8.1 Introduction

Quantum mechanical tunnelling is one of the earliest studied phenomena that is qualitatively different in quantum mechanics and classical mechanics. It was first proposed for electrons by Hund in his studies of electrons or atoms penetrating a potential barrier,\textsuperscript{1-3} by several others in molecular and solid-state physics,\textsuperscript{3} and most famously by Gamow in the context of the escape of $\alpha$ particles from the nucleus,\textsuperscript{3-6} but it was slow to be widely appreciated in chemical kinetics. In a broad review of general and physical chemistry in 1932, it was stated\textsuperscript{7} that “Chemical differences between hydrogen and diplogen [as the authors called deuterium] are to be expected on theoretical grounds for several reasons, but the dominating factor is likely to be the smaller half-quantum of zero-point energy possessed by diplogen in virtue of its greater mass.” Notably there is no mention of tunnelling in this generalization. A similar review in 1934 has large sections on deuterium and kinetics, but tunnelling is again not mentioned.\textsuperscript{8} The role of tunnelling in chemical reactions seems to have first been speculated by Applebey and Ogden in 1936\textsuperscript{9} and discussed more fully by Bell at a Faraday Discussion in 1938,\textsuperscript{10} who advanced it as a possible cause of nonlinear
Arrhenius plots. His discussion, however, takes no cognizance of transition state theory, and he discusses tunnelling as a one-dimensional phenomenon governed by the raw potential energy surface, an oversimplification that continued to be made by many workers for many decades into the future. Bell stated, “So far no evidence has been produced that directly demonstrates the presence of the ‘tunnel effect’.” At the same Faraday Discussion, Wigner explained how tunnelling could lead to a leveling out of Arrhenius plots at low temperature, leading to a smaller pre-exponential factor, but concluded that “apart from reactions involving H, the tunneling effect cannot be made responsible at ordinary temperatures for any large decrease of the temperature independent factor.”

Nevertheless, there was valuable theoretical work on tunnelling in the 1930s. Already, in 1932, Wigner showed how classical transition state theory, with the transition state at the saddle point, could be quantized to the lowest order, $O(\hbar^2)$, in $\hbar$, including the leading term due to tunnelling. And Bell, in 1933, derived the tunnelling probability for a one-dimensional Eckart barrier and concluded that calculations in the literature on ortho-para conversion in hydrogen were “probably in error” due to neglect of tunnelling. Bell said “The barrier is treated as one-dimensional. This is permissible, since there is in general a definite direction of approach of two reacting molecules for which the activation energy is a minimum,” which is generally not a good approximation, as discussed below. He also said “It is concluded that a quantum-mechanical treatment is necessary for any reaction involving the motion of a hydrogen atom or proton, while heavier atoms may be considered to behave classically,” which is very true as a general rule, noting that general rules usually have some exceptions (and there indeed some “exceptional” cases of non-hydrogenic tunnelling), although it took decades before the community routinely took this into account.

Moving ahead in time, I note a 1970 review on proton transfer reactions, where Caldin said “The question is not whether tunnelling occurs, but whether it is detectable.” In the same article Caldin said, “The first definite indication that the tunnel effect was significant in controlling the rate of a reaction in solution was published in 1956” by Bell et al. in their study of the deuterium kinetic isotope effect in the base-catalyzed bromination of 2-carbethoxy cyclopentanone. Caldin also concluded that “The tunnel effect is by now the likeliest interpretation of the evidence on several reactions,” with which I agree. However, he added “It remains true that such reactions appear to be exceptional and that in most proton-transfer reactions the tunnelling corrections can probably be ignored,” with which I disagree. In my opinion, theoretical work has now established that it is most likely that any hydrogen-atom transfer or proton transfer with a barrier of about 5 kcal mol$^{-1}$ or higher is probably dominated by tunnelling at room temperature and below (and often too much higher temperatures), and the burden of proof in interpreting experiments on such reactions should shift from “Is there any evidence of tunnelling?” to asking if there is any reason to doubt that the reaction is dominated by tunnelling, since that is the usual finding when such reactions...
are examined with modern theoretical analysis. The goal of this chapter is to review the methods used to calculate tunnelling probabilities realistically.

8.2  Theory

In 1970, LeRoy et al.\textsuperscript{15} showed that the tunnelling probability for realistic shapes of potential energy barriers could differ significantly from tunnelling by an Eckart barrier or the even simpler parabolic approximation. The inadequacy of an Eckart barrier is also discussed in other papers.\textsuperscript{16,17} Nevertheless, one still sees papers in the current literature using the approximation of an Eckart barrier. However, in order to obtain a more accurate tunnelling probability, it is necessary not just to use a realistic shape of the potential energy barrier, but also to go beyond the one-dimensional tunnelling model to obtain an effective potential or an ensemble of effective potentials.

