Derivation of a general three-dimensional crack-propagation law: 
A generalization of the principle of local symmetry

Jennifer A. Hodgdon* 
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853

James P. Sethna† 
Laboratory of Applied Physics, Technical University of Denmark, DK-2800 Lyngby, Denmark and NORDITA, DL-2100 Copenhagen Ø, Denmark 
(Received 24 June 1992)

We derive a general crack-propagation law for slow brittle cracking, in two and three dimensions, using discrete symmetries, gauge invariance, and gradient expansions. Our derivation provides explicit justification for the “principle of local symmetry,” which has been used extensively to describe two-dimensional crack growth, but goes beyond that principle to describe three-dimensional crack phenomena as well. We also find that there are materials properties needed to describe the growth of general cracks in three dimensions, besides the fracture toughness and elastic constants previously used to describe cracking.

I. INTRODUCTION

There are many aspects of the problem of crack growth that have received a lot of attention recently. For instance, there has been much interest in dynamic fracture and the accompanying crack bifurcation and other instabilities. The transition between failure due to percolation of a network of many small cracks and failure due to a single dominating crack has also been explored, as well as the transition between brittle and ductile cracking. Pattern formation in multiple cracking has also been of interest. In light of all the interest in these rather complex phenomena of fracture, it is somewhat surprising to find that little is known about the growth laws for even slow-growing, single three-dimensional cracks, though there has been some work done on calculating the paths of cracks in two and three dimensions, and many measurements and calculations of the crack growth rate for simple two-dimensional geometries.

The problem of finding a growth law for cracks would seem to be of fundamental interest; so, in this paper, we apply standard tools of theoretical physics—gradient expansions, symmetry, and gauge symmetry—to find the most general possible growth law for a three-dimensional crack growing slowly in a homogeneous, isotropic medium. Since, with today’s computers, it is possible to make precise numerical computations of the elastic fields for arbitrary three-dimensional geometries in a matter of hours, we do not consider the related problem of finding the stress state of the material containing the crack, but consider it to be completely known.

We also compare the crack growth law we derive here to previously derived and measured properties of cracks in two and three dimensions. There is one parameter f in our growth equation which is expected to be very large for real materials. We will see that our growth equation is simplified in the $f \to \infty$ limit, so that it becomes equivalent to previously known growth laws, under some loading conditions, but that, under other conditions, there is more interesting behavior which has yet to be explored fully (work in progress).

II. SIMPLIFICATIONS

We begin by simplifying the problem of crack propagation using length and time scale considerations. First, we smooth our crack problem over the length scale $l_s$ which characterizes the size of inhomogeneities and anisotropies in the material containing the crack. Although for a single-cystal sample, removing the anisotropy by smoothing is impossible, so that $l_s$ is effectively larger than the sample, for many situations of practical relevance $l_s$ is much smaller than the size of the body. For example, in a glass, $l_s$ is a few atomic sizes; in a polycrystal, it is the grain size; in concrete, it is the size of or distance between the pebbles it contains. In those cases, we can smooth the crack problem over $l_s$ without losing much information, making the crack a smooth surface, and the material containing the crack continuous, homogeneous, and isotropic. For this paper, we assume $l_s$ is small.

A second length scale in crack propagation problems arises because every material has some stress above which it fails to have linear elastic properties (e.g., it begins to flow, plasticity or viscously, it emits dislocations, its bonds break, or it has a martensitic transformation). For some materials, this stress is very low, and there is essentially no linear elastic regime. For others, linear elasticity is valid except very near the crack tip, where the stress is much higher than in the bulk of the body. For these materials, there is a length scale $l_{nl}$ which characterizes the size of the nonlinear process zone around the crack tip; $l_{nl}$ can range from a few angstroms in glass to tens of centimeters in concrete. In this paper, we consider only materials for which $l_{nl}$ is small compared to the length of the crack and the size of the body, so that the bulk of the ma-
terial can be considered linear elastic. This work, then, describes materials usually considered linear and brittle, as well as materials exhibiting viscoelasticity, plasticity, and martensitic transformation toughening, as long as the length scale for these behaviors is sufficiently small. In principle, the nonlinear properties of such materials could be included to extend applicability to smaller length scales.

A third length scale relevant to crack propagation is associated with the degree of translational invariance along the crack front. For many crack systems studied in the past, every plane perpendicular to the crack front is equivalent, which means that the problems are effectively two dimensional. On the other hand, some practical crack problems are not two dimensional, but instead have crack-front curvature or stresses which vary along the crack front. If this is the case, then there is a length scale, which we call the dimensional crossover length \( l_{dc} \), above which the problem is effectively three dimensional; \( l_{dc} \) is either one of the geometric lengths associated with the crack geometry (such as the radius of curvature of the crack front), or is associated with the stress gradient: \( l_{dc} \approx \sigma / \sqrt{\pi \sigma} \). For this paper, we assume that \( l_{dc} \) is large, though not necessarily as large as the size of the body containing the crack, and we expand in powers of quantities which are inversely proportional to \( l_{dc} \) (i.e., gradients).

