Evolving Crystal Forms: Frank’s Characteristics Revisited

J W Cahn, J E Taylor and C A Handwerker

1 INTRODUCTION

Models of interface-controlled crystal growth or dissolution that assume that the velocity of growth \( v \) is a function only of the surface normal \( n \) and system parameters, such as temperature and matrix concentration, have a long history (see for example the historical overview of dissolution kinetics of minerals by Litsakes and Ney (1985)). An early proponent of this model was Gross (1918), who explored some of its predictions. In 1958 Frank discovered a simple set of predictions for the evolution of crystal shapes under these circumstances, based on the method of characteristics. He studied step motion using an analogy between steps moving across a surface and the mathematics of traffic flow. In the limit of small steps, this method is valid, for then crystal growth can be dealt with as being continuous in time and a function of normal direction only. The theory was further explored by Chernov (1963, 1984) in the USSR. In 1972, Frank returned to this problem and treated it without using steps, assuming only that some \( v(n) \) was given. He derived all his results from scratch, without tying them into a more general mathematical framework. We show here what this framework is.

The advantage of putting this problem into a larger context is (i) that the types of extension of the problem become clear, and (ii) the ‘reason’ that the results hold is seen not to be something special to this problem but to be part of a greater whole.

We will also address the issue of growth versus dissolution, both in for-
hazardous. Whenever there are shocks (see below) present in growth with positive time, there is non-uniqueness for the crystal in negative time. Letting \( u' \) be \(-v\) and running time forward is not the same as recovering the previous growth history of a form containing shocks.

There is also the question of limiting shapes. Chernov (1984) observed that 'the steady state growth shape is described by the envelope of the family of planes given by the equation \( n \cdot x = v(n)t \).' This is indeed the only steady state growth shape that is bounded at each time; this is a consequence of a more general theorem (whose proof we intend to provide in a later paper) that any bounded initial seed crystal which grows with positive normal velocity \( v \) and which at time 0 contains \( s_0 W_0 \) and is contained in \( s_1 W_0 \) will, for all positive \( t \), contain \( (s_0 + t) W_0 \) and be contained in \( (s_1 + t) W_0 \), where

\[
W_0 = \{ x : x \cdot n \leq v(n) \} \quad \text{for every unit vector } n.
\]

Note that \( W_0 \) is (by definition!) convex. There are, however, a number of infinite-extent surfaces (many of which are neither completely convex nor completely concave) that also retain their shape on growth.

Chernov (1984) asserts that 'from the geometric procedure of constructing Wulff’s envelope (\( W_0 \) here) it follows that the higher the growth rate of a face (or of a rounded area), the smaller its size on the convex steady state growth shape must be.' This is not quite accurate; the areas depend on the values of \( n \) on neighbouring directions. In fact, if one defines a new growth velocity function \( v_n \) by \( v_n(n) = v(n) + n \cdot b \), then the \( W_0 \) for the two growth velocities differ only by translation, and the limiting end shapes for growth of finite seeds are the same. Chernov goes on to say 'Analogously, faces with maximal growth rates should form the concave growth shape' (i.e. with the cavity decreasing in size). But there is no analogy! The concave growth shapes depend strongly on the shape of the initial seed (and the details of the dependence of \( v \) on \( n \)); there are exceptions to most general rules one might try to formulate, as we will demonstrate. In particular, there are a number of concave bounded-hole steady state growth shapes.

2 FORMULATION AND SOLUTION WITHIN THE FRAMEWORK OF PDE

Assume that \( v \) is a function defined on unit vectors \( n \) giving the normal velocity of growth of a crystal. That is, if the crystal at a certain time \( t \) and at a point \( x = (x_1, x_2, x_3) \) on its surface has exterior unit normal \( n \), then the instantaneous velocity of growth in direction \( n \) at \( x \) is \( v(n) \) (or more generally, \( v(x, t, n) \)). In order to convert this statement into a differential equation let \( \tau(x) \) be the time that the surface of the crystal arrives at \( x \), so
that \( \tau(x) = t_1 \) is an implicit equation for the shape (the set of \( x \) in the surface) at time equal to \( t_1 \), \( \tau(x) = t_2 \) is the equation of the shape at time equal to \( t_2 \), etc. (For the moment, let us restrict our attention to a region and time interval where \( v \) is always positive or always negative so that \( \tau \) is a uniquely defined function of \( x \) there.) By regarding \( \tau \) as a function of \( x \), a time sequence of crystal forms, such as is shown in figure 1, can be considered as a contour map of the function \( \tau \). If \( v \) is positive, then the gradient of \( \tau \) at any point \( x \) where that gradient is defined is a vector whose direction is that of the oriented surface normal \( n \) (since the surfaces are level surfaces in \( \tau \)) and whose magnitude is inversely proportional to \( v(n) \). Thus \( \nabla \tau / |\nabla \tau| = n \) and \( v(n) = 1 / |\nabla \tau| \), i.e.

\[
|\nabla \tau| \cdot v(\nabla \tau) = 1.
\]

It is convenient, especially for the PDE which follows, to extend \( v \) to a function on all vectors by defining \( v(rn) = rv(n) \) for all \( r > 0 \). When we do this, the above statement becomes

\[
v(\nabla \tau) = 1.
\]

(Note that \( v \) has become dimensionless by this extension.) If \( v \) is negative, then \( \nabla \tau \) is in the direction of \(-n\) and its magnitude is inversely proportional to \(-v(n)\), so that

\[
-v(-\nabla \tau) = -|\nabla \tau| \cdot v(n) = 1.
\]

If we let \( v' = -v(-n) \) when \( v \) is negative, then the above statement is just

\[
v'(\nabla \tau) = 1.
\]

The mathematical equivalence of growth and dissolution problems should now be completely clear; in order to treat the negative-velocity case, one in fact switches to the equivalent positive-velocity problem.

**Figure 1** A sequence of crystal forms can be considered a contour map in which each contour is the surface at a constant time. The gradient of the time when the surface reaches a point \( x \) is \( n/\nabla \tau \).