Johnston and Rapp\textsuperscript{18} were the first to try to include multidimensional effects in tunnelling, and Johnston described their method as resulting from “several bold assumptions about the appropriate tunnelling path or paths.”\textsuperscript{19} This is very stimulating work, but the method itself is not quantitative and is only of historical interest.

Attempts to treat quantum mechanical tunnelling more accurately were held back by two factors: (i) the lack of appreciation of how to include quantum effects in transition state theory so that one could test predictions against experimental rate constants and kinetic isotope effects (or accurate theoretical results, when they became available) and (ii) insufficient knowledge of potential energy surfaces. The first impediment involved understanding what we now call quasiclassical transition state theory. Quasiclassical transition state theory refers to calculating all partition functions quantum mechanically but treating the reaction coordinate (which does not appear in the transition state partition function because it is the degree of freedom missing in the transition state) classically; quantizing the partition functions of bound modes was done already by Wigner in 1932\textsuperscript{12} and was assumed in Eyring’s classic paper of 1935.\textsuperscript{20} A convenient name for the set of modes excluding the reaction coordinate is the “modes transverse to the reaction coordinate”, and we can say that Eyring quantized the transverse modes.

The above-mentioned papers dealt with canonical transition state theory, \textit{i.e.}, the evaluation of rate constants for a canonical ensemble characterized by a temperature. To treat tunnelling rigorously one needs at least a microcanonical theory where the dynamics are calculated for each total energy (which is a parameter in the time-independent Schrödinger equation) and then thermally averaged. More specifically one must ask the question of what reaction probabilities are implied by quasiclassical transition state theory, and then one could ask how to modify these reaction probabilities by including energy-dependent tunnelling probabilities. It was shown, first partially\textsuperscript{21,22} and then more completely,\textsuperscript{23} that transition state theory corresponds to an adiabatic treatment of all motions except the reaction coordinate; adiabaticity in this context means that their quantum numbers are
conserved as the system proceeds along the reaction path (where the reaction coordinate measures progress along the reaction path). In fact, it was later shown that a calculation assuming such adiabaticity leads to microcanonical variational transition state theory, in which the transition state dividing surface is chosen to minimize the flux through the transition state for each total energy, rather than conventional transition state theory, where the transition state dividing surface is located at the saddle point.²⁴,²⁵

It is conventional to add tunnelling to transition state theory by multiplying by a transmission coefficient that makes up for the simplifications inherent in quasiclassical transition state theory. In general there can be several contributions to the transmission coefficient, including the correction for the breakdown of the equilibrium assumption of transition state theory, correction for the breakdown of the no-recrossing approximation of transition state theory (i.e., that the net flux through the transition state dividing surface is equal to the one-way flux, which in classical mechanics would mean that trajectories do not recross the transition state dividing surface), and correction for the classical treatment of the reaction coordinate. One can consider the overall transmission coefficient to be a product of three correction factors corresponding to these effects.²⁶ The third factor is called the tunnelling transmission coefficient, and it is the one of primary interest here. The tunnelling transmission coefficient may be defined as the ratio of the reaction rate calculated when reaction-coordinate motion is treated quantum mechanically to that calculated when it is treated classically. The considerations in the previous paragraph imply that the tunnelling correction to transition state theory should involve barrier heights determined by adding zero point energy in modes transverse to the reaction coordinate to the potential energy surface along reaction paths. The simplest general version of such a theory is as follows. One starts by calculating a steepest descents path down from the saddle (both toward reactants and toward products) in isoinertial coordinates, which are coordinates scaled so that the reduced mass is the same in all directions.¹⁶,²⁵ We call such a path the minimum-energy path²⁵ (MEP); it is also often called the intrinsic reaction coordinate²⁷ (IRC, although a more semantically correct name would be intrinsic reaction path). One usually then makes (often only implicitly) the ground-state tunnelling approximation, which is simply that the tunnelling transmission coefficient is calculated for ground-state reactions but used for all reactions, even those that do not emanate from the ground state of the reactants. For reaction from the ground state, one adds the local zero point energy in transverse modes all along the path, which replaces the potential energy barrier by the vibrationally adiabatic ground-state barrier. [The potential energy along the reaction path is the Born–Oppenheimer potential along that path and is called $V_{\text{MEP}}(s)$; the highest point of $V_{\text{MEP}}(s)$ occurs at the saddle point, and subtracting the value at reactants yields what is called classical barrier height. The barrier maximum on the ground-state vibrationally adiabatic potential curve, which is called $V_0^G(s)$, is not necessarily at the saddle point, and the barrier height, again obtained by
subtracting the value at reactants, may be greater or smaller than the classical barrier height. Given the vibrationally adiabatic barrier, one then calculates the tunnelling as if the reaction path were rectilinear, which is an approximation since the reaction path is curved and the kinetic energy operator is different in curvilinear coordinates. This simple scheme, first reported in 1971, has been called by several names, such as the vibrationally adiabatic zero-curvature approximation, the vibrationally adiabatic ground-state approximation, and the minimum-energy-path adiabatic approximation; I here call it the zero-curvature tunnelling (ZCT) approximation, which is a name for it that we have used in many papers since 1991.