A fourth type of length scale that enters the crack problem is a crossover length \( l_c \), where the linear elastic solution for the displacement \( u \) near a crack, \( u = C r^{1/2} f(\theta) \), where \( (r, \theta) \) are the polar coordinates centered at the crack tip, \( f \) is a known trigonometric function, and \( C \) is a constant, crosses over to behavior with a different power law in \( r \). (There is a corresponding crossover in the stress and strain; both stress and strain can be found from the displacement.) This arises because the elastic equations around a crack allow many solutions of the form \( u = C r^{1/2} f_c(\theta) \). Asymptotic analysis shows that the only allowed values of \( b \) are integers and half integers. The integer \( b \) solutions turn out to be unimportant for fracture because they have the same displacement on both crack surfaces, and hence no crack opening. The half integer \( b \) solutions do have crack openings, and could all represent fractures; the usual practice, however, is to use only the \( b = \frac{1}{2} \) solution. We can justify this practice using a length scale argument; closer to the crack tip than a length \( l_c^+ \), the \( b = \frac{1}{2} \) solution dominates over larger \( b \) solutions; farther away from the crack tip than a second length \( l_c^- \), the \( b = \frac{1}{2} \) solution dominates over smaller \( b \) solutions. Normally, \( l_c^- \) is only a few times larger than the nonlinear process zone, and \( l_c^+ \) is only a few times smaller than the crack length, so that, given the above assumptions, we do expect the \( b = \frac{1}{2} \) solution to dominate all other solutions. We assume that is the case in this paper.

Finally, we also simplify the crack propagation problem by considering cracks that are growing slowly enough that inertial and relativistic (close to the speed of sound) effects are unimportant. Some of these effects could be included in future work, but the present analysis suffices for cracks that arrest after growing a certain distance, such as when a wedge is driven into a crack; cracks that grow at a constant speed, such as under constant displacement loading; and cracks that may eventually speed up, but which are currently growing slowly, as in the cases of fatigue cracks, subcritical cracking, and the first stages of growth under constant force loading.

### III. RELEVANT VARIABLES

The knowledge of length scales from the previous section simplifies our problem to a nearly two-dimensional smooth crack, in an isotropic, homogeneous, linear-elastic, continuous bulk medium. Still, at first glance it appears that there are many variables which could influence the propagation of the crack; for example, the load on the surface of the body, type of material, temperature, ambient atmosphere, and the stress and fracture history. However, we are concerned here with the propagation of a crack given the elastic fields in the body, not the precise conditions that produce those fields. Also, many variables, such as stress history and temperature, can be included implicitly in materials constants, which we also assume are known. So, for this work, the relevant variables are the elastic fields near the crack tip, materials constants, and the current configuration of the crack.

Now, in effectively two-dimensional problems satisfying the assumptions of the previous section, it is well known that the elastic fields (e.g., stress, strain, and displacement) near the tip of the crack—which is the only area we expect to influence crack growth—are determined by three stress-intensity factors (SIF's). For instance, the displacement obeys

\[
U_r(r, \theta) = \sum_a K_a \sqrt{2\pi r} J_a^0(\theta),
\]

(1)

where \( (r, \theta) \) are the polar coordinates in the two-dimensional plane, the \( K_a \) are the mode I, II, and III SIF's, and the \( J_a^0 \) are known trigonometric functions. [There are similar expressions for the stress and strain fields, each SIF corresponds to a physical mode of crack opening (see Fig. 1.)] So, the relevant variables for the elastic field near a two-dimensional crack, instead of being a displacement vector at every point in the plane, are reduced to just three numbers, the SIF's.

In three dimensions, we need more variables for the elastic fields, since each plane in the material is different. In situations where the variation from plane to plane in the material is strong, that would involve solving the full three-dimensional elasticity equations, but the variation is weak under the assumptions of the previous section. So, to good approximation, each plane satisfies two-dimensional elasticity, and in each one we can characterize the elasticity by three SIF's. Therefore, for the elasticity in three dimensions, we need three SIF's at each point on the crack-front curve. As in the two-dimensional case, this is a large reduction in variables from a displacement vector at every point in the body.

Another variable that is relevant to crack growth is the starting shape of the crack, which consists of two-
IV. THE CRACK GROWTH LAW

Now we are ready to derive a crack growth law in our relevant variables: materials constants, the SIF's, and the three unit vectors associated with the crack geometry. That is, we are ready to derive an expression for the time evolution of the crack front curve \( \mathbf{x}(\lambda) \) as an expansion in the gradient \( \partial / \partial s \equiv (\partial s / \partial \lambda)^{-1} \partial / \partial \lambda \); \( \partial / \partial s \) is small under the assumptions above, and strictly zero in two dimensions.

Under the assumption that the crack surface is smooth, the time derivative of \( \mathbf{x} \) must be in the plane of the crack, so we can immediately write that

\[
\frac{\partial \mathbf{x}}{\partial t} = A \hat{n} + B \hat{t},
\]

for some \( A \) and \( B \). Since \( \hat{t} \) is determined by the crack-front curve \( \mathbf{x} \), while \( \hat{n} \) is not, if we want to know more than the instantaneous change in \( \mathbf{x} \), we will also have to know \( \partial \hat{n} / \partial t \). Since \( \hat{n} \) is a unit vector, \( \partial \hat{n} / \partial t \) cannot have a component along \( \hat{n} \), so we can write that

\[
\frac{\partial \hat{n}}{\partial t} = C \hat{t} + D \hat{b},
\]

for some \( C \) and \( D \). Equations (2) and (3) are now enough to specify the complete time evolution of the crack, given expressions for \( A, B, C, \) and \( D \).

There are several considerations we can use to find these coefficients. First, gauge symmetry and the requirement that \( \hat{t}, \hat{n}, \) and \( \hat{b} \) are mutually perpendicular (see Appendix A) reduce Eqs. (2) and (3) to

\[
\frac{\partial \mathbf{x}}{\partial t} = v \hat{n} + w_n \hat{t},
\]

\[
\frac{\partial \hat{n}}{\partial t} = \left[ \frac{- \partial v}{\partial s} + w_n \frac{\partial \hat{n}}{\partial s} \right] \hat{t} + \left[ E + w_n \frac{\partial \hat{b}}{\partial s} \right] \hat{b},
\]

where \( E \) and \( v \) are physical functions, \( w_n \) is a nonphysical function which characterizes the gauge, and all quantities are implicit functions of \( \lambda \) and \( t \). This means that the growth of the crack is now determined by only two physical functions.

