared to ω_0. The ensuing statements are true to lowest nontrivial order in ω_0/ω_c, cf. Appendix A.) Then the standard elementary circuit calculation shows that the equation of motion of the circuit (where for later convenience we take as our basic variable the magnetic flux Φ passing through it) is modified to

$$C\ddot{\Phi} + \dot{\Phi}/R + \Phi/L = 0. \tag{2.9}$$

Hence the effect of incorporating the resistive element in the circuit is to give the resonance frequency an imaginary part and also, in the case of under-critical

damping, to shift its real part (downwards). This shift, however, should be carefully distinguished from the one occurring in the above case of the nucleus; unlike the latter, it is entirely determined by the phenomenological dissipative coefficient (here R) and does *not* require any modification of the "potential" term in the Lagrangian (here $\Phi^2/2L$). Thus, on elementary physical grounds we do *not* in this case expect the coupling of the circuit to the dissipative element to result in a "frequency-renormalization effect" as we have defined the term. It follows, of course, that in choosing an effective Lagrangian to describe the coupling of the "system" (circuit) to the "environment" (resistor) we must take care to ensure this result. Actually, it turns out that a sufficiently careful analysis of the standard Lagrangian technique as applied to electromagnetic interactions will automatically guarantee this, and we carry out such an analysis in Appendix A.

Another case in which we do not, on physical grounds, expect a frequency-renormalization effect to occur is that of adiabatic coupling. That is, if we start from the dynamics of a system interacting strongly with its environment but described by the zeroth-order adiabatic approximation (which allows no dissipation), and then add as the dissipation-producing interaction the terms omitted in this approximation, then we would expect that these terms do not lead to a frequency-renormalization effect, since any such effect should already have been taken into account in the zeroth-order approximation. We shall demonstrate explicitly in Appendix C that this expectation is correct (to lowest nontrivial order in the departure from adiabaticity).

We see, therefore, that the question "How is quantum tunnelling affected by dissipation?" is not necessarily equivalent to the question "How is quantum tunnelling affected by the interaction with the environment which produces the dissipation?" In considering the latter question it may, depending on the physical nature of the system considered, be necessary to take into account frequency-renormalization effects. Since it is the first question we wish to consider in this paper, it is convenient to be able to treat all cases in a unified way, irrespective of whether or not they show physical frequency-renormalization effects (which in many macroscopic cases of practical interest are likely to be experimentally unobservable even if they occur). The obvious way to do this is to treat the $V(q)$ which appears in Eq. (2.1) not as the original "bare" potential $V_b(q)$ seen by the isolated system, but as the renormalized potential, that is the quantity $V_b(q) - \frac{1}{2}M|\Delta\omega^2|q^2$ (or the appropriate nonlinear generalization, see Appendix C). In cases where there is no physical frequency-renormalization effect, then obviously $V(q) \equiv V_b(q)$. Consequently, our restrictions concerning smoothness, barrier height, etc. should always be understood as referring to the renormalized potential.

To sum up, we wish to compare the tunnelling characteristics of a system whose quasiclassical dynamics is given by the damped equation of motion (2.8) with those of a reference system described by (2.7), *for the same potential function* $V(q)$, (and, of course, the same mass) irrespective of whether or not the potential seen by the real (dissipative) system contains a contribution from frequency-renormalization effects.

There is, however, one respect in which our formulation of the problem is still ambiguous. What, precisely, do we mean by the statement that Eq. (2.8) holds "under

appropriate conditions?" Evidently it should hold provided that the characteristic frequency of motion of the system (as driven by the external force, if necessary) is sufficiently low. But this is still not quite unambiguous, for the following reason. Generally speaking the dissipative mechanism will have associated with it not only some characteristic frequency ω_c (see next section) but also some characteristic *velocity* v_c. We shall see in Appendix C that certain kinds of dissipative mechanism produce reactive terms in the equation of motion (2.8) which are of order $(\omega/\omega_c)(v/v_c)^2$. However low ω, these terms will become important if v (or the amplitude q) is large enough. Thus, if q denotes a typical amplitude of the motion and ω a typical frequency, it is necessary to distinguish two (in general, different) statements: (a) that for fixed q we can find an ω small enough that (2.8) holds without appreciable correction, (b) that we can find an ω small enough that (2.8) holds for any q without appreciable correction. (Statement (b) of course implies (a) but not vice versa.) We shall refer to cases (a) and (b) as cases of "quasi-linear" dissipation and of "strictly linear" dissipation, respectively. (Strictly speaking, of course, one should talk of a quasi-linear/strictly linear *dissipative mechanism*, since the lowest-order terms which distinguish the two cases are reactive in nature.) Since these terms are of a different form from those occurring in the original undamped equation of motion (2.7) (i.e., they cannot in general be mimicked by an adjustment of the potential $V(q)$) the two cases are in principle experimentally distinguishable. However, there may well be some cases of physical interest in which it is in practice possible to verify condition (a) but not condition (b): i.e., we can always verify that the dissipative mechanism is quasi-linear, but not necessarily that it is strictly linear. As we shall see, the principal quantitative results of this paper, and in particular Eq. (4.27), if taken as equalities, are valid only for strictly linear dissipation: for the case where the dissipation is quasi-linear but not strictly linear we get related inequalities (Section 4). Fortunately, for a reason we shall see at the end of Section 3, the condition of strict linearity (or an equivalent condition) does hold, at least approximately, for most of the cases currently of practical interest.

It may be helpful to conclude this section by posing a few questions we should like to answer:

(1) Does the dissipation increase or decrease the tunnelling probability?

(2) Is the effect uniquely determined (for a given potential $V(q)$) by a single parameter, the friction coefficient η, or is it model-dependent?

(3) What is the asymptotic formula for the factor by which the tunnelling probability is multiplied in the weak-damping limit $(\eta/M\omega_0 \to 0)$?

(4) What is the corresponding formula in the heavily overdamped limit $(\eta/M\omega_0 \to \infty)$?

In the next two sections we shall set up a method of answering all these questions.

3. Description of the Mechanism of Dissipation

In our statement of the problem, we have demanded that the (experimentally verified) equation of motion of our system in the quasiclassical metastable regime should be given by Eq. (2.8). We then wish to ask, how does a quantum system described by this dissipative quasiclassical equation of motion behave with respect to quantum tunnelling? To answer this question we must clearly have some description of dissipation in a quantum-mechanical system.

The quantization of a dissipative system moving in a stable potential is of course a very old problem and has been tackled on a number of levels (cf. [38]). In the first place, there have been many attempts to write down modified Schrödinger equations or the equivalent for the system by itself, without explicit reference to any external agency. Examples of such approaches are the time-dependent Hamiltonian theory of Kanai [39], the nonlinear Schrödinger equation used by Kostin [40] and Yasue [41] and the complex quantization procedure of Dekker [42]. These descriptions are usually justified a posteriori by demonstrating that they reproduce known results for certain special cases such as the harmonic oscillator in the limit of weak damping; however, their theoretical foundations are sufficiently unclear that we feel it would be unwise to use them in the present context, which is about as far as it is possible to get from that limit. If, then, we reject such quasi-phenomenological descriptions of the system in isolation, it follows that we must enquire explicitly into the physical mechanism of dissipation. Here again there are (at least) two possible approaches (cf. [43, p. 1995]). On the one hand, we could consider the environment as acting on the system by means of random forces which are specified in some statistical manner. Such a description is of course very widely used in classical statistical mechanics, in the form of the Langevin equation or related equations, and there have been a number of attempts (see [44]) to apply it also to dissipative quantum systems: see especially the work of Koch *et al.* [45]. The alternative approach, which we shall use in this paper, is to regard the "system" and its environment as together forming a closed system (the "universe," as we shall denote it for present purposes) which can be described by a Lagrangian or Hamiltonian, to solve (in principle!) for the motion of the whole and to derive from this solution a description of the properties of the system (which, of course, would now more properly be called a subsystem). In this picture the phenomenon of dissipation is simply the transfer of energy from the single degree of freedom characterising the "system" to the very complex set of degrees of freedom describing the "environment;" it is implicitly assumed that the energy, once transferred, effectively disappears into the environment and is not recovered within any time of physical interest (i.e., one treats the mathematical existence of Poincaré recurrences as physically irrelevant). Formally, one assumes that the number of degrees of freedom of the environment tends to infinity; this assumption is implicit in the replacement of sums by integrals which we shall carry out without further comment at appropriate stages in the calculation. Such an approach to quantum dissipation is of course already widely used in various areas of physics, notably in the

theory of damping of electromagnetic radiation in a cavity [46] and of deep inelastic collisions of nuclei [47].

It is necessary, then, to formulate an explicit Lagrangian for the interacting system and environment in such a way that Eq. (2.8) will result in the appropriate limit. Since, as discussed in the Introduction, we do not always in practice know very much about the physical mechanism of dissipation, it is desirable that the model should be as general as possible subject only to the condition that it generates Eq. (2.8) when required. However, to make the subsequent calculations tractable it is necessary to impose one important restriction on it, namely, that *any one environmental degree of freedom is only weakly perturbed by its interaction with the system.* For most cases of interest, at least when the system variable is macroscopic, this assumption is physically reasonable; in that case the environment is usually also (geometrically) macroscopic and the interaction of the system with any one environmental degree of freedom is generally proportional to the inverse volume, while the characteristic energy of such a degree of freedom is volume-independent. Naturally, the wave functions of the environment may be extremely strongly perturbed by the system over a local region (as in the case of a body moving in a liquid, for example); but this can usually be handled by means of the adiabatic approximation, and the effective residual coupling which is due to the corrections to adiabaticity (see below) then usually does have the required property [47]. It cannot be over-emphasized that the condition that any *one* environmental degree of freedom is only weakly perturbed in no way implies that the interaction is "weak" from the point of view of the system (which interacts with a very large number of degrees of freedom); indeed it is quite compatible with very strong damping.

Now, the motion of any physical system $\mathscr{E}$ which is only weakly perturbed around its equilibrium state can always be adequately represented (at $T=0$ at least) by regarding that system as equivalent to a set of simple harmonic oscillators. Since this point has been given considerable discussion in the literature (see, in particular, [47, Sect. 2]) we shall not spend time on it here, but relegate the details to Appendix C. There we show that the most general type of Lagrangian we need to consider is obtained by adding to (2.1) first the unperturbed Lagrangian of the "environmental" oscillators, namely,

$$L_{\text{osc}} = \sum_j (\tfrac{1}{2} m_j \dot{x}_j^2 - \tfrac{1}{2} m_j \omega_j^2 x_j^2) \tag{3.1}$$

and secondly an interaction term which may be taken *without loss of generality* (barring pathological cases: see Appendix C) to be of the form[7]

$$L_{\text{int}} = -\sum_j F_j(q)\, x_j + \Phi(q) \tag{3.2}$$

[7] Strictly speaking Φ can also be a function of p. This possibility introduces only trivial complications and for the sake of simplicity of presentation we ignore it here.

where the function $\Phi(q)$ may depend on the details of the oscillator spectrum (i.e., on the m_j and ω_j) but not on their dynamical variables $x_j, \dot{x}_j$.

The function $\Phi(q)$ is related to the question of cancellation of the frequency (or potential) renormalization mentioned in Section 2. To see this, we assume for the moment that $\Phi(q)$ is zero and ask what is the minimum value of the potential energy of the "universe" (system plus environment) which can be attained for given q. Clearly to attain this minimum we need to set $x_j = F_j(q)/m_j\omega_j^2$ for all j, and the resulting "effective potential" $V_{\text{eff}}(q)$ is then given by[8]

$$V_{\text{eff}}(q) = V(q) - \sum_j F_j^2(q)/2m_j\omega_j^2. \tag{3.3}$$

In the special case $F_j(q) \propto q$ (cf. below) this is equivalent to a negative shift $\Delta\omega^2$ in the (squared) frequency of small oscillations of the system as described in Section 2. If now we include in the interaction Lagrangian (3.2) a nonzero $\Phi(q)$ given by

$$\Phi(q) \equiv -\sum_j F_j^2(q)/2m_j\omega_j^2, \tag{3.4}$$

then the effect is, trivially, simply to cancel this effect and reduce $V_{\text{eff}}(q)$ to $V(q)$, i.e., we ensure that the system cannot lower its potential energy below the original uncoupled value by moving off the q axis in the many-dimensional space whose axes are q and the x_j's.

In all cases of physical interest we either find that $\Phi(q)$ is given by the expression (3.4) or that it is zero. The first situation arises where there is no physical frequency (or potential)-renormalization effect; we show in Appendix A that this is so for electromagnetic coupling of the type found (e.g.) in SQUIDs, and also in the case of adiabatic coupling which is formally identical to it. The second situation is characteristic of cases where the frequency-renormalization effect is real, e.g., the nuclear physics example discussed in Section 2. Nevertheless, for the reason discussed in Section 2 it is convenient also in this case to add a term of the form (3.4) to L_{int} while subtracting it from the "unperturbed" Lagrangian (2.1); in this way we achieve our goal of comparing damped and undamped systems for the same observable potential (or frequency).

Thus, the most general Lagrangian with which we need to deal when considering a system weakly coupled (in the sense specified above) to its environment and whose phenomenological equation of motion is (2.8) is of the form

$$\begin{aligned} L = {} & \tfrac{1}{2}M\dot{q}^2 - V(q) + \tfrac{1}{2}\sum_j (m_j\dot{x}_j^2 - m_j\omega_j^2 x_j^2) \\ & - \sum_j F_j(q)\, x_j - \sum_j F_j^2(q)/2m_j\omega_j^2. \end{aligned} \tag{3.5}$$

[8] In a nuclear-physics context this is often known as the "adiabatic potential."

The calculations of Section 4 go through for any Lagrangian of this general form (see the end of that section). However, to render the results useful we clearly have to know something about the distribution of the quantities m_j, ω_j, and $F_j(q)$. In considering any specific physical system we can, in principle, approach this problem in either of two ways.

First, we might have sufficient confidence in a particular microscopic model of the system in question to feel sure we know the relevant parameters a priori. Such a situation might, for example, arise in the case of tunnelling transitions in a ring laser, where the interactions should be obtainable directly from quantum electrodynamics, or in an ideal oxide-layer Josephson tunnel junction, where the resistive mechanism is intrinsic and believed to be well described in terms of Bogoliubov quasiparticles [48]. In such cases the problem reduces to transforming the known Lagrangian into a representation of the form (3.5).

The second possibility—which is the one emphasized in our formulation of the problem in Section 2—is that we might have a phenomenological description of the *classical* motion of the system, without necessarily having a particular microscopic model as the basis for it. In this case the interesting question is whether this knowledge of the classical motion will determine the parameters entering L to the extent necessary to make useful predictions about tunnelling rates. (We shall see below that the answer is yes.) Of course intermediate cases are also possible: we might have a classical phenomenological description plus *some* constraints on possible microscopic models. An example is the case of a SQUID with its Josephson junction physically shunted by a normally resistive piece of wire: here we have one important piece of a priori knowledge, namely, that the coupling to the resistor must be linear in the flux (see Appendix A), i.e., that in (3.5) we can put $F_j(q) = qC_j$; on the other hand we certainly do not know the C_j's, m_j's, and ω_j's in detail except insofar as they enter the (experimentally measurable) resistance.

Let us therefore ask what constraints are imposed on the choice of the parameters m_j, ω_j, and $F_j(q)$ in (3.5) by a knowledge that the classical motion of the system is governed by a dissipative equation of the form (2.8) (the case of more general dissipative equations is mentioned below). This question is discussed in Appendix C, with the following conclusions. If we know that the dissipative mechanism is quasilinear but not that it is strictly linear (see the end of Section 2), then the only constraint we can impose is the relation

$$\frac{\pi}{2} \sum_j \frac{1}{m_j \omega_j^2} \left(\frac{\partial F_j}{\partial q} \right)^2 \delta(\omega - \omega_j) = \eta \tag{3.6}$$

for $\omega \ll \omega_c$, where ω_c is the characteristic frequency at which the phenomenological equation (2.8) begins to break down, and η is the phenomenological friction coefficient appearing in (2.8). Equation (3.6) is compatible with (e.g.) a distribution of $F_j(q)$'s which are each individually of very short range in q; certainly they need not be linear in q. If on the other hand we require that the dissipation be *strictly* linear, we find that we *must* have

$$F_j(q) = qC_j \tag{3.7}$$

so that (3.6) then becomes a constraint on the spectral density $J(\omega)$ defined by

$$J(\omega) \equiv \frac{\pi}{2} \sum_j (C_j^2/m_j\omega_j)\, \delta(\omega - \omega_j), \tag{3.8}$$

namely,

$$J(\omega) = \eta\omega. \tag{3.9}$$

We will see in the next section that while the constraint (3.9) is sufficient to determine the exact tunnelling rate, the weaker constraint (3.6) is adequate only to determine a lower bound on it. Naturally, there are cases where we can assert (3.7), and hence (3.9), not because we know from experiment that the dissipation is strictly linear but because of some a priori knowledge of the interactions involved: see, e.g., the SQUID example discussed above. Since (3.7) (with (3.6)) implies strictly linear dissipation as well as vice versa, we shall refer to such cases also as cases of strictly linear dissipation.

The simple classical dissipative equation of motion (2.8), with its single amplitude- and frequency-independent friction coefficient η, is of course by no means the most general classical equation which can describe dissipation. In the most general case there is little[9] that can be usefully said about the parameters entering the Lagrangian (3.5). However, it is worth considering a case which, while not the most general possible, is nevertheless considerably more general than (2.1), namely, that in which the coefficient η is allowed to depend on amplitude: $\eta = \eta(q)$. For this case we show in Appendix C that the parameters are constrained by a relation which is the obvious generalization of (3.6)

$$\frac{\pi}{2} \sum_j \frac{1}{m_j\omega_j^2} \left(\frac{\partial F_j}{\partial q}\right)^2 \delta(\omega - \omega_j) = \eta(q). \tag{3.10}$$

The special case of a "separable" interaction, namely, one satisfying the condition

$$F_j(q) = C_j f(q) \tag{3.11}$$

with $f(q)$ independent of j, is of particular interest. In Section 4 we shall show that if (3.11) is satisfied, we can derive an exact expression for the tunnelling rate in terms of $\eta(q)$; if it is not, we can derive only a lower bound (this result exactly parallels what happens in the linear case).

In the rest of this paper we shall concentrate primarily on the case of strictly linear

[9] More accurately, there is plenty we can say but it is too complicated to be worth saying (cf. Appendix C).

dissipation, i.e., the case in which (3.7) holds. In other words, we shall usually take as our Lagrangian the specific form of (3.5) obtained by inserting (3.7)

$$L(q, \dot{q}: \{x_j, \dot{x}_j\}) = \tfrac{1}{2}M\dot{q}^2 - V(q) + \tfrac{1}{2}\sum_j m_j(\dot{x}_j^2 - \omega_j^2 x_j^2)$$
$$- q\sum_j C_j x_j - \tfrac{1}{2}M|\Delta\omega^2|\, q^2, \tag{3.12}$$

where the "frequency shift" $\Delta\omega^2$ (cf. Section 2) is defined by

$$\tfrac{1}{2}M\,\Delta\omega^2 \equiv -\sum_j C_j^2/2m_j\omega_j^2. \tag{3.13}$$

Moreover, we shall usually assume that the dissipation is not only strictly linear but also frequency-independent in the region of interest for tunnelling (see Section 4), i.e., that Eq. (3.9) holds for all frequencies of interest to us.

It might at first sight be thought that the model so defined is of rather limited interest. We believe, however, that in the context of the real-life macroscopic quantum tunnelling problem it will cover the vast majority of cases of interest to us. Let us consider separately the question of strict linearity (i.e., Eq. (3.7)) and of frequency-independence (Eq. (3.9)). With regard to the first, we note (a) that in many cases (cf. the examples quoted above) we know from general a priori considerations that the coupling must indeed have the form (3.7), and (b) that even when it does not, we often know enough about it to be sure that, under the conditions of a real-life tunnelling experiment, Eq. (3.7) is not a bad approximation. To illustrate the latter point, let us consider an ideal tunnel oxide Josephson junction. According to the standard microscopic model of such a system [48],[10] the dissipation in this case arises from the normal quasiparticles present, and depends on the phase difference $\Delta\varphi$ of the condensate on the two sides of the junction because the parameters of the Bogoliubov transformation which defines the quasiparticle eigenstates are periodic functions of the condensate phase. Thus, quite independently of the details of the formulation of the problem in the language of Eq. (3.5) (which is, indeed, not an entirely trivial operation) we expect that the quantities $F_j(q)$ will have a range in q (i.e., $\Delta\varphi$) which is of order (say) $\pi/2$. Since most experiments to look for macroscopic quantum tunnelling in Josephson systems have operated[11] (and are likely to continue to operate) in a regime such that the width of the potential barrier $U(\Delta\varphi)$ (Eq. (1.3)) along the $\Delta\varphi$ axis is small compared to 1, we see that to approximate the true coupling by the linear form (3.7) should involve little error. As to the question of frequency-independence, we note that we can very often state on the basis of

[10] Of course for the $T = 0$ case studied in this paper this model predicts zero dissipation. The ensuing remarks therefore apply, strictly speaking, only to a possible finite-temperature generalization of our results.

[11] We note that experiments to look for macroscopic quantum coherence in such systems, if they are ever to have any chance at all of success, will almost certainly have to operate in this regime; see [2, Sect. 5].

experiments conducted in the classically accessible regime that the relevant coefficient (junction conductance, etc.) is approximately independent of frequency; and that in any case the modification to our results necessary to allow for frequency-dependence is trivial (see Section 4) and should not affect them qualitatively. Thus, while we shall give results for more general cases, we make no apology for concentrating from now on primarily on the model described by Eqs. (3.7) and (3.9).

4. A Formula for the Tunnelling Rate

Our principal goal in this section is to calculate, for a system described by the Lagrangian (3.12), the rate of tunnelling out of a metastable minimum of the potential $V(q)$. In principle it should be possible to do this by any of a variety of methods; for example, some features of the results may be straightforwardly obtained by application of the "many-dimensional WKB method" [50] (see end of this section). However, the most convenient method for our purposes is a generalization of the "instanton" technique which was originally formulated, in the context of a discussion of classical thermodynamic metastability, by Langer [51] and has been applied to the calculation of the decay of metastable states in field theory by Stone [52], Callan and Coleman [53], and many subsequent authors. We will therefore start with a brief review of the use of the technique for noninteracting systems; the discussion follows closely the paper of Callan and Coleman [53], to which the reader is referred for further details (see also [54]).

We consider a one-dimensional system described by a coordinate $q(\tau)$ and a Lagrangian

$$L = \tfrac{1}{2}M\dot{q}^2 - V(q), \tag{4.1}$$

where we assume that there are no velocity-dependent forces in the problem. The potential $V(q)$ is assumed to have a local minimum, taken as zero of potential, at the point $q = 0$. It can be shown (see, e.g., Feynman and Hibbs [55, Sect. 10.2]) that the density matrix

$$\rho(q_i, q_f; \beta) \equiv \sum_n \psi_n^*(q_i)\, \psi_n(q_f) \exp -\beta E_n \tag{4.2}$$

(where as usual β^{-1} is Boltzmann's constant times the temperature) can be represented as a path integral

$$\rho(q_i, q_f; T) = \int_{q(0)=q_i}^{q(T)=q_f} \mathscr{D}q(\tau) \exp\left(-\int_0^T L_E(q, \dot{q})\, d\tau/\hbar\right), \tag{4.3}$$

where $T \equiv \beta\hbar$ and the "Euclidean" Lagrangian $L_E(q, \dot{q})$, which in the present case is identical to the Hamiltonian, is given by

$$L_E(q, \dot{q}) \equiv \tfrac{1}{2}M\dot{q}^2 + V(q). \tag{4.4}$$

As is explicitly indicated in Eq. (4.3), the path integral runs over all paths taken by the coordinate q as a function of the (imaginary) "time" τ which leave the point q_i at time zero and arrive at q_f at time T. By taking the limit $T \to \infty$ in Eq. (4.3) we obtain an expression for the wave functions and energy of the groundstate. Although formula (4.3) is quite general, we shall consider for the moment only the special case $q_i = q_f = 0$.