Note that although the ZCT approximation involves a one-dimensional tunnelling calculation, it is actually a multidimensional tunnelling approximation because the vibrationally adiabatic potential used along the reaction path involves contributions from the transverse modes. Furthermore, these contributions vary with the distance \( s \) along the reaction path. If the reaction coordinate were separable, as assumed in quasiclassical transition state theory, these contributions would, by the definition of separability, be independent of \( s \). Hence the tunnelling transmission coefficient corrects not just for quantum effects but for non-separability. (Whereas the tunnelling transmission coefficient corrects for quantal non-separability, one may consider that the recrossing transmission coefficient, which is the second factor mentioned above, corrects for classical non-separability since trajectories would not recross the dividing surface if the reaction coordinate were separable.)

The second impediment to progress in tunnelling theory for chemical kinetics began to disappear when accurate calculations of the \( \text{H}^+\text{H}_2 \) potential energy surface finally appeared. Early work on the potential energy surface for this reaction was clouded by the question of whether there was a single symmetric barrier or two asymmetric barriers on either side of a well sometimes called “Lake Eyring”; predictions of asymmetric saddle points persisted as late as 1967, as reviewed elsewhere. Nevertheless, the Porter–Karplus surfaces of 1964, with a collinear symmetric saddle point, correctly became accepted as qualitatively correct. Another important advance critical to the development of accurate methods for quantum mechanical tunnelling was the calculation of accurate quantum mechanical rate constants for the collinear \( \text{H}^+\text{H}_2 \) reaction, which is a two-dimensional problem. These rate constants now allowed the testing of approximate tunnelling methods against accurate quantum mechanics in more than one dimension. One immediately found that ZCT tunnelling is inaccurate, leading to rate constants underestimated by factors of 19 and 3.4 at 200 and 300 K, respectively. This was a serious impasse for theory, and it was argued that it was due either to the breakdown of the adiabatic approximation or the assumption of zero curvature. (We now know that the latter is the chief culprit.)

The impasse began to be resolved in 1977 when Marcus and Coltrin proposed a new tunnelling path for collinear \( \text{H}^+\text{H}_2 \). This path corresponded to a path along the vibrational turning points on the concave side of the reaction path, and they justified this path based on a semiclassical
argument (the Wentzel–Kramers–Brillouin (WKB) approximation) in which the decay of the probability density $|\psi|^2$ in a classically forbidden region is $e^{-\theta}$ where $\theta$ is an action integral given by

$$\theta = \int_{\text{tunnelling region}} d\xi |p_{\xi,\text{eff}}|$$

(8.1)

$$\theta = \int_{\text{tunnelling region}} d\xi \sqrt{2\mu (E - V_{\text{eff}}(\xi))}$$

where $\xi$ measures distance (in the isoinertial coordinate system) along the tunnelling path, $p_{\xi,\text{eff}}$ is the $\xi$-component of the effective momentum, $\mu$ is the reduced mass to which the isoinertial coordinates are scaled, $E$ is the energy, and $V_{\text{eff}}(\xi)$ is the effective potential. Since $V_{\text{eff}}(\xi) > E$ in a tunnelling region, $p_{\xi,\text{eff}}$ is imaginary, and this is an imaginary-action integral. The semiclassical argument is basically that any path within the zero-point vibrational amplitude (transverse to the MEP) has the same effective potential for tunnelling, namely the ground-state vibrationally adiabatic potential curve, also called $V^G_a(s)$. Therefore, in isoinertial coordinates, the path with the least exponential decay, i.e., the dominant path, is the shortest path, i.e. (for smooth MEPs with smoothly changing local zero-point energy), the path is along the concave-side vibrational turning points. Later work showed that it is best to calculate the tunnelling along such a dominant path rather than averaging over many paths within the width of the vibrational wave packet.\(^{32}\) The Marcus–Coltrin path was generalized to treat the $H + H_2$ reaction in the real three-dimensional world,\(^{33}\) and again good results were obtained.

Because this approach involves a path on the concave side of the reaction path, it may be called a corner-cutting tunnelling approximation. Path curvature in classical mechanics leads to a system veering to the convex side of the path, as occurs at the first turn of a bobsled run. The direction of the effect gets turned around in quantum tunnelling, leading to motion on the concave side of the MEP. This is sometimes called the quantum bobsled effect.