Consideration of the discrete symmetries of the crack problem further constrains these two functions. We consider symmetry operations, at some point \( \lambda_0 \) on the crack front, which leave the unit vectors at \( \lambda_0 \) fixed and reflect or rotate the material, preserving the physical properties that \( \hat{t}(\lambda_0) \) is the tangent to \( \mathbf{x} \), and that \( \hat{n}(\lambda_0) \) is the direction of crack growth at \( \lambda_0 \). (This type of transformation is equivalent to leaving the material fixed and transforming the coordinates, but we choose to leave the coordinates fixed to avoid questions of the parity of the coordinate system.) There are two independent symmetry operations, which we take to be (a) 180° rotation about \( \hat{n}(\lambda_0) \), and (b) reflection in the \( n(\lambda_0)-t(\lambda_0) \) plane (see Fig. 2).

We need the transformation properties of \( E \) and \( v \) in
Eq. (4), as well as the variables and derivatives at $\lambda_0$, all quantities either remain the same or change sign under the two transformations. First, $v$, which is the rate of crack growth, must remain unchanged under both transformations, since a change of sign would change crack growth into crack healing. In contrast, $E$ must change sign under both transformations, since it multiplies $\hat{b}$ in Eq. (4), and while $\hat{b}$, the local normal to the crack plane, is invariant under the transformations, the physical normal to the crack plane changes sign under both.

Next, we need the transformation properties of the SIF's. $K_1$ transforms like the stress component $\sigma_{ab}$, $K_{II}$ like $\sigma_{bb}$, and $K_{III}$ like $\sigma_{ab}$, as can be seen from Fig. 1. The stress $\sigma$ is a tensor that transforms with the material; symmetry operation (a) takes material at $(x_n, x_y, x_z)$; symmetry operation (b) takes material at $(x_n, -x_y, -x_z)$ to $(x_n, -x_y, x_z)$. Therefore, $K_1 \rightarrow K_1$ under both (a) and (b), $K_{II} \rightarrow -K_{II}$ under both, and $K_{III} \rightarrow -K_{III}$ under (a) and $-K_{III}$ under (b).

We also need the transformation properties of the SIF gradients. Consider a case where one of the SIF's, before transformation, is greater in absolute value for $x_i > 0$ than for $x_i < 0$, so that $\partial|K|/\partial s > 0$. Then, under transformation (a), the material at $x_i > 0$, which has the greater $|K|$, moves to $x_i < 0$, so that $\partial|K|/\partial s < 0$ after the transformation; transformation (b) leaves $x_i$ unchanged, so that $\partial|K|/\partial s$ remains $> 0$. These results can be combined with the transformation properties of the SIF's themselves to give the transformation properties of the gradients of the SIF's; the results are in Table I.

We also need the transformation properties of gradients of the unit vectors. We are interested only in non-vector quantities for use in $v$ and $E$; to make the vector gradients into pseudoscalar quantities, we use the combinations $\partial \hat{a}/\partial s \cdot \hat{b}$, where $a$ and $b$ stand for $[n, b, t]$. We use the term pseudoscalar because, while these are not vector quantities, they do not necessarily transform as scalars under operations (a) and (b), but may change sign. The SIF's are also pseudoscalars.] Noting that

$$\frac{\partial (\hat{a} \cdot \hat{b})}{\partial s} = 0 = \frac{\partial \hat{a}}{\partial s} \cdot \hat{b} + \frac{\partial \hat{b}}{\partial s} \hat{a},$$

for all $a$ and $b$, since $\hat{a}$, $\hat{b}$, and $\hat{c}$ are mutually orthogonal unit vectors, we can see that there are only three independent combinations to consider, which we take to be $(\partial \hat{a}/\partial s) \cdot \hat{c}$, $(\partial \hat{b}/\partial s) \cdot \hat{c}$, and $(\partial \hat{b}/\partial s) \cdot \hat{b}$. From Fig. 3, we can see that the transformation properties of these three quantities are as shown in Table I.

The only relevant variables left to consider transformation properties of are the materials constants, but they are by definition true scalars, and cannot change sign under any transformation. Also, their gradients $\partial /\partial s$ are zero, as they are constant throughout the material.

With these symmetry properties, we can now derive a law for crack growth. First, we examine the case of a crack in two dimensions, where both $\partial /\partial s$ and $w$ are zero; the fact that $\partial /\partial s = 0$ means that the only nonzero scalar and pseudoscalar quantities are combinations of the SIF's and materials constants. Under symmetry operation (a), $K_1$, $K_{III}$, and the materials constants remain the same, while $K_{II}$ changes sign; under operation (b), both $K_{II}$ and $K_{III}$ change sign, and everything else remains the same (see Table I). Since the three SIF's have different transformation properties, and only $K_{II}$ transforms like $E$, we see that the two-dimensional crack growth law must have the form

<table>
<thead>
<tr>
<th>Quantity</th>
<th>(a) $180^\circ$ rotation</th>
<th>(b) $n-t$ reflection</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_1$</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>$K_{II}$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$K_{III}$</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>$\partial K_1/\partial s$</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>$\partial K_{II}/\partial s$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\partial K_{III}/\partial s$</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>$\partial (\hat{a} \cdot \hat{b})/\partial s$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$E$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$v$</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

FIG. 2. Transformations (a) rotation about $\hat{a}$ and (b) reflection in the $n-t$ plane on (c) the untransformed crack.
\[ \frac{\partial \mathbf{x}}{\partial t} = v\mathbf{n}, \]
\[ \frac{\partial \mathbf{n}}{\partial t} = -fK_{\Pi} \mathbf{t}, \tag{6} \]

where both \( v \) and \( f \) are true scalars; i.e., functions of materials constants, \( K_1, K_{\Pi}^2, \) and \( K_{\Pi}^2 \). The minus sign makes \( f > 0 \) correspond to the observed direction of crack growth under mode-II loading; this is discussed in the next section.