In the semiclassical limit ($\hbar \to 0$) the functional integral is dominated by the paths for which $L_E(q, \dot{q})$ is an extremum. These are precisely the paths which describe the allowed motions of the classical particle in the *inverted* potential $\tilde{V}(q) \equiv -V(q)$. If the local minimum at $q = 0$ were an absolute minimum, then the only possible "classical" path for which $q(0) = q(T) = 0$ would be the trivial one in which the particle sits at the origin for all τ. The small fluctuations around this path then give the correct expression for $\rho(0, 0: T)$, namely,

$$\rho(0, 0: T) = |\psi_0(0)|^2 \exp -\omega_0 T/2, \tag{4.5}$$

where ω_0 is the small oscillation frequency; the exponential incorporates the effect of the finite ground-state energy, $\hbar\omega_0/2$.

If the local minimum of $V(q)$ at $q = 0$ is not an absolute minimum (see Fig. 1), then a second type of classical path in the inverted potential $\tilde{V}(q)$ is possible; this is the type called a "bounce" (instanton) by Callan and Coleman, in which the particle starts to roll off the local maximum of $\tilde{V}$ at $\tau = 0$, moves across to the point $q = q_0$ (which corresponds to the exit point B in Fig. 1), turns around there and returns by time T to the origin. We note that such a classical motion is still possible in the limit $T \to \infty$ because the point $q = 0$ is a local maximum of $\tilde{V}(q)$; at the beginning and end of the motion $q(\tau)$ approaches zero exponentially. (That is, if we shift the origin of time so that $q(0) = q_0$, then for large $|\tau|$ we have $q(\tau) = \text{const} \exp -\omega_0 |\tau|$.) Such a "bounce" type of trajectory corresponds not to a minimum but to a saddlepoint of the Euclidean action, and it therefore contributes a small imaginary part to the groundstate energy (for the argument on this point, see Callan and Coleman [53]). If B denotes the action of the bounce in the limit $T \to \infty$, that is,

$$B \equiv \int_{-\infty}^{\infty} L_E(q_{\text{cl}}(\tau), \dot{q}_{\text{cl}}(\tau))\, d\tau, \tag{4.6}$$

where $q_{\text{cl}}(\tau)$ denotes the classical bounce trajectory, then the formula for the tunnelling decay rate Γ derived in this way reads

$$\Gamma = \mathcal{A}^{-1/2}(B/2\pi\hbar)^{1/2} \exp(-B/\hbar) \times (1 + \mathcal{O}(\hbar)), \tag{4.7}$$

where Δ, a quantity with the dimensions of (frequency)$^{-2}$, is the ratio of two determinants and does not contain $\hbar$; in general it is of order ω_0^{-2}. We do not elaborate these results further here as they will emerge later as a special case of the results to be derived below.

We shall now generalize this technique to the problem of interest to us. For simplicity we shall deal explicitly with the case of strictly linear friction (the extension to the more general case is straightforward and is given at the end of this section); we therefore start from the Lagrangian (3.12), that is,[12]

$$L = \tfrac{1}{2}M\dot{q}^2 - V(q) + \sum_\alpha (\tfrac{1}{2}m_\alpha \dot{x}_\alpha^2 - \tfrac{1}{2}m_\alpha \omega_\alpha^2 x_\alpha^2)$$
$$- q\sum_\alpha C_\alpha x_\alpha - \tfrac{1}{2}M|\Delta\omega^2|\, q^2, \tag{4.8}$$

where the "frequency shift" $\Delta\omega^2$ appearing in the last term is given by Eq. (3.13). As discussed in Section 2, the potential $V(q)$ is assumed to have a local minimum, taken as the zero of potential, at the origin and to become zero again at some finite value q_0 of q, beyond which it is always negative. We define the density matrix of the "universe" (system plus environment) by

$$\rho(q_i, \{x_{\alpha i}\}: q_f, \{x_{\alpha f}\}: \beta) \equiv \sum_n \psi_n^*(q_i, \{x_{\alpha i}\})\, \psi_n(q_f, \{x_{\alpha f}\}) \exp -\beta E_n \tag{4.9}$$

(where $x_{\alpha i}, x_{\alpha f}$ are "initial" and "final" values of the coordinate x_α) and the *reduced* density matrix of the system by

$$K(q_i, q_f: \beta) \equiv \int \prod_\alpha dx_{\alpha i} \rho(q_i, \{x_{\alpha i}\}: q_f, \{x_{\alpha i}\}: \beta). \tag{4.10}$$

Thus, introducing $T \equiv \beta\hbar$ as above, we have

$$K(q_i, q_f: T) \equiv \int \prod_\alpha dx_{\alpha i} \sum_n \psi_n^*(q_i, \{x_{\alpha i}\})\, \psi_n(q_f, \{x_{\alpha i}\}) \exp(-E_n T/\hbar). \tag{4.11}$$

It is clear that, as in the one-particle case, the imaginary part of the grounddstate energy will give the tunnelling rate out of the metastable state, and may be obtained from a study of $K(0, 0: T)$ in the limit $T \to \infty$. It must be strongly emphasized that, despite what one might at first sight be tempted to infer from the form of (4.10), the tunnelling rate so calculated is the *total* rate of tunnelling out of the metastable groundstate, irrespective of whether the tunnelling results in real excitation of the environment or not (i.e., whether the process is "elastic" or "inelastic").

Because the Lagrangian (4.8) contains no velocity-dependent terms, the full density

[12] To avoid possible notational confusion we now label the different degrees of freedom of the environment by the Greek subscript α.

matrix (4.9) can be written as a path integral in the usual way, and hence K can be written

$$K(q_i, q_f; T) = \int \prod_\alpha dx_{\alpha i} \int_{q(0)=q_i}^{q(T)=q_f} \mathscr{D}q(\tau) \prod_\alpha \int_{x_\alpha(0)=x_{\alpha i}}^{x_\alpha(T)=x_{\alpha i}} \mathscr{D}x_\alpha(\tau)$$
$$\times \exp\left(-\int_0^T L_E(q, \dot{q}; \{x_\alpha, \dot{x}_\alpha\})\, d\tau/\hbar\right), \tag{4.12}$$

where the "Euclidean" Lagrangian is given by

$$L_E(q, \dot{q}; \{x_\alpha, \dot{x}_\alpha\}) \equiv \tfrac{1}{2}M\dot{q}^2 + V(q) + \tfrac{1}{2}\sum_\alpha m_\alpha(\dot{x}_\alpha^2 + \omega_\alpha^2 x_\alpha^2)$$
$$+ q\sum_\alpha C_\alpha x_\alpha + \tfrac{1}{2}M|\Delta\omega^2|\, q^2. \tag{4.13}$$

Now, as the result of the fact that only linear and quadratic powers of x_α and $\dot{x}_\alpha$ occur in the Lagrangian, the functional integrations over the paths $\{x_\alpha(\tau)\}$ and then the integrations over the coordinates $\{x_{\alpha i}\}$ can be performed in closed form. The procedure, while somewhat tedious, is of course entirely straightforward and we simply quote the standard result ([56, p. 82]: put $if \equiv Cq$)

$$\int dx_i \int_{x(0)=x_i}^{x(T)=x_i} \mathscr{D}x(\tau) \exp -\frac{1}{\hbar}\left\{\int_0^T |\tfrac{1}{2}m(\dot{x}^2 + \omega^2 x^2) + xCq(\tau)|\, d\tau\right\} \equiv Q(T)$$
$$= I(0) \exp\left\{+\frac{C^2}{4m\hbar\omega}\int_0^T d\tau \int_0^T d\tau' \frac{q(\tau)\, q(\tau') \cosh \omega(|\tau - \tau'| - T/2)}{\sinh \omega T/2}\right\}, \tag{4.14}$$

where $I(0)$ is the value of the expression for $C = 0$, which by Eqs. (4.3) and (4.2) is just $\frac{1}{2}$ cosech $\omega T/2$. This expression can be formally simplified somewhat if we agree to define $q(\tau)$ outside the range $0 \leqslant \tau < T$ by the prescription[13] $q(\tau + T) \equiv q(\tau)$; then Eq. (4.14) becomes

$$Q(T) = \tfrac{1}{2}\,\mathrm{cosech}(\omega T/2) \exp\left\{\frac{C^2}{4m\hbar\omega}\int_{-\infty}^{\infty} d\tau' \int_0^T d\tau \exp(-\omega|\tau - \tau'|)\, q(\tau)\, q(\tau')\right\}. \tag{4.15}$$

Substituting this result into (4.12), we find

$$K(q_i, q_f; T) = K_0(T) \int_{q(0)=q_i}^{q(T)=q_f} \mathscr{D}q(\tau)(\exp -S_0/\hbar) \exp \Lambda/\hbar, \tag{4.16}$$

[13] We note, although it is not important for the present problem, that this introduces some complications when we wish to apply Eq. (4.23) to calculate $K(q_i, q_f; \beta)$ for $q_i \neq q_f$; cf. Appendix B.

where

$$K_0(T) \equiv \prod_\alpha (\tfrac{1}{2} \operatorname{cosech} \omega_\alpha T/2), \tag{4.17}$$

$$S_0[q(\tau)] \equiv \int_0^T \{\tfrac{1}{2}M\dot{q}^2 + V(q)\}\, d\tau, \tag{4.18}$$

$$\Delta[q(\tau)] \equiv \int_0^T \tfrac{1}{2}M\,\Delta\omega^2 q^2(\tau)\, d\tau + \sum_\alpha \left\{ \frac{C_\alpha^2}{4m_\alpha \omega_\alpha} \int_{-\infty}^{\infty} d\tau' \int_0^T d\tau \right.$$
$$\left. \times \exp(-\omega_\alpha |\tau - \tau'|)\, q(\tau)\, q(\tau') \right\}. \tag{4.19}$$

The form of $\Delta[q(\tau)]$ can be simplified if we use the identity

$$q(\tau)\, q(\tau') \equiv \tfrac{1}{2}\{q^2(\tau) + q^2(\tau') - (q(\tau) - q(\tau'))^2\}. \tag{4.20}$$

Substituting this into (4.19), we integrate the term in $q^2(\tau)$ over τ and vice versa.[14] Then, using Eq. (3.13), we see that these two terms exactly cancel the first term in Δ. Thus, we obtain

$$\Delta = -\tfrac{1}{2} \int_{-\infty}^{\infty} d\tau' \int_0^T d\tau\, \alpha(\tau - \tau')\{q(\tau) - q(\tau')\}^2, \tag{4.21}$$

where the quantity $\alpha(\tau - \tau')$ is defined by

$$\alpha(\tau - \tau') \equiv \sum_\alpha \frac{C_\alpha^2}{4m_\alpha \omega_\alpha} \exp -\omega_\alpha |\tau - \tau'|$$
$$\equiv \frac{1}{2\pi} \int_0^\infty J(\omega) \exp(-\omega |\tau - \tau'|)\, d\omega \geqslant 0 \tag{4.22}$$

with the spectral density $J(\omega)$ defined by Eq. (3.8). Thus, finally, the reduced density matrix of the system can be written in the fairly compact form

$$K(q_i, q_f; T) = K_0(T) \int_{q(0)=q_i}^{q(T)=q_f} \mathscr{D}q(\tau) \exp(-S_{\text{eff}}[q(\tau)]/\hbar), \tag{4.23}$$

[14] Since Eq. (4.19) is not formally symmetric with respect to interchange of τ and τ', it is easiest to perform this maneouvre by going over to the form of (4.19) obtained from the explicitly symmetric expression (4.14) rather than (4.15).

where the effective action $S_{\text{eff}}[q(\tau)]$ is given by the expression

$$S_{\text{eff}}[q(\tau)] \equiv \int_0^T \{\tfrac{1}{2}M\dot{q}^2 + V(q)\}\, d\tau + \tfrac{1}{2}\int_{-\infty}^{\infty} d\tau' \int_0^T d\tau\, \alpha(\tau - \tau')\{q(\tau) - q(\tau')\}^2 \qquad (4.24)$$

with the positive quantity $\alpha(\tau - \tau')$ given by (4.22).

The result (4.23), with the definitions (4.24) and (4.22), is applicable to a wide range of problems connected with tunnelling (or more generally classically forbidden behaviour) in a dissipative system. We emphasize again that in evaluating the double integral in (4.24) it is necessary to bear in mind that $q(\tau)$ is to be continued outside the range $0 \leqslant \tau < T$ by the prescription $q(\tau + T) = q(\tau)$; thus, in general considerable care must be taken in going to the limit $T \to \infty$. This point is essential when we evaluate $K(q_i, q_f; T)$ for q_i and/or q_f nonzero (cf. Appendix B, where this is shown explicitly for the case of the damped simple harmonic oscillator). Fortunately, in the case of a bounce, where $q(\tau)$ tends to zero sufficiently fast at both ends of the interval $0 \leqslant \tau < T$ (cf. below), the additional terms arising from the region outside this interval tend to zero as T tends to infinity: we will therefore be able to take the limit $T \to \infty$ without any problem. We note at this point that a result somewhat similar in appearance to (4.24) has been given in the work of Sethna [29].

From Eqs. (4.22)–(4.24) we can immediately draw one important qualitative conclusion: since $\alpha(\tau - \tau')$ is positive definite, the contribution of the last term in (4.24) is always positive and therefore, to the extent that the tunnelling probability is dominated by the saddlepoint value of (4.24) (i.e., in the WKB limit—see below) the existence of dissipation always tends to suppress tunnelling.[15] Note that this result is completely independent of the form of the spectral density $J(\omega)$. This answers the first question posed in Section 2.

To proceed further we need to invoke the assumption that the characteristic frequency scale ω_c (see Section 3) over which the spectral density $J(\omega)$ is equal to the low-frequency form $\eta\omega$ to a good approximation is much larger than the characteristic frequencies of the problem. In the present context the relevant characteristic frequency is the inverse of the "bounce time" (characteristic length of the bounce); in the next section we shall define this quantity more precisely and show that for the cubic potential explicitly considered there it is of order ω_0^{-1} and $\gamma\omega_0^{-2}$ for the underdamped and overdamped regimes, respectively. It is clear that similar order-of-magnitude estimates will be valid for any potential $V(q)$ whose shape is not too pathological. Thus, for a macroscopic system (and for some microscopic systems also) the condition is likely to be very well fulfilled. As an example, $J(\omega)$ might have the Drude form

$$J(\omega) = \frac{\eta\omega}{1 + \omega^2\tau_D^2} \qquad (4.25)$$

[15] It is at first sight tempting, but incorrect, to conclude more generally that $K(q_i, q_f; T)$ is always smaller for the damped system than for the undamped one. This is connected with both the choice of zero of energy and the normalization of the functional integral, which fortunately we do not need to go into here.

and the condition is then obviously fulfilled provided that ω_0^{-1} is much greater than the Drude relaxation time τ_D.

It is now evident that the main contribution to the last term in Eq. (4.24) will come from values of $|\tau - \tau'|$ which are of the order of the bounce time: the region of small $|\tau - \tau'|$ gives no specially important contribution, since $\alpha(\tau - \tau')$ tends to a finite value[16] for $|\tau - \tau'| \to 0$. It then follows that for the purpose of evaluating (4.24) it is legitimate to replace $\alpha(\tau - \tau')$ by its asymptotic form for "large" $|\tau - \tau'|$; that is, we put

$$\alpha(\tau - \tau') \equiv \frac{1}{2\pi}\int_0^\infty J(\omega)\, e^{-\omega|\tau-\tau'|}\, d\omega \cong \frac{1}{2\pi}\int_0^\infty \eta\omega e^{-\omega|\tau-\tau'|}\, d\omega$$

$$= \frac{\eta}{2\pi}\frac{1}{(\tau - \tau')^2}. \tag{4.26}$$

The error incurred by substituting (4.26) into (4.24) will be at most of the order of ω_0/ω_c, where ω_c is the characteristic scale of structure of $J(\omega)$. It is easy to convince oneself (cf. Section 5) that in the context of the tunnelling problem this substitution does not lead to any divergences. In the more general case, and in particular if we wish to calculate $K(q_i, q_f; T)$ for $q_i \neq q_f$, we may find that it leads to a logarithmic divergence in certain integrals (cf. Appendix B); in any such case we must of course return to the true form of $\alpha(\tau - \tau')$ (or $J(\omega)$), which in general will be equivalent to cutting off the integral at an upper limit of order ω_c. Subject to this proviso, it is easy to see that the substitution (4.26) is quite generally valid (not just in the context of the tunnelling problem) provided only that T is not too small compared to ω_0^{-1}; the relevant characteristic "times" in the functional integral are then automatically of order ω_0^{-1} or greater. We note that after the substitution (4.26) the whole functional integral, and hence a fortiori the tunnelling probability, is a function (apart from the mass M and potential $V(q)$) only of the friction coefficient η; this answers the second question posed in Section 2. (It should be stressed, however, that this feature is peculiar to the strictly linear case: cf. below).

With the above provisos, therefore, the effective action which enters the functional integral (4.23) is given by the simple expression

$$S_{\text{eff}}\{q(\tau)\} = \int_0^T [\tfrac{1}{2}M\dot{q}^2 + V(q)]\, d\tau + \frac{\eta}{4\pi}\int_{-\infty}^{\infty} d\tau \int_0^T d\tau' \frac{\{q(\tau) - q(\tau')\}^2}{(\tau - \tau')^2}. \tag{4.27}$$

From here on the argument runs closely parallel to that developed by Callan and Coleman [53] for the undamped case. First, we can verify explicitly that substitution of (4.27) in (4.23) gives the same result for the density matrix in the harmonic region

[16] Actually, for the simple Drude model (4.25) the quantity $\alpha(\tau - \tau')$ diverges as $\ln|\tau - \tau'|$ for small values of its argument. This divergence clearly does not invalidate the argument, since it is integrated over.

near the metastable minimum at $q = 0$ as is given by more elementary arguments; this is done in Appendix B, where it is shown that for $q_i \neq q_f$ it is essential to take account of the periodicity condition $q(\tau + T) = q(\tau)$. Secondly, we can establish the existence of a "bounce" solution as in the undamped case; this follows simply from the fact that for sufficiently small $q(\tau)$ all the terms in the effective action $q(\tau)$ are positive and increasing functions of q, while for large q the quantity S_{eff} can be made negative and arbitrarily large by staying long enough in the region $q > q_0$, where the potential is negative. Thus, there must be at least one saddlepoint separating the two types of solutions; this is the "bounce." (A formal proof of the existence of a bounce is given for the quadratic-plus-cubic potential in the next section.) As in the undamped case, the bounce solution must satisfy an "equation of motion" which now contains a term nonlocal in time

$$M\ddot{q} = \frac{\partial V}{\partial q} + \frac{\eta}{\pi}\int_{-\infty}^{\infty} d\tau' \frac{(q(\tau) - q(\tau'))}{(\tau - \tau')^2}, \tag{4.28}$$

where the integral is to be interpreted as its principal value. We see that the function $q(\tau)$ is symmetric around the value of τ at which it reaches its maximum amplitude.[17] Let us now specialize to the case $q_i = q_f = 0$, let T become very large and redefine the origin of time to lie at $T/2$ (so that those integrals in (4.27) which are not infinite in range run from $-T/2$ to $T/2$). Consider a single bounce which has its centre (point of maximum amplitude) at $\tau = 0$. We expect intuitively that "most of" the bounce will be contained in a region of time around $\tau = 0$ which is of order ω_0^{-1}, γ^{-1} or some combination of them ($\gamma \equiv \eta/2M$); this expectation is confirmed in the next section. For large $|\tau|$ the bounce amplitude does not fall off exponentially as in the undamped case; in fact it can be seen from (4.28) that in this limit we have

$$q(\tau) \cong Cq_0/\omega_0^2\tau^2, \tag{4.29}$$

where the dimensionless coefficient C is of the order of γ times the "length" of the bounce. Fortunately, this difference has no great effect on the problem provided we are prepared to let T become arbitrarily large: in particular, the contribution to the integral over $d\tau$ in (4.27) from the region $|\tau| > T/2$ (where, we recall $q(\tau)$ has to be continued periodically) is negligible in this limit, so that we can let $q(\tau)$ equal zero for $|\tau| > T/2$ without appreciable error.

Now consider the possibility of a series of bounces between $-T/2$ and $T/2$ with widely separated centres. There is now a further difference with the undamped case: not only does $q(\tau)$ approach zero between the bounces only as a power law, not exponentially, but there is an effective "attractive potential" between two bounces which is proportional to the inverse square of the distances between their centres; this arises from the last term in (4.27). Again, fortunately nothing is lost by ignoring this

[17] This is not immediately obvious from (4.28) itself: it follows, however, when we minimize S as a functional of the Fourier transform $q(\omega)$ (cf. next section).

effect and treating the bounces as forming a noninteracting "dilute gas." To see this, let us consider for convenience the case where γ and ω_0 are of the same order of magnitude, so that the "length" of the bounce is of order ω_0^{-1} (the generalization is straightforward). Then the contribution to the action (4.27) of the interaction[18] of two bounces whose centres are separated by τ, where $\omega_0^{-1} \ll \tau \ll T$, is of order $-B/\omega_0^2\tau^2$, where B is the action of an isolated bounce. On the other hand the mean "density" ρ of bounces on the τ axis (defined by the statement that the contribution of n bounces to the functional integral (4.27) is proportional to $(\rho T)^n/n!$) is seen from the formula below to be of order $\omega_0(B/2\pi\hbar)^{1/2} \exp -B/\hbar$; it is, in fact, nothing but the tunnelling rate Γ, to within factors of order unity. Thus the bounces do indeed effectively form a dilute gas, with mean spacing of order Γ^{-1}. Thus the correction to the functional integral due to the inter-bounce interaction is at most of order Γ/ω_0 relative to the "free" contribution: since this factor is at most of the order of the exponentially small quantity $e^{-B/\hbar}$, it must for consistency be neglected in the WKB limit.[19]

Thus, the situation is for all practical purposes exactly analogous to that which holds in the undamped case, the sole difference being that the action of a single bounce contains the additional nonlocal term in η (Eq. (4.27)). From here on we can simply take over the argument of Callan and Coleman [53]; all considerations relating to the summation over multi-bounce configurations, analytic continuation of the expressions into the complex plane, etc., are identical to those for the undamped case. (The details are given in [7]). The final result for the decay rate Γ is the following:

$$\Gamma = A \exp(-B/\hbar), \tag{4.30}$$

where the bounce exponent B and prefactor A are given by the following expressions:

$$B \equiv \int_{-\infty}^{\infty} \{\tfrac{1}{2}M\dot{q}^2 + V(q)\}\, d\tau + \frac{\eta}{4\pi}\int_{-\infty}^{\infty} d\tau \int_{-\infty}^{\infty} d\tau' \left(\frac{q(\tau) - q(\tau')}{\tau - \tau'}\right)^2, \tag{4.31}$$

where the integral is evaluated along the "classical" bounce path, that is, for the function $q(\tau)$ which satisfies the conditions $q(-\infty) = q(\infty) = 0$ and corresponds to the saddlepoint value of the right-hand side of (4.30) (regarded as a functional of $q(\tau)$). (In the limit $\eta \to 0$, this trajectory is just that corresponding to classical motion in the *inverted* potential $\tilde{V}(q) \equiv -V(q)$). The prefactor A is given, to lowest order in $\hbar$, by the expression

$$A = (B/2\pi\hbar)^{1/2} \left| \left(\frac{\det \hat{\mathscr{D}}_0}{\det' \hat{\mathscr{D}}_1}\right) \right|^{1/2}, \tag{4.32}$$

[18] That is, the term arising from the last term in (4.27).