---

**Evolving Crystal Forms**

Thus the statement that the normal velocity is given by \( v(n) \) becomes equivalent when \( v > 0 \) to the statement

\[
v(\nabla \tau) - 1 = 0
\]

and when \( v < 0 \) to

\[
v'(\nabla \tau) - 1 = 0.
\]

This equation may look much more cryptic than the equivalent statement in words, but it clearly presents the problem as a first-order partial differential equation of the form

\[
F(x, \tau, \nabla \tau) = 0.
\]

In fact, one sees that if \( v \) depends on the position \( x \) of the point on the surface and the arrival time \( \tau \) as well as on the normal direction \( n \) at that point, the resulting equation is

\[
F(x, \tau, \nabla \tau) = v(x, \tau, \nabla \tau) - 1 = 0
\]

(or the same with \( v' \), if \( v \) is negative), which is again just a first-order PDE.

If the normal velocity is not always of one sign, one has to let \( \tau(x) \) be the set of arrival times at which the crystal surface passes through \( x \). Sometimes (e.g. if this set of arrival times at \( x \) contains no intervals or accumulation points) \( \tau \) can still be regarded locally as a function of \( x \) if time is restricted to an appropriate interval. In any case, one always has to be careful about and treat separately the places where \( v = 0 \) and \( \nabla \tau \) is undefined.

The theory of first-order PDE (John 1975) asserts (under conditions of smoothness of the given function \( F(x, \tau, p) \) in all its variables and of the initial data to be specified shortly) that the initial value problem can be solved by the use of curves called characteristics. These curves, parametrised by some variable \( t \) and thus written as \( x(t) \) (and with \( \tau = \tau(t) \) and \( \nabla \tau = \nabla \tau(t) \) at \( x(t) \)), emanate from each point in the initial surface and are determined by the following set of seven ordinary differential equations:

1. \( dx_i/dt = \partial F/\partial p_i \) for \( i = 1, 2, 3 \) (i.e. \( dx_i/dt = \nabla_p F(x, \tau, p) \))
2. \( d\tau/dt = \sum_i p_i \partial F/\partial p_i \) (i.e. \( d\tau/dt = p_i \nabla_p F(x, \tau, p) \))
3. \( dp_i/dt = -\partial F/\partial x_i - p_i \partial F/\partial \tau \) for each \( i \)
   (i.e. \( dp_i/dt = -\nabla_p F \cdot \nabla \tau \)).

Because \( p \cdot \nabla_p F = p \cdot \nabla_p v = v(p) = 1 \) (as is discussed further in the appendix), (2) above will always reduce to \( d\tau/dt = 1 \) for us, and so the parameter \( t \) along the characteristics can be always (for this crystal growth problem) be taken to be the arrival time \( \tau \).

Note that since only derivatives of \( F \) appear in these equations, we can replace \( F \) by \( v \) (or \( v' \), if \( v \) is negative) in the above equations.
If \( v \) does not depend explicitly on \( x \) then (3) becomes

\[
(3') \quad \frac{d \ln(p_t)}{dt} = \frac{d \mathcal{V}}{d \mathcal{S}} \text{ (unless } p_t \text{ is initially 0, in which case } p_t \text{ stays 0 for all time on that characteristic) for each } \mathcal{S}.
\]

Since \( \ln(p_t) \) has the same derivative for each \( \mathcal{S} \) (or \( p_t = 0 \)), we conclude that the ratios of the \( p_t \) are constant along characteristics, and thus that the normal \( n = p_t / |p_t| \) to the crystal stays constant along characteristics (though \( p_t \) may change in magnitude). Thus if this no-spatial-dependence case, a characteristic is the trajectory of a point with a given normal as the crystal grows. (Such is not the case when \( v \) depends on \( x \), as the temperature-gradient example below will demonstrate.)

If \( v \) does not explicitly depend on either the arrival time \( \tau \) or spatial position \( x \), then (3) becomes

\[
(3^a) \quad \frac{dx}{dt} = \mathcal{V} \mathcal{S}, \quad \text{which says that } p = \text{as well as } n = p / |p| \text{ is constant along characteristics, and (1) becomes }
\]

\[
(1^a) \quad \frac{dx}{dt} = \mathcal{V} \mathcal{S}, \quad \text{which says that characteristics are straight lines of the form}
\]

\[ x = x_0 + \mathcal{V} \mathcal{S} (n_0) \]

where \( x(t) \) is on the surface of the crystal at time \( t \). (Here \( \mathcal{V} \mathcal{S} (n_0) \) means that we are evaluating the gradient at \( n_0 \), not that we are taking any kind of surface gradient. As discussed in the appendix, \( \mathcal{V} \mathcal{S} \) is constant in radial directions.)

The crystal shape at time \( t \) is the locus of all the points \( x(t) \) on all the characteristics from initial points \( x_0 \). When \( v \) depends only on \( n \) and not on \( x \) or \( t \), if we follow an element of surface of a given orientation \( n \), it will travel with constant velocity, here given by \( \mathcal{V} \mathcal{S} (n) \). Since the plot of \( \mathcal{V} \mathcal{S} (n) \) is the level set \( v(p) = 1 \) for the function \( v(p) \), and since \( v \cdot \mathcal{V} \mathcal{S} (p) = v(p) \), this is equivalent to Frank's result (1958), which said that the element moves in the direction \( n' \) of the normal to Frank's polar plot of the slowness vectors \( n / v(n) \) and with a speed given by \( v(n)(n \cdot n') \).