In three dimensions, the gradient \( \partial / \partial s \) is not strictly zero, and there are therefore more scalar and pseudoscalar quantities to consider than in two dimensions (see Table I). From the transformation properties of these quantities we can see that the physical crack growth law, to first order in \( \partial / \partial s \), has the general three-dimensional form
\[ \frac{\partial \mathbf{x}}{\partial t} = v\mathbf{n} + w\mathbf{t}, \tag{7} \]
\[ + \left[ -fK_{\Pi} + g_1K_{\Pi} \frac{\partial K_1}{\partial s} + g_{\Pi}K_{\Pi}K_{\Pi} \frac{\partial K_1}{\partial s} \right] \mathbf{t} \]
\[ + \left[ g_{\Pi} \frac{\partial K_{\Pi}}{\partial s} + h_{ib} \frac{\partial \mathbf{t}}{\partial s} \right] \mathbf{b} + \left[ h_nK_{\Pi} \frac{\partial \mathbf{n}}{\partial s} \right] \mathbf{t}, \]
\[ + \left[ (h_{ib}K_{\Pi}K_{\Pi} + w) \frac{\partial \mathbf{n}}{\partial s} \right] \mathbf{t}, \]
\[ \text{where the } f, g_{\alpha}, \text{ and } h_{ij} \text{ are scalars, zeroth order in } \partial / \partial s \text{ (functions of materials constants, } K_1, K_{\Pi}^2, \text{ and } K_{\Pi}^2 \text{); and } \]
v \text{ is a scalar, up to first order in } \partial / \partial s \text{ (a function of materials constants, } K_1, K_{\Pi}^2, K_{\Pi}^2, (K_{\Pi}K_{\Pi}K_{\Pi} / \partial s), (K_{\Pi}K_{\Pi} \partial K_1 / \partial s), (K_{\Pi} \partial K_{\Pi} / \partial s), [K_{\Pi} \partial (\mathbf{t} / \partial s) \cdot \mathbf{b}], [\partial (\mathbf{n} / \partial s) \cdot \mathbf{t}], \text{ and } [K_{\Pi} \partial (\mathbf{n} / \partial s) \cdot \mathbf{b}]. \]

Equation (7) can easily be expanded to higher order in \( \partial / \partial s \). Also, we note here that while the SIF’s have been used in this section as the variables describing the elastic fields in the crack, Eq. (7) is actually applicable to other solutions to the elasticity near a crack. That is because, in an isotropic body, it turns out that the coefficients appearing in the other solutions have the same symmetry as the SIF’s. So, if it is ever determined that one of the other solutions is more important for a particular problem, Eq. (7) can still be used, by substituting the proper coefficients for the \( K_{\alpha} \).

V. THE UNDETERMINED FUNCTIONS IN THE CRACK GROWTH LAWS

Now, in Eqs. (6) and (7) we have general forms for the crack growth law for two- and three-dimensional cracks. These equations contain unspecified functions—\( v, f, g_{\alpha}, \) and \( h_{ij} \)—which must be determined from considerations other than the symmetry considerations we used to find Eqs. (6) and (7). Probably the most important of these functions is the crack growth rate \( v \), a function of materials constants \( K_1, K_{\Pi}^2, \) and \( K_{\Pi}^2 \) in two dimensions. This function has been measured for mode-I cracks\(^{13}\) and generally has the form of the solid line in Fig. 4(a), with no growth for \( K_1 \) below some value \( K_{1C} \), which depends on the material, and a monotonically increasing growth rate above \( K_{1C} \). This schematic form has also been found in a theoretical calculation for a viscoelastic system.\(^{16}\) However, both in theory\(^{20–23}\) and in very clean experimental systems,\(^{21–22}\) the growth rate has the form in Fig. 4(b), with a negative value (crack healing) when \( K_1 < K_{1C} \). This means that the growth rate is a continuous function which passes through zero at \( K_{1C} \), so that for SIF’s near \( K_{1C} \) (i.e., for slow growth), we can expand \( v \) as

![Fig. 3. Illustrations for the transformation properties of the three unit vector combinations. (a) Top view of the surface of a planar crack with a curved crack front, which makes \((\partial \mathbf{n} / \partial s) \cdot \mathbf{t} \neq 0\), and \((\partial \mathbf{n} / \partial s) \cdot \mathbf{b} = (\partial \mathbf{t} / \partial s) \cdot \mathbf{b} = 0\) \((\mathbf{b} \text{ is constant}). \)

Under rotation of the material about \( \mathbf{t} (\mathbf{t}_o) \), keeping the vectors fixed [symmetry operation (a)], it can be seen that the value of \((\partial \mathbf{n} / \partial s) \cdot \mathbf{t} \neq 0\), and \((\partial \mathbf{n} / \partial s) \cdot \mathbf{b} = (\partial \mathbf{t} / \partial s) \cdot \mathbf{b} = 0\) \((\mathbf{b} \text{ is constant}). \)