[19] It should be carefully noted that these arguments are *not* valid in the case of a discussion of quantum *coherence* (see Section 6).

where the differential operators $\hat{\mathscr{D}}_0$, $\hat{\mathscr{D}}_1$ are given, respectively, by

$$\hat{\mathscr{D}}_0 q(\tau) \equiv \left(-\frac{d^2}{d\tau^2} + \omega_0^2\right) q(\tau) + \frac{\eta}{\pi M} \int_{-\infty}^{\infty} \frac{q(\tau) - q(\tau')}{(\tau - \tau')^2} d\tau', \tag{4.33}$$

$$\hat{\mathscr{D}}_1 q(\tau) \equiv \left(-\frac{d^2}{d\tau^2} + M^{-1} V''[q_{\mathrm{cl}}(\tau)]\right) q(\tau) + \frac{\eta}{\pi M} \int_{-\infty}^{\infty} \frac{q(\tau) - q(\tau')}{(\tau - \tau')^2} d\tau'. \tag{4.34}$$

Here $q_{\mathrm{cl}}(\tau)$ is the classical bound path as described above, the integral is to be taken as its principal part, and the prime in Eq. (4.32) indicates that the zero eigenvalue, which corresponds to the uniform translation of the bounce along the τ axis, is to be omitted. Equations (4.30)–(4.35) constitute, in principle, the solution to our problem.

We will now generalize these results to an arbitrary Lagrangian of the form (3.5). The calculation proceeds exactly as above, the only difference being that the quantity $C_j q$ is everywhere replaced by $F_j(q)$. The equation analogous to (4.24) is

$$S_{\mathrm{eff}}[q(\tau)] = \int_0^T \{\tfrac{1}{2} M \dot{q}^2 + V(q)\}\, d\tau + \tfrac{1}{2} \int_{-\infty}^{\infty} d\tau' \int_0^T d\tau\, Z[q(\tau)], \tag{4.35}$$

where

$$Z[q(\tau)] \equiv \sum_j \frac{e^{-\omega_j |\tau - \tau'|}}{4 m_j \omega_j} \{F_j(q(\tau)) - F_j(q(\tau'))\}^2. \tag{4.36}$$

This result, while quite general, is not of much interest unless we can relate the parameters of the model to some simple dissipative coefficient. Let us therefore specialize to the case, discussed in Section 3, where the dissipation is adequately described by a coefficient $\eta(q)$. Most generally, the relation between the parameters and $\eta(q)$ is given by Eq. (3.10). In the special case of a "separable" interaction (Eq. (3.11)) we can use (3.10) to express $Z[q(\tau)]$ uniquely in terms of $\eta(q)$

$$Z[q(\tau)] = \frac{1}{2\pi(\tau - \tau')^2} \left[\int_{q(\tau)}^{q(\tau')} \eta^{1/2}(q)\, dq\right]^2. \tag{4.37}$$

So in this case the result for the effective action is a simple generalization of Eq. (4.27) obtained by the replacement

$$\eta\{q(\tau) - q(\tau')\}^2 \to \left[\int_{q(\tau)}^{q(\tau')} \eta^{1/2}(q)\, dq\right]^2. \tag{4.38}$$

In the general case there is *no* unique relation between the quantity $Z[q(\tau)]$ and the low-frequency friction coefficient $\eta(q)$. However, by using the obvious inequality

$$\begin{aligned}
&\sum_j \frac{1}{m_j\omega_j^2}\,\delta(\omega-\omega_j)\{F_j(q(\tau))-F_j(q(\tau'))\}^2 \\
&\quad\leqslant \left\{\int_{q(\tau)}^{q(\tau')}\left[\sum_j \frac{1}{m_j\omega_j^2}\left(\frac{\partial F_j}{\partial q}\right)^2\delta(\omega-\omega_j)\right]^{1/2} dq\right\}^2 \\
&\quad= \frac{2}{\pi}\left(\int_{q(\tau)}^{q(\tau')}\eta^{1/2}(q)\,dq\right)^2
\end{aligned} \tag{4.39}$$

we see that Eq. (4.37) constitutes an upper limit on $Z[q(\tau)]$. Thus we can make the general statement that an upper limit on the effective WKB exponent B, and hence a lower limit on the tunnelling rate, is given by the inequality

$$B \leqslant \int_{-\infty}^{\infty} d\tau\{\tfrac{1}{2}M\dot{q}^2 + V(q)\} + \frac{1}{4\pi}\int_{-\infty}^{\infty} d\tau \int_{-\infty}^{\infty} d\tau'(\tau-\tau')^{-2}\left\{\int_{q(\tau)}^{q(\tau')}\eta^{1/2}(q)\,dq\right\}^2 \tag{4.40}$$

with the integral being taken as always along the "bounce" (saddlepoint) path which makes it an extremum. The inequality (4.40) becomes an equality only if the interaction is separable (i.e., if (3.11) holds).

The case of linear dissipation is clearly a special case of the above. We see that expression (4.31) constitutes an exact expression for the bounce exponent if and only if the dissipation is strictly linear; in the more general (quasilinear) case it constitutes only an upper limit. Hence, the calculations of the next section, which are based on Eq. (4.31), will give in this case only a lower limit on the tunnelling rate. (An *upper* limit is clearly given by the WKB expression calculated without account of damping, since $Z[q(\tau)]$ is obviously positive). Naturally, if we have a specific microscopic model for the nonseparable interaction, we can obtain an exact expression for B as a function of the parameters of the model by direct substitution into (4.36). The importance of the general result (4.40), however, is that it guarantees us that, provided the classical dissipative behaviour can indeed be described by a single coefficient $\eta(q)$, then no mechanism of this dissipation, however complicated, can lower the tunnelling rate below a value which is calculable directly in terms of this coefficient. The implications of this are noted in Section 7.

Everything we have said above generalizes very trivially to the case where η is frequency-dependent (provided it is amplitude-independent). For example, in the strictly linear case all we have to do is to replace the simple expression (4.26) for $\alpha(\tau-\tau')$ by the more general form obtained from (4.22) when $J(\omega)$ is given by $\omega\eta(\omega)$. The more general case can be discussed similarly: it is not worth writing out the explicit expressions here.

Finally, since the techniques used in the derivation of Eqs. (4.30)–(4.34) are not yet universally familiar, it is worth remarking that, as already observed by Sethna [58], many of the features of these results can actually be seen rather directly by using the conceptually simpler "many-dimensional WKB" approach. In such an approach the complete "universe" is visualized as tunnelling along the easiest path in the many-dimensional space whose axes are q and the x_α, and the exponent $B/\hbar$ in the tunnelling formula is given by the expression

$$B/\hbar = \hbar^{-1} \int [2MV(q, x_\alpha)]^{1/2}\, ds, \tag{4.41}$$

where the integral is taken between the origin and the nearest point in the many-dimensional space for which V again becomes zero, and ds is an appropriate metric in the space ($ds^2 = dq^2 + \sum_\alpha (m_\alpha/M)\, dx_\alpha^2$). One interesting feature is that once we have incorporated the counter-term (3.4) in the potential, the actual minimum value of V attained on any hyperplane of constant q is the same as for the isolated system (but of course occurs for $x_\alpha \neq 0$); the reduction of tunnelling by dissipation can therefore be viewed as arising entirely from the fact that the path length in the many-dimensional space is increased. It is of course possible to derive the formula (4.31) by this method: what we have to do is to note that the value of $B/\hbar$ given in (4.41) is simply the minimum value of the action taken over any path which goes from the origin to the barrier edge and back, take the functional derivative of the action with respect to the x_α's and use the resulting equation to eliminate them. (See Sethna [58].) The resulting expression for the action is then equivalent to (4.31). However, it seems to us that there is a rather subtle point in this manœuvre: the expression (4.31) (or Sethna's Eq. (4.9)) only seems to result if we impose boundary conditions ($x_\alpha = 0$) on the x_α at the beginning and end of the bounce, not at the turning point. (If we were to do the latter, we would apparently get a more complicated expression). Thus, we are implicitly allowing the x_α to take any values when $q = q_0$, not necessarily the values which actually bring the system out of the barrier. That this manœuvre gives the right answer appears to be connected with the special properties of harmonic oscillators, and we believe it may be a little dangerous to try to generalize it to other problems (such as the quantum coherence problem, see Section 6) without thorough examination. For this and other reasons we feel that the method developed in this section, though based on perhaps less familiar techniques, may be simpler to apply in general.

5. Analysis of the General Formula

In this section we shall carry out, as far as possible, a quantitative analysis of the formula (4.30) which gives the tunnelling rate in the presence of dissipation. Since in the WKB limit (which is the only limit in which the formula is valid anyway) the influence of dissipation is overwhelmingly through the bounce exponent, we shall

devote most of our effort to the analysis of the quantity B (Eq. (4.31)) and discuss the prefactor A more briefly and less generally. Obviously it would be possible to obtain completely quantitative values of B as a function of damping by solving the Euler–Lagrange equation resulting from the variation of (4.31) numerically, but we shall see that we can get a great deal of insight into its general behaviour by purely analytic means. For definiteness we shall carry out the analysis explicitly for a quadratic-plus-cubic potential of the form (2.5), which is likely to be the case of prime interest for macroscopic systems; however, the generalization to the case where the anharmonic term is of the form q^n, n any integer greater than 2, is completely straightforward, and we can then if required use the results obtained to set bounds on the tunnelling rate for a general potential (cf. below).

It is convenient to introduce the following dimensionless quantities:

$$u \equiv \omega_0 \tau, \tag{5.1a}$$

$$z(u) \equiv q(u)/q_0, \tag{5.1b}$$

$$\alpha \equiv \eta/2M\omega_0 \equiv \gamma/\omega_0, \tag{5.1c}$$

$$\sigma\{z(u)\} \equiv S_{\text{eff}}\{q(\tau)\}/\tfrac{1}{2}M\omega_0 q_0^2, \tag{5.1d}$$

$$b(\alpha) \equiv B/\tfrac{1}{2}M\omega_0 q_0^2. \tag{5.1e}$$

Then the dimensionless bounce $b(\alpha)$ is the saddlepoint value of the quantity

$$\sigma\{z(u)\} \equiv \int_{-\infty}^{\infty} du \left\{ \left(\frac{dz}{du}\right)^2 + (z^2 - z^3) \right\} + \frac{\alpha}{\pi} \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} du' \left(\frac{z(u) - z(u')}{u - u'}\right)^2 \tag{5.2}$$

subject to the boundary conditions $z(0) = z(\infty) = 0$. We notice that any excursion to negative values of z can only increase the value of σ, so it is convenient to impose the condition $z(u) \geqslant 0$ explicitly.

Let us define z_0 as the maximum value attained by the function $z(u)$; in the limit $\alpha = 0$ it follows from the definition (5.1b) that z_0 is 1, and it is obvious (cf. below) that for $\alpha > 0$ we have $z_0 > 1$. Then we write

$$z(u) \equiv z_0 f(\omega u) \equiv z_0 f(t), \qquad t \equiv \omega u, \tag{5.3}$$

where the quantity ω is for the moment arbitrary. We also introduce the positive dimensionless quantities ($\tilde{A}$ and $\tilde{B}$ bear no relation to the prefactor A and bounce exponent B):

$$\tilde{A}[f] \equiv \int_{-\infty}^{\infty} (df/dt)^2 \, dt, \tag{5.4a}$$

$$\tilde{B}[f] \equiv \int_{-\infty}^{\infty} f^2(t) \, dt, \tag{5.4b}$$

$$\tilde{C}[f] \equiv \int_{-\infty}^{\infty} f^3(t)\,dt, \tag{5.4c}$$

$$\tilde{D}[f] \equiv \frac{1}{\pi}\int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dt' \left(\frac{f(t)-f(t')}{t-t'}\right)^2. \tag{5.4d}$$

From its definition the function $f(t)$ must satisfy the inequalities $0 \leqslant f(t) \leqslant 1$; moreover, it will become obvious from what follows (and the trivial inequality $\tilde{C} \leqslant \tilde{B}$) that no interest attaches to functions $f(t)$ for which any of the quantities (5.1a)–(5.1d) fail to exist, so we confine ourselves to functions for which $\tilde{A},\ldots,\tilde{D}$ are well defined. We note that the combinations

$$\kappa[f] \equiv \tilde{B}^2\tilde{C}^{-2}(\tilde{A}\tilde{B})^{1/2}, \tag{5.5a}$$

$$\lambda[f] \equiv \tilde{D}(\tilde{A}\tilde{B})^{-1/2} \tag{5.5b}$$

are invariant under changes of length and time scale separately, as of course are any combination of them. Substituting (5.3) and (5.4) into (5.2), we obtain

$$\sigma(z_0, \omega, \tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}) = z_0^2(\omega\tilde{A} + \omega^{-1}\tilde{B} + \alpha\tilde{D}) - \omega^{-1}\tilde{C}z_0^3. \tag{5.6}$$

Evidently, for any given choice of ω and $f(t)$, σ will have a maximum at the value

$$z_0 = \tfrac{2}{3}\omega(\omega\tilde{A} + \omega^{-1}\tilde{B} + \alpha\tilde{D})/\tilde{C} \tag{5.7}$$

and this value is

$$\sigma_0[\omega, f] = \tfrac{4}{27}\omega^2(\omega\tilde{A} + \omega^{-1}\tilde{B} + \alpha\tilde{D})^3/\tilde{C}^2. \tag{5.8}$$

Now, we know that for $\alpha = 0$ the functional (5.8) is bounded below by a value corresponding to the undamped bounce. Since all the quantities in (5.8) are positive, it is therefore immediately clear that for $\alpha \neq 0$ this functional is bounded below and thus attains a minimum value. The point at which this value is attained corresponds to a saddlepoint of the functional $\sigma\{z(\omega)\}$ and hence gives the bounce trajectory.[20] It follows immediately from these considerations that the bounce $b(\alpha)$ is a monotonically increasing function of α, as we should expect.

The ensuing analysis is simplified if we work in terms of the scale-invariant quantities $\kappa[f]$ and $\lambda[f]$ defined in Eq. (5.5), and introduce also the notation

$$\tan\theta \equiv \alpha\lambda[f]/\sqrt{5}. \tag{5.9}$$

We minimise the functional (5.8) with respect to ω and eliminate the resulting value

[20] This argument of course cannot exclude the possibility of more than one saddlepoint (in which case the one corresponding to minimum σ would dominate).

of ω. After some straightforward algebra we obtain for the functional $\sigma_{00}[f]$ so defined

$$\sigma_{00}[f] = \frac{32}{25\sqrt{5}}\kappa[f]\,F(\theta) \tag{5.10}$$

$$F(\theta) \equiv \frac{(1+\frac{3}{2}\sin\theta)^3}{(1+\sin\theta)^2\cos\theta}. \tag{5.11}$$

The dimensionless bounce $b(\alpha)$ is then the absolute minimum of $\sigma_{00}[f]$ as a functional of f. What we have done is effectively to choose the "time scale" of the bounce ω^{-1} to have its best value; once this is done the variation of σ_{00} with the "shape" of the bounce, that is with the form of the function f, is fairly weak. It should be noticed, however, that the notion of a time-scale is not well defined until we impose some normalization on the function f; a natural choice is to set $\tilde{B} \equiv 1$, which is equivalent to the condition

$$\omega^{-1} \equiv \int_{-\infty}^{\infty} z^2(u)\,du/z_0^2. \tag{5.12}$$

For the moment, however, it is convenient to leave the normalization arbitrary; clearly nothing depends on this, since according to Eqs. (5.9)–(5.11) σ_{00} is a function only of the invariants $\kappa[f]$ and $\lambda[f]$, which do not depend on the normalization.

In the limit of zero damping we know of course the exact form of the function $f(t)$ (namely, $\operatorname{sech}^2 t$) and can compute all the quantities (5.4) (see Appendix D). The relevant value of κ is $5\sqrt{5}/6$, and the resulting value of the dimensionless bounce exponent b is

$$b(0) \equiv b_0 = \tfrac{16}{15}. \tag{5.13}$$

Since according to (5.10)–(5.11) in this limit the function $f(t)$ must minimize $\kappa[f]$, it follows that for general f we have

$$\kappa[f] \geqslant 5\sqrt{5}/6. \tag{5.14}$$

To proceed further we need the following inequalities, which are proved in Appendix D:

$$F(\theta) \geqslant (125/32)\tan\theta, \tag{5.15}$$

$$1 + \tfrac{5}{2}\tan\theta \leqslant F(\theta) \leqslant 1 + \tfrac{35}{8}\tan\theta, \tag{5.16}$$

$$\kappa[f]\,\lambda[f] \geqslant 8\pi/9. \tag{5.17}$$

To obtain a *lower* limit on $b(\alpha)$ which is valid for all α and reduces to b_0 for $\alpha \to 0$ we combine the first of the inequalities (5.16) with (5.14) and (5.17) to obtain

$$b(\alpha) \geqslant b_0 + \frac{128\pi}{225}\alpha. \tag{5.18}$$

Another general limit, which is stronger for large α (and tends to the exact result in the limit $\alpha \to \infty$) is

$$b(\alpha) \geqslant \frac{8\pi}{9}\alpha. \tag{5.19}$$

To obtain an *upper* limit on $b(\alpha)$, we need only choose some plausible trial form of $f(t)$ and substitute the values of κ and λ obtained from it into Eqs. (5.9)–(5.11). Some relevant forms of $f(t)$ are tabulated in Appendix D together with their values of κ and λ. In the limits of weak and strong damping we know the exact limiting forms of the bounce trajectory (the first and second functions in the table in Appendix D, respectively). In each case we obtain results exact to the next order by evaluating the "perturbation" (that is, the terms in $\sigma_{00}[f]$ of order α in the weak-damping and of order α^{-2} in the strong-damping limit) along the limiting trajectories. In this way we obtain

$$b(\alpha) = b_0 + \frac{48\zeta(3)}{\pi^3}\alpha + O(\alpha^2) \qquad (\alpha \to 0) \tag{5.20}$$

and

$$b(\alpha) = \frac{8\pi}{9}\alpha + \frac{2\pi}{9}\alpha^{-1} + O(\alpha^{-3}) \qquad (\alpha \to \infty). \tag{5.21}$$

These two forms of $f(t)$, or others (cf. Appendix D) can be inserted into Eqs. (5.9)–(5.11) to give general upper limits. We notice in particular the limit obtained by using the strong-damping form and the second inequality of (5.16)

$$b(\alpha) \leqslant \frac{256\pi}{225\sqrt{10}} + \frac{224\pi}{225}\alpha. \tag{5.22}$$

Finally, returning to dimensional variables, we can summarize the above results by saying that the *correction* ΔB to the WKB exponent B arising from damping is of the general form

$$\Delta B = \Phi(\alpha)\,\eta q_0^2, \tag{5.23}$$

where the dimensionless function $\Phi(\alpha)$ (the quantity called A or $\varphi(\alpha)$ in our earlier paper [28]) is always of order 1, has the exact value $12\zeta(3)/\pi^3$ for weak damping and the exact value $2\pi/9$ for strong damping (in accordance with the conjecture made in our earlier work [28]), and for intermediate values of α can be bracketed in the way described above. In Fig. 2 we show the asymptotic values of and the bounds on $b(\alpha)$ obtained above as well as some typical values obtained from trial functions.

Before leaving the subject of the semiclassical bounce we should discuss the

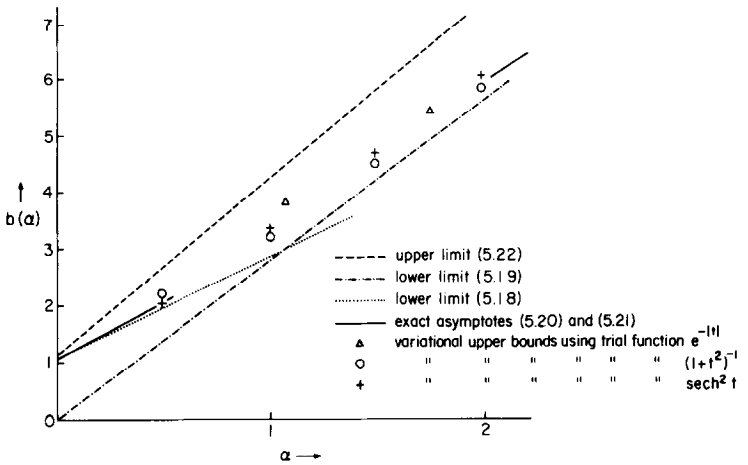

FIG. 2. Limits on the dimensionless bounce $b(\alpha)$.

question of its length and "time" scales. Quite generally we find that the maximization with respect to z_0 and the minimization with respect to ω carried out above leads to the results

$$z_0 = \frac{4}{5}\frac{\tilde{B}}{\tilde{C}}\left(\frac{1+\frac{3}{2}\sin\theta}{1+\sin\theta}\right), \tag{5.24}$$

$$\omega = \left(\frac{\tilde{B}}{5\tilde{A}}\right)^{1/2}(\sec\theta - \tan\theta). \tag{5.25}$$

The length scale z_0 is unambiguously defined; it is the maximum excursion in the damped bounce relative to that in the undamped one. However, as remarked above, the definition of ω requires some normalization of $f(t)$; let us choose $\tilde{B} \equiv 1$. Then, using the inequalities

$$(1 + 2\tan\theta)^{-1} \leqslant \sec\theta - \tan\theta \leqslant (1+\tan\theta)^{-1} \tag{5.26}$$

we see that ω is of order $\tilde{A}^{-1/2}[F(0)]^{-1}$. From (5.10) and (5.15) we then have for the "time-scale" $T_0 \equiv 2\pi/\omega$ and length scale z_0 the order-of-magnitude estimates

$$z_0 \sim \tilde{C}^{-1}, \qquad T_0 \sim \tilde{C}^2 b(\alpha). \tag{5.27}$$

For weak or moderate damping we clearly have $\tilde{C} \sim 1$, so the length and time scales are of the same order as for the undamped bounce. In the limit of strong damping we can use the exact solution $(f(t) = (1+t^2)^{-1}$, see Appendix D); then we find that z_0 is $\frac{4}{3}$ and T_0 is $(\tilde{D}/10\tilde{A}) = \alpha/5\pi$. Going back to dimensional units, we find that the bounce path $q(\tau)$ in this limit has the form

$$q(\tau) = \frac{(4/3)\,q_0}{1 + (\omega_0^2/2\gamma)^2\,\tau^2}. \tag{5.28}$$

Thus, even in the limit of very strong damping the spatial excursion in the bounce trajectory is still finite; however, the time scale diverges as the "slow" relaxation time γ/ω_0^2. We have not proved explicitly that the behaviour of z_0 and T_0 as a function of α is monotonic, but this seems extremely probable.

We now turn more briefly to the analysis of the prefactor A in the expression (4.30) for the tunnelling rate Γ; A is given by Eq. (4.32). In the ensuing analysis it is convenient to take the total time region T allowed for the bounce to be very large but finite, thereby allowing (where necessary) an expansion of $q(\tau)$ as a Fourier sum rather than an integral, and to take the limit $T \to \infty$ at the end of the calculation. It is also convenient to take, as above, the centre of the bounce to lie at $\tau = 0$. We note that while we are required to continue $q(\tau)$ periodically outside the region $-T/2 \leqslant \tau \leqslant T/2$ by the prescription $q(\tau + T) \equiv q(\tau)$ (or, alternatively, to use a more complicated form of the last term in (4.33) and (4.34)), the physical problem is also constrained by the boundary conditions $q(T/2) = q(-T/2) = 0$. We shall, however, assume that the value of A is negligibly affected if we simply omit these boundary conditions. The physical justification for this is that in calculating the factor A we are comparing fluctuations about the bounce path with those about the trivial path $q(\tau) \equiv 0$; since the two paths differ only in a time region of order T_0 (see above), where $T_0/T \to 0$, any extra freedom arising from the behaviour near the boundaries ($\tau = \pm T/2$) should cancel out in the final result.[21] In the following we will also, where convenient, take complex Fourier transforms even though $q(\tau)$ is physically constrained to be real; working with sine and cosine transforms would of course give the same results but is more cumbrous.