2.1 A simple two-dimensional example

We illustrate the use of characteristics in figure 2 by taking the two-dimensional example of an initial circle and growth velocities \( v \) which are of the form

\[ v_{a, b}(n) = a + b(n \cdot (1, 0))^2 = a + bn_1^2. \]

For \( a > 0 \) and \( b > -a \), \( v \) will be positive; \( b = 0 \) is the isotropic case, and the case \( b = -a \) will be dealt with below under discontinuous coarsening. Multiplying by \( |p| \) to create \( v(p) \), we obtain

\[
v(p) = a |p| + (b |p|)(p \cdot (1, 0))^2 = a |p| + (b |p|)(p_1)^2.
\]

Setting \( a = 1 \), polar plots of \( v \) are circles for \( b = 0 \) which pinch towards the centre when \( b \) is negative as shown in figure 2. Taking the gradient, we obtain

\[
\nabla v = (an_1 + bn_1(2 - n_1^2), an_2 - bn_2n_1^2).
\]

The plot of \( \nabla v \) is a circle for \( b = 0 \), but becomes distorted when \( b \) is non-zero. For \( b < -\frac{1}{2} \) or \( b > a \) it becomes self-intersecting and develops two swallowtails, as shown in figure 2. For \( b > a \), the second component of \( \nabla v \) changes sign at \( n_1 = \pm (a/b)^{1/2} \) and \( \pm 1 \); there are six crossings of the \( x_1 \) axis. Similarly, for \( -a < b < -\frac{1}{2} \), there are six crossings of the \( x_2 \) axis. The tips of the swallowtails are spindles; the curves turn back on themselves without a change of slope.

![Image](image.png)

**Figure 2** Example of the use of characteristics when the growth rate \( v \) is of the form \( v_{a, b}(n) = a + bn_1^2(= a + b \cos^2 \theta) \) for various ratios of \( b/a \). (a) Polar diagrams of \( v(n) \). (b) Diagrams of \( \nabla v(n) \). (c) and (d) Evolving crystal forms from an initial circle for \( b/a = -\frac{3}{4} \) and \( -\frac{1}{2} \).
For \(-\frac{1}{2}a < b < a\), \(V\nu\) is smooth and does not intersect itself.

The solutions are shown for times \(t > 0\) by connecting the equal-time points on the characteristics. For the curve with \(b/a = -\frac{1}{2}\), the outward characteristics diverge and the solution is valid for all times. However, for the curve for \(b/a = -\frac{3}{2}\), the characteristics converge and cross. Continuing along the characteristics leads to a swallowtail shape as a perfectly valid mathematical solution to the PDE. However, the physics of crystal growth does not allow this; a point \(x\) enters the crystal only once, unless \(v\) changes sign and the crystal melts back through \(x\). This suggests that characteristics that cross must be terminated, in a way that will be discussed next under shocks.

Note that if we assume that the \(v\) of this example is for inward growth from the circle, characteristics that diverged for outward growth converge for inward growth, and vice versa. The orientations on the circle that eventually disappear on outward growth by having their characteristics terminate in a shock do not disappear in a similar shock on inward growth (though they may disappear due to their characteristics intersecting with an entirely different set of characteristics).

2.2 The question of regularity: shocks and fans

An important theorem that holds with minor modification in any dimension is clearly and simply presented in two dimensions by John (1975): If \(F\) has continuous second derivatives with respect to its variables \(x, \tau, p\), and if along an initial curve \(x(s) = x_0(s), 0 \leq s \leq 1\), the value \(\tau(x) = \tau_0\) is assigned, with \(x_0\) having continuous second derivatives, and if one also has \(p_0(s)\) which is continuously differentiable and satisfies \(F(x_0(s), \tau_0, p_0(s)) = 0\) and \(p(s) \cdot dx/ds = 0\) for each \(s\), and if \(dx/ds\) and \(\zeta_p F\) are never zero and never parallel, then in some neighborhood of the initial curve there will exist one and only one solution \(\tau(x)\) of \(F(x, \tau, p) = 0\) with \(\tau(x_0(s)) = \tau_0\) and \(\zeta_\tau \tau_0 = p_0(s)\) for all \(0 \leq s \leq 1\).

This solution is the one constructed by the method of characteristics; it extends for precisely the open time interval such that characteristics do not cross. An examination of the proof shows that the condition \(\zeta_\tau F \neq 0\) is used only in the proof that the characteristics together form a surface with no gaps or overlaps for short times; one can easily see that the motion of the surface remains solvable by characteristics even if there are points or intervals along which \(\zeta_\tau F = 0\), if we leave such points stationary. The requirement that \(dx/ds \neq 0\) together with the requirement that \(x_0\) be twice continuously differentiable is the requirement that the curve itself be smooth. The condition that \(dx/ds\) and \(\zeta_p F\) are never parallel is automatic from the fact that the normal component of \(\zeta_p F\) has magnitude \(1/|p| = v(\alpha)\). Remembering that \(\tau(x)\) is the arrival time of the crystal at the point \(x\) (i.e., the time at which the point \(x\) is in the interface), from this one can prove in our case that if \(v\) has continuous second derivatives and

if the entire initial crystal shape has a normal vector field which is continuous and piecewise-continuously differentiable, then the evolving crystal shape is uniquely defined for some open interval of time containing the initial time and each shape in this continuum will also have a continuous and piecewise-continuously differentiable normal vector field. Entirely analogous results hold in three dimensions, both in terms of the theorem (John 1975) and its application to crystal growth.