From this crack configuration, it can be seen that \((\partial \mathbf{n} / \partial s) \cdot \mathbf{b} \text{ changes sign under symmetry operation (a) and (b)}. \)
\[ \nu(K_1) = v_0 \frac{K - K_C}{K_C} \] with \( v_0 \) a material-dependent constant. For mode-II and -III cracks, as well as mixed-mode cracks, since the elastic energy released per unit area of crack surface (the “energy release rate”) is proportional to \( \left[ K_1^2 + K_II^2 + K_III^2 / (1 - \nu) \right] \), where \( \nu \) is Poisson’s ratio, we expect that a growth-rate function valid for all modes of cracking can be expanded as

\[ \nu(K_1, K_{II}, K_{III}) \approx v_0 \frac{K - K_C}{K_C}, \]

where

\[ K = \left[ K_1^2 + K_{II}^2 + K_{III}^2 / (1 - \nu) \right]^{1/2}. \]

Also, note that for fatigue cracking, where our growth laws must still hold (on time scales long compared to the load cycle), the growth rate generally does not go to zero sharply at \( K_C \), but has the more gradual turn-on behavior of the dotted line in Fig. 4(a).

Now, we saw in the previous section that the crack growth rate in three dimensions can also depend on gradient quantities, besides the SIFs; the dependence of the crack growth rate on these quantities has not been measured, to our knowledge. However, there is no reason to suppose that the dependence on these quantities has special behavior (e.g., zero crossing or very strong dependence on SIFs) near \( K_C \), the value of the SIF where the crack growth rate becomes positive. So, for slow growth, where \( K \approx K_C \), we expect that we can approximate the dependence of the crack growth rate on the gradient quantities by a constant, and absorb it into \( v_0 \).

Similarly, we expect that the seven functions \( \gamma, g_{\alpha \beta}, \) and \( h_{ij} \) in Eqs. (6) and (7), which are allowed by symmetry to be functions of materials constants \( K_1, K_{II}, \) and \( K_{III}, \) can be approximated as constants when the growth rate is small. This means that to find the material-specific form of the crack growth law, for slow growth, it is a reasonable approximation to measure only the linearized dependence of the crack growth rate on \( K \), and the constant parts of \( \gamma, g_{\alpha \beta}, \) and \( h_{ij} \). So, with only a small number of experimental data points, our general formulation can be used to predict the growth of real cracks.

We can also use the above discussion of length scales to make an order-of-magnitude estimate of some of these constants. Under our assumptions, the only time scale in crack growth comes from the crack growth rate, which we write as \( \nu = -\partial a / \partial t \), where \( a \) is the length of the crack along a trajectory which has constant \( \lambda_n \) in the reference gauge (see Appendix A). If we divide the second part of Eq. (7) by \( \nu \), we have

\[
\frac{\partial \hat{n}}{\partial a} = - \left[ \frac{1}{\nu} \frac{\partial \nu}{\partial s} + \frac{w}{\nu} \frac{\partial \nu}{\partial s} \hat{n} \right] \hat{t} \\
+ \left[ \frac{F}{\nu} K_{II} + \frac{g_1}{\nu} K_{III} \frac{\partial K_1}{\partial s} + \frac{g_{II}}{\nu} K_{II} \frac{\partial K_{II}}{\partial s} + \frac{g_{III}}{\nu} K_{III} \frac{\partial K_{III}}{\partial s} + \frac{h_{ib}}{\nu} \frac{\partial h_{ib}}{\partial s} \hat{n} + \frac{h_{ii}}{\nu} K_{II} \frac{\partial \hat{n}}{\partial s} \hat{n} \right] \hat{b},
\]

and

\[
\hat{b} = \frac{\partial \hat{n}}{\partial a} / \left( \frac{\partial \hat{n}}{\partial a} \right). \]

FIG. 4. (a) Form of the crack growth rate \( \nu \) as a function of stress intensity factor \( K \) in ordinary crack experiments, where the growth rate is zero below a critical value \( K_C \) of the stress-intensity factor, and then has a sharp turn-on. The dotted line shows the behavior under fatigue, where the growth rate increases more gradually. (b) Form of the crack growth rate in very clean experiments, where crack healing can take place.
The left-hand side now has dimensions of inverse length, as must each term on the right-hand side. This means, in particular, that \( f/v \) contains an inverse length, and since the only length scales in the problem are microscopic, under our assumptions, \( f/v \) must be nearly infinite. (\( f/v \) also contains an inverse SIF scale; the natural scale for SIF's is \( K_C \).)

On the other hand, all other terms on the right-hand side of Eq. (9) contain the gradient operator \( \partial / \partial s \), which has dimensions of inverse length, so their coefficients \( \tilde{g}_a / v \) and \( h_{ij} / v \) do not contain length scales. For that reason, we do not expect these functions to be either very large or very small, but instead to be of order 1 (with appropriate factors of \( K_C \)).

VI. PREDICTIONS OF THE TWO-DIMENSIONAL CRACK GROWTH LAW

Let us now examine the two-dimensional crack growth law, Eq. (6):

\[
\begin{align*}
\frac{\partial x}{\partial t} &= v \hat{n} , \\
\frac{\partial \hat{n}}{\partial t} &= -f K_{II} \hat{b} .
\end{align*}
\]

(10)

When \( K_{II} = 0 \), this equation says that the crack grows in a straight line (since \( \partial \hat{n} / \partial t = 0 \)), in agreement with the "principle of local symmetry" generally used to predict crack growth in two dimensions. However, the principle of local symmetry also says that \( K_{II} = 0 \) is maintained at all times by the propagating crack—in effect, that the crack curves in such a way as to keep \( K_{II} = 0 \). Our law, in contrast, says that it is only a nonzero \( K_{II} \) which can make the crack curve, but that (with \( f > 0 \)) the crack curves in such a way as to make \( K_{II} \) smaller (see Fig. 5).