Two technical issues needed to be resolved to make the method practical and accurate for general polyatomic reactions. First, the distance to the vibrational turning point sometimes exceeds the distance at which curvilinear coordinates attached to the reaction path become multivalued. This was solved\(^{34}\) by reformulating the theory in terms of an effectively reduced mass so that the integration is carried out along the MEP instead of along the actual dominant tunnelling path:

$$\theta = \int_{\text{tunnelling region}} ds |p_{s,\text{eff}}|$$

(8.2)

$$\theta = \int_{\text{tunnelling region}} ds \sqrt{2\mu_{\text{eff}}(s)(E - V^G_a(s))}$$

where $\mu_{\text{eff}}$ is smaller than $\mu$ in regions where the reaction path is curved. Because the tunnelling path can lie in a region where the coordinate system
based on the MEP is multivalued, the actual tunnelling path is not well defined in this method, but $\mu_{\text{eff}}$ is well defined and is formulated such that eqn (8.1) and (8.2) give the same result in the limit where the MEP has small curvature; in such a limit the concave side vibrational turning point is in the region describable by the curvilinear coordinates based on $s$. However, $\mu_{\text{eff}}$ is also defined such that it remains well defined and smooth in all cases and is generally applicable.\textsuperscript{34}

The three-dimensional $\text{H} + \text{H}_2$ reaction has three transverse vibrational modes, but the curvature (which is a vector) has only a single nonzero component because the reaction path is collinear and so it never curves into the bending coordinate directions. The reaction path of a general polyatomic reaction, though, has multidimensional curvature. The second technical issue to be resolved in turning the Marcus–Coltrin idea into a general scheme concerned the physically correct calculation of the amount of corner cutting when there is multidimensional reaction path curvature. When this was done properly,\textsuperscript{35,36} the method was called the centrifugal-dominant small-curvature semiclassical adiabatic ground-state approximation, but the shorter name of small-curvature tunnelling (SCT) approximation soon replaced the long one.

Marcus and Coltrin had pointed out that their physical model would break down for the common mass combination where a hydrogen or proton tunnels between two massive molecular subsystems,\textsuperscript{31} and later Babamov and Marcus\textsuperscript{37,38} (following earlier qualitative considerations by Marcus\textsuperscript{39}) presented another approximation designed to treat the transfer of a hydrogen atom or proton between two heavier particles in a collinear reaction using polar coordinates. In order to understand why a different treatment is needed for this case, it is useful to consider what the potential energy surface and MEP for a bimolecular atom transfer reaction look like when plotted in isoinertial coordinates. Figure 8.1 illustrates the key points. On the left we see that the reaction valley and the product valley are perpendicular when plotted vs. the bond distances of the forming and breaking bonds. However, this is not the case in isoinertial coordinates. The angle between the reactant valley and the product valley in isoinertial coordinates is called the skew angle $\beta$, and for a reaction of the form $\text{A} + \text{BC} \rightarrow \text{AB} + \text{C}$, where $\text{A}$, $\text{B}$, and $\text{C}$ are atoms, molecules, or molecular fragments ($\text{B}$ is a hydrogen atom or a proton for the cases of most interest here) it is given by

$$\beta = \arccos \left( \frac{m_X m_C}{m_{\text{AB}} m_{\text{BC}}} \right)$$

(8.3)

where $m_X$ is the mass of $X$. We see that $\beta = 60^\circ$ for the case of equal masses (center panel of Figure 8.1), but it gets very small ($\cos \beta \approx 1$) when $m_B \ll m_A$ and $m_B \gg m_C$ (rightmost panel of Figure 8.1). Figure 8.1 shows that a large skew angle corresponds to low curvature of the MEP, and a small skew angle corresponds to large curvature of the MEP. With this in mind, it is also useful to review the motivation for using isoinertial coordinates. If we did not use isoinertial coordinates, the reduced mass in eqn (8.1) would be a function of the location on the curvilinear tunnelling path. To find the optimized tunnelling path, one would have to find the best compromise of a
path with small reduced mass (to minimize the radicand), a path with a low barrier (also to minimize the radicand), and a short path (to minimize the length of the interval over which the integration is carried out). When we use isoinertial coordinates, all paths have the same reduced mass and we just need a compromise of low barrier and short path. In the small-curvature limit, any path shortening due to cutting the corner more than the outer vibrational turning point of the vibration transverse to the reaction path would involve an increase in effective potential that would outweigh the advantage of a shorter path, and in this way we arrive at the SCT tunnelling approximation. When the skew angle is small, one can obtain a much shorter path by severe corner cutting because the product valley is close to the reactant valley. This is illustrated by the nearly vertical path in the right panel of Figure 8.1. In the limit of large reaction-path curvature, the distance criterion totally dominates, and the optimum tunnelling path should be a straight line from the reactant valley to the product one.