We thus see that there are several ways in which this theorem can fail to cover cases we would like to consider. First is that we want to consider growth velocities \(v\) that may be continuous with respect to normal direction but not have a continuous gradient. (As we will see, such discontinuities in the gradient of \(v\) do occur at facet orientations of \(W_{\alpha}\).) The second is that we want crystal growth to be determined for all positive time (including when characteristics cross) and for shapes with corners and edges. These issues are resolved by the use of shocks and fans, provided we assume that both our initial surface and \(v\) are continuous everywhere and piecewise \(C^2\) (i.e., first and second derivatives are piecewise continuous). Shocks must be used whenever characteristics cross. A shock is a discontinuity along the surface, in this case in \(V\nu\) and in \(n\) as well, and thus gives rise to a corner or an edge on the crystal surface. The jump condition at a shock is determined from the physics of the problem, not the PDE; here the physics requires that the crystal surface (as determined by the points \(x(t)\) the characteristics reach at time \(t\)) be continuous, and (provided \(v\) is positive) that once the crystal has grown past a point, that point should remain part of the crystal. Thus a shock occurs when two characteristics arrive at the same point at the same time. Differentially, our shock condition at a point \(x\) along a shock between normals \(n_\alpha\) and \(n_\beta\) is, in two dimensions, that

\[
\frac{dx}{dt} = s_\alpha t_\alpha + \nu v(p_\alpha) = s_\beta t_\beta + \nu v(p_\beta)
\]

where \(t_\alpha\) is a unit tangent vector to the \(\alpha\) surface and is thus perpendicular to \(p_\alpha\), \(t_\beta\) is defined similarly for the \(\beta\) surface, and \(s_\alpha\) and \(s_\beta\) are numbers determined by solving the vector equation (which is a system of two linear equations in those two variables). In the special case where each gradient lies in the other plane, this becomes

\[
\frac{dx}{dt} = \nu v(p_\beta) + \nu v(p_\alpha).
\]

In three dimensions, at a point in an edge of a growing crystal we have three naturally defined tangent vectors, one along that edge and the other two perpendicular to it and into the two surfaces meeting along that edge. The equation for \(dx/dt\) is similar to the above, except there are two \(s\) for the \(\alpha\) side and two for the \(\beta\) side. The three-dimensional vector equation thus yields three linear equations in four unknowns, but the extra degree of freedom corresponds to the fact that one has a shock surface rather than
a shock line emanating from an edge, and the degree of freedom is in the
direction of the tangent line to the edge. At a corner where three surfaces
meet, there are six \( s \) and two three-dimensional vector equations, yielding
six equations in six unknowns, and thus a differential equation of the
propagation of the corner. If a corner with more than three surfaces is to
propagate, then special relations must exist among the \( v(n) \) in order to
enable a solution to the equations to exist; this can happen in crystals
because of their symmetry. (In order for such a shock to form with more
than three surfaces related by symmetry, there must be appropriate
symmetry in the initial crystal shape as well.)

Shocks in particular appear when a non-convex crystal grows so that two
different portions of surfaces come into contact. At that instant, the shock
starts from the contact point(s) and spreads out all around them.

If contact occurs on a whole piece of surface at a given time, then all the
characteristics going into that piece of surface from both sides terminate.
(It is convenient to put the crystal surface mathematically into the region
which grows with positive velocity when it is called the ‘crystal’, so that
if contact occurs on a whole piece of surface, that portion automatically
disappears into the interior of the crystal.) Similarly, all remaining character-
istics collide and terminate at the instant that a hole in a crystal (or a
dissolving crystal) disappears.

Fans of characteristics are used at points where characteristics are not
uniquely defined. One way that this can happen is that there is a point on
the growing crystal with normal \( n_0 \) and \( v \) has a discontinuous gradient at
\( n_0 \). Then in place of \( \mathcal{V} \) one uses all the convex combinations of the limits
of \( \mathcal{V}(n) \) as the normals \( n \) approach \( n_0 \). An example where this happens is
\( \mathcal{V}_{	ext{cube}}(n) = |n_1| + |n_2| + |n_3| \); see the appendix. This results in a facet
developing there (or propagating, if that point is already in a facet): note
that if one has a facet with a direction \( n_0 \) where \( \mathcal{V} \) is discontinuous, then
there are fans of characteristics emitted from each point, and fans from
neighbouring points in the facet cross each other.

It is as if there is a whole continuum of shocks. However, they all in fact
give the same two items of information: the facet is moving forward at
velocity \( v(n_0) \), and is spreading out no further than the rate allowed by
the outer characteristics of the fans (and perhaps less, if there are shocks
at the edges of the facet).

Another way that characteristics can be non-uniquely defined is at a
corner or edge of the crystal surface. Here one puts in characteristics
corresponding to all normals in the convex hull of the normals around the
edge or corner (using the extended definition above if \( \mathcal{V} \) is undefined for
some of those normals). Note that it is quite possible that \( \mathcal{V} \) is the same
on all those normals (for example, this can happen with \( v = \mathcal{V}_{	ext{cube}} \)), so that
the fan from all the limiting normals around the edge or corner reduces to
one uniquely defined characteristic after all. It is also possible that none of

those characteristics overlap; in that case, the corner becomes rounded
instantly as growth proceeds. Finally, some of the members of the fan of
characteristics may overlap other members; this can become quite complex,
and the above shock condition of continuity of the surface is often
not adequate to determine the way the corner grows because there may be
more than one solution which produces a continuous crystal surface. The
appropriate condition seems to be that the crystal should grow as fast as
possible, subject to the condition that the crystal surface, as determined
by the points the characteristics reach at time \( t \), be continuous. We believe
that this is the correct shock condition because we believe that this is also
the condition for such a solution to be stable under perturbations (in the sense
that a perturbation to such a solution will not grow without bound).

That is, look at the tangent cone to the initial surface \( C_0 \) at a point \( x_0 \)
which is on an edge or in a corner of \( C_0 \); this cone consists of rays through
the origin and is

\[
\{ y : \exists s_k \perp 0 \text{ and } y_k \rightarrow y \text{ with } x_0 + s_k y_k \text{ in } C_0 \}.
\]