Since we are most interested in the strong-damping limit, it is convenient to introduce the dimensionless time variable $t \equiv \omega_0 \tau/2\alpha$. We denote the frequency conjugate to t by ω. Then the operation of summing over the states allowed by the periodic boundary conditions goes over into an integral according to the prescription

$$\sum_j \to \frac{\Omega}{2\pi} \int d\omega \qquad (\Omega \equiv \omega_0 T/2\alpha). \tag{5.29}$$

Let us denote the dimensionless classical bounce trajectory, measured in units of q_0, by $\tilde{q}_{\mathrm{cl}}(t)$. Then we can write Eq. (4.30) in the form

$$A = \omega_0 (B/2\pi\hbar)^{1/2} K^{1/2}, \tag{5.30}$$

where the dimensionless factor K is given by

$$K \equiv \left| \frac{\det \hat{H}_0}{\det'(\hat{H}_0 + \hat{V})} \right|. \tag{5.31}$$

Here the operators $\hat{H}_0$ and $\hat{V}$ (so denoted to help one's intuition by analogy with a

[21] It should be observed that this argument would not necessarily go through if we were comparing the fluctuations around the damped path with those around the undamped one.

scattering problem) are defined, in the representation whose basis is normalized plane waves with frequency ω, by the relations

$$\hat{H}_0 \psi_n(\omega) \equiv (1 + |\omega| + \omega^2/4\alpha^2)\, \psi_n(\omega), \tag{5.32}$$

$$\hat{V}\psi_n(\omega) \equiv \int_{-\infty}^{\infty} V(\omega - \omega' : \alpha)\, \psi_n(\omega')\, d\omega', \tag{5.33}$$

with

$$V(\omega) \equiv -\frac{3}{2\pi} \int_{-\infty}^{\infty} \tilde{q}_{cl}(t)\, e^{-i\omega t}\, dt. \tag{5.34}$$

The quantity $V(\omega)$ depends on α because the bounce trajectory $q_{cl}(t)$ is a function of α (both in scale and in shape). The prime in Eq. (5.31) denotes, as always, that the zero eigenvalue is to be omitted.

Consider the spectrum of the operator $\hat{H}_0 + \hat{V}$. We know that it has one[22] negative eigenvalue λ_0, with associated eigenfunction $\psi_0(\omega)$, which corresponds to the steepest descent from the saddlepoint in function space, and one zero eigenvalue, with eigenfunction $\psi_1 = \text{const}\, \omega V(\omega)$, corresponding to uniform translation of the bounce along the time axis. It may or may not have other "bound state" (discrete) eigenvalues lying between 0 and 1, but in any case will certainly have a quasi-continuous spectrum (i.e., with spacing vanishing as Ω^{-1} in the limit $\Omega \to \infty$) starting at $\lambda = 1$. In the states of the quasi-continuous spectrum the effect of $\hat{V}$ is of order Ω^{-1} relative to that of $\hat{H}_0$.

Using the relation $\ln(\det A) = \mathrm{Tr}(\ln A)$, we write the logarithm of the factor K (Eq. (5.31)) in the form

$$\begin{aligned} \ln K = {} & \langle \psi_0 | \ln \hat{H}_0 | \psi_0 \rangle + \langle \psi_1 | \ln \hat{H}_0 | \psi_1 \rangle - \ln |\lambda_0| \\ & + \mathrm{Tr}''(\ln \hat{H}_0) - \mathrm{Tr}''\ln(\hat{H}_0 + \hat{V}). \end{aligned} \tag{5.35}$$

Here the operator $\ln \hat{H}_0$ is defined by its matrix elements $(\ln(1 + |\omega| + \omega^2/4\alpha^2))$ in the ω-representation, in which it is diagonal, and the double prime on the trace indicates that it is taken over the (infinite-dimensional) subspace orthogonal to ψ_0 and ψ_1. If now we work in the representation defined by the eigenfunctions of $\hat{H}_0 + \hat{V}$, then for all states n of the continuous spectrum we can make the replacement

$$\langle \psi_n | \ln \hat{H}_0 | \psi_n \rangle - \langle \psi_n | \ln(\hat{H}_0 + \hat{V}) | \psi_n \rangle = -\langle \psi_n | \hat{H}_0^{-1} \hat{V} | \psi_n \rangle. \tag{5.36}$$

(The legitimacy of this manoeuvre may be verified by noting that the second term is just $\ln \lambda_n$ and writing λ_n explicitly as a bilinear form in $\psi_n(\omega)$.) If, therefore, we define the quantity

$$C_0 \equiv \sum_{\substack{j \\ (0 < \lambda_j < 1)}} \{\langle j | \ln \hat{H}_0 | j \rangle - \ln \lambda_j + \langle j | \hat{H}_0^{-1} \hat{V} | j \rangle, \tag{5.37}$$

[22] That there is no more than one follows from the fact that the zero eigenvalue corresponds to an eigenfunction with a single node (cf. [50]).

where the sum over j runs only over the bound states with $\lambda_j > 0$ (if any), and take into account that $\hat{H}_0^{-1}\hat{V}\psi_1 = -\psi_1$, we can rewrite (5.35) in the form

$$\ln K = D_0(\alpha) - \operatorname{Tr} \hat{H}_0^{-1}\hat{V}, \tag{5.38}$$

where the quantity $D_0(\alpha)$, which depends only on the bound-state spectrum, is given by

$$D_0(\alpha) \equiv \langle \ln \hat{H}_0 \rangle_0 + \langle \ln \hat{H}_0 \rangle_1 - \ln |\lambda_0| + C_0 + \langle \hat{H}_0^{-1}\hat{V} \rangle_0 - 1. \tag{5.39}$$

The unrestricted trace which appears in (5.38) can be evaluated directly in the ω-representation

$$\begin{aligned} \operatorname{Tr} \hat{H}_0^{-1}\hat{V} &= V(0) \int_{-\infty}^{\infty} \frac{d\omega}{1 + |\omega| + \omega^2/4\alpha^2} \\ &= -\frac{3}{\pi} \int_{-\infty}^{\infty} \tilde{q}_{cl}(t)\, dt \times \frac{1}{\sqrt{\alpha^2 - 1}} \ln\left(\frac{\alpha + \sqrt{\alpha^2 - 1}}{\alpha - \sqrt{\alpha^2 - 1}}\right) \qquad (\alpha \geqslant 1) \end{aligned} \tag{5.40}$$

with a similar easily calculable expression for $\alpha < 1$. (cf. Eqs. (29) of Appendix B). Substituting (5.40) in (5.38) and the latter in (5.30) we finally obtain the prefactor A in the tunnelling rate.

In the limit $\alpha \to \infty$ this programme can be carried out explicitly. Using Eq. (5.28), we see that the expression (5.40) reduces simply to $8 \ln 2\alpha$. The quantity D_0 does not depend on α in this limit and is estimated in Appendix D to be about $-1 \cdot 5$. Thus the final expression for A in this limit is

$$A = 4(\exp(D_0/2))\, \alpha^4 \omega_0 (B/2\pi\hbar)^{1/2}. \tag{5.41}$$

Bearing in mind that B is related to its "undamped" value B_0 by a factor $(5\pi/6)\alpha$, (cf. Eqs. (5.20)–(5.21)) and using Eqs. (2.3) and (2.6)), we can therefore finally relate A to the "undamped" prefactor A_0 by

$$A/A_0 = c\alpha^{9/2}, \qquad c \equiv (2\pi e^{D_0}/9)^{1/2}. \tag{5.42}$$

Thus in the limit of strong damping ($\gamma \to \infty$) the prefactor is proportional to $\gamma^{9/2}\omega_0^{-7/2}$. For intermediate damping we can obtain approximate values of the prefactor by substituting approximate forms of $\tilde{q}_{\mathrm{cl}}(t)$ in (5.40).

In conclusion, we note that the techniques of this section, although specifically applied to a problem where the potential is of the form (2.5), can obviously be straightforwardly adapted to any potential of the more general form

$$V(q) = \tfrac{1}{2}M\omega_0^2 q^2 - \lambda q^n. \tag{5.43}$$

(However, the argument which leads to the inequality (5.17)—see Appendix D—does not have a straightforward generalization to this case, and it would be necessary to find other ways of limiting the quantity $\kappa\lambda$ if we wish to produce a general lower limit on $b(\alpha)$). Moreover, since it is easy to show that if the condition $V_1(q) \geqslant V_2(q)$ holds for all q, then the tunnelling rate in V_1 is less than that in V_2 (cf. Eq. (4.27)), we can

obtain bounds on the rate for any potential which is *everywhere* bounded by two potentials of the type (5.43). It is, however, essential to appreciate that it is not sufficient for the bounds to hold everywhere "under the barrier," that is, to the left of the point B in Fig. 1; in the damped case the system is still in some sense in the tunnelling regime beyond the point B. Once must also be careful not to assume that sharp cutoffs can be put on otherwise smooth potentials without drastic effects: see the discussion of the "truncated harmonic oscillator" in the next section.*

6. Discussion

We now turn to the physical interpretation of the results presented in this paper. A simple way of summarizing them qualitatively is to say that in a damped system the tunnelling probability can be written in the form

$$P \propto \exp - \text{const}(V_0/\hbar\omega_{\text{eff}}), \tag{6.1}$$

where ω_{eff} is an effective frequency of the system which varies from the undamped frequency ω_0 for weak damping to a quantity of the order of the "slow" relaxation frequency ω_0^2/γ for strong damping (cf. [59, 60]). Thus, the effect of damping is to increase the WKB exponent by a factor of order $(1 + \gamma/\omega_0)$. There are a number of ways of interpreting this result intuitively.

First, we note that for a simple damped harmonic oscillator the groundstate probability density can always be written in the form (see Appendix B)

$$\rho(q) = \text{const} \exp - q^2/2\langle q^2 \rangle, \tag{6.2}$$

where the mean-square displacement $\langle q^2 \rangle$ is given by Eq. (B.28), that is,

$$\langle q^2 \rangle = \frac{\hbar}{2M\omega_0} f(\gamma/\omega_0), \tag{6.3}$$

where the expression $f(\gamma/\omega_0)$, given by Eq. (B.29), tends to 1 for small γ and for large γ is proportional to $(\omega_0/\gamma)\ln(\gamma/\omega_0)$. Now, we might suppose that to estimate the order of magnitude of the effect of damping on tunnelling it is legitimate to replace the true barrier by a "model" potential which has its small-q (harmonic) form up to some (γ-independent) cutoff q_0' of the order of the true barrier width q_0 and then drops away discontinuously to a negative value. It is now tempting to argue that the tunnelling probability for the model potential is simply proportional to the value of $\rho(q_0')$ calculated for the simple oscillator, so that the WKB exponent is $q_0'^2/2\langle q^2 \rangle$ with $\langle q^2 \rangle$ given by Eq. (7.3); and that therefore the exponent for the real problem should

**Note added in proof.* The method given in this section to calculate the prefactor $A(\alpha)$ is not quite correct, in that (5.36) omits a term $(\ln\langle\psi_n|\hat{H}_0|\psi_n\rangle - \langle\psi_n|\ln\hat{H}_0|\psi_n\rangle)$ which may in general be of the same order as those kept. This does not affect the α-dependence for strong damping but does change the numerical factor. A numerical computation of $A(\alpha)$ has recently been carried out by L.-D. Chang and S. Chakravarty, who have also obtained more accurate values of the exponent B; we thank them for helpful discussions on this point.

be given by the same expression to within a constant of order unity. Such an argument evidently gets the leading-order effect right but for strong damping underestimates the true suppression of the exponent by a factor of order $[\ln(\gamma/\omega_0)]^{-1}$. At first sight this seems to contradict the theorem stated at the end of Section 5 concerning the monotonic dependence of the tunnelling probability on potential, since if we choose $q_0' \geqslant 4q_0/3$ the model potential is *greater* than the true one for all values of q which are reached by the (true) instanton trajectory. The fallacy lies in the assumption that the tunnelling probability for the model potential is proportional to the quantity $\rho(q_0')$ calculated for the simple oscillator. In fact, consider a potential of the form const. $\{(q/q_0')^2 - (q/q_0')^n\}$ of which the model potential can be considered to be the limit as $n \to \infty$. The larger the value of n, the steeper the last part of the instanton trajectory and the faster the system will tend to move along it. For the undamped system, this accelerated motion cannot increase the total action beyond the value it would have for a smoother potential, but for the damped system it can and does increase the effective action (4.29), because of the last term. To the extent that this last, steep part of the trajectory contributes significantly, the quantity $\rho(q_0')$ (which is essentially determined by the rest of the trajectory) clearly badly underestimates the suppression factor. These remarks suggest more generally that for a damped system it may be a very bad approximation to replace the true potential by one which has a discontinuity—at least unless special care is taken in handling the latter.

A second way of looking at our results is in terms of the formal quantum theory of measurement. A well-known feature of the latter is that any "measurement" of a quantum system projects its state on to the relevant eigenfunction of the measured quantity and hence may inhibit its evolution. In our case we can regard the environment as subjecting the system to repeated position measurements. Now, intuitively speaking, the frequency with which the position is measured to within an accuracy of the order of the zero-point excursion $(\hbar/2m\omega_0)^{1/2}$ is of the order of the relaxation time from the first excited state to the groundstate, that is, of order γ^{-1}. (This crude argument can of course be made more precise.) If the system is repeatedly examined to see whether it is within the zero-point limits, then by the usual arguments (see, e.g., [31]) its probability of propagating outside them, and hence of tunnelling, will be reduced. Thus we should expect the "observation" by the environment to suppress tunnelling appreciably as soon as the frequency γ of this observation becomes comparable to the natural propagation frequency, that is to ω_0—a result which agrees qualititatively with Eq. (6.1). An advantage of this way of looking at the problem is that we can see immediately that quantum *tunnelling* may be a very much more difficult phenomenon to suppress than quantum *coherence* (see Section 1), since the relevant time scale for the latter is many orders of magnitude longer. To be concrete, let ω_0 be as above the characteristic frequency of small oscillations, which to within a factor of order unity is the "attempt frequency" for penetrating the barrier. The actual tunnelling rate ω_τ (whether expressed in terms of an amplitude or a probability) is smaller relative to ω_0 by an exponential factor. Quantum coherence, in the sense in which it is defined in the Introduction, requires

that the system propagate undisturbed (hence unobserved) for a time comparable to ω_τ^{-1} (cf. [31, 32]); quantum tunnelling, as we have just seen, merely requires that it be unobserved for the much shorter time ω_0^{-1}. Thus, *the existence of quantum tunnelling in a given system* (macroscopic or otherwise) *is no guarantee that that system will also show quantum coherence.* For further details of the argument see [2, Sect. 5].[23]

Next let us examine the question of the limits of validity of our results. It is essential to our argument as formulated that the system, once it has tunnelled through the barrier, has no appreciable probability of returning and interfering with itself, and this is guaranteed if the potential beyong the exit point is everywhere negative as assumed in Section 2. While this is a relatively realistic description of the potential felt by (e.g.) an α-particle tunnelling out of a nucleus, there are many cases of practical interest on both the microscopic and the macroscopic scale where the tunnelling is between two finite wells. (In particular this is always the case for a SQUID). The question then arises, under what circumstances can we apply our results to such a case?

Intuitively speaking, we should expect it to be a sufficient condition to be able to treat the lower well as effectively unbounded if the probability of the system being reflected at the far end and returning to the barrier with sufficient energy to get back into the metastable state is negligible. To be more quantitative, it is certainly sufficient that it should have lost in the process of going to and fro across the lower well an energy comparable to the effective zero-point energy of the metastable state. For weak or moderate damping this quantity is of order $\hbar\omega_0/2$, while for heavy damping it certainly cannot be greater than a quantity of order $\hbar\gamma/2$. Suppose for definiteness that the lower well is approximately harmonic in shape, with oscillation frequency ω_1 and classical relaxation time τ_1, and let the energy (relative to the bottom of the lower well) with which the system emerges from the barrier be of order $N\hbar\omega_1$; thus, N is large to the extent that the motion in the lower well is semiclassical. Then it is easy to see that the energy lost in transit across the lower well is of order $N\hbar\tau_1^{-1}$, and correspondingly a sufficient condition to treat it as infinite is, for weak or moderate damping in the upper well

$$N \gtrsim \omega_0 \tau_1. \tag{6.4}$$

For strong damping in the upper well this condition should be replaced by $N \gtrsim \gamma\tau_1$, which for most practical cases ($\gamma \sim \tau_1^{-1}$) is simply the condition for the motion in the lower well to be semiclassical. Actually this condition may well be too stringent. Suppose that ω_1 is much less than ω_0, as may well be the case in practice. (In particular, it will often be the case for a highly hysteric *rf* SQUID, since the effective

[23] In making this argument quantitative it is of course essential to bear in mind that the spectrum of environment fluctuations which couple to the system (the spectrum of "observations") may itself be frequency-dependent. For this reason, while the last two sentences on p. 96 of [2] are correct, the second sentence on p. 97 is somewhat misleading (the quantity γ is not in fact $1/2CR$). In fact, the question of quantum coherence in the case considered in this paper ($J(\omega) \propto \omega$) is quite subtle (cf. [61, 62]); we hope to discuss it elsewhere.

ω_1 is then the *average* "harmonic" frequency of the potential (1.2) ignoring the sinousoidal corrugations, that is $(LC)^{-1/2}$, whereas ω_0 involves the much larger term $(I_c\varphi_0/C)^{1/2}$.). Now, it is easy to show, for a simple harmonic oscillator potential, that coherence between a wave packet newly emerging from the barrier (treated in this context as a phenomenological δ-function source of particles) and one which has made one double transit of the well is destroyed as soon as N is large compared to $\omega_1\tau_1$. Since our calculation of the tunnelling probability is affected by reverse transitions only to the extent that they are coherent with the original process, it follows that the possibly weaker condition

$$N \gtrsim \omega_1\tau_1 \tag{6.5}$$

is also a sufficient condition for the validity of our results. In most cases of practical interest both the conditions (6.4) and (6.5) are satisfied.

It should be emphasized that the fulfilment of the criteria (6.4) and/or (6.5) merely justifies the general technique used in Section 4 of this paper; it does not, in itself, guarantee that an instanton path actually exists. In fact, in view of the results of Section 5 for the specific case of a cubic potential (see especially Eq. (5.28)) we see that an instanton path may well not exist if the lower well is insufficiently deep relative to the upper one; crudely speaking, we need that the depth of the lower well relative to the upper should be comparable to the height of the barrier. In the case $\omega_0 \ll \gamma \ll V_0/\hbar$ this is a more stringent condition than (6.4) or (6.5); nevertheless, in most practical cases it is again usually well fulfilled.

Finally we examine briefly the question of corrections to the WKB approximation. As is well known, for an undamped system the condition for validity of this approximation is that the barrier height V_0 should be large compared to $\hbar\omega_0$. This is, roughly speaking, equivalent to the statement that the zero-point fluctuation amplitude should be small compared to the barrier width, or that the WKB exponent should be large compared to unity. Whichever of the latter formulations we use, it is clear that for a damped system the relevant criterion becomes $V_0 \gg \hbar\omega_{\text{eff}}$, and in particular that for the overdamped limit it is only necessary that $V_0 \gg \hbar\omega_0^2/\gamma$. Hence if the WKB approximation is good for an undamped system, it is certainly good for the corresponding damped one. In particular it should be good for almost all systems which are likely to show macroscopic-scale tunnelling.

Thus, the theory developed in this paper should cover the overwhelming majority of systems which can be used experimentally to investigate macroscopic quantum tunnelling. The minority of cases which violate one or more of the above conditions, in particular cases where the relative depth of the two wells is small compared to the barrier height (cf. [8]) is of course also of considerable interest, as are the much more common examples where this situation occurs on the microscopic scale (e.g., the case of defect tunnelling in solids [29]). Such cases can obviously be handled by suitable application of the general formulae developed in Section 4, in particular Eqs. (4.23) and (4.27), but the subject is clearly too large to discuss here. However, it is worth making one remark to illustrate the fact that dissipation can actually modify the behaviour qualitatively. Suppose that we are interested in quantum *coherence* in the

sense defined in the Introduction, that is, in the behaviour of a particle moving between two *equivalent* minima. For an undamped system the relevant calculation in the instanton formalism is standard (see, e.g., [54]) and involves considering a trajectory in which the particle leaves one well at $\tau = -\infty$ and ends up in the other at $\tau = \infty$; that is, effectively, half a bounce. Suppose for definiteness the damping is weak: then, at first sight at least, it should be possible to a first approximation to neglect it when considering the instanton trajectory $q(\tau)$. If, having done this, we then treat the last term in (4.27) as a perturbation, then the contribution of this term diverges logarithmically; consequently, we should expect the coherence behaviour to be qualitatively different from that of the undamped system—a result which is indeed found by phenomenological treatments of the "detuning" effect (see, e.g., [31]). Note that this effect appears only because the spectral density $J(\omega)$ (Eq. (3.28)) is proportional to ω in the low-frequency limit; it has no analog, for example, in the case of defect tunnelling in solids [29], where $J(\omega)$ is proportional to ω^3. (In the language of quantum measurement theory, we are considering the case where the frequency of observations which effectively localize the system on one side or other of the barrier is proportional to the tunnelling frequency ω_r in the limit $\omega_r \to 0$, whereas in the defect case it tends to zero faster than ω_r in this limit. We hope to discuss this and related questions in the future.

To conclude this section we briefly outline some generalizations and problems for further work. In the bulk of this paper we have considered the case of strictly linear frequency-independent dissipation. The generalization to the quasilinear and nonlinear cases is fairly trivial and is given at appropriate points. To obtain the generalization to frequency-dependent dissipation, we simply use Eqs. (4.24) and (4.22) with the appropriate form of $J(\omega)$ (or the analogous quantity, see Eq. (3.6)). Provided only that $J(\omega)$ is proportional to ω in the limit $\omega \to 0$, and for $\omega \to \infty$ falls off sufficiently fast ($J(\omega) \leqslant \text{const}\, \omega^{-n}, n > 0$) to ensure the convergence of the expression (3.13) for the frequency shift $\Delta\omega^2$, then the results are qualitatively unchanged and can be crudely expressed by replacing η in Eq. (5.23) by an effective friction coefficient for frequencies of the order of the characteristic bounce frequency (i.e., ω_0 in the weak-damping case and ω_0^2/γ in the strong-damping case). (cf. [16]).

A somewhat less obvious generalization is to the case of dissipation in a field theory (cf. [27]), where the dissipative mechanism may involve spatial gradients of the fields. We are confident that a generalization of the method of this paper to this case is possible in principle, but have not at the time of writing constructed one. Another important question concerns the generalization of the results of this paper to finite temperatures. While for an undamped system the crossover between quantum tunnelling and classical thermal barrier hopping takes place, as the temperature is raised, at a temperature of the order of $\hbar\omega_0/k_B$, it is very plausible, in view of the results given in Appendix B, that for a heavily damped system the crossover should taken place at a much lower temperature, of order [60] $\hbar\omega_0^2/\gamma k_B$. However, it is clearly highly desirable to constract a detailed quantitative theory of tunnelling in damped systems at finite temperature. Finally there is the complex of questions, mentioned above, associated with the finite extent of the lower well and with quantum

coherence effects—questions which, while primarily of interest in the context of microscopic tunnelling systems, may in certain special circumstances also be relevant to the macroscopic case. We are currently working on some of these problems.

7. Conclusion

In this paper we have attempted to motivate, define, and discuss the question: What is the influence of dissipation on quantum tunnelling? Specifically, we have studied the quantum tunnelling behaviour of a system whose semiclassical dynamics is given by the dissipative Eq. (2.8), or a generalization of this. Our main conclusions are as follows: (1) The presence of dissipation always tends to suppress quantum tunnelling. (2) In the case of strictly linear dissipation (or more generally of a separable interaction), the suppression factor can be uniquely related to the phenomenological dissipation coefficient: in the general case a lower limit on this factor can be similarly related to it. (3) In the experimentally important case of strictly linear dissipation in a potential with cubic anharmonicity, the dominant part of the suppression factor can be written in the form $\exp - \Phi(\alpha)\, \eta q_0^2/\hbar$, where η is the dissipation coefficient, q_0 the distance to be traversed under the barrier, and $\Phi(\alpha)$ is a function of order 1 which we can estimate as explained in Section 5.