Now, we can resolve the differences between the principle of local symmetry and our crack propagation law by dividing the crack growth rate \( v = \partial a / \partial t \) and by writing \( \hat{n} \) and \( \hat{b} \) in terms of the angle \( \theta \) that \( \hat{n} \) makes with the \( x \) axis. With these changes, Eq. (6) becomes

\[
\begin{align*}
\frac{\partial x}{\partial a} &= \cos \theta \hat{x} + \sin \theta \hat{y} , \\
\frac{\partial \theta}{\partial a} &= -(f/v) K_{II} .
\end{align*}
\]

(11)

In principle, \( f \), \( v \), and \( K_{II} \) are functions of \( x \) and \( \theta \). However, in the case of a small amount of growth at the end of a long crack, we expect \( f \) and \( v \) to be nearly constant as the crack grows, since the SIF's only change by a small amount during the growth (see Appendix B). Also, when \( \theta \) differs from the angle that makes \( K_{II} = 0 \) by only a small amount \( \Delta \theta(x) \), we can approximate \( K_{II} \) as

\[
K_{II}(x) = K_{I}(0) \Delta \theta(x)/2 ,
\]

(12)

neglecting terms of order \((x/a)^{1/2}\), where \( a \) is the length of the original long crack and \( K_{I}(0) \) is the starting value of \( K_{I} \) (see Appendix B). Substituting Eq. (12) in Eq. (11), we see that

\[
\frac{\partial \Delta \theta}{\partial a} = -\left[ \frac{fK_{I}(0)}{2v} \right] \Delta \theta ,
\]

(13)

which we can immediately solve, taking \( f \) and \( v \) constant,

\[
\Delta \theta(a) = \Delta \theta(0) e^{-fK_{I}(0)a/2v} .
\]

(14)

That is, if we start a crack with a small deviation \( \Delta \theta(0) \) from the direction predicted by the principle of local symmetry, then our crack propagation law says that the deviation decays with a characteristic distance of \( 2v/(fK_{I}) \), and since we saw in the previous section that \( f/v \) is very large, this distance is very short. In the limit that the distance is zero, or \( f/v \rightarrow \infty \), our crack propagation law for two dimensions agrees with the principle of local symmetry.

VII. CONCLUSIONS

We have seen that from symmetry principles we can derive a crack growth law for both two- and three-dimensional geometries. The laws we derived, Eqs. (6) and (7), contain several functions, such as the crack growth rate, which are not determined by symmetry alone and must be measured in controlled cracking experiments or atomic simulations; we expect that for slow
crack growth, only one value for each of the functions should suffice for predicting crack paths.

Our crack growth law agrees with the principle of local symmetry\(^{25}\) in the limit \(f \to \infty\) that the microscopic length scales in the crack problem are truly zero, in two dimensions. In three dimensions, it will require a more careful asymptotic analysis to understand completely the predictions of the crack growth law in this limit. Initial simulation results indicate that for a crack that is slightly perturbed from the two-dimensional plane, the principle of local symmetry is enough to determine the growth under mode-I loading. However, even using a large value of \(f\) in our simulations does not prevent \(K_n\) from becoming large in the vicinity of sharp features which tend to occur under mode-III loading, in violation of the principle of local symmetry. So, the principle of local symmetry is not enough to determine the growth of all cracks, even in the \(f \to \infty\) limit, and our full growth equation—subject to asymptotic analysis—provides an alternative.

By using symmetry principles, separation of length scales, gauge invariance, and gradient expansions, then, we have derived effective, macroscopic equations governing the growth of cracks in three dimensions. We have coarse-grained the problem so that microscopic details—such as atomic bond breaking, crystalline grain morphology, deformation near the crack tip in response to strain, and surface effects—are on such small length scales that they cannot affect the macroscopic crack growth. Understanding the microscopic origins of our effective growth equations, and describing crack growth on very small length scales, where our crack growth law is not valid, will demand calculations that include these microscopic details.

ACKNOWLEDGMENTS

We acknowledge the support of U.S. DOE Grant No. DE-FG02-88-ER45364; J.A.H. was also funded by the NSF and DOE. We would like to thank P. Wash Wawrzynek, Dave Potyondy, and Tony Inglesfield for introducing us to the problem of crack propagation, for many discussions about the subject, and for allowing us to use their two-dimensional finite element analysis code (FRANC); Jim Krumhansl for his continuing support and critical evaluation of this research; and Andy Ruina, Robb Thompson, and Jim Rice for their comments on earlier versions of this work. J.P.S. would like to thank Stephen Langer for introducing him to the gauge invariance of curves. We would also like to thank the Technical University of Denmark and NORDITA for support and hospitality; J.A.H. also thanks Dave Srolovitz at the University of Michigan for hospitality.