In our own work, we developed a method to treat the large-curvature limit for general polyatomic reactions by using straight-line tunnelling paths as motivated above. Early versions were called the large-curvature ground-state method, and the theory achieved its final form in version 4. The version 4 approximation (LCG4) is called the large-curvature tunnelling (LCT) approximation. It differs from SCT in three key ways: (1) The tunnelling path is not restricted to the region inside the ground-state vibrational amplitude where one may use the vibrationally adiabatic potential curve, and large-curvature approximations use a diabatic effective potential in the region beyond the outer turning points of the transverse vibrational motion. (2) A given tunnelling path need not conserve the transverse vibrational quantum numbers, but rather a system beginning in the ground vibrational state of the reactants may tunnel into excited vibrational states of the product (when the reaction is considered in the exothermic or thermoneutral direction).

Figure 8.1 Three views of the potential energy contours and minimum energy path (MEP) for A + BC → AB + C reactions. The saddle point (SP) is also labeled. Left: plotted vs. the lengths of the forming and breaking bonds. Center: plotted vs. two mass-scaled coordinates (x and y) for the case where A, B, and C have the same mass. Right: plotted vs. two mass-scaled coordinates (x and y) for the case where the mass of B is much less than the masses of A and B. The right-hand plot is for an exothermic reaction, and it also shows two possible straight-line paths along which tunnelling might occur.
(3) The tunnelling, even for a given initial and final vibrational state, is not dominated by a single tunnelling path for each energy but rather there is an average over a sequence of tunnelling paths, each starting somewhere on the MEP in the reactant valley and proceeding along a straight line to a point (determined by a quantization criterion) on the MEP in the product valley. Plots showing examples of large-curvature tunnelling paths for tunnelling from the ground state of reactants into either the ground state or excited vibrational states of the product are available in previous papers.43,44,47

Two points about curvature of the reaction path can be emphasized. First of all, it is not the skew angle alone that determines the effect of reaction-path curvature on tunnelling because the skew angle is a global characteristic of the path. The key issue is how big is the curvature in the region of the barrier, and an exothermic reaction may have the barrier very early where the reaction path is not yet very curved. In such a case, the effect of curvature on the tunnelling may be small or negligible. For bimolecular hydrogen-transfer or proton-transfer reactions, the region of largest curvature tends to be where the forming bond is about half made and the breaking bond is about half broken; this is also the region where thermoneutral and nearly thermoneutral reactions have their barriers and also where intrinsic barriers tend to be largest (the intrinsic barrier for a symmetric reaction is the same as the barrier; the intrinsic barrier for a non-thermoneutral reaction is the barrier in the exoergic direction) and hence where tunnelling effects can be most significant. The second point to be made is that it is not as easy to guess the curvature for unimolecular reactions as it is for concerted bimolecular ones,48 because the skew angle in a bimolecular reaction is calculated from the cosine of the angle between the relative translation of the reactants and the relative translation of the products, but the directions of entrance and exit for a unimolecular reaction are very case specific.

Because the LCT approximation requires more extensive calculations than the SCT approximation, it was originally much more expensive, but now efficient interpolation algorithms are available to make it very affordable.49,50

While the SCT and LCT formalisms represent two limiting cases, both of them seem to be accurate for intermediate curvature, so they cover all possible cases. A simple procedure to cover all cases is simply to perform both kinds of calculation and choose whichever gives a larger tunnelling probability.51 When this is done for each value of the energy the result is called microcanonical optimized multidimensional tunnelling (μOMT).52

In classical mechanics, Hamilton’s principle states that solving the equations of motion are equivalent to finding the least-action path.53 Eqn (8.1) shows that the dominant tunnelling path is the one with the least imaginary action, and this may be considered an analytic continuation of Hamilton’s principle to complex momenta (the classical momentum is imaginary or complex when the classical kinetic energy is negative, as it would be in a tunnelling region where the potential energy is greater than the total energy). It is natural to ask if this can be used to obtain a more general approximation to the tunnelling probability than the small-curvature and large-curvature limiting approximations discussed above. In a sense, this is done by classical S matrix theory, where one
calculates a trajectory in imaginary time with complex coordinates and momenta,\textsuperscript{54} but this is impractical for real systems because it is hard to make the trajectory satisfy the correct real boundary conditions. However, one can find a variational approximation to the tunnelling trajectory by finding the least-imaginary-action path from among a sequence of paths that satisfy the correct boundary conditions. One major problem that needs to be solved in such an approach is determining the individual components of the complex momentum since the action integral involves $p \cdot dq$, where $q$ is a vector of coordinates, and therefore we need the components of $p$ in order to calculate the exponential decay. This was accomplished by using the same kind of diabatic model as used in the LCT method, and this results in a least-action ground-state tunnelling approximation that is also called the least-action tunnelling (LAT) approximation.\textsuperscript{55,56} Usually we expect that the LAT approximation is not needed, and it is sufficient to use the $\mu$OMT approximation, but in a set of cases where it was applied to polyatomic reactions,\textsuperscript{57} namely the reaction of CF$_3$ with CH$_4$, CD$_3$H, and CD$_4$, it did give kinetic isotope effects in consistently better agreement with experiment. Least-action tunnelling approximations have also been proposed by Taketsugu and Hirao\textsuperscript{58} and Tautermann \textit{et al.}\textsuperscript{59}