To the extent that the behaviour of a given *rf* SQUID in the classical regime can be adequately described by the resistively shunted junction (RSJ) model, our results can be trivially transposed to it by the replacements $q \to \Phi$, $V(q) \to U(\Phi)$, $M \to C$, $\eta \to \sigma_n \equiv R_n^{-1}$. Our results can also be transposed if we assume that the correct description of the SQUID is some simple generalization of the RSJ model, e.g., such as to incorporate a nonlinear or frequency-dependent normal conductance σ_n. To go beyond this assumption would require an explicit discussion of the justification (for (a possibly generalised) Eq. (1.4), which in turn would lead us into detailed questions concerning the microscopic model of the junction, etc. These questions are unlikely, in our opinion, to affect the general nature of the results concerning the influence of dissipation on tunnelling, and are sufficiently technical that we have not attempted to discuss them here.

Finally, it should of course be emphasized that all the calculations of this paper have been carried out within the conventional framework of quantum mechanics, that is, under the assumption that this framework can indeed be extrapolated to the macroscopic scale in the sense discussed in the Introduction. Should it eventually turn out that for a particular type of physical system quantum tunnelling is not observed under conditions where the theory predicts it should be, no doubt the most obvious inference would be that the calculations, or the model on which they are based, are wrong; however, an alternative inference, which it would be unwise to exclude totally a priori, would be that quantum mechanics cannot in fact be extrapolated in this way.

In the course of this research, which has extended over the last four years, we have had many fruitful discussions with many colleagues both at the University of Sussex

and elsewhere, some of whom are acknowledged in [28]. Over the last two years we have particularly benefited from discussions and/or correspondence about this problem with John Clarke, R. de Bruyn Ouboter, Roger Koch, Juhani Kurkijärvi, Gerd Schön, J. P. Sethna, Richard Voss, David Waxman, and Richard Webb. We are also grateful to John Bardeen for communicating to us his work on the related problem of tunnelling of charge density waves in quasi-one-dimensional solids, and Allen Goldman for suggestions concerning possible quantum tunnelling of the vortex–antivortex complex. We thank A. Widom and T. D. Clark for a preprint which induced us to elucidate the question of the "anomalous" cases mentioned in Appendix C. We are particularly grateful to Gabriel Barton for a critical reading of the first draft of the manuscript and helpful comments on it. One of us (AJL) gratefully acknowledges the hospitality of the Laboratory of Atomic and Solid State Physics at Cornell University during April 1980, when an important part of this work was done; the other (AOC) acknowledges financial support from CAPES (Coordenecao de Aperfeicaomento de Pessoal de Nivel Superior) and from the Royal Society under the exchange agreement with CNPq (Conseilho Nacional de Desenvolvimento Cientifico e Tecnologico).

APPENDIX A:
FORM OF THE LAGRANGIAN FOR ELECTROMAGNETIC SYSTEMS

In this appendix we shall discuss a point which, while basically a matter of elementary classical Lagrangian electrodynamics, nevertheless seems capable of causing a certain amount of confusion in discussions of the effect of dissipation on quantum tunnelling. It concerns the correct choice of Lagrangian in the case where, with the "natural" choice of coordinates for the problem, the coupling is by velocity-dependent forces. The most obvious example of such a problem is in electromagnetic problems, for example in a SQUID, where the basic variable, the magnetic flux, is coupled to a quantity with the dimension of electric current. However, a similar situation arises in the general case of adiabatic coupling, since as we see in Appendix C this can be made isomorphic to the electromagnetic problem.

We start with some very simple considerations concerning velocity-dependent forces in general. Let us imagine a system described by coordinates x, y whose scale is chosen in such a way that the Lagrangian becomes

$$L_0 = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x, y). \tag{A.1}$$

We now add to this a velocity-dependent term of the form

$$\Delta L_1 = \varepsilon \dot{x} y. \tag{A.2a}$$

Because addition of a total time derivative to the Lagrangian changes the classical

equations of motion not at all (and the quantum-mechnical transition amplitudes only trivially) we may equally well replace (A.2) by either of the equivalent forms

$$\Delta L_2 = -\varepsilon x\dot{y} \tag{A.2b}$$

$$\Delta L_3 = \frac{\varepsilon}{2}(\dot{x}y - x\dot{y}). \tag{A.2c}$$

Going over to the Hamiltonian formalism in the usual way, we see that the Hamiltonian of the system can be wtitten in any of three equivalent forms corresponding to (A.2–A.4), respectively,

$$H_1 = \frac{(p_x - \varepsilon y)^2}{2m} + \frac{p_y^2}{2m} + V(x, y), \tag{A.3a}$$

$$H_2 = \frac{(p_y - \varepsilon x)^2}{2m} + \frac{p_x^2}{2m} + V(x, y), \tag{A.3b}$$

$$H_3 = \frac{(p_x - \varepsilon y/2)^2}{2m} + \frac{(p_y + \varepsilon x/2)^2}{2m} + V(x, y) \tag{A.3c}$$

$$= H_0 + \frac{\varepsilon L_z}{2m} + \frac{\varepsilon^2}{4}\frac{1}{2m}(x^2 + y^2),$$

where H_0 is the Hamiltonian derived from (A.1) and $L_z \equiv (xp_y - yp_x)$ is the angular momentum in the xy plane. The most natural physical interpretation of Eqs. (A.3a)–(A.3c) would of course be as a decription of a charged particle moving in the plane under the influence of a magnetic field perpendicular to the plane.

If the Hamiltonian H_0 represents an anisotropic harmonic oscillator, that is, if

$$V(x, y) = \tfrac{1}{2}m(\omega_1^2 x^2 + \omega_2^2 y^2) \qquad (\omega_1 \neq \omega_2 \text{ in general}), \tag{A.4}$$

then it is very straightforward to diagonalize the Hamiltonian (A.3b) (most easily by carrying out first a canonical transformation which interchanges the roles of y and p_y, and then a simple rotation of coordinates), and to check that the value of the mean-square displacement x^2 is *decreased* relative to its value for H_0 alone, quite independently of the ratio of ω_1 and ω_2. Since the exponential tail of the wave function is determined entirely by $\langle x^2 \rangle$, it follows that the whole probability distribution is shrunk up by comparison with the unperturbed state (cf. Appendix B). If now we consider a tunnelling problem, where, let us say, the escape is in the x direction and is obtained by adding to (A.4) a suitable function of x (e.g., $-\lambda x^n$, $n > 2$), then it is intuitively obvious that the effect of the magnetic field, or more generally of a velocity-dependent coupling of the type (A.2), will be to *depress* the tunnelling probability (at least in the WKB limit)[24] Thus the results which we are

[24] This result can be made quite rigorous in the special case $\omega_1 = \omega_2$, since the conversion of $-\lambda x^n$ to $-\lambda(x^2 + y^2)^{n/2}$ giving a spherically symmetric problem, cannot decrease the tunnelling rate.

about to obtain by a more roundabout route in this appendix and Section 4 are physically very reasonable (cf. also the remarks in Section 6).

We now turn to explicit consideration of electromagnetic coupling. In what follows it is helpful to remember that in practice it will be the *electromagnetic* coordinates (flux, etc.) which are the analog of the principal coordinate q used in the text of the paper, and the particle-like coordinates (positions, currents, etc.) which are the analog of the environment; thus the situation is the reverse of that of, say, an atom interacting with the electromagnetic radiation field in a cavity.

Consider first the case of a simple one-dimensional harmonic oscillator with a charge e which interacts with the electromagnetic field. It is consistent for present purposes to set $\varphi = \nabla \cdot \mathbf{A} = 0$ (cf. below). If $x(t)$ denotes its coordinate and $E(t) \equiv -\partial A(t)/\partial t$ the electric field, its classical equation of motion will be

$$\ddot{x}(t) + \omega_0^2 x = -\frac{e}{m}\frac{\partial A(t)}{\partial t}. \tag{A.5}$$

In the present context it is important to describe correctly not only the influence of the field on the particle but also that of the particle on the field. Strictly speaking of course, the latter is described by a vector potential $\mathbf{A}(\mathbf{r}t)$ which obeys the usual equation (in units such that $c = 1$)

$$\nabla^2 \mathbf{A} - \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mathbf{j}(\mathbf{r}, t) \tag{A.6}$$

and is subject to certain boundary conditions (imposed, for example, by the walls in which the system is enclosed). However, for present purposes it is adequate to ignore these complications and treat $\mathbf{A}(t)$ in (A.5) as the vector potential (strictly, its component along the direction of the particle motion) averaged over a region large compared to the scale of particle motion; in that case is sufficient to require[25] that the equation of motion should be given by

$$\frac{\partial^2 A}{\partial t^2} = e\dot{x}(t), \tag{A.7}$$

where, for simplicity, we have considered a region of unit volume.

We now require a Lagrangian $L(q_i, \dot{q}_i; t)$ such that the Lagrange equations for the various generalized coordinates q_i, namely,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \tag{A.8}$$

[25] We emphasize that it is very straightforward to handle the problem taking into account the spatial variation of $\mathbf{A}$. The effect would be simply to add a "potential energy" proportional to $\mathbf{A}^2$ to the Lagrangian, where $\mathbf{A}$ is now the amplitude of the relevant normal mode. The only reason we do not do so is to avoid obscuring the very simple point made here.

should reproduce equations (A.5) and (A.7). First we must choose our generalized coordinates. One of them we shall always take in the present discussion to be the vector potential $A(t)$, and the "natural" choice of the other is the particle coordinate $x(t)$. With these choices we find the standard textbook Lagrangian (see, e.g., [63])

$$L(x, \dot{x}; A, \dot{A}) = \tfrac{1}{2}m\dot{x}^2 - \tfrac{1}{2}m\omega_0^2 x^2 + \frac{\dot{A}^2}{2} + e\dot{x}A. \tag{A.9}$$

If we prefer to use a Hamiltonian formulation, we get by the standard prescriptions the form (where $p_x \equiv \partial L/\partial \dot{x}$, $p_A \equiv \partial L/\partial \dot{A}$)

$$H = \frac{1}{2m}(p_x - eA)^2 + \tfrac{1}{2}p_A^2 + \tfrac{1}{2}m\omega_0^2 x^2. \tag{A.10}$$

Naturally, had we chosen, we could have used a Lagrangian with a coupling term $-ex\dot{A}$ rather than $e\dot{x}A$; then the functional form of the Hamiltonian expressed in terms of the new canonical momenta (which are of course different from the above ones) would be different: cf. the earlier discussion.

Now suppose that for some reason it is convenient to choose the second quantity $q_i(t)$ which is to play the role of "coordinate" in our Lagrangian formalism to be directly related, not to the particle position $x(t)$ but to its *velocity* $v(t) \equiv \dot{x}(t)$. It is tempting simply to take our original Lagrangian (A.9) and try to re-express it in terms of $v(t)$ and *its* derivative. However, this is incorrect; in general, even if such a re-expression is possible, such a procedure does not result in the correct Lagrangian, i.e., if $v(t)$ is treated as a "coordinate" and inserted into (A.8) with L given in this way, the correct equations of motion (A.5) and (A.7) do not result. This is because Lagrange's equations, while invariant under ordinary coordinate transformations, are *not* (unlike Hamilton's equations) invariant under contact transformations which mix the q_i and $\dot{q}_i$ together [63]. Instead we must proceed as follows. Introduce a quantity $y(t)$ defined by

$$y(t) \equiv \omega_0^{-1}(\dot{x}(t) + eA(t)/m) \tag{A.11}$$

which will now play the required role of coordinate q_i. In terms of this variable the required equations of motion (A.8) and (A.7) take the form (we differentiate (A.5) once)

$$\ddot{y} = -\omega_0^2(y - eA(t)/m\omega_0) \tag{A.12}$$

$$\ddot{A} = e\omega_0(y - eA(t)/m\omega_0) \tag{A.13}$$

The sole criterion for the choice of a Lagrangian function of y, $\dot{y}$, A, and $\dot{A}$ is that insertion of it in Eqs. (A.8), with $q_1 \equiv A$, $q_2 \equiv y$, should yield Eqs. (A.12) and (A.13). It is easy to verify that a possible choice is

$$L = \tfrac{1}{2}m\dot{y}^2 + \tfrac{1}{2}\dot{A}^2 - \tfrac{1}{2}m\omega_0^2\left(y - \frac{eA}{m\omega_0}\right)^2. \tag{A.14}$$

It is further easily verified that the Hamiltonian obtained from (A.14), when expressed (with the aid of the equations of motion) in terms of x, $\dot{x}$, etc., is identical to the form (A.10)—as indeed we should expect, since the Hamiltonian formalism *is* invariant under contact transformations. The crucial point is that if we wish to use as our "coordinate" for the purposes of the Lagrangian formalism a quantity related to the particle velocity, then the coupling of such a term to the electromagnetic field must include not only a linear term ($\sim yA$) *but also a quadratic term* $-e^2A^2/2m$, which is equivalent to a positive addition to the harmonic potential energy.

It is now relatively straightforward to extend the argument to cases of real physical interest. Consider for example, as in Section 2, a simple LC-circuit made out of superconducting materials. If we choose as our basic "coordinate" the flux Φ threading the self-inductance of the circuit, a suitable "free" Lagrangian L_0 to describe its behaviour is

$$L(\Phi, \dot{\Phi}) = \frac{1}{2} C\dot{\Phi}^2 - \frac{1}{2L}\Phi^2. \tag{A.15}$$

Now imagine that we connect a piece of normal metal (resistor) in parallel with the capacitor. The microscopic expression for the correction to the Lagrangian which we would write down in ordinary electrodynamics is

$$\Delta L = \int \mathbf{j}(\mathbf{r}) \cdot \mathbf{A}(\mathbf{r})\, dV, \tag{A.16}$$

where $\mathbf{j}(\mathbf{r}) \equiv \sum_i e\mathbf{v}_i \delta(\mathbf{r} - \mathbf{r}_i)$ is the electric current density operator of the electrons (labelled by subscript i) in the resistor, and the integral is taken over its volume. If l denotes the distance along the resistor, then (A.16) can be rewritten

$$\Delta L = \int \mathbf{A} \cdot \mathbf{dl}\, I_n(l), \tag{A.17}$$

where $I_n(l)$ is the current flowing across a surface in the wire perpendicular to $\mathbf{dl}$ at point l. Neglecting plasma effects, we can say that $I_n(l)$ will be independent of l, and moreover the integral $\int \mathbf{A} \cdot \mathbf{dl}$ is, apart from a constant which is of no interest to us and can be eliminated by a suitable choice of gauge, equal to the total flux Φ threading the circuit.[26] (This follows because the exclusion of electric fields from the interior of the superconducting wires ensures that $\mathbf{A}$ is constant in them.) So we can rewrite ΔL in the simpler form

$$\Delta L = I_n \Phi, \tag{A.18}$$

[26] We assume that the resistor is not so far from the capacitor that retardation effects are appreciable. This is certainly true for the RSJ model of a SQUID, where the "resistor" is not a separate element but is built into the junction itself.

where I_n is the current flowing through the resistor. It is also possible, if we wish, to define a charge Q_n by the equation

$$Q_n(t) \equiv \int_0^t I_n(t')\, dt' \tag{A.19}$$

and then we can rewrite ΔL in the equivalent form (apart from a total time derivative)

$$\Delta L = -Q_n \dot{\Phi}. \tag{A.20}$$

Now, provided we are prepared to treat I_n as a quantity whose dynamics we somehow know a priori, (A.18) (or (A.20)) is a perfectly satisfactory interaction term in the Lagrangian. Indeed, adding it to (A.15) and writing down the Lagrange's equation for Φ yields immediately the required circuit equation

$$C\ddot{\Phi} - I_n + \Phi/L = 0. \tag{A.21}$$

If we were now to invoke the normal phenomenological equation for the resistor, namely, Ohm's law, we would of course reduce this to

$$C\ddot{\Phi} + \sigma_n \dot{\Phi} + \Phi/L = 0, \tag{A.22}$$

where σ_n is the normal conductance of the resistor. However, our goal is to *justify* this result (or equivalently Ohm's law) from a microscopic Lagrangian which includes the resistor degrees of freedom explicitly. To do this it is convenient to take the quantity Q_n (Eq. (A.19)) and express it in terms of the normal coordinates of the resistor Q_α as was done in Section 3 (As in that section, we implicitly assume here that the motion of the flux will perturb the resistor sufficiently weakly that the response is linear). Naturally, in the presence of impurity scattering, phonons, etc. the physical interpretation of the Q_α is likely to be very complicated; nevertheless we can write quite generally

$$Q_n = \sum_\alpha \tilde{C}_\alpha Q_\alpha. \tag{A.23}$$

If, therefore, we wish to treat as "coordinates" the quantities Q_α themselves (the "natural" choice), then the correct form of the Lagrangian is that derived from (A.18), that is

$$\Delta L = \Phi \sum_\alpha \tilde{C}_\alpha \dot{Q}_\alpha \tag{A.24}$$

If on the other hand we wish to use as "coordinates" quantities with the dimensions of current and thereby avoid explicitly velocity-dependent couplings, then the

argument developed above for the simple harmonic oscillator goes through almost verbatim: the simplest choice is the quantity

$$x_\alpha \equiv \omega_\alpha^{-1}(\dot{Q}_\alpha + \tilde{C}_\alpha \Phi/m_\alpha) \tag{A.25}$$

(where m_α is the "mass" of the normal mode as discussed in Section 3) and the correct form of the Lagrangian is now ($C_\alpha \equiv \omega_\alpha \tilde{C}_\alpha$)

$$\begin{aligned} L = {} & \tfrac{1}{2}C\dot{\Phi}^2 - \Phi^2/2L + \sum_\alpha \{\tfrac{1}{2}m_\alpha \dot{x}_\alpha^2 - \tfrac{1}{2}m_\alpha \omega_\alpha^2 x_\alpha^2\} \\ & - \Phi \sum_\alpha C_\alpha x_\alpha - \Phi^2 \left\{\tfrac{1}{2}\sum_\alpha C_\alpha^2/m_\alpha \omega_\alpha^2\right\} \end{aligned} \tag{A.26}$$

in agreement with (3.5). Thus, correct treatment of the electromagnetic coupling automatically introduces a counter-term of the form (3.4) which eliminates the unphysical frequency-renormalization effects which would otherwise occur. The argument can clearly be generalized to more complicated cases such as that of a SQUID in the RSJ model quite trivially: all we need do is to add to (A.15), and hence to (A.26), the extra Josephson locking energy. Moreover, we see that since the case of adiabatic coupling is formally identical to that of electromagnetic coupling (see the discussion in Appendix C), the counter-term must be included in this case also.

We finally comment on two related points. First, the quantity $J(\omega)$ (Eq. (3.8)) is in this case, not surprisingly, nothing other than the imaginary part of the resistor current–current correlation function $\langle\langle I_n : I_n \rangle\rangle(\omega)$. In fact, using the fact that Φ in Eq. (A.25) is a c-number with respect to the environmental oscillator states and denoting by $|\alpha\rangle$ the first excited stated of the oscillator labelled α, we find

$$\begin{aligned} \mathrm{Im}\langle\langle I_n : I_n \rangle\rangle(\omega) &\equiv \pi \sum_n |\langle 0 | I_n | n \rangle|^2 \, \delta(E_n - E_0 - \hbar\omega) \\ &= \frac{\pi}{\hbar} \sum_\alpha \tilde{C}_\alpha^2 \, |\langle 0 | \dot{Q}_\alpha | \alpha \rangle|^2 \, \delta(\omega - \omega_\alpha) \\ &= \frac{\pi}{\hbar} \sum_\alpha C_\alpha^2 \, |\langle 0 | x_\alpha | \alpha \rangle|^2 \, \delta(\omega - \omega_\alpha) \\ &= \frac{\pi}{2} \sum_\alpha \frac{C_\alpha^2}{m_\alpha \omega_\alpha} \delta(\omega - \omega_\alpha) \equiv J(\omega). \end{aligned} \tag{A.27}$$

Thus for small ω we have a relation analogous to (3.9)

$$J(\omega) = \sigma_n \omega, \tag{A.28}$$

where σ_n is the conductance of the shunt.

Second, we may ask how the above analysis applies to the case of a superconducting shunt, a case in which we certainly do expect a frequency shift: in this case,

as the phenomenological equations show, there is a kinetic inductance $m/N_A e^2$ which is effectively in parallel with the electromagnetic inductance L (as can be seen by redrawing the circuit appropriately); here, N_A is the number of superconducting electrons per unit area of the shunt. Thus the frequency shift is *positive* in this case ($1/L \to 1/L + N_A e^2/m$). In the context of our discussion this is a case where the assumption of smoothness of the function $J(\omega)$ is spectacularly violated: the quantity $\omega^{-1}J(\omega)$ has in fact a delta-function singularity at $\omega = 0$, and this leads to a negative real part of $K(\omega)$ in (C.35) which produces the required shift.

Finally, we note that everything we have said in this appendix generalizes very simply to the case of nonlinear damping: see Appendix C.

The patient reader who has persevered thus far may well feel that in this appendix we have managed to make a mountain out of a molehill. After all, it is perfectly easy to treat the problem of coupling of the flux to a resistor (and to demonstrate the absence of a frequency shift if the resistor is normal) by using the conventional Lagrangian or Hamiltonian formalism without making the rather exotic choice of "coordinate" embodied in Eq. (A.25). The reason we have nevertheless gone through the above argument explicitly is twofold. First, it is convenient to be able to use, as we do in Section 3, a formalism which deals in a unified manner with a general coupling which leads to dissipation, irrespective of whether or not this coupling also produces a physical frequency shift. Second, while the normal choice of Lagrangian coordinates is perfectly convenient for work in "real time," it is somewhat inconvenient when we intend to carry out a Wick rotation as in Eq. (4.3), since it will mean that even after this transformation the argument of the exponent in the functional integral is not real.

APPENDIX B: THE DAMPED SIMPLE HARMONIC OSCILLATOR

In this Appendix we shall establish the form of the thermal equilibrium density matrix of the damped harmonic oscillator by a simple and direct method, and then compare the results with those obtained by using the formalism of this paper.[27] We assume, as always, that the dissipation arises from a (strictly) linear coupling to a bath of environmental "oscillators," i.e., that the appropriate Lagrangian is given by Eq. (3.12) with $V(q) \equiv \frac{1}{2}M\omega_0^2 q^2$. As we have argued in Section 3, this assumption is very much less restrictive then it looks. The purpose of this appendix is, first, to check the method of this paper against an exactly soluble model and in particular to reinforce further the identification of the quantity $\omega^{-1}J(\omega)$ with the phenomenological friction coefficient η, and secondly to provide some results which help us to guess the likely temperature-dependence of the escape probability (cf. Section 6).

[27] The question of frequency shifts is ignored in this Appendix: thus, ω_0 is always the renormalized frequency.