If \( C_0 \) is piecewise \( C^2 \), then this tangent cone is composed of planar seg-
ments. For each \( y \) in the tangent cone such that there is a normal \( n_y \) to
the tangent cone at \( y \), plot \( y + t v(n_y) \) for \( t = 1 \). At each point \( y \) in an edge or
corner of the tangent surface, plot \( y + t v(n) \) for \( t = 1 \) and all \( n \) omitted
at that edge or corner. (Note the similarity of Huygens’ wavefront con-
struction (Morse and Feshbach 1953); see below for another analogy of this
crystal-growth problem to wavefront propagation.) The resulting surface
will be continuous, because of the convention regarding discontinuities in
\( \mathcal{V}(n) \), but will possibly be self-intersecting. If along the edges or at the
corner it has portions which stick out beyond the rest of the surface, in the
sense that the boundary of the solid formed by using \( 0 \leq t \leq 1 \) rather than
\( t = 1 \) has portions that are not in the surface obtained with \( t = 1 \) or the
initial surface, then the corresponding \( \mathcal{V}(n) \) must not be used to construct
a characteristic at \( x_0 \) in the original problem. This continuity condition
means that any ‘ears’ which stick out beyond the resulting planar surfaces
will not be able to be used for convex corners, but portions of them will
be able to be used for concave corners; for saddle-shaped corners with both
concave and convex edges, it is quite possible that some parts of the ‘ears’
will be used but others will not. This construction determines which values
of \( \mathcal{V} \) are to be used in constructing characteristics on the original surface.
It will coincide with choosing among possible growth paths the one that is
stable, because it inherently includes all perturbations of the corner by
including all values of \( \mathcal{V} \) at that corner (values of \( \mathcal{V} \) where \( n \) is not a direc-
tion omitted at the corner or edge cannot contribute to growth—they are
automatically behind the surfaces obtained from the plane segments of the
tangent cone).

Solutions that are not piecewise \( C^2 \) but have curvatures approaching
infinity in some of their pieces, e.g. with portions of circular cones, do 
arise, either as part of C₀ or as growth features. One example occurs 
in figure 2(d), if we assume that these solutions are for a problem of 
cylindrical symmetry; the tangent cone at the shock for \( h/a = -\frac{1}{2} \) is then a 
circular cone that is not piecewise \( C^2 \). Another example occurs at the 
termini of a shock surface, places on the surface where the sharp crest of 
an edge becomes rounded. Such singular points are part of swallowtail 
development in \( R^1 \). If \( C₀ \) is not piecewise \( C^2 \), then the decision of which 
characteristics to use is more complicated, but is still based on fastest 
growth subject to continuity.

We do not know of any other physical model in which a shock condition 
is imposed that necessarily (at time \( t = 0 \)) involves both continuity and 
shortest time, even though characteristics have long been used. But the 
right way mathematically to describe the solution is probably as the viscosity 
solution of the PDE (Chen et al 1989), rather than through specifying 
the details of the shocks and fans.

3 EXAMPLES

3.1 Some examples of growth of finite seed crystals

The isotropic case, \( v(n) = \text{constant} \), arbitrarily set to 1, is intuitively simple 
and is useful for introducing shocks and fans. It leads to a \( v(p) = |p| \), and 
thus to a \( \nabla v(n) = n \), which is constant in magnitude and always parallel to 
\( n \). The limiting outward growth form \( W_∞ \) is a unit sphere; any starting 
finite shape growing outward will eventually tend to a sphere. For example, 
the faces of a starting cube will grow by translation without an area 
increase, the fans on edges and corners will round the edges into growing 
cylinders, and the eight corners will grow octants of spheres. After a large 
enough elapsed time the corner spherical pieces will dominate, so that it 
will look like a sphere, but the eight spherical pieces will have different 
centres (the corners of the original cube), and there will be cylindrical strips 
separating the eight octants and six little flat facets (of the same size as the 
faces of the original cube) at the ends of the cylindrical strips.

An initial crystal consisting of two balls of different radii \( r₁ \) and \( r₂ \) joined 
by a thin cylinder of radius \( r₀ \) about the axis joining their centres will grow 
as balls of radius \( r₁ + t \) and \( r₂ + t \) joined by a cylinder of radius \( r₀ + t \) until 
the spheres touch; the spheres will meet the cylinder along circles in a sharp 
edge. The shock surfaces will early on consist of the union of all these 
circles, forming surfaces which are surfaces of revolution of parabolas. 
When the balls touch, the cylinder is entirely eliminated and the crystal 
continues to grow as intersecting balls, the shock surfaces having collided 
and formed a new shock surface which happens to be a rotated hyperbola.

The shock will get weaker and weaker (i.e. angle between the normals at 
the discontinuity will approach zero) but will never entirely disappear.

For \( v = v_{\text{cube}} = |n₁| + |n₂| + |n₃| \), starting from any bounded seed the 
end form will be a cube (since \( W_∞ \) is the cube with corners \((±1, ±1, ±1)\)), 
but it is instructive to examine how an initial sphere will evolve in time. 
Because \( \nabla v_{\text{cube}} \) is constant (it is \((±1, ±1, ±1)\)) in each of the eight open 
octants, the eight spherical segments in each of the open octants translate 
without distortion or expansion in these directions to become rounded corners of the cube; the fans emanating from the six points along the 
\( \langle 100 \rangle \) directions create square facets expanding in size, and fans from 
the circular arcs at the octant boundaries generate pieces of circular cylinders, 
not growing in arc length but lengthening along their axis to become 
curved edges of the cube (see figure 3). Ultimately the facets dominate and 
the form is a growing cube with rounded edges and corners whose radii are 
those of the starting sphere. If \( v \) were any other normal velocity function 
with the same \( W_∞ \), but containing ears in the plot of \( v \), then the crystal 
might grow faster than the cylindrical and spherical portions, possibly 
forming corners or edges there due to shocks, but it would never grow 
beyond the growing cube initially surrounding it. An example of the effect of 
'ears', which are also called 'swallowtails', in growth was shown in 
figure 2 for \( v_{\text{cube}} \).