APPENDIX A: GAUGE SYMMETRY AND CRACKS

There are many cases where the natural mathematical description of a problem introduces fictitious degrees of freedom with no physical relevance. The most well-known example is in electromagnetism, where all physical quantities are unchanged when the gradient of an arbitrary function \(\chi(r)\) is added to the vector potential \(A\). The transformation \(A \to A + \nabla \chi\) is an example of a "gauge transformation"; the invariance of physical quantities under such a transformation is called "gauge invariance." The strong and weak forces of particle physics also have gauge transformations associated with them.\(^{26}\) In general relativity, the choice of coordinate system for space-time is arbitrary; this gauge invariance can be used to derive momentum and energy conservation.\(^{26}\) Another use of gauge symmetry is in site-disorder spin glasses,\(^{25}\) where a gauge transformation is used to show that the certain forms of disorder do not result in spin-glass properties, but in ferromagnetic behavior.

The term "gauge" is particularly appropriate for the gauge symmetry of our problem,\(^{28}\) where the parametrization \(\lambda\) of the crack-front curve \(x(\lambda, t)\) is arbitrary: how one "gauges" (measures) the points along the curve cannot affect the growth of the crack. There are two different types of gauge symmetry for cracks. The first type is the freedom to change the parametrization at any one time, which we call the "one-time gauge symmetry." The second is the freedom to choose how the parametrization at some time is related, through the growth equation, to the parametrization at a later time; we call this the "time-dependent gauge symmetry." Crack growth laws must satisfy both gauge symmetries; that is, neither the one-time nor the time-dependent gauge transformation can change the physical crack growth equation.

To satisfy the one-time gauge symmetry, the crack growth equation must be invariant under any change in the parameter \(\lambda\). This means that the equation cannot have any direct dependence on the value of \(\lambda\) at a point on the crack-front curve, but instead must depend on physical quantities, which are implicit functions of \(\lambda\). Also, derivatives along the crack-front curve appearing in the equation cannot be \(\partial / \partial \lambda\), but must instead be in terms of the arc length \(s\), because \(\partial / \partial s\) is gauge invariant, up to a choice of a minus sign; the equation must also be invariant under reversal of this sign.

The time-dependent gauge symmetry is slightly more complicated. Knowing \(\partial x(\lambda, t) / \partial t\) tells us how a point with parameter value \(\lambda\) evolves in time. This means that the time evolution of the parametrization is implicit in the crack growth equation, and time-dependent gauge transformations change its appearance, unlike one-time gauge transformations. This change happens in a well-defined way: if we have the crack growth law of Eqs. (2) and (3),

\[
\frac{\partial x(\lambda, t)}{\partial t} = A \hat{n} + B \hat{t},
\]

\[
\frac{\partial \hat{n}(\lambda, t)}{\partial t} = C \hat{t} + D \hat{b},
\]

where the right-hand sides are implicit functions of \(\lambda\) and \(t\), when we introduce a time-dependent gauge transformation to a new parameter \(\mu(\lambda, t)\), then the crack growth law becomes

\[
\frac{\partial x(\mu, t)}{\partial t} = A \hat{n} + B \hat{t} + \frac{\partial x(\lambda, t)}{\partial \lambda} \frac{\partial \mu}{\partial t},
\]

\[
\frac{\partial \hat{n}(\mu, t)}{\partial t} = C \hat{t} + D \hat{b} + \frac{\partial \hat{n}(\lambda, t)}{\partial \lambda} \frac{\partial \mu}{\partial t},
\]
where the right-hand sides are now implicit functions of \( \mu \) and \( t \), except where indicated. Writing \( \partial / \partial \lambda = (\partial s / \partial \lambda) (\partial / \partial s) \) (with \( s \) the arc length), defining a new function \( w = (\partial s / \partial \lambda) (\partial \mu / \partial t) \), and using the definition of \( \hat{\mathbf{t}} = \partial \mathbf{x} / \partial s \), we can write Eq. (A2) as

\[
\frac{\partial \mathbf{x} (\mu, t)}{\partial t} = A \hat{\mathbf{n}} + (B + w_2) \hat{\mathbf{t}}, \tag{A3}
\]

\[
\frac{\partial \hat{\mathbf{n}} (\mu, t)}{\partial t} = \left[ C + \frac{\partial \mathbf{x}}{\partial s} \frac{\partial \mathbf{n}}{\partial \mathbf{x}} \right] \hat{\mathbf{t}} + \left[ D + \frac{\partial \mathbf{n}}{\partial \mathbf{x}} \hat{\mathbf{b}} \right] \hat{\mathbf{b}},
\]

where \( w \) characterizes the time dependence of the gauge transformation.

There are three particular time-dependent gauges that we have found to be of special use in our study of crack growth. First, there is the “reference gauge,” where \( B + w \) in Eq. (A3) is zero. In this case, curves of constant parameter \( \lambda_\alpha \) are the integral curves of \( \hat{\mathbf{n}} \), and the form of the growth equation is found to be

\[
\frac{\partial \mathbf{x} (\lambda_\alpha, t)}{\partial t} = v \hat{\mathbf{n}}, \tag{A4}
\]

\[
\frac{\partial \mathbf{n} (\lambda_\alpha, t)}{\partial t} = - \frac{\partial v}{\partial s} \hat{\mathbf{t}} + \mathbf{E} \hat{\mathbf{b}}.
\]

The \( \partial \mathbf{v} / \partial s \hat{\mathbf{t}} \) term comes from the requirement that \( \hat{\mathbf{n}} \cdot \hat{\mathbf{t}} = 0 \) be preserved at all times; the functions \( \mathbf{E} \) and \( \mathbf{v} \) are free, as far as time-dependent gauge symmetry is concerned.