The (reduced) thermal density matrix for the oscillator is defined by Eq. (4.11), that is,

$$K(q_i, q_f; \beta) \equiv \int \prod_{\alpha} dx_\alpha \, \psi_n^*(q_i, \{x_\alpha\}) \, \psi_n(q_f, \{x_\alpha\}) \exp -\beta E_n, \tag{B.1}$$

where $\beta \equiv (k_B \theta)^{-1}$, θ being the real (physical) temperature. (In this appendix, unlike Section 4, we will keep β finite unless otherwise stated). Now, the Lagrangian (3.7) is bilinear in q and the x_α, and hence can in principle be exactly diagonalized. The resulting energies of the universe are the sums of the energies of the normal modes, whose coordinates y_j are related to q_0 and the x_α by some linear transformation; moreover, the thermal density matrix is the product of the density matrices for each normal mode which have the standard form (see [56, p. 51])

$$\rho(y_j, y_j') = \left(\frac{\mu_j \nu_j}{2\pi\hbar \sinh \beta\hbar\nu_j}\right)^{1/2} \exp -\left\{\frac{\mu_j \nu_j}{2\hbar \sinh \beta\hbar\nu_j} \times [(y_j^2 + y_j'^2) \cosh \beta\hbar\nu_j - 2y_j y_j']\right\}, \tag{B.2}$$

where ν_j is the oscillation frequency of the jth normal mode and μ_j its "mass." The only point we need to note about Eq. (B.2) is that the exponent is of Gaussian form and symmetric with respect to the interchange of y_j and y_j'. Since q and the x_α's are related to the y_j by some linear transformation, it follows from the properties of Guassian integrals that $K(q_i, q_f; \beta)$ is also Gaussian and symmetric with respect to interchange of q_i and q_f. Introducing the "centre-of-mass" and "relative" variables $X \equiv (q_i + q_f)/2$, $\xi \equiv q_i - q_f$, we see then that the most general allowable form of $K(X, \xi; \beta)$ is

$$K(X, \xi, \beta) = C_0 \exp -\tfrac{1}{2}(\lambda^{-1} X^2 + \mu \xi^2), \tag{B.3}$$

where the coefficients C_0, λ, and μ are functions both of β and of the parameters of the problem. To determine the form of λ and μ we can use the following simple argument. We first note that these two quantities are nothing but the thermal mean-square values of coordinate and (apart from a constant) momentum

$$\langle q^2 \rangle \equiv \int X^2 K(X, 0; \beta) \, dX = \lambda, \tag{B.4}$$

$$\langle p^2 \rangle \equiv \int dX \int dp \, p^2 \int \exp(ip\xi/\hbar) \, K(X, \xi; \beta) \, d\xi = \hbar^2 \mu, \tag{B.5}$$

where we used the normalization condition

$$\int K(X, 0; \beta) \, dX = 1. \tag{B.6}$$

On the other hand, we may obtain expressions for $\langle q^2 \rangle$ and $\langle p^2 \rangle$ in terms of the parameters of the model, as follows. As is well known, the response of the (quantum-mechanical expectation value of the) coordinate of a simple (undamped) harmonic oscillator is identical in classical and quantum mechanics, and therefore the response function $\chi(\omega) \equiv (\delta x/\delta F)_\omega$ is identical in the two cases. Applying this theorem to the true eigenmodes of the system and then performing the linear transformation to q and the x_α, we conclude that the response function $\chi_q(\omega)$ for the system variable q may itself be calculated by classical arguments. But we know that the classical equation of motion is simply (2.8) with $V(q) \equiv \frac{1}{2}M\omega_0^2 q^2$, and hence[28]

$$\chi_q = \frac{1}{M(\omega_0^2 - \omega^2 - 2i\gamma\omega)} \qquad (\gamma \equiv \eta/2M). \tag{B.7}$$

But, by the fluctuation-dissipation theorem we have

$$\langle q^2 \rangle = \frac{\hbar}{\pi} \int_0^\infty \coth(\beta\hbar\omega/2)\, \mathrm{Im}\, \chi_q(\omega)\, d\omega \tag{B.8}$$

and hence, identifying $\langle q^2 \rangle$ with λ as above,

$$\lambda = \frac{\hbar}{M\pi} \int_0^\infty \frac{\coth(\beta\hbar\omega/2)\, 2\gamma\omega}{(\omega_0^2 - \omega^2)^2 + 4\gamma^2\omega^2}\, d\omega. \tag{B.9}$$

The expression is convergent at the upper limit as it stands, and can be evaluated explicitly in the low- and high-temperature limits (cf. below).

In a similar way, since the momentum response function $\chi_p(\omega)$ is related to $\chi_q(\omega)$ by

$$\chi_p(\omega) = M^2\omega^2\chi_q(\omega) \tag{B.10}$$

we find the result

$$\mu = \frac{M}{\pi\hbar} \int_0^\infty \frac{\coth(\beta\hbar\omega/2)\, 2\gamma\omega}{(\omega_0^2 - \omega^2)^2 + 4\gamma^2\omega^2}\, d\omega. \tag{B.11}$$

Unlike (B.9), this expression is logarithmically divergent. The physical reason for this is that the phenomenological Eq. (2.8) has been implicitly assumed to apply for arbitrarily short times, whereas in reality there will be some "microscopic" time beyond which the inertia of the environment, etc., will come into play (cf. [57, Chap. 2]). To put it another way, the assumed form (3.9) of the environment spectral density is not valid to arbitrarily high frequencies (see the discussion in Appendix C). Thus we must introduce some form of high-frequency cutoff on the integral in (B.11) which takes effect for $\omega \sim \omega_c$, and its value is then of the order of $\gamma \ln(\omega_c/\omega_0)$ or $\gamma \ln(\omega_c/\gamma)$ according as the damping is weak or strong. (This assumes $\beta\hbar\omega_c \gg 1$: cf. below).

[28] Our convention for the definition of response functions follows that of Pines and Nozières [64], apart from a minus sign.

We now turn to an alternative derivation of some of the above results using the formalism employed in the body of the paper. Apart from the prefactor, the reduced density matrix is given exactly by the expression

$$K(q_i, q_f; \beta) \equiv K(x, \xi; \beta) = \text{const} \exp - S_{cl}(q_i, q_f; \beta)/\hbar, \tag{B.12}$$

where S_{cl} is the "action" (4.31) evaluated along the classical path which starts at q_i at "time" zero and ends at q_f at β. Let us take a Fourier transform of the path $q(\tau)$, writing

$$q(\tau) = \sum_n q_n e^{i\omega_n \tau}, \qquad \omega_n = 2\pi n/\beta. \tag{B.13}$$

Then one easily varifies that

$$S_{cl} = \beta \sum_n \tfrac{1}{2} M(\omega_n^2 + \omega_0^2 + 2\gamma|\omega_n|)\,|q_n|^2. \tag{B.14}$$

Now, any path $q(\tau)$ leaving q_i at time zero and arriving at q_f at time β can be decomposed into the sum of a "symmetric" part $q_s(\tau)$ which goes from $(q_i + q_f)/2 \equiv X$ back to X, and an "antisymmetric" part $q_a(\tau)$ which runs from $(q_i - q_f)/2 \equiv \xi/2$ to $-\xi/2$. It is obviouus from (B.14) that the action of the original path is just the sum of those due to the two paths separately, and moreover that these contributions are proportional to X^2 and ξ^2, respectively. Hence we can immediately write (renaming T as β)

$$K(x, \xi; \beta) = \text{const} \exp - \tfrac{1}{2}\{\lambda'^{-1} X^2 + \mu' \xi^2\}, \tag{B.15}$$

where the constants λ' and μ' are to be determined from the action of the symmetric and antisymmetric paths, respectively.

The determination of λ' is straightforward. The symmetric path is determined by minimizing expression (B.14) subject to the boundary condition $q(0) \equiv q(\beta) = X$. This gives

$$q_n = K(\omega_n^2 + \omega_0^2 + 2\gamma|\omega_n|)^{-1}, \tag{B.16}$$

where

$$K \equiv X/\sum_n (\omega_n^2 + \omega_0^2 + 2\gamma|\omega_n|)^{-1}. \tag{B.17}$$

Substituting (B.16) back into (B.14) and comparing the result with (B.15), we can make the identification

$$\lambda' = \frac{\hbar}{M} \sum_n \frac{1}{\omega_n^2 + \omega_0^2 + 2\gamma|\omega_n|}. \tag{B.18}$$

By rewriting the expression (B.9) in the equivalent form

$$\lambda = \frac{\hbar}{M\pi} \operatorname{Im} \int_{-\infty}^{\infty} \frac{d\omega}{\omega_0^2 - \omega^2 - 2i\gamma\omega} \left(\frac{1}{\exp \beta\omega - 1} \right) \tag{B.19}$$

and evaluating the integral by contour integration,[29] we see that $\lambda \equiv \lambda'$ as we should expect.

The determination of μ' is slightly more delicate because of the discontinuity of the antisymmetric path at the ends of the region of integration. For convenience we shall first shift the origin of the "time" axis so that this region is $-\beta/2 \leqslant \tau \leqslant \beta/2$. Then we replace the boundary conditions $q(\beta/2) = -q(-\beta/2) = \xi/2$ by the stipulation that $q(\tau)$ is periodic with period β plus the condition

$$q\left(\frac{\beta - \varepsilon}{2}\right) - q\left(-\frac{(\beta - \varepsilon)}{2}\right) = \xi, \tag{B.20}$$

where ε is a small positive quantity which will be allowed to tend to zero at the end of the calculation. It is evident that these conditions produce a "semiclassical" path which is smooth throughout most of the region of integration but has a steep part at the ends (with slope $\sim\xi/\varepsilon$) which in the limit $\varepsilon \to 0$ tends to a straight line.

We now proceed as above by minimizing the action (B.14) subject to the boundary condition (B.20). This leads to the result

$$q_n = (-1)^{n+1} \frac{iL \sin(\omega_n \varepsilon/2)}{\omega_n^2 + \omega_0^2 + 2\gamma|\omega_n|}, \tag{B.21}$$

where the Lagrange multiplier L is determined by

$$\begin{aligned} L &= \xi \Big/ 2 \left[\sum_n \sin^2(\omega_n \varepsilon/2)/(\omega_n^2 + \omega_c^2 + 2\gamma|\omega_n|) \right] \\ &\equiv \xi/Q(\varepsilon), \end{aligned} \tag{B.22}$$

where we note that $Q(\varepsilon)$ is proportional to ε in the limit $\varepsilon \to 0$. Substituting (B.21) back into (B.14), we find

$$S_{\mathrm{cl}} = \tfrac{1}{4}\beta M \xi^2 / Q(\varepsilon) \tag{B.23}$$

so that S_{cl} diverges as ε^{-1} in the limit $\varepsilon \to 0$. It is clear that this divergence is entirely due to the "steep" parts of the path; in fact, we easily verify that in this limit the expression becomes exactly the contribution of the straight line

[29] Recall that the sum in (B.18) involves all n, whereas a contour in the upper half-plane will include only positive n.

$q(\tau) = -(\xi/\varepsilon)(\tau - \beta/2)$, $(|\tau - \beta/2)| \leqslant \varepsilon/2)$, namely, $\frac{1}{2}M\xi^2/\varepsilon$. This contribution should clearly not be counted in the physical action which enters (B.12). If we write

$$Q(\varepsilon) = \frac{\beta}{2}(\varepsilon - A\varepsilon^2 + O(\varepsilon^3)) \qquad (\varepsilon \to 0), \tag{B.24}$$

then the explicit expression for A is

$$A = \beta^{-1} \sum_n \left(1 - \frac{\omega_n^2}{\omega_n^2 + \omega_0^2 + 2\gamma|\omega_n|}\right) \tag{B.25}$$

and the physical action is $\frac{1}{2}M\xi^2 A$. Substituting this in (B.12) and comparing with (B.15), we find

$$\mu' = \frac{M\beta^{-1}}{\hbar} \sum_n \left(1 - \frac{\omega_n^2}{\omega_n^2 + \omega_0^2 + 2\gamma|\omega_n|}\right). \tag{B.26}$$

A suitable contour integration of (B.11) (where we must remember that part of the "contour at infinity" will contribute a finite amount) shows that $\mu' = \mu$. Thus the method of this paper yields the correct expression for the complete groundstate density matrix of the damped harmonic oscillator, as of course it must. It is interesting to note that while the contribution to the kinetic-energy term which arises from the discontinuity of $q(\tau)$ at the ends of the region of integration is spurious as discussed above, the contribution of the "damping" term (i.e., the last term in (4.27)) from this discontinuity[30] is real and is indeed necessary to ensure the correct logarithmic divergence of (B.11) at the upper limit. (More realistically, of course, the $(\tau - \tau^1)^{-2}$ behaviour of the integrand of this term is cut off at a value of $|\tau - \tau^1|$ of the order of ω_c^{-1}, thereby producing a large but finite value of the integral as discussed in connection with (B.11)).

Finally we comment briefly on the temperature-dependence of the quantities λ and μ (see Eq. (B.3)). First, using the familiar Kramers–Kronig and longitudinal sum rules for $\mathrm{Im}\,\chi_q(\omega)$ (see, e.g.,. Nozières and Pines [64] we see that in the high-temperature limit where $\coth(\beta\hbar\omega/2)$ tends to $2/\beta\hbar\omega$, λ tends to $(\beta M\omega_0^2)^{-1}$, and μ tends to $M/\hbar^2\beta$; thus in this limit the density matrix becomes (writing the temperature $(k_B\beta)^{-1}$as θ)

$$K(X, \xi; \theta) = \text{const} \exp -\frac{1}{2}\left(\frac{M\omega_0^2}{k_B\theta}X^2 + \frac{k_B\theta M}{\hbar^2}\xi^2\right) \tag{B.27}$$

independently of the degree of damping. It should be noted, however, that whereas λ achieves essentially its asymptotic form for $k_B\theta \gg \hbar\omega_0^2/\gamma \ln(\gamma/\omega_0)$ (cf. below), μ does

[30] That is, the contribution from the regions close to the discontinuity on opposite sides, where $q(\tau')$ is different by $\sim\xi$ from $q(\tau)$.

not do so until $k_B\theta \gg \hbar\gamma \ln(\omega_0/\gamma)$. In the zero-temperature limit the integral (B.9) can be evaluated explicitly and we find

$$\lambda \equiv \langle q^2 \rangle = \frac{\hbar}{2M\omega_0} f(\alpha) \qquad (\alpha \equiv \gamma/\omega_0), \tag{B.28}$$

where

$$f(\alpha) = \frac{1}{\sqrt{1-\alpha^2}} \left[1 - \frac{2}{\pi} \tan^{-1} \frac{\alpha}{\sqrt{1-\alpha^2}} \right], \qquad \alpha \leqslant 1, \tag{B.29a}$$

$$f(\alpha) = \frac{1}{\sqrt{\alpha^2-1}} \frac{1}{\pi} \ln \left| \frac{\alpha + \sqrt{\alpha^2-1}}{\alpha - \sqrt{\alpha^2-1}} \right|, \qquad \alpha \geqslant 1. \tag{B.29b}$$

We note that $f(\alpha)$ attains the value $2/\pi$ for $\alpha = 1$ and for large α has the asymptotic form $(2/\pi\alpha) \ln 2\alpha$. Thus the probability density distribution of the oscillator in coordinate space is severely narrowed in the overdamped limit, while the corresponding momentum distribution is broadened.

A final question of interest concerns the order of magnitude of the temperature at which the form of λ changes over from the zero-temperature value (B.29) to the high-temperature form $(\beta M\omega_0^2)^{-1}$. In the underdamped limit this temperature is obviously of order $\hbar\omega_0/k_B$; in the overdamped limit we see by studying the contribution to the integral (B.9) from the regions $\omega \ll \omega_0^2/\gamma$, $\omega_0^2/\gamma \ll \omega \ll \gamma$, and $\omega \gg \gamma$, respectively, that the crossover temperature is of order $((\hbar\omega_0^2/k_B\gamma) \ln(\gamma/\omega_0))$. That is, for a highly damped system the crossover to a classical description of the density distribution takes place at a much lower temperature than for the corresponding undamped system.

APPENDIX C:
The Coupling Mechanism in a Dissipative System

In this appendix our goals are (1) to make plausible the statement that the Lagrangian (3.5) can be taken as the most general description necessary for a dissipative system obeying Eq. (2.8) (or the generalization to a q-dependent η) (2) to justify Eq. (3.6) (or Eq. (3.10)) which relates the parameters of the model to the phenomenological friction coefficient (3) to elaborate somewhat the distinction made in Section 2 between "quasi-linear" and "strictly linear" dissipation. The first two parts of this programme, at least, to some extent duplicate or parallel various results already in the literature (cf. in particular [47]) and at first sight might even seem rather pointless, since in almost all cases of current practical interest in the context of the tunnelling problem one knows a priori, on the basis of some specific model, that the system-environment interaction is indeed of the form (3.2) (and indeed in many cases that it is of the "strictly linear" form (3.7)). However, since one of the goals of the present work is to put us in a position to infer as far as possible the effects of damping on tunnelling from a knowledge of the phenomenological quasiclassical equations alone, we feel there is some point in exploring how far these equations constrain the possible coupling mechanism.

We consider a system described in terms of a variable q whose classical equation of motion is of the form

$$M\ddot{q} + \eta(q)\dot{q} + \frac{\partial V}{\partial q} = F_{\text{ext}}(t), \tag{C.1}$$

where we have allowed the friction coefficient η to be amplitude-dependent but not, for the moment, to depend on frequency. It is to be emphasized that $F_{\text{ext}}(t)$ in Eq. (C.1) is taken to be a "true" force which acts on the system alone (and not, for example, also on the dissipative mechanism); to be precise, the total work done per unit time by the external force is taken to be simply $\dot{q}F_{\text{ext}}(t)$. Equivalently, we can say that the rate of dissipation of energy by the system into its environment is given by the expression $\eta(q)\,\dot{q}^2$. Although at first sight this might seem obviously true, it is desirable to be quite explicit about it, for the following reason: It is in fact possible to construct examples of tunnelling problems[31] in which the energy dissipated into the environment is not proportional to the squared velocity $\dot{q}^2$ of the *tunnelling* coordinate q, but to some other quantity, e.g., to the squared rate of change of the momentum conjugate to q. In such cases it may be possible formally to cast the equation of motion of the tunnelling variable in the form (C.1), but only at the expense of replacing $F_{\text{ext}}(t)$ on the right-hand side by a quantity which, although a given function of time controlled by the experimenter, is not a "true" force is the sense specified above. Moreover, such cases have the anomalous property that (at least in the examples we have studied) the (pseudo) "friction coefficient" $\eta(q)$ appearing in Eq. (C.1) is actually *negative* over much of the tunnelling region. It is therefore quite reasonable to expect that even the qualitative effect of dissipation in such cases may be quite different [65] from that obtained in this paper (cf. the results of Appendix B: for the case of a simple damped harmonic oscillator, dissipation proportional to $\dot{q}^2$ increases the mean square momentum $\langle p^2 \rangle$ and conversely dissipation proportional to $\dot{p}^2$ increases $\langle q^2 \rangle$). Although this "anomalous" case may possibly be relevant to some systems which can be studied in the future in the context of macroscopic quantum tunnelling, it does not appear to be relevant to any system of current practical interest and we therefore explicitly exclude it from consideration in this paper;[32] we hope to consider it elsewhere in the future.*

We now turn to the justification[33] of the "oscillator-bath" representation of the environment used in Section 3. We make essential use, here, of the condition that the environment $\mathscr{E}$ is only weakly perturbed by the motion of the system (though we will

[31] We thank A. Widom and T. D. Clark for a preprint [65] which, by explicitly constructing such a model problem, induced us to elucidate this point.

[32] Note that in the case of linear friction such explicit exclusion is redundant.

[33] The ensuing considerations refer only to the $T = 0$ case relevant to the work of this paper. The finite-temporation case needs separate discussion.

* *Note added in proof.* A paper on tunnelling in the presence of arbitrary linear dissipation, which includes the "anomalous" case as a special case, is currently in preparation for submission to *Physical Review B*.

eventually make a considerable generalization of this). Let the coordinates (not necessarily one dimensional) of $\mathscr{E}$ be collectively denoted ξ, and let its Hamiltonian when isolated be $\hat{H}_0$. The eigenfunctions $\psi_n(\xi)$ satisfy the equation

$$\hat{H}_0 \psi_n(\xi) = E_n \psi_n(\xi). \tag{C.2}$$

Let the interaction of $\mathscr{E}$ with an external agency described by some set of coordinates q be given by some real function V of ξ, q and their conjugate momenta, and define

$$V_{nm}(p, q) \equiv \int \psi_n^*(\xi)\, \hat{V}(p, q, p_\xi, \xi)\, \psi_m(\xi)\, d\xi = V_{mn}^*(p, q), \tag{C.3}$$

where p is the momentum conjugate to q and p_ξ that conjugate to ξ. We suppose now that the following condition is satisfied for all relevant values of p and q:

$$|V_{nm}(p, q)/(E_n - E_m)| \sim \varepsilon \ll 1. \tag{C.4}$$

If the system $\mathscr{E}$ then starts in its groundstate, the probability of a process in which n factors of V enter will be proportional to ε^{2n}. Hence, if we neglect terms of order ε^3 and higher, the only matrix elements we need be concerned with are the elements V_{j0}. We define Hermitian operators $\hat{x}_j$ and $\hat{p}_j$ in the space of the eigenfunctions $\psi_j(\xi)$ by their matrix elements

$$\langle j | \hat{x}_j | 0 \rangle = \langle 0 | \hat{x}_j | j \rangle = (\hbar/2m_j\omega_j)^{1/2}, \tag{C.5a}$$

all other elements zero,

$$\langle j | \hat{p}_j | 0 \rangle = -\langle 0 | p_j | j \rangle = i m_j \omega_j \langle j | x_j | 0 \rangle, \tag{C.5b}$$

all other elements zero, where $\omega_j \equiv (E_j - E_0)/\hbar$ and the "masses" m_j are arbitrary. We now add, for each index j, a (possibly fictitious) infinite set $|n_j\rangle$, $n = 0, 1, 2, \ldots$, of oscillator states: in this notation the groundstate of $\mathscr{E}$ is a product of $|0_j\rangle$, and the states called j above are now denoted $|1_j\rangle \prod_{k \neq j} |0_k\rangle$. The operators $\hat{x}_j$ and $\hat{p}_j$ are extended into the new space so constructed by assigning to them the conventional oscillator matrix elements ($\langle n_j + 1 | \hat{x}_j | n_j \rangle = (n_j + 1)^{1/2} (\hbar/2m_j\omega_j)^{1/2}$, etc.). We also construct an effective Hamiltonian

$$\hat{H}_{\text{eff}} = \sum_j \left(\frac{\hat{p}_j^2}{2m_j} + \frac{1}{2} m_j \omega_j^2 \hat{x}_j^2 \right). \tag{C.6}$$

Finally we add to $\hat{H}_{\text{eff}}$ an interaction Hamiltonian of the form

$$\hat{H}_{\text{int}} = -\sum_j (F_j(p, q)\, \hat{x}_j + G_j(p, q)\, \hat{p}_j), \tag{C.7}$$

where

$$F_j(p,q) \equiv \left(\frac{2m_j\omega_j}{\hbar}\right)^{1/2} \mathrm{Re}\, V_{j0}(p,q),$$

$$G_j(p,q) \equiv \left(\frac{2}{\hbar m_j\omega_j}\right)^{1/2} \mathrm{Im}\, V_{j0}(p,q). \tag{C.8}$$

Since the matrix elements of $\hat{H}_{\mathrm{int}}$ and the eigenvalues of $\hat{H}_{\mathrm{eff}}$ reproduce the true matrix elements $V_{nm}(q)$ and true energy level differences $E_n - E_0$ for all "singly excited" states, the combination $\hat{H}_{\mathrm{eff}} + \hat{H}_{\mathrm{int}}$ will generate the correct dynamics of the system to the extent that double excitation is negligible. It is only to the extent that this condition holds that the response of $\mathscr{E}$ will be linear in the external perturbation, so we can say that for our purposes the replacement of the true Hamiltonian by that appropriate to a set of harmonic oscillators is effectively equivalent to the assumption that the response of the environment to the perturbation exerted by the "system" is linear (or more precisely, that the response to a classical perturbation whose magnitude is less than or equal to the maximum value of the interaction function $\hat{H}_{\mathrm{int}}(p,q)$ obtained for values of p and q of interest in the tunnelling problem, is linear). Whether or not this condition holds for any specific physical system considered is of course a matter for detailed argument in each case; however, it should again be stressed that the condition of linearity of environment response is quite different from (and much weaker than) the condition of weak damping—just as the fact that a beam of light may be almost totally absorbed in a gas by no means implies that nonlinear effects have to be taken into account.