![Figure 3](image_url)

Figure 3 Evolving crystal forms for \( v = v_{\text{square}} = |n₁| + |n₂| \), starting from a 
circle and from a rounded square. In each of the four open quadrants, \( \nabla v_{\text{square}} \) is 
constant and the characteristics are parallel. The four circular segments in each of 
the open quadrants translate without distortion upon expansion to become rounded 
corners of the square. (a) If \( C₀ \) is a circle, fans emanating from the \( \langle 10 \rangle \) directions 
create facets. (b) If \( C₀ \) is a rounded square, fans emanating from the facets overlap 
and lead to translation and expansion of the facets.
3.2 Examples of shapes of inward growth from a finite form

Simply because crystal growth in this model remains well defined until the surface entirely disappears, there is a limit to the shape of any particular hole in a crystal as the hole volume goes to zero, or to the shape of any crystal as it dissolves, given \( v \) (we orient the surface inward in either case so that we can use a positive velocity \( v \)). This shape is just the set of limit points of the surface as time approaches the instant of disappearance—the points on the characteristics at the final moment of collision. One can also often characterise the shape of a piece just before it disappears. But these shapes are not independent of the initial shape. A non-concave starting shape will often become disconnected on shrinking, with separate pieces having different limiting shapes as they disappear, usually at different times. Also, dissolving forms need not be more rounded than growth forms, as simple examples will show. It is furthermore the case that the actual value of \( v \) on all directions matters a great deal, rather than just the values on the directions that appear in \( W_\infty \). We illustrate some of what can happen by example.

Again let us consider the case where \( v(n) = 1 \), but we are now letting the ‘crystal’ be on the outside of some finite region, so that the exterior normals point into the finite region. If the initial shape \( C_0 \) is exactly spherical, the shape of the hole will stay spherical. But a shape close to that of a sphere deviates increasingly from spherical, developing shocks. A starting parallelepiped remain a parallelepiped with the same angles; the shocks from the intersecting characteristics from the faces will maintain the corners and edges, and, because they grow too slowly, the inward-directed fans from the corners and edges will be eliminated. When the shortest altitude has shrunk to zero, the limiting form will be a two-dimensional parallelogram. An initial shape \( C_0 \) that is a convex polyhedron remains a convex polyhedron, with surviving faces parallel to the faces of \( C_0 \) as it decreases in size. Shocks due to intersecting characteristics from the faces will maintain the corners and edges; nothing will be contributed by the inward-directed fans from the corners and edges. Intersecting shocks will disappear into a new shock, causing corners and edges to combine into new corners and edges; some faces and edges may disappear. A cube or a rhomboid remains a cube or a rhomboid, but a rectangular solid that is not a cube becomes increasingly elongated; the end form is a rectangular plate unless the two shorter sides were initially of the same length, in which case the end form is a line segment. These simple counterexamples disprove the general statement that dissolving forms are more rounded than growth forms.

In general, when one starts with an initial shape \( C_0 \) exactly similar to the central inversion of \( W_\infty \), and when \( v(n) \) has the smallest values it can to have that \( W_\infty \) (which is equivalent to saying that \( v(p) \) is a convex function, or that the plot of the function \( n/v(n) \) is the boundary of a convex region), then the shape does not change as it shrinks, and the end form is the same as that of the initial shape. Track the points inward along the \( \nabla v \) vector(s) corresponding to the normal at the point and the shape stays the same; the omitted directions move exactly at the speed so as to stay omitted at the corners. But it is not just the initial shape that matters. If \( v \) is a little bigger on the omitted directions and the starting shape is the central inversion of \( W_\infty \), then some of those directions will appear, both immediately and in the limiting shape, thereby blunting the edges and/or corners and possibly even entirely consuming all the directions occurring in the original hole. But even when the starting shape is the central inversion of \( W_\infty \), it is not always the case that ‘faces with maximal growth rates should form the concave growth shape’ (since the growth rates on omitted normals of \( W_\infty \) are usually larger than on the directions which occur). Rather new directions can show up here if their growth rates are greater than that minimally allowed by \( W_\infty \).

Of course, things can be quite different if one has different starting shapes. If the velocity function \( v \) is \( v_{\text{cube}} \) as defined above, any parallelepiped oriented so its normals point inward (whether its faces have special orientations—those where \( \nabla v \) is discontinuous—or not) will have each of its faces translate inward with constant speed with new no new orientations appearing at the corners and edges. The fans of characteristics introduced at some of the edges and corners may not give rise to shocks at all (when \( \nabla v \) is the same on all adjoining faces of the parallelepiped), whereas the fans at the other edges and corners predict a motion slower than the shock between the adjoining faces and are eliminated. The limiting form for a starting parallelepiped is thus a line segment or a 2D rhombus, unless the initial shape is a rhomboid (in which case it is a point approached via the same rhomboid). An initial sphere dissolving with this same velocity function \( v(n) \) will be composed of eight segments of spheres of constant radius but decreasing extent, translating inward along the \( \{111\} \) directions at constant speed and getting chopped off by shocks which form along the octant boundaries. The limiting form as this originally spherical particle disappears into a point is an octahedron.

The only rule is that if the characteristics corresponding to a given direction are excluded by crossing characteristics at some time, then they disappear, whether one has ‘growth’ or ‘dissolution’. Characteristics with directions corresponding to ‘ears’ are not necessarily excluded at concave corners (as they are at convex corners) and so regions with these normals can grow in a concave shape if the corners are not too sharp. Only in this sense can corners be said to round on dissolution. These new orientations, if they form and grow, can crowd out slower growth directions if the characteristics from another corner cross them. But a simple counterexample of an initial shape that is very flattened shows that there is no general theorem here—the end-form can consist primarily of slow-growth
directions. These simple examples illustrate some of the simpler anisotropic growth results. They demonstrate that while the limiting forms from outward growth are independent of initial form, the forms from inward growth always reflect the initial form. They also illustrate how shocks and fans operate in easily visualised cases.