It has not been emphasized so far, but the tunnelling transmission coefficient also includes quantum mechanical effects on reaction-coordinate motion when the energy exceeds the effective barrier height. Just as a quantum mechanical system shows non-classical transmission across a barrier at energies below the barrier top, it also shows non-classical reflection for energies above the barrier top (this may be thought of as a diffraction effect). The situation is particularly clear for a parabolic barrier (i.e., a purely quadratic barrier), where the quantum results can be obtained analytically. If $P(E)$ is the transmission probability at energy $E$, and $V_{\text{max}}$ is the maximum potential energy of the parabolic barrier, one finds\textsuperscript{57}

$$P(E) = \frac{1}{1 + e^{2\theta}}, E < V_{\text{max}}$$ \hspace{1cm} (8.4)

$$P(V_{\text{max}}) = 0.5,$$ \hspace{1cm} (8.5)

$$P(V_{\text{max}} + \Delta) = 1 - P(V_{\text{max}} - \Delta), \Delta > 0.$$ \hspace{1cm} (8.6)

For non-parabolic barriers, the probability is not necessarily 0.5 at the barrier top, and these equations do not hold precisely; nevertheless, they are enforced in all of our tunnelling approximations discussed above with $V_{\text{max}}$ replaced by the maximum of $V_{\text{SG}}(s)$. This maximum is called $V_{\text{AG}}$. Note that eqn (8.4) reduces to the WKB approximation, $e^{-2\theta}$, when $\theta$ is large. Replacing $e^{-2\theta}$ by $1/(1 + e^{2\theta})$ is called parabolic uniformization; it can also be derived by phase integral methods.\textsuperscript{60}

Equations (8.4)–(8.6) show that the effect of tunnelling is to broaden the reaction threshold. Although the semiclassical tunnelling approximations we are discussing here correspond to reactions emanating in the ground state (when the reaction is considered in the exothermic or thermoneutral
direction), this broadening actually occurs at all reaction thresholds (where each threshold is associated with a quantized energy level of the transition state).\textsuperscript{61–65} The broadening may be understood as energy–time uncertainty broadening because the transition state energy level has a finite lifetime. The amount of broadening at each threshold may be understood quantitatively in terms of the lifetime of a quantum mechanical resonance state associated with the reaction threshold, and this lifetime may be correlated with the width of the barrier on the vibrationally adiabatic potential energy curve. The ground-state approximation that we have made may be recast in this language as the assumption that the broadening is the same at all thresholds. It is not, but this assumption can still be used to get useful tunnelling transmission coefficients.

Note that if we used eqns (8.4) and (8.5), and if we assumed no tunnelling, \textit{i.e.}, if we set $\theta = 0$, the probability of passage through the transition state dividing surface would be a Heaviside function rising from 0 to 1 at $E = V_{\text{max}}$; that is, the rate would be controlled by $V^A_G$ rather than by $V^G_a(s)$ at the saddle point, as in conventional transition state theory. Since the low-temperature limit of variational transition state theory has the variational transition state at the maximum of $V^G_a(s)$ rather than at the saddle point, the correct inclusion of tunnelling in transition state theory is more consistent with variational transition state theory than with conventional transition state theory. The fact that the variational transition sate is not exactly at the maximum of $V^G_a(s)$ at finite temperature leads to a small consistency factor called the classical adiabatic ground-state correction,\textsuperscript{42,66} or the consistency may be enforced by using improved canonical variational theory (ICVT),\textsuperscript{42,50} but these are small effects and the technical details need not concern us here.

Another point worth mentioning is that the transmission coefficient is not necessarily larger for deuterium-substituted system than for a protium system, in contrast to what one would find if tunnelling were one-dimensional. The simplest-to-understand example of the counterintuitive case occurs as a special case for the common situation where the zero point effects make the vibrationally adiabatic ground-state barrier height smaller than the classical barrier height because the zero-point energy of the transition state is smaller than the zero-point energy of the reactants. Then, for a small classical barrier, the vibrationally adiabatic barrier might be low but above the zero-point energy of reactants for deuterium transfer but even lower and below the zero-point energy of reactants for protium transfer. So there would be tunnelling for deuterium but not for protium. Further discussion of this counterintuitive case is given elsewhere.\textsuperscript{66a,66b}

### 8.3 Validation

The methods described above (ZCT, SCT, LCT, $\mu$OMT, and LAT) have been widely tested against exact quantum mechanical solutions of the Schrödinger equation for the rate constants of many collinear atom–diatom reactions and a few three-dimensional atom–diatom reactions, and they have also been widely used for more complicated reactions where they can
be compared to experiment. The tests against exact quantum mechanics are
direct tests of the accuracy of the calculated rate constant that are of special
interest because the transmission coefficients and the accurate quantum
mechanical rate constants are calculated for the same potential energy
surface, so the comparison is not compounding the error due to the dif-
ference in the surfaces with the error due to the semiclassical tunnelling
calculation. However, since the comparisons are done for the rate constant,
they also test the underlying transition state theory to which the tunnelling
transmission coefficient is applied. For the tests reported here the under-
lying transition state theory is anharmonic ICVT.