In Eq. (C.7) we have kept only terms of first order in the environment variables. Should we also keep (a) terms of second and higher order (b) terms of zeroth order? With regard to (a), this is a somewhat delicate question. Consider, for example, a term of the form

$$\hat{H}'_{\mathrm{int}} = \sum_{jj'} U_{jj'}(q)\, x_j x_{j'}. \tag{C.9}$$

If such a term exists, it will have an effect comparable to the original effective Hamiltonian H_{eff} (Eq. (C.6)) and will in general lead to a considerable rearrangement of the environment levels. Thus the weak-coupling condition is not satisfied and in such a case the adiabatic approximation must be used (see below). Thus for the moment we omit terms of the second and higher order. On the other hand terms of zeroth order in p_j and x_j may exist in the coupling (as, for example, in the case of electromagnetic coupling, where q is magnetic flux or something similar: see Appendix A), and will then influence the motion of the system, generally by changing either the effective potential or the effective mass or both. Thus we should generalize the interaction Hamiltonian slightly to read

$$\hat{H}_{\mathrm{int}} = -\sum (F_j(p,q)\,\hat{x}_j + G_j(p,q)\,\hat{p}_j) + \Phi(p,q), \tag{C.10}$$

where the form of Φ may depend on the parameters $(m_j, \omega_j, F_j(pq), G_j(pq))$ of the environment but not on the dynamical variables p_j, x_j.

It would no doubt in principle be quite possible to proceed with all our tunnelling calculations on the basis of a general Hamiltonian of the form (C.10). However, the Lagrangian (3.5) is so convenient that it is useful to see under what assumptions (C.10) can be reduced to (3.5) (or the equivalent Hamiltonian). The reasoning here will be plausible rather than totally rigorous: we suspect that rigorous proofs probably can be constructed for at least some of the statements made below, but they are likely to be extremely tedious and there seems little point in trying to produce them here.

We first observe that the phases of the excited-state wave functions $\psi_n(\xi)$ relative to the groundstate are arbitrary, and hence (in view of Eqs. (C.5)) the choice of the variables x_j and p_j contains an element of freedom. Moreover, in all cases of physical interest the variable q has a definite parity under the operation of time reversal (TR). Let us then use the above freedom to choose the "coordinate" x_j to have the same parity under TR as q; p_j then automatically has the opposite parity. Since the interaction Hamiltonian (C.10) must be invariant under TR, it then immediately follows that if q is even, $F_j(pq)$ must be an even function of p and G_j an odd one. (with no restrictions on their q-dependence). If q is odd, on the other hand, F_j must be an odd function of q and G_j an even one, with no restriction on the p-dependence. Further symmetries of the system may constrain the functions F_j and G_j further: e.g., if parity is conserved and q is odd under both TR and space reflection (as, for example, in the case of magnetic flux), then $F_j(G_j)$ must be even (odd) in p.

The ensuing argument is somewhat delicate. It is based on the fact that, in general, any dependence of the F_j or G_j on p is liable to produce a form of dissipation which cannot be cast in the form $\eta(q)\,\dot{q}^2$. This can be seen, for example, by considering the simple case $F_j = 0$, $G_j = C_j p$. Indeed, at first sight it is tempting to argue that $F_j(qp)$ and $G_j(qp)$ must be functions of q only, in which case it would follow immediately that in the case where symmetry constrains G_j to be an odd function of p, as above, $G_j \equiv 0$, and the form (3.5) is established. This is actually rather too simple. To obtain a more convincing argument we introduce the notation

$$(m_j\omega_j)^{-1/2}\,(F_j + im_j\omega_j G_j) = R_j e^{i\theta_j}, \tag{C.11}$$

where R_j is real. We then evaluate the second-order perturbation-theoretic expression for the energy dissipated by the system into its environment, taking the system dynamical variables $p(t)$ and $q(t)$ to be given functions of time and writing for the Fourier components $F_\omega \equiv \int F[p(t), q(t)]\, e^{i\omega t}\, dt$ the relations $F_\omega = (i\omega)^{-1}\,[\dot{F}]_\omega = (i\omega)^{-1}\,\{(\partial F/\partial p)\dot{p} + (\partial F/\partial q)\dot{q}\}$, etc. (For the details of the operations involved, cf. below.) In general this procedure gives us terms in the dissipation which are of the form $\eta(q)\,\dot{q}^2$, but also nonvanishing terms involving p and its derivatives. A necessary (though not sufficient) condition for these to be convertible into terms of the form $\eta(q)\,\dot{q}^2$ is that the system satisfies the relation

$$p = M\dot{q} \tag{C.12}$$

to within some function which, while it need not vanish identically, is guaranteed to be zero for any forced motion of the environment. The function in question is, of course, from Hamilton's equations just $\partial \hat{H}_{\text{int}}/\partial p$ (where any contribution from $\Phi(p, q)$ can be ignored, since it simply renormalizes the effective mass). The condition that $\partial \hat{H}_{\text{int}}/\partial p$ be zero for any forced motion of the environment yields, after a little algebra, the condition

$$\int_0^\infty \frac{d\omega'}{\omega'^2 - \omega^2} \{\omega' K_1(p, q: \omega) + i\omega K_2(p, q: \omega)\} = 0 \tag{C.13}$$

where

$$K_1(p, q: \omega) \equiv \sum_j \frac{1}{2} \frac{\partial}{\partial p} R_j^2 (p, q)\, \delta(\omega - \omega_j), \tag{C.14a}$$

$$K_2(p, q: \omega) \equiv \sum_j R_j^2 \frac{\partial \theta_j}{\partial p} (p, q)\, \delta(\omega - \omega_j). \tag{C.14b}$$

Barring, possibly, some sort of pathological conspiracy between the behaviour of the R_j and θ_j in different frequency regions, Eq. (C.13) implies

$$K_1(p, q: \omega) = K_2(p, q: \omega) = 0 \tag{C.15}$$

for all p, q, and ω. Evidently the most natural way to satisfy (C.15) is to set each of the $F_j(pq)$ and $G_j(pq)$ individually independent of p for all q. It then immediately follows from the considerations of the above paragraph that $G_j \equiv 0$ if q is either even under TR or odd under both TR and space reflection. This covers most of the cases of practical interest.

We cannot, of course, absolutely exclude the possibility that the F_j and G_j depend individually on p in such a way that Eq. (C.15) is satisfied (or, in cases satisfying neither of the above invariance coditions, that G_j is some nonzero even function of q). However, the above argument suggests that any such dependence would have to be rather pathological in nature (and, we suspect, might well be further restricted, if not totally excluded, by further arguments along the same lines). Thus, while we have no rigorous proof that the Lagrangian (3.5) is the most general which can give rise to the equation of motion (C.1), we feel that we have made it sufficiently plausible to use that Lagrangian as a basis for the work of this paper.

It should be noted that the argument leading to Eq. (C.12) above rests essentially on the fact that the quantities C_j, m_j, ω_j characterizing the environment do not already contain the phenomenological viscosity explicitly. Were this not the case, then the replacement of Eq. (C.12) by a relation of (for example) the form

$$p = M\dot{q} + \eta q \tag{C.16}$$

would in principle be acceptable, since the ω-dependent relationship between the Fourier components of $p(t)$ and $q(t)$ could be cancelled by a suitable distribution of

the environment spectral density, etc. Such a case would arise, for example, if we take an adiabatic-type coupling (cf. below) of the form

$$\hat{H}_{\text{int}} = \left(p - \sum_j c_j p_j\right)^2 \Big/ 2M - p^2/2M \tag{C.17}$$

and reduce it to the form (C.10) by incorporating the terms bilinear in the p_j in the (so-called) "unperturbed" Hamiltonian of the environment. The point is that the resulting description does not conform to the conditions specified above, since the environment is now *already* strongly perturbed by the interaction (in fact, the relevant spectral densities will now contain the damping η explicitly). While such a procedure may be relevant to the "anomalous" type of problem mentioned above, in the case of the canonical problem dealt with in this paper Hamiltonians of the form (C.17) are much more simply handled by another method: see below.

We next discuss how to incorporate into our scheme the important case of a system whose interaction with its environment is well described by the adiabatic approximation. In such a case let q and ξ represent, respectively, the "slow" degree of freedom and the collection of "fast" degrees of freedom; in the context of a macroscopic wave function description (see Introduction) these will be the macroscopic and microscopic coordinates, respectively. Let us schematically indicate a typical frequency of the slow degree of freedom by ω_0 and that of the fast degrees of freedom by ω_c; then by hypothesis we have $\omega_0/\omega_c \ll 1$. To zeroth order in the small parameter ω_0/ω_c the energy eigenfunctions have the form

$$\Psi_{ik}(q, \xi) = \varphi_{ik}(q)\, \chi_k(\xi; q), \tag{C.18}$$

where the functions $\chi_k(\xi; q)$ satisfy a Schrödinger equation in which q enters as a parameter

$$(\hat{T}(\xi) + V(\xi, q))\, \chi_k(\xi; q) = U_k(q)\, \chi_k(\xi; q), \tag{C.19}$$

where $\hat{T}(\xi)$ schematically represents the (q-independent) kinetic-energy or similar terms in the "fast" Hamiltonian. The functions $\varphi_{ik}(q)$ are solutions of the equation

$$\left\{-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial q^2} + U_k(q) - E_{ik}\right\} \varphi_{ik}(q) = 0, \tag{C.20}$$

where M is the coefficient of $\frac{1}{2}\dot{q}^2$ in the Lagrangian and may or may not have the significance of a physical mass (see Introduction). The functions $\Psi_{ik}(q, \xi) \equiv |ik\rangle$ form a complete orthonormal set, and the correction $\Delta\hat{H}$ to the Hamiltonian arising from the terms neglected in the zeroth-order adiabatic approximation (which is of order ω_0/ω_c) can be expressed in terms of its matrix elements in this basis

$$\langle ik\,|\Delta\hat{H}|\,jl\rangle = -\frac{\hbar^2}{2M}\int dq \int d\xi \left\{2\varphi_{ik}^*(q)\frac{\partial}{\partial q}\varphi_{jl}(q)\,\chi_k^*(\xi;q)\frac{\partial}{\partial q}\chi_l(\xi;q)\right.$$
$$\left. + \varphi_{ik}^*(q)\,\varphi_{jl}(q)\,\chi_k^*(\xi;q)\frac{\partial^2}{\partial q^2}\chi_l(\xi;q)\right\}. \tag{C.21}$$

Now we note a special feature of the adiabatic approximation as applied to macroscopic systems: The microscopic states $\chi_k(q)$ which are coupled together by $\Delta\hat{H}$ are in general different only in the behaviour of a small fraction ($\sim 1/N$) of the macroscopically many particles involved. As a consequence, the effective potential $U_k(q)$ felt by the macroscopic degree of freedom q depends on the "microscopic" index k only to order $1/N$, and so the states $\varphi_{ik}(q)$ are effectively independent of k. This simplifies the ensuing formulae considerably. We contruct a linear operator[34] $\hat{K}(q)$ in the space spanned by the functions $\chi_k(\xi;q)$ by giving its matrix elements in this basis:

$$\langle k\,|\hat{K}(q)|\,l\rangle \equiv \int \chi_k^*(\xi;q)\left(-i\hbar\frac{\partial}{\partial q}\right)\chi_l(\xi;q)\,d\xi. \tag{C.22}$$

The operator so defined is Hermitian and is parametrically dependent on q. Since it is a linear operator and the basis set $\chi_k(\xi, q)$ is complete for given q, it follows after a little algebra that the perturbation terms (3.15) can be written

$$\langle ik\,|\Delta\hat{H}|\,jl\rangle = \frac{1}{2M}\langle i\,|\,p\hat{K}_{kl}(q) + \hat{K}_{kl}(q)p + (\hat{K}^2)_{kl}\,(q)|\,j\rangle, \tag{C.23}$$

where as usual

$$\langle i\,|\,p\,|j\rangle \equiv \int \varphi_i^*(q)\left(-i\hbar\frac{\partial}{\partial q}\right)\varphi_j(q)\,dq \tag{C.24}$$

and we have written for clarity $\langle k\,|\hat{K}(q)|\,l\rangle \equiv \hat{K}_{kl}(q)$, etc. Thus the total Hamiltonian can be written

$$\hat{H} = \frac{1}{2M}(p + \hat{K}(q))^2 + U(q) + \hat{H}_{\text{env}}, \tag{C.25}$$

where $\hat{H}_{\text{env}}$ is an operator which refers exclusively to the "environment" (microscopic) coordinates ξ, and $\hat{K}(q)$ is another operator on these coordinates which is parametrically dependent on q. The cross-term in brackets is symmetrized as in (C.23). Apart from a factor of the coupling constant $(-e)$ the expression (C.25) is formally identical to the familiar Hamiltonian of a charged particle interacting with an electromagnetic vector potential in one dimension.

[34] For clarity in the next few lines we put circumflexes on those qualities (only) which are operators with respect to the microscopic coordinates ξ.

The argument that the environment can be treated as a set of harmonic oscillators now proceeds similarly to that given above for the weak-coupling case, with the operator $\hat{K}$ now playing the role of the system-environment interaction. If the Born–Oppenheimer expansion is to be a good approximation as we are assuming, then we can argue, as above, that as regards the term linear in $\hat{K}$ we need only the matrix elements of $\hat{K}$ between the ground and excited states, and that these can be put in one–one correspondence with oscillator matrix elements. If as above we choose the environment coordinates x_j to have the same parity under time-reversal as q, then since $\hat{K}$ is not a function of p the only possible form of $\hat{K}$ linear in the environment variables is

$$\hat{K}(q) = \sum_j K_j(q)\, p_j. \tag{C.26}$$

Suppose we insert this expression in Eq. (C.25). Is it guaranteed to give the term in $\hat{K}^2$ correctly? Here we invoke again the fact that, in view of the macroscopic nature the system, the microscopic states $\chi_k(q;\xi)$ differ only in the behaviour of a small number of particles. To the extent that this is so, the approximation

$$\langle k\,|\hat{K}^2|\,l\rangle \cong \langle k\,|\hat{K}|\,0\rangle\langle 0\,|\hat{K}|\,l\rangle \tag{C.27}$$

should be well justified, and we substitution (C.26) is legitimate. Thus we finally end up with an adiabatic Hamiltonian of the form

$$\hat{H} = \frac{1}{2M}\left(p + \sum_j K_j(q)\, p_j\right)^2 + U(q) + \hat{H}_{\text{env}} \tag{C.28}$$

of which (C.17) is a special case. Here $\hat{H}_{\text{env}}$ has the standard oscillator form (C.6).

It is now straightforward to show (most simply, though not most elegantly, by the procedure $m_j\omega_j x_j \rightleftarrows p_j$, addition of a total time derivative to the Lagrangian to give an interaction term proportional to the $\dot{x}_j$, and use of the procedure of Appendix A) that the description (C.28) is canonically equivalent to the Lagrangian (3.5), with $F_j(q)$ given by

$$F_j(q) = (m_j\omega_j)^2 \int_0^q K_j(q')\, dq'. \tag{C.29}$$

Thus the case of adiabatic coupling may be treated—albeit perhaps somewhat artificially—as a special case of our formalism, and the automatic occurrence of the counter-term in Eq. (3.5) indicates that in this case there is no physical frequency shift, as indeed we should expect.

The upshot of the above rather messy and inelegant argument is that for all practical purposes we can take the Lagrangian for our problem without loss of generality to be given by Eq. (3.5), that is,

$$L = \tfrac{1}{2}M\dot{q}^2 - V(q) + \tfrac{1}{2}\sum_j m_j(\dot{x}_j^2 - \omega_j^2 x_j^2) - \sum_j F_j(q)\, x_j$$
$$- \sum_j F_j^2(q)/2m_j\omega_j^2, \tag{C.30}$$

where, as in the main text, we shall abbreviate the last term as $\frac{1}{2}M\,\Delta\omega^2 q^2$. We now enquire what constraints are imposed on the parameters of the model by the condition that the motion of the system is to obey the phenomenological damped equation (C.1). For orientation we first discuss the separable linear case ($F_j(q) = C_j q$) and prove Eq. (3.6), that is,

$$J(\omega) \equiv \frac{\pi}{2}\sum_j \frac{C_j^2}{m_j\omega_j}\,\delta(\omega - \omega_j) = \eta\omega. \tag{C.31}$$

It is straightforward to verify that the condition (C.31) must hold if a classical treatment of the Lagrangian (C.30) is to produce the equation of motion (2.8). Adding to (C.30) a term $qF_{\text{ext}}(t)$ to describe the action of a possible external force, we obtain the equations of motion

$$M\ddot{q} = -\frac{\partial V}{\partial q} + M\,\Delta\omega^2 q + F_{\text{ext}}(t) - \sum_j C_j x_j, \tag{C.32}$$

$$m_j\ddot{x}_j = -m_j\omega_j^2 x_j - C_j q. \tag{C.33}$$

Taking the Fourier transforms of Eqs. (C.32)–(C.33), eliminating the x_j and using (C.30), we find that the Fourier transform $q(\omega)$ of the system trajectory satisfies the equation

$$-M\omega^2 q(\omega) = -\left(\frac{\partial V}{\partial q}\right)(\omega) + F_{\text{ext}}(\omega) + K(\omega)\, q(\omega), \tag{C.34}$$

where

$$K(\omega) \equiv \sum_j \frac{C_j^2\omega^2}{m_j\omega_j^2(\omega_j^2 - \omega^2)}. \tag{C.35}$$

We now make the crucial assumption (cf. above) that for all values of ω of interest for the problem under consideration the spectrum of the environmental oscillators may be treated as continuous. Defining as above

$$J(\omega) \equiv \frac{\pi}{2}\sum_j \frac{C_j^2}{m_j\omega_j}\,\delta(\omega - \omega_j) \tag{C.36}$$

we have

$$K(\omega) = \frac{2\omega^2}{\pi}\int_0^\infty \frac{J(\omega')\, d\omega'}{\omega'(\omega'^2 - \omega^2)}. \tag{C.37}$$

Now, to get agreement with the Fourier transform of the phenomenological equation (2.8) we need the conditions $\mathrm{Im}\, K(\omega) = \eta\omega$, $\mathrm{Re}\, K(\omega) = 0$ to hold for all values of interest. (To define the imaginary part of K we attribute, as usual, an infinitesimal positive imaginary part to ω.) The first condition obviously requires the relation (C.31) to hold for all relevant values of ω. Let us assume that (C.31) in fact holds to a good approximation for values of ω much less than some characteristic frequency ω_c of the environment. (For example, in a metal ω_c might be of the order of the inverse Drude relaxation time.) Then the real part of $K(\omega)$, for small ω, is easily seen to be of order $\eta\omega^2/\omega_c$, that is, smaller by a factor ω/ω_c than the imaginary part. Thus, provided that all the relevant values of ω are small compared to ω_c, this term will be negligible and the effect of the coupling to the dissipative mechanism (environment) will simply be to modify the classical equation of motion in a way described by the friction coefficient η (as in Eq. (2.8)). (Any frequency-renormalization effect has already been taken into account when $V(q)$ is taken to be the renormalized potential). In the more general case it is clear that the imaginary part (J) and real part of $K(\omega)$ are, respectively, the real and imaginary parts of the complex frequency-dependence admittance describing the dissipative element, which can in principle be obtained from experiment.

What are the "relevant" values of the frequency ω? We shall see in Section 5 that for a lightly damped system the frequencies of importance in the tunnelling rate are of the order of the classical small-oscillation frequency ω_0, while for the heavily over-damped case they are of the order of the "slow" relaxation frequency ω_0^2/γ ($\gamma \equiv \eta/2M$). Thus, if the condition $\omega \ll \omega_c$ is fulfilled for the undamped system it is certainly fulfilled for the same system subject to arbitrarily heavy damping. Actually, it is characteristic of macroscopic systems that the typical frequency of the macroscopic coordinate is very much lower than the frequencies of the environment which provides the dissipation, so that our condition is usually very well fulfilled.[35] Were this not so, the whole long-established tradition of describing macroscopic behaviour by simple phenomenological equations involving frequency-independent friction coefficients, conductivities, etc., would of course be ill-founded.

Naturally, it is trivial to extend the above results, if necessary, to cases where the relevant system frequencies are not negligibly small compared to ω_c: all that is necessary is to equate $J(\omega)$ to the imaginary part of the appropriate impedance $Z(\omega)$. Of course in such cases we expect the real part of $K(\omega)$ to be nonzero.

We now return to the general case, in which $F_j(q)$ is not necessarily linear in q. We might as well handle immediately the case in which η may be amplitude-dependent (though we assume that it is frequency-independent). We will establish a relation analogous to (C.31), namely,

$$\eta(q) = \frac{\pi}{2} \sum_j \frac{1}{m_j \omega_j^2} \left(\frac{\partial F_j}{\partial q} \right)^2 \delta(\omega - \omega_j). \tag{C.38}$$

[35] It is possible that it may break down in some Josephson junctions (or SQUIDS), where the capacitance is small and the small-oscillation frequency therefore high. If so, the the classical equation (2.8) has to be generalized.

This generalization of (3.24) is almost intuitively obvious, since we would expect quite generally that the "local" friction coefficient should be related to the correlations of the local fluctuating forces, which are proportional to the quantities $\partial F_j/\partial q$. To establish (C.38) formally we proceed as follows: First, we derive from the Lagrangian (C.30), with the interaction term and the counterterm as given, equations of motion for q and the x_j, and eliminate the latter. This gives the following equation for q:

$$M\ddot{q}(t) = -\frac{\partial V}{\partial q}(t) + F_{\text{ext}}(t) - \sum_j \frac{1}{m_j\omega_j^2}\frac{\partial F_j}{\partial q}(t)\frac{\hat{\mathscr{D}}^2}{\hat{\mathscr{D}}^2+\omega_j^2}F_j[q(t)], \tag{C.39}$$

where $\hat{\mathscr{D}}^2$ is shorthand for the operator d^2/dt^2. We can rewrite the last term in the form

$$\sum_j \frac{1}{m_j\omega_j^2}\frac{\partial F_j}{\partial q}[q(t)]\frac{\hat{\mathscr{D}}}{\hat{\mathscr{D}}^2+\omega_j^2}\left\{\frac{\partial F_j}{\partial q}[q(t)]\,\dot{q}(t)\right\}. \tag{C.40}$$

We now take the Fourier transform of (C.40) and try to equate it to that of a phenomenological dissipative term of the form $\eta(q)\dot{q}$. To get an exact equality we require the relation

$$[\eta(q)]_\omega = \sum_j \frac{1}{m_j\omega_j^2}\int_{-\infty}^{\infty} d\omega' \left(\frac{\partial F_j}{\partial q}\right)(\omega''-\omega'+\omega)\frac{\partial F_j}{\partial q}(\omega'-\omega'')\frac{i\omega'}{\omega_j^2-\omega'^2} \tag{C.41}$$

which must hold for arbitrary ω''. We first examine the part of this expression which comes from the pole of the expression $i\omega'/(\omega_j^2-\omega'^2)$. This is, if interpreted in the usual way,

$$\frac{\pi}{2}\sum_j \frac{1}{m_j\omega_j^2}\int_{-\infty}^{\infty} d\omega' \left(\frac{\partial F_j}{\partial q}\right)(\omega''-\omega'+\omega)\left(\frac{\partial F_j}{\partial q}\right)(\omega'-\omega'')$$
$$\times\{\delta(\omega_j-\omega')+\delta(\omega_j+\omega')\}. \tag{C.42}$$

By changing the variable of integration in (C.42) to $\omega''-\omega'$ and using the fact that this expression must (to get agreement with (C.41)) be independent of ω'', we see that it is legitimate to replace the ω' in the δ-functions by any (sufficiently small) frequency whatever. The integral over ω' is then a simple convolution, and a Fourier transformation back to the time variable immediately yields the condition (C.38). We note carefully that so far we have seen nothing to suggest that a purely linear dissipation ($\eta(q)\equiv\eta=\text{const}$) cannot be obtained from a general (nonseparable) interaction of the form (C.30).