3.3 Additional growth rules
An edge or a corner in a growing crystal is the result of intersecting or overlapping characteristics. If the corner or edge is convex and identical to one in \( W_\infty \), then it will stay the same shape, whether the plot of \( \nabla v \) has ‘ears’ or not. If the corner or edge is convex and sharper than the corresponding region of \( W_\infty \) (so that it omits normals that occur in \( W_\infty \)) then a fan of characteristics will emerge from the corner, and the corner will immediately round out or, if \( W_\infty \) also omits some of the normals omitted by the corner, it will partially round out and thereby become less sharp. If the corner or edge is convex and \( W_\infty \) omits all of the normals that the corners omits plus more, then the corner will stay the same shape only if the surfaces meeting at the corner are planes or if the plot of \( \nabla v \) has no ‘ears’ at that corner; if the surfaces meeting at the corner have any positive curvature and if \( v(n) \) is larger than required by \( W_\infty \) for the directions present in the corner of the crystal but in \( W_\infty \), then these directions will ‘grow out’, their characteristics will be cut off by shocks, and the corner will sharpen.

Concave corners and edges will stay the same shape unless the plot of \( x_0 + \nabla v(n) \) for the directions \( n \) omitted at the edge or corner at \( x_0 \) extends past the plot of \( x + \nabla v(n(x)) \) for points \( x \) near \( x_0 \). In particular, sharp enough concave corners and edges always remain the same shape. Less sharp concave corners and edges can become more rounded if the omitted directions grow fast enough, or can become more sharp if the surfaces which meet at the corner or edge are not planar but also concave.

Saddle-shaped regions can easily become more sharply curved in some tangent directions, even developing edges and corners where none were previously, and can simultaneously become less sharply curved in others.

It should be noted that if in an initial convex polyhedron there is a planar facet with a direction in a concave part of an ‘ear’ of the plot of \( \nabla v \), then not only will that direction disappear by growing out, but while it is growing out, the adjacent facets will usually partially round out at their edges with this facet.

3.4 Example of extension to case of constant temperature gradient
Suppose that the temperature is \( T_m - bx \), where \( T_m \) is the melting temperature and \( b > 0 \), so that growth occurs for a seed crystal in the right half-space. Suppose \( u(x, r, n) = bx_1 u(n) \); as before, extend \( u \) and \( v \) to be defined on all vectors \( p \) rather than just unit vectors \( n \). We will give solutions in the two-dimensional isotropic case of \( u(n) = 1 \) for all \( n \).

As before, we define \( u'(x, r, p) = v(x, r, p) \) where \( u > 0 \) and \( u'(x, r, p) = -v(x, r, p) \) where \( u < 0 \), and then the PDE is \( u' - 1 = 0 \). In our example, \( u'(x, r, p) = b|x_1||p| \) and so the PDE is

\[
F(x, r, p) = b|x_1||p| - 1 = 0.
\]

If we let \( s = \text{sign}(v) = \text{sign}(x_1) \) and \( b' = sb \), then the PDE is equivalent to

\[
b'|x_1||p| = 1.
\]

Since \( p \) is \( \nabla T \) in the PDE, and \( \nabla T = \text{sign}(v) \nabla T \) \( n \) for \( n \) the normal to the surface, we have \( p = s||p||n = n\langle bx \rangle \). If an initial point \( x_0 = (x_{10}, x_{20}) \) has normal \( n_0 \), then \( p_0 = (p_{10}, p_{20}) = (n_{10}bx_{10}, n_{20}bx_{10}) \).

Since \( \delta p/|p| \delta p_1/|p| = sn_0 \), the equations for the characteristics become

\[
\begin{align*}
\frac{dx_1}{dt} &= b'|x_1||p| = (bx_1n_1) \\
\frac{dx_2}{dt} &= b'|x_2||p| = (bx_1n_2) \\
\frac{dp_1}{dt} &= -sb'|p| = -b'(p_1^2 + p_2^2)^{1/2} \\
\frac{dp_2}{dt} &= 0
\end{align*}
\]

so along characteristics, \( p_2 \) is constant. If \( p_{20} \) is non-zero, the equation for \( p_1 \) integrates to

\[
p_1 = -|p_{20}| \sinh(b't - a)
\]

where \( a \) depends on \( n_0 \) and \( x_0 \) and satisfies \( \sinh a = p_{10}/|p_{20}| = sn_{10}/|n_{20}| \) (and thus \( \cosh a = 1/|n_{20}| \)). If \( p_{20} = 0 \), then

\[
p_1 = p_{10} \exp(-\text{sign}(p_{10})b't).
\]

Using these formulae for \( p_1 \), the equation for \( x_1 \) becomes separable, but it is easier to use the original PDE and the solution for \( p \) to get, for \( p_{20} \neq 0 \),

\[
x_1(t) = 1/|b'| |p| = 1/|b' |p_{20}| \cosh(b't - a)
\]

\[
= x_{10}|p_0|/|p_{20}| \cosh(b't - a) = x_{10}/|n_{20}| \cosh(b't - a)
\]

and, for \( p_{20} = 0 \),

\[
x_1(t) = x_{10} \exp(\text{sign}(p_{10})b't) = x_{10} \exp(\text{sign}(n_{10})b't).
\]

The differential equation for \( x_2 \) is thus

\[
\frac{dx_2}{dt} = b'|x_2||p| = |x_{10}|b'| |n_{20}\cosh^2(bt - a)|
\]

so

\[
x_2(t) = (1/b'p_{20})\tanh(b't - a) + x_{20} + x_{10}p_{10}/p_{20}
\]

\[
= (|x_{10}|/|n_{20}|)\tanh(b't - a) + x_{20} + x_{10}n_{10}/n_{20}.
\]
Using the hyperbolic function identities (or integrating $dx_3/dx_1 = p_3/p_1$) one obtains
\[ x_1^2(t) + (x_2(t) - C)^2 = x_{10}/n_{20}^2 \]
the equation of a circle with centre $(0, C)$ and radius $\sqrt{x_{10}/n_{20}}$, where $C$ is $x_{20} + x_{10}n_{10}/n_{20}$. Thus all characteristics are arcs of circles.