We summarized a series of such tests comparing $k_{\text{approx}}$ and $k_{\text{acc}}$ (where
$k_{\text{approx}}$ is the approximate rate constant computed by ICVT with the given
tunnelling transmission coefficient, and $k_{\text{acc}}$ is the accurate quantum
mechanical rate constant) in a review article.\textsuperscript{67} The tests include up to
52 collinear and three-dimensional rate constants for various atom–diatom
reactions at each temperature (the precise number for each temperature
is in Table 8.1). In order to allow the errors to be averaged without bias
toward overestimates or underestimates and without cancellation of errors
between overestimates and underestimates, test results were averaged in
terms of positive percentage error, where positive percentage error is
defined, for example, such that it is $+50\%$ when $k_{\text{approx}}$ and $k_{\text{acc}}$ differ by a
factor of 1.5 (in contrast to the signed percentage error, which is $+50\%$
when $k_{\text{approx}}/k_{\text{acc}} = 1.5$ but $-33\%$ when $k_{\text{acc}}/k_{\text{approx}} = 1.5$). The logarithmically
averaged positive percentage errors are shown in Table 8.1.

Table 8.1 shows that we obtain the best results with the LAT approxi-
mation and that the results are only ever so slightly less accurate on average
with $\mu$OMT. Table 8.1 also shows that the results are greatly improved as
compared to ZCT when one includes reaction-path curvature, as in all four of
the methods to the right of ZCT. This is an especially important result when
one considers that essentially all one-dimensional tunnelling results in the
literature may be considered to be approximations to ZCT.

The results at 200 K are especially striking because the error without
tunnelling is very large; an average positive percentage error of 1600\% cor-
responds to an underestimate of the rate constant by a factor of 17, whereas

<table>
<thead>
<tr>
<th>$T$ (K)</th>
<th>$N$\textsuperscript{b}</th>
<th>No tunnelling</th>
<th>ZCT</th>
<th>SCT</th>
<th>LCT</th>
<th>$\mu$OMT</th>
<th>LAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>41</td>
<td>1576</td>
<td>332</td>
<td>63</td>
<td>49</td>
<td>41</td>
<td>40</td>
</tr>
<tr>
<td>250</td>
<td>44</td>
<td>488</td>
<td>167</td>
<td>43</td>
<td>32</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>300</td>
<td>52</td>
<td>273</td>
<td>116</td>
<td>34</td>
<td>24</td>
<td>23</td>
<td>21</td>
</tr>
<tr>
<td>400</td>
<td>53</td>
<td>124</td>
<td>75</td>
<td>36</td>
<td>22</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>600</td>
<td>41</td>
<td>45</td>
<td>33</td>
<td>21</td>
<td>18</td>
<td>17</td>
<td>16</td>
</tr>
</tbody>
</table>

\textsuperscript{a}From Table 77 of ref. 67.

\textsuperscript{b}$N$ is the number of comparisons to exact results at this temperature; this is limited by the
number of exact rate constants available for comparison.
an average positive percentage error of 40% corresponds to an error of a factor of 1.4. The error is particularly large because the 41 tests of the semiclassical theory at 200 K are dominated by hydrogen transfers; the 41 cases at 200 K include 23 protium transfers (17 collinear and six 3D), 11 transfers of muonium, deuterium, or tritium (10 collinear and one 3D), and seven transfers of a halogen (4 collinear and three 3D).

The improvement continues up to higher temperatures. For example, the average positive percentage errors at 300 K in Table 8.1 correspond to errors by a factor of 3.7 for ZCT, 2.2 for ZCT, 1.3 for SCT, and 1.2 for LCT, µOMT, or LAT. This illustrates a general point, namely that the LCT method (and therefore also the µOMT and LAT methods) has better across-the-board accuracy than the SCT method. Nevertheless, the much simpler SCT method is usually good enough; its average error is higher mainly because it can greatly underestime the tunnelling in cases with very small skew angles. 67