Now, however, we consider the rest of the expression (C.41), which comes from the principal part of the integral over ω'. It is convenient to consider for definiteness

the special case $\omega = 0$ and to introduce the function (which is of course in general a function of the particular trajectory $q(t)$ we are considering)

$$K(\zeta, z) \equiv \sum_j \frac{1}{m_j \omega_j^2} \left| \left(\frac{\partial F_j}{\partial q} \right) (\zeta) \right|^2 \delta(\omega_j - z). \tag{C.43}$$

Then the equation of motion of q, in Fourier-transformed form, acquires an extra term $\Delta\eta_{\omega''}\dot{q}_{\omega''}$, where $\Delta\eta_{\omega''}$ is defined by

$$-i\,\Delta\eta_{\omega''} = \mathscr{P} \int_{-\infty}^{\infty} d\zeta \int_{-\infty}^{\infty} dz\, \frac{K(\zeta, z)(\zeta + \omega'')}{z^2 - (\zeta + \omega'')^2}. \tag{C.44}$$

In the special case of a linear separable interaction of the type (3.7), $K(\zeta)$ is proportional to $\delta(\zeta)$ and (C.11) just gives back the reactive linear term of order $\eta\omega^2/\omega_c$ mentioned above. Let us, however, consider the more general case. If $K(\zeta, z)$ has structure on a scale $\sim\omega_0$ with respect to z, then in general we will get a term of order $\eta\omega^2/\omega_0$ which, barring accidents of cancellation, will be nonlinear, (cf. below), and if $\omega_0 \lesssim \omega_c$ will be larger than that obtainable from the "optimum" assumption that the scale of variation of $K(\zeta, z)$ in z is the same as that of its integral over ζ (which is essentially $\eta(z)$), i.e., the microscopic frequency ω_c mentioned in Section 3. Let us therefore make this "optimum" assumption, namely, that $K(\zeta, z)$ is essentially constant, (equal to $K(\zeta)$) with respect to z for $z \lesssim \omega_c$: and let us further consider the case in which ω'' is much less than ω_c (though not zero). Now, the principal part integral over z will in general be of order $\omega_c^{-1}(1 + O(\zeta^2/\omega_c^2) + \cdots)$. The ζ-integral then gives

$$-i\,\Delta\eta_{\omega''} \sim \omega_c^{-1} \int K(\zeta)(\zeta + \omega'')(1 + \text{const}\,\zeta^2/\omega_c^2 + \cdots). \tag{C.45}$$

The terms independent of ω'' vanish because $K(\zeta)$ from its definition is an even function of ζ. The term in ω'' coming from the 1 in the brackets gives a linear reactive term of the same form as that coming from (C.37). The first nonlinear term comes from the ζ^2 term in the bracket: it is proportional to the expression

$$\frac{\omega''}{\omega_c} \int \frac{\zeta^2 K(\zeta)}{\omega_c^2}\, d\zeta. \tag{C.46}$$

Now, the integral of $K(\zeta)$ over ζ is proportional to the friction coefficient η, so that we can write this term as

$$(\Delta\eta_{\omega''})_{\text{n.l.}} = \text{const}\,\eta\, \frac{\omega''}{\omega_c} \frac{\bar{\omega}^2}{\omega_c^2}, \tag{C.47}$$

where $\bar{\omega}$ is the typical frequency represented in $K(\zeta)$. Suppose that the typical velocity of the system is v and the typical range of the interaction potentials $F_j(q)$ is

a. Then, clearly, the quantity $\bar{\omega}$ is of order v/a, and finally, the first nonlinear term due to the dissipative effects, which is reactive in nature, is of order $\eta(\omega/\omega_c)(v^2/v_c^2)$, where $v_c \equiv \omega_c a$ is a typical velocity associated with the dissipative mechanism. Hence for any finite frequency, however small, it will eventually become important if we let the amplitude of motion (hence the velocity) become large enough. This completes the demonstration that the *only* form of interaction compatible with *strictly* linear dissipation (or, more precisely, with strictly linear effects of the dissipative mechanism) is the separable form (3.7).

The above considerations may take on a little more flesh if we consider a specific model for the nonseparable interaction. Let us, for example, consider the following model, which might serve as a crude model for certain types of friction between solid surfaces (through *not* for conventional "sliding friction")

$$F_j(q) = b_j \exp - (q - R_j)^2/a_j^2, \tag{C.48}$$

where the sites R_j are randomly distributed along the q axis and the frequencies ω_j of the oscillators associated with them are uncorrelated to the parameters b_j, a_j, and R_j (which may or may not be correlated with one another). For such a model we can of course evaluate the quantity $K(\zeta, z)$ in a closed form and demonstrate the truth of Eq. (C.47), where $\bar{\omega} \sim v/\bar{a}$ ($\bar{a}$ is a typical value of the a_j's).

Let us summarize the results of this appendix. First, in the context of the problem of interest to us, as defined in the second paragraph, we find that barring possible rather pathological cases (which we have not been able to exclude rigorously but have argued are unlikely to be relevant in real life) the most general Lagrangian we ever need to consider is (3.5). Second, we have shown that in the most general case the parameters of the model are constrained by the relation (C.38) (or more generally (3.10)). The case of linear dissipation $\eta(q) = \eta$ is a special case of this. Third, in the special case of strictly linear friction we have shown that $F_j(q)$ must have the separable form qC_j. It should be emphasized once more that in most practical cases of current interest the considerations of this appendix are superfluous, since one can establish the form of the coupling directly from some microscopic model.

APPENDIX D:
SOME MATHEMATICAL RESULTS NEEDED IN SECTION 5

We first prove the results (5.15)–(5.16) concerning the function

$$F(\theta) \equiv \frac{(1 + (3/2)\sin\theta)^3}{(1 + \sin\theta)^2 \cos\theta} \qquad (0 \leqslant \theta \leqslant \pi/2). \tag{D.1}$$

Consider the quantity

$$G(\theta) \equiv (1 + \sin\theta)^2 \cos\theta [F(\theta) - 1] \tag{D.2}$$

and write it in the form

$$G(\theta) = (1 + \tfrac{3}{2}\sin\theta)^3 - (1 + \sin\theta)^3 + (1 + \sin\theta)^2\,(1 + \sin\theta - \cos\theta)$$
$$= \tfrac{3}{2}\sin\theta\left[1 + \tfrac{5}{2}\sin\theta + \tfrac{19}{12}\sin^2\theta\right] + (1 + \sin\theta)^2\,(1 + \sin\theta - \cos\theta). \quad \text{(D.3)}$$

Now the expression in the square brackets is evidently bounded below by $(1 + \sin\theta)^2$ and above by $\frac{19}{12}(1 + \sin\theta)^2$; similarly, the expression $(1 + \sin\theta - \cos\theta)$ is bounded below and above by $\sin\theta$ and $2\sin\theta$, respectively. Thus we have

$$\tfrac{5}{2}\sin\theta(1 + \sin\theta)^2 \leqslant G(\theta) \leqslant \tfrac{35}{8}\sin\theta(1 + \sin\theta)^2 \quad \text{(D.4)}$$

from which it follows immediately that

$$1 + \tfrac{5}{2}\tan\theta \leqslant F(\theta) \leqslant 1 + \tfrac{35}{8}\tan\theta \quad \text{(D.5)}$$

as stated in the text (Eq. (5.16)). Finally Eq. (5.15) may be obtained directly from the fact that the expression

$$(1 + \tfrac{3}{2}\sin\theta)^3 - \tfrac{125}{32}\sin\theta(1 + \sin\theta)^2 \quad \text{(D.6)}$$

is, as easily demonstrated, positive everywhere in the range $0 < \theta < \pi/2$.

We now turn to the inequality (5.17). We define Fourier transforms by the prescription

$$f(\omega) \equiv (2\pi)^{-1/2}\int_{-\infty}^{\infty} f(t)\exp(-i\omega t)\,dt. \quad \text{(D.7)}$$

Then the quantities (5.4) are given by the expressions

$$\tilde{A} = \int_{-\infty}^{\infty} \omega^2\,|f(\omega)|^2\,d\omega, \quad \text{(D.8a)}$$

$$\tilde{B} = \int_{-\infty}^{\infty} |f(\omega)|^2\,d\omega, \quad \text{(D.8b)}$$

$$\tilde{C} = (2\pi)^{-1/2}\int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' \int_{-\infty}^{\infty} d\omega''\,f(\omega)f(\omega')f(\omega'')\,\delta(\omega + \omega' + \omega''), \quad \text{(D.8c)}$$

$$\tilde{D} = 2\int_{-\infty}^{\infty} |\omega|\,|f(\omega)|^2\,d\omega. \quad \text{(D.8d)}$$

We immediately see (though it is not needed for present purposes) that a simple Schwarz inequality yields $\lambda \equiv \tilde{D}/(\tilde{A}\tilde{B})^{1/2} \leqslant 2$.

We now proceed to demonstrate the inequality (5.17), that is,

$$\kappa[f]\,\lambda[f] \equiv \tilde{D}\tilde{B}^2/\tilde{C}^2 \geqslant 8\pi/9. \tag{D.9}$$

By exploiting the symmetry of the expression (D.8c), we can rewrite it in the form

$$\begin{aligned}\tilde{C} &= \frac{6}{\sqrt{2\pi}} \int_0^\infty d\omega \int_0^\infty d\omega' f(\omega) f(\omega') f(-\omega-\omega') \\ &\leqslant \frac{6}{\sqrt{2\pi}} \int_0^\infty d\omega \int_0^\infty d\omega' |f(\omega)|\,|f(\omega')|\,|f(\omega+\omega')|\end{aligned} \tag{D.10}$$

(where the inequality follows from the fact that $f(\omega) \equiv f^*(-\omega)$). We now apply a simple Schwarz inequality to produce the result

$$\begin{aligned}\tilde{C} &\leqslant \frac{6}{\sqrt{2\pi}} \left\{ \int_0^\infty d\omega \int_0^\infty d\omega' |f(\omega)|^2 |f(\omega')|^2 \right\}^{1/2} \\ &\quad \cdot \left\{ \int_0^\infty d\omega \int_0^\infty d\omega' |f(\omega+\omega')|^2 \right\}^{1/2} \\ &\equiv \frac{6}{\sqrt{2\pi}} \frac{B}{2} \left\{ \int_0^\infty d\omega \int_0^\infty d\omega' |f(\omega+\omega')|^2 \right\}^{1/2}.\end{aligned} \tag{D.11}$$

The double integral J in the braces may be evaluated by rotating the axis in the $\omega - \omega'$ plane through 45°. This gives

$$J = \frac{1}{2} \int_{-x}^{x} dy \int_0^\infty dx\, |f(x)|^2 = \int_0^\infty x\, |f(x)|^2\, dx \equiv \frac{\tilde{D}}{4} \tag{D.12}$$

and hence

$$\tilde{C} \leqslant \frac{3}{\sqrt{8\pi}} \tilde{B}\tilde{D}^{1/2} \tag{D.13}$$

from which the inequality (D.9) follows immediately. We note that for the inequality to become an equality, $f(\omega)$ must be real (implying an $f(t)$ symmetric around $t = 0$) and must satisfy the condition

$$f(\omega+\omega') = f(\omega) f(\omega') \qquad (\omega, \omega' \geqslant 0). \tag{D.14}$$

This immediately implies that $f(\omega) = \exp - \text{const}\,|\omega|$, which in turn implies that the form of $f(t)$ is $(1+t^2)^{-1}$. That this function does indeed yield an extremum of the quantity $\kappa\lambda$ can be verified by noting that when suitably scaled (cf. Eq. (5.28)) it solves the Euler–Lagrange equation corresponding to (5.2) in the limit $\alpha \to \infty$. (In view of Eqs. (5.9)–(5.11) and (5.16), this is evidently equivalent to minimizing $\kappa\lambda$).

We tabulate for reference (Table I) the values of the scale-invariant quantities κ, λ,

TABLE I

	Function f	$\kappa\|f\|$	$\lambda\|f\|$	$\kappa\lambda$
1	$\operatorname{sech}^2 t$	$5\sqrt{5}/6$	$18\sqrt{5}\,\zeta(3)/\pi^3$	$75\zeta(3)/\pi^3$
2	$(1+t^2)^{-1}$	$8\pi/9\sqrt{2}$	$\sqrt{2}$	$8\pi/9$
3	$e^{-\|t\|}$	$9/4$	$4/\pi$	$9/\pi$

and $\kappa\lambda$ for three functions which give, respectively, the minima of the functional $\sigma_{00}[f]$ (Eq. (5.10)) for very small and very large α and a reasonably good value for $\alpha \sim 1$.

Finally, let us calculate the value of $D_0(\alpha)$ given by Eq. (5.39) in the extremely overdamped regime ($\alpha \to \infty$). In order to do this we need to know all the bound states of the operator $\hat{H}_0 + \hat{V}$ (see (5.32)–(5.34)) with $\tilde{q}_{cl}(t) = \frac{4}{3}(1+t^2)$. In other words, we need to solve the following eigenvalue problem in the ω-representation

$$(1+|\omega|)\,\psi_n(\omega) - 2\int_{-\infty}^{\infty} e^{-|\omega-\omega'|}\psi_n(\omega')\,d\omega' = \lambda\psi_n(\omega), \tag{D.15}$$

where we have dropped terms of order α^{-1}.

We list below the normalized eigenfunctions and eigenvalues we have found which can be shown to satisfy (D.15) by direct substitution.

Eigenfunction	Eigenvalue
$\psi_0(\omega) = \left(\frac{2}{5+2\sqrt{5}}\right)^{1/2}\left[\left(\frac{1+\sqrt{5}}{2}\right)+|\omega|\right]e^{-|\omega|}$,	$\lambda_0 = \left(\frac{1+\sqrt{5}}{2}\right)$,
$\psi_1(\omega) = \sqrt{2}\,\omega e^{-|\omega|}$,	$\lambda_1 = 0$,
$\psi_2(\omega) = \left(\frac{2}{5-2\sqrt{5}}\right)^{1/2}\left[\left(\frac{1-\sqrt{5}}{2}\right)+|\omega|\right]e^{-|\omega|}$,	$\lambda_2 = \frac{\sqrt{5}-1}{2}$.

As one can see, we could find only one bound state apart from $\psi_0(\omega)$ and $\psi_1(\omega)$. This has eigenvalue lying between 0 and 1 and, consequently, must contribute to C_0 defined by (5.37). Whether $\psi_2(\omega)$ is the only eigenfunction with $0 < \lambda < 1$ is still an open question. Although we have not proved this result we shall take it for granted for the present calculation. Therefore, we can write $D_0(\alpha)$ as

$$D_0(\alpha) = \sum_{n=0}^{2} \langle \ln \hat{H}_0 \rangle_n - \ln|\lambda_0| - \ln\lambda_2 - 1 + \langle \hat{H}_0^{-1}\hat{V}\rangle_0 + \langle \hat{H}_0^{-1}\hat{V}\rangle_2. \tag{D.16}$$

The matrix elements of the operators $\ln \hat{H}_0$ and $\hat{H}_0^{-1}\hat{V}$ can be easily evaluated in the ω-representation as

$$\langle \ln \hat{H}_0 \rangle_n = \int_{-\infty}^{\infty} d\omega \ln(1+|\omega|)\,\psi_n^2(\omega) \tag{D.17}$$

and

$$\langle \hat{H}_0^{-1}\hat{V}\rangle_n = -2\int_{-\infty}^{\infty} d\omega \frac{\psi_n(\omega)}{1+\omega}\int_{-\infty}^{\infty} e^{-|\omega-\omega'|}\psi_n(\omega')\,d\omega'. \tag{D.18}$$

Then, inserting (D.17), (D.18), and the tabulated values of $\psi_n(\omega)$ and λ_n in (D.16) we get

$$D_0 = -1 + 2\int_0^{\infty} d\omega \ln(1+\omega)\, e^{-2\omega}[2 - 4\omega + 6\omega^2]$$

$$-4\int_0^{\infty} d\omega(1+\omega)^{-1} e^{-2\omega}[1+\omega-\omega^2+2\omega^3]. \tag{D.19}$$

Now, defining

$$L_n(\lambda) \equiv \int_0^{\infty} d\omega \ln(1+\omega)\, e^{-\lambda\omega}\omega^n = (-1)^n \frac{d^n}{d\lambda^n} L_0(\lambda),$$

$$\Delta_n(\lambda) \equiv \int_0^{\infty} d\omega \frac{e^{-\lambda\omega}}{1+\omega}\omega^n = (-1)^n \frac{d^n}{d\lambda^n}\Delta_0(\lambda),$$

and noticing that $L_0(\lambda) = \lambda^{-1}\Delta_0(\lambda)$ (which follows from an integration by parts of $L_0(\lambda)$) we can write D_0 as

$$D_0 = -\tfrac{17}{2} + 19e^2 E_1(2),$$

where $E_1(x)$ is defined as [66]

$$E_1(x) \equiv \int_x^{\infty} \frac{e^{-t}}{t}\,dt \qquad (x > 0).$$

This function is tabulated in many mathematical tables (see, for example, [66]) and the final result for D_0 reads

$$D_0 \approx -1.5.$$

References

1. E. Schrödinger, Die Naturwissenschaften **23**, (1935), 807.
2. A. J. Leggett, *Prog. Theor. Phys. supplement* no. 69 (1980), 80.
3. A. J. Leggett, *J. Phys. Colloq.* **39** (1978), C6–1264.
4. O. V. Lounasmaa, "Experimental Principles and Methods below 1*K*," Chap. 7, Academic Press, New York, 1974.
5. D. J. Scalapino, *in* "Tunnelling Phenomena in Solids" (E. Burstein and S. Lundqvist, Eds.), p. 477, Plenum, New York, 1969.
6. J. Kurkijärvi, *Phys. Rev. B* **6** (1972), 832.

7. A. O. CALDEIRA, Ph. D. thesis, University of Sussex, 1980, unpublished.
8. W. DEN BOER AND R. DE BRUYN OUBOTER, *Physica B + C* **98**, (1980), 185.
9. YU. M. IVANCHENKO AND L. A. ZIL'BERMAN, *Zh. Eksp. Teor. Fiz.* **55** (1968), 2395; translation, *Soviet Phys. JETP* **28**, (1969), 1272.
10. M. M. NIETO, *Phys. Rev.* **167** (1968), 416.
11. A. J. LEGGETT, *Prog. Theor. Phys.* **36** (1966), 901.
12. R. DE BRUYN OUBOTER AND D. BOL, *Physica* **112B** (1982), 15.
13. R. J. PRANCE *et al.*, *Nature* **289** (1981), 542.
14. T. A. FULTON AND L. N. DUNKELBERGER, *Phys. Rev. B* **9** (1974), 1760.
15. T. A. FULTON AND L. N. DUNKELBERGER, *Bull. Amer. Phys. Soc.* **19** (1975), 205.
16. R. F. VOSS AND R. A. WEBB, *Phys. Rev. Lett.* **47** (1981), 265.
17. L. D. JACKEL *et al.*, *Phys. Rev. Lett.* **47** (1981), 697.
18. A. J. LEGGETT, *in* "Proceedings, 6th International Conference on Noise in Physical Systems," Gaithersburg, Md. 1981, NBS Special Publication No. 614, p. 355.
19. J. BARDEEN, *Phys. Rev. Lett.* **42** (1979), 1498.
20. J. BARDEEN, *Phys. Rev. Lett.* **45** (1980), 1978.
21. G. GRUNER, A. ZETTL, W. G. CLARK, AND J. BARDEEN, *Phys. Rev. B* **24** (1981), 7247.
22. K. MAKI, *Phys. Rev. Lett.* **39** (1977), 46.
23. A. M. GOLDMAN, private communication.
24. P. LETT, W. CHRISTIAN, SURENDRA SINGH, AND L. MANDEL, *Phys. Rev. Lett.* **47** (1981), 1892.
25. I. M. LIFSHITZ AND YU. KAGAN, *Zh. Eksp. Teor. Fiz.* **62** (1972), 385; translation, *Soviet Phys. JETP* **35** (1972), 206.
26. S. V. JORDANSKII AND A. M. FINKEL'SHTEIN, *Zh. Eksp. Teor. Fiz.* **62** (1972), 403; translation, *Soviet Phys. JETP* **35** (1972), 215.
27. D. BAILIN AND A. LOVE, *J. Phys. A* **13** (1980), L271.
28. A. O. CALDEIRA AND A. J. LEGGETT, *Phys. Rev. Lett.* **46** (1981), 211.
29. J. P. SETHNA, *Phys. Rev. B* **24** (1981), 698.
30. YU. KAGAN AND L. A. MAKSIMOV, *Zh. Eksp. Teor. Fiz.* **65** (1973), 622; translation: *Soviet Phys. JETP* **38** (1974), 307.
31. M. SIMONIUS, *Phys. Rev. Lett.* **40** (1978), 980.
32. R. A. HARRIS AND L. STODOLSKY, *J. Chem. Phys.* **74** (1981), 2145.
33. A. WIDOM, G. MEGALOUDIS, J. E. SACCO, AND T. D. CLARK, Nuovo Cimento 61*B* 112 (1981).
34. A. WIDOM AND T. D. CLARK, *Phys. Rev. Lett.* **48** (1982), 63.
35. A. WIDOM AND T. D. CLARK, *Phys. Rev. Lett.* **48** (1982), 1572.
36. J. KURKIJÄRVI, *J. Low. Temp. Phys.* **45** (1981), 37.
37. A. O. CALDEIRA AND A. J. LEGGETT, *Phys. Rev. Lett.* **48** (1982), 1571.
38. A. O. CALDEIRA AND A. J. LEGGETT, *Physica B* and *C*, in press.
39. E. KANAI, *Prog. Theor. Phys.* **3** (1948), 440.
40. M. D. KOSTIN, *J. Chem. Phys.* **57** (1972), 3589.
41. K. YASUE, *Ann. Phys.* **114** (1978), 479.
42. H. DEKKER, *Phys. Rev. A* **16** (1977), 2116.
43. O. PENROSE, *Reps. Prog. Phys.* **42** (1979), 1937.
44. R. BENGURIA AND M. KAC, *Phys. Rev. Lett.* **46** (1981), 1.
45. R. H. KOCH, D. J. VAN HARLINGEN, AND J. CLARKE, *Phys. Rev. Lett.* **45** (1980), 2132.
46. I. R. SENITZKY, *Phys. Rev.* **119** (1960), 670.
47. K. MÖHRING AND U. SMILANSKY, *Nucl. Phys. A* **338** (1980), 227.
48. V. AMBEGAOKAR, U. ECKERN, AND G. SCHÖN, *Phys. Rev. Lett.* **48** (1982), 1795.
49. K. L. SEBASTIAN, *Chem. Phys. Lett. (Netherlands)* **81** (1981), 14.
50. T. BANKS, C. M. BENDER, AND T. T. WU, *Phys. Rev. D* **8** (1973), 3346; 3366.
51. J. S. LANGER, *Ann. Phys. (NY)* **41** (1967), 108.
52. M. STONE, *Phys. Lett. B* **67** (1977), 186.
53. C. G. CALLAN AND S. COLEMAN, *Phys. Rev. D* **16** (1977), 1762.

54. S. COLEMAN, "The Uses of Instantons," International School of Subnuclear Physics "Ettore Majorana" 1977.
55. R. P. FEYNMAN AND A. HIBBS, "Quantum Mechanics and Path Integrals," McGraw–Hill, New York, 1965.
56. R. P. FEYNMAN, "Statistical Mechanics," p. 82, Benjamin, New York, 1972.
57. D. FORSTER, "Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions," Benjamin, New York, 1975.
58. J. P. SETHNA, *Phys. Rev. B* **25** (1982), 5050.
59. J. KURKIJÄRVI, *Phys. Lett.* **88A** (1982), 241.
60. K. K. LIKHAREV, *in* "Proceedings, 16th International Conference on Low Temperature Physics," p. 1079, Los Angeles, Calif., 1981.
61. S. CHAKRAVARTY, *Phys. Rev. Lett.* **49** (1982), 681.
62. A. J. BRAY AND M. A. MOORE, *Phys. Rev. Lett.* **49** (1982), 1545.
63. H. GOLDSTEIN, "Classical Mechanics," Addison–Wesley, Reading Mass., 1950.
64. D. PINES AND P. NOZIERES, "The Theory of Quantum Liquids," Benjamin, New York, 1966.
65. A. WIDOM AND T. D. CLARK, preprint.
66. M. ABRAMOWITZ AND I. A. STEGUN, "Handbook of Mathematical Functions," National Bureau of Standards, Washington D.C., 1964.