Since we know the growth law more explicitly in this gauge than in the generic gauge, it is a natural reference gauge for discussing other time-dependent gauges. In fact, comparing Eq. (A4) to Eqs. (A1) and (A3), we see that the growth law in a general time-dependent gauge can be written as

\[
\frac{\partial \mathbf{x} (\mu, t)}{\partial t} = v \hat{\mathbf{n}} + w_2 \hat{\mathbf{t}}, \tag{A5}
\]

\[
\frac{\partial \mathbf{n} (\mu, t)}{\partial t} = \left[ - \frac{\partial v}{\partial s} + w_2 \frac{\partial \mathbf{n}}{\partial \mathbf{x}} \right] \hat{\mathbf{t}} + \left[ E + w_2 \frac{\partial \mathbf{n}}{\partial \mathbf{x}} \hat{\mathbf{b}} \right] \hat{\mathbf{b}}.
\]

Note that \( w_2 = (\partial s / \partial \lambda_\alpha) (\partial \mu / \partial t) \) is a nonphysical gauge function characterizing the time dependence of the chosen gauge; \( E \) and \( v \), on the other hand, are physical functions describing the crack growth.

Another particular time-dependent gauge of interest is the “arc-length gauge,” where the parameter \( \lambda \) is always equal to the arc length \( s \) along the crack-front curve, measured from some starting point (such as the edge of the body). Since the arc length along the crack front is \( s(\lambda) = \int_{\lambda}^{\lambda} d\mathbf{x} / \partial \lambda \cdot d\lambda \) if we are initially in the arc-length gauge, then we find that we can remain in that gauge by choosing

\[
w_2 (\lambda) = - \int_{\lambda}^{\lambda} \left[ \frac{\partial \mathbf{n}}{\partial \lambda'} \hat{\mathbf{t}} \right] d\lambda'. \tag{A6}
\]

Although arc length is physically the most natural parametrization for a curve, the arc-length gauge is not usually very convenient, since \( w (\lambda) \) is a nonlocal function—as the crack grows, if the arc length of a section near some point \( \lambda_\gamma \) stretches (or shrinks), \( \lambda_\gamma \) must shift upwards (downwards) for all points with \( \lambda_\gamma > \lambda_\gamma \).

A third time-dependent gauge, the “\( z \) gauge,” is useful for cracks that have crack fronts which are nearly aligned along some axis, which we take to be the \( z \) axis. In this case, it is natural to use a gauge where the parameter \( \lambda_z \) of the crack front is the \( z \) coordinate of \( x(\lambda, t) \). To achieve this, the \( z \) component of \( \partial \mathbf{x} / \partial t \) in the crack growth equation must be zero, which makes

\[
w_2 (\lambda_z) = \pm \int_{\lambda}^{\lambda} \frac{\partial \mathbf{x}}{\partial t} \cdot \hat{\mathbf{t}} \cdot \frac{\partial \mathbf{n}}{\partial \lambda'} \hat{\mathbf{t}} \cdot \hat{\mathbf{b}} \cdot \hat{\mathbf{b}}, \tag{A7}
\]

Similar ideas can be used to make special gauges for cracks that are nearly circular, parametrizing with the angle \( \phi \) from the \( x \) axis, for instance, and for other common crack geometries.

**APPENDIX B: APPROXIMATION OF \( K_\Pi \) IN TWO DIMENSIONS**

In this appendix, we derive the approximate equation (12),

\[
K_\Pi (x) = K_\Pi (0) \left[ \frac{\Delta \theta (x)}{2} \right], \tag{B1}
\]

which gives \( K_\Pi \) after the crack has grown to \( x \) from the end of a long crack of length \( a \), in terms of the deviation \( \Delta \theta (x) \) of \( \theta \) (the angle that \( \hat{\mathbf{n}} \) makes with the \( x \) axis) from the angle that makes \( K_\Pi (x) = 0 \). First, we can use the results of Cotterell and Rice \(^{11}\) to find that as a function of the \( x \) coordinate of \( x \), measured from the end of the original long crack,

\[
K_\Pi (x) = K_\Pi (0) + \frac{1}{2} \theta (x) K_\Pi (0) \left[ \frac{2}{\pi} \right]^{1/2} \left[ \frac{2}{x} \right]^{1/2} \left[ \int_0^x \frac{\theta (x')}{(x-x')}^{1/2} dx' \right], \tag{B2}
\]

where \( T \) is the nonsingular (constant) part of the tensile stress at the end of the crack. Now, when the principle of local symmetry is satisfied, \( \theta (x) \) has the value which makes \( K_\Pi (x) = 0 \); we assume that \( \theta \) differs from this value by a small amount \( \Delta \theta (x) \). Furthermore, since \( T \), the constant tensile stress in the body, is the source for \( K_\Pi \), we can take \( T = b [K_\Pi (0) / \sqrt{a}] \), with the appropriate geometric factor \( b. \) Then we find that

\[
K_\Pi (x) = K_\Pi (0) \Delta \theta (x) / 2 \left[ 1 - 4b / (2x) \pi a \right]^{1/2} \left[ \int_0^x \Delta \theta (x') \left[ \frac{x}{x-x'} \right]^{1/2} dx' \right], \tag{B3}
\]

Now, if \( \Delta \theta (x) \) is constant, then the integral on the right-hand side is equal to 1; small variations of \( \Delta \theta (x) \) from the con-
stant value leave the integral approximately 1. Noting that the integral is multiplied by \((x/a)^{1/2}\), which is small by assumption, we can therefore neglect it, and approximate \(K_\Pi\) as

\[ K_\Pi(x) = K_1(0)[\Delta \theta(x)/2], \]

which is Eq. (B1). We also note that under the approximations of this appendix, and with further results of Cotterell and Rice, \(K_1\) is approximately constant.