After the review article, a unique opportunity arose to test multidimensional semiclassical tunnelling methods against an accurate quantum mechanical rate constant (for a given potential energy surface) for a polyatomic reaction,68,69 in particular the reaction H+CH$_4$ → H$_2$ + CH$_3$, for which accurate quantum mechanical rate constants were obtained by very difficult calculations.70 (Actually, the accurate quantum calculations for this case, unlike those used for tests described above, do involve an approximation, namely that rotation and vibrations are separable. That assumption is also made in the calculations to which we compare, so it does not affect the validity of the comparison.) The ICVT/µOMT calculations were found to agree with the accurate quantum ones within 22% (maximum error) over the full range of temperatures from 200 to 1000 K, over which the rate constants vary by eight orders of magnitude.69

### 8.4 Extensions

In applications to molecules that have torsions, the transition state may have several conformations with only small barriers between them. For such a case we developed multipath variational transition state theory where one includes tunnelling along the lowest-energy path through each conformation of the transition state.71,72

In addition to gas-phase reactions discussed so far, the multidimensional semiclassical tunnelling methods have also been used for diffusion at gas–solid interfaces,73–76 for bulk diffusion and transport from a gas–solid interface to the bulk,76 for reactions in solid matrices,77 for the reactions in liquid-phase solutions, and for enzyme-catalyzed reactions.78–87

For solid-state reactions, it is important to pay special attention to the low-temperature limit. In the low-temperature limit, all the reaction is in the exothermic direction and it proceeds from the lowest energy state of the reactants. Therefore the rate constant becomes independent of temperature. This is easily accommodated in the semiclassical theory by considering tunnelling only at the quantized energies of the reactants.73–77
For reactions in liquids and enzymes, the number of minimum energy paths may become essentially uncountable and one must use statistical methods, in particular potentials of mean force, rather than an unaveraged potential energy surface, to calculate the reaction rate.\textsuperscript{88–91} If implicit solvation models are used to calculate the free energies of the reacting solute, one does not have a minimum energy path, but one can calculate a minimum free energy path, and by making a practical approximation called the canonical mean shape zero-order (CMS-0) approximation, the tunnelling probabilities may be calculated using the potential of mean force rather than the potential energy surface.\textsuperscript{88} The potential of mean force is also called the free energy surface.\textsuperscript{92,93}

However, if the solvent is treated explicitly by molecular dynamics simulation with weighted histogram analysis\textsuperscript{94} (WHAM) of the free energy, one may employ a liquid-state multipath tunnelling approximation by selecting a statistical sample of paths from the variational transition state slice through the WHAM free energy profile along the reaction coordinate. This is called ensemble-averaged variational transition state theory, and it has been applied successfully to many enzyme reactions.\textsuperscript{78–84} A key finding is that the tunnelling transmission coefficient can depend strongly on ensemble averaging.\textsuperscript{95,96}

### 8.5 Concluding Remarks

Computational details of the methods discussed above are given elsewhere.\textsuperscript{42,97–99}

The present review has focussed on the development of semiclassical multidimensional tunnelling methods that are available in the Polyrate\textsuperscript{100} and Pilgrim\textsuperscript{101} computer programs. Reviews of recent work on tunnelling effects in chemistry are available elsewhere.\textsuperscript{83,102–108}

### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMS-0</td>
<td>Canonical mean shape zero-order</td>
</tr>
<tr>
<td>ICVT</td>
<td>Improved canonical variational theory</td>
</tr>
<tr>
<td>LAT</td>
<td>Least-action tunnelling</td>
</tr>
<tr>
<td>LCT</td>
<td>Large-curvature tunnelling</td>
</tr>
<tr>
<td>MEP</td>
<td>Minimum-energy path</td>
</tr>
<tr>
<td>µOMT</td>
<td>Microcanonical optimized multidimensional tunnelling</td>
</tr>
<tr>
<td>SCT</td>
<td>Small-curvature tunnelling</td>
</tr>
<tr>
<td>WHAM</td>
<td>Weighted histogram analysis method</td>
</tr>
<tr>
<td>WKB</td>
<td>Wentzel–Kramers–Brillouin</td>
</tr>
<tr>
<td>ZCT</td>
<td>Zero-curvature tunnelling</td>
</tr>
</tbody>
</table>

### Acknowledgments

This work was supported in part by the U.S. Department of Energy, Office of Basic Energy Sciences under Award DE-SC0015997.
References


103. W. T. Borden, Reactions that involve tunneling by carbon and the role that calculations have played in their study, *WIREs Comput. Mol. Sci.*, 2016, 6, 20–46.


---


**Page correction**

262 near end of middle paragraph
there indeed some  ->  indeed there are some

267 third line below equation (8.3)
$m_B \gg m_C$  ->  $m_B << m_C$

268 Fig. caption 8.1, third last line
masses of A and B.  ->  masses of A and C.

271 middle of page
**correction**,\(^{42,66}\)  ->  **correction**,\(^{42,66a}\)

271 end of section 8.2
elsewhere.\(^{66a,66b}\)  ->  elsewhere.\(^{66b}\)

273 line 9
3.7 for ZCT,  ->  3.7 for no tunneling,