If the initial crystal is entirely in the region $x_1 > 0$, then the size of the crystal grows exponentially, and it approaches the $x_2$ axis very slowly. If the crystal is entirely in the region $x_1 < 0$, it would disappear in finite time. If it straddles the $x_2$ axis, then part will grow and part will shrink, but it will always straddle the $x_2$ axis. For example, if the initial crystal is the unit circle centred at the origin, then the characteristics are portions of non-intersecting circles which fill the plane; trajectories in the left half of the plane are inward along the circles towards the portion of the $x_2$ axis within the initial crystal (corresponding to that part of the crystal shrinking), whereas in the right half of the plane they are outward along the circles and back to the portions of the $x_2$ axis which lie outside the original crystal. The points $(0, 1)$ and $(0, -1)$, where $v = 0$, are stationary. Furthermore, one can show that for each point $(x_1(t), x_2(t))$ on a characteristic from any initial point $(x_{10}, x_{20})$ on that unit circle,
\[ (x_1(t) - \sinh(bt))^2 + x_2^2(t) = \cosh^2(bt) \]
so that, at time $t$, the surface of the growing crystal is itself a circle, one with centre $(\sinh(bt), 0)$ and radius $\cosh(bt)$.

Since crystals start out as circles remain as circles, one might ponder whether this spatially dependent velocity were equivalent to a velocity $v'$ of the form $v'(t, n) = a(t) + b(t) \cdot n$. Indeed, if one uses
\[ v'(t, n) = b \sinh(bt) + n \cdot (b \cosh(bt), 0) \]
then an initial circle of radius 1 centred at the origin does grow as circles with centres $(\sinh(bt), 0)$ and radii $\cosh(bt)$, and one checks that indeed at any point $(x_1(t), x_2(t)), v'(t, n) = v(n) = bx_1(t)$. Note that the crystal shapes—the level sets of the arrival time function—are the same for the two velocities, but that the characteristics for the two different PDEs are not. Also, $v'$ reproduces the same growth as $v$ only for initial circles centred at the origin; for an initial circle centred at $(a, 0)$ of radius $r$, one would have to use the different velocity
\[ v^* = ab \cosh(bt) + rb \sinh(bt) + n \cdot (ab \sinh(bt) + rb \cosh(bt), 0) \]
as the velocity to reproduce the growth given by $v$ (circles with centres $(acosh(bt) + r \sinh(bt), 0)$ and radii $a \sinh(bt) + r \cosh(bt)$), and for initial forms that are not circles, in general no growth law depending only on time will reproduce the same growth as $v(x, n) = bx_1$.

The idea of ‘limiting shape’ is also much looser here. If one uses barrier arguments (as will be proven in our subsequent paper), putting a circle inside and outside an arbitrary shape and growing each, the circles do not get closer together when translated back to the origin and rescaled, but rather their radii remain in the same constant ratio as they had initially, if they initially had the same centre. Thus one cannot say that the limiting form is a circle, for an arbitrary initial finite seed with at least part of it in the right half-plane.

### 3.5 Example of $v(n)$ determined by anisotropic elasticity

When stresses develop in multicomponent crystals because of diffusion from the surrounding medium into the crystals of some component that changes the lattice parameters, the local solubility of the crystals in the surrounding medium changes in a way that depends primarily on the local stress at the surface. As a result, the crystals can grow or dissolve in response to these stresses. Because the elastic stresses depend on anisotropic elastic moduli, and the details of the stress patterns depend on the normal to the surface $n$ but not on the details of the diffusion, elasticity theory can be used to predict a growth velocity $v$ depending only on the normal direction $n$.

This situation has been observed in a variety of settings, usually involving two crystals, one dissolving, the other growing (Handwerker et al 1985, Yoon et al 1986). It is called diffusion-induced grain boundary migration (DIBM) when the surfaces are pre-existing grain boundaries, and diffusion-induced recrystallisation (DIR) when new crystals nucleate, usually at the interface between an existing crystal and a fluid, and the new crystals grow into the old as in DIBM. Migration resulting from self-stresses is known as liquid film migration (LFM) when there is a liquid film between the two crystals; the pair of crystal—liquid surfaces can be treated as a single surface.

For cubic crystals, the coherency strain energy is a function only of the strain $\delta$ and an orientation-dependent elastic parameter, $Y(n)$.

\[ f_{\text{coherent}} = Y(n)\delta^2 \]

When the change in lattice parameter $a$ with composition is constant (Vegard’s law)

\[ f_{\text{coherent}} = Y(n)\eta^2(C^s - C_0)^2 \]

where $C^s$ is the composition of the solid at the interface, $C_0$ is the initial composition and $\eta = (da/dC)/a(C_0)$.

Hilliard (1970) has given an expression for $Y$ for cubic crystals. Cubic elasticity is determined by three elastic constants $c_{11}, c_{12}$ and $c_{44}$; it is useful to rearrange his expression to be in terms of the modulus of compressibility
$K = c_{11} + 2c_{12}$, which is isotropic, and two ratios that best describe the orientation dependence of the elastic properties for cubic crystals: the 'anisotropy', $A = (c_{11} - c_{12})/2c_{12}$, and $B = c_{12}/c_{44}$, which is 1 for any solid composed of particles held by central pairwise forces (Cauchy's condition):

$$Y(n) = 2K \frac{3(A-1)^2\theta + A(A-1)\phi + A}{4(2A + 3B + 1)(A-1)^2\theta + 4(A-1)(A+B)\phi + 2A+B}$$

where the two cubic invariants of $n$, $\theta$ and $\phi$ are given by $\theta = n_1^2n_2^2n_3^2$, $\phi = n_1^2n_2^2 + n_2^2n_3^2 + n_3^2n_1^2$, where the coordinates of $n$ are defined relative to the cube axes.

In liquid film migration, DIR and DGM, there are two solid grains, 1 and 2, of the same phase and the same initial composition but differing by rotation $R$. Since there is a single surface (or two parallel surfaces) between the grains, the outward normals of grain 1 and grain 2 are in opposite directions. If we let the common coordinate system be relative to the cube axes of 1, then the normal relative to the cube axes of 2 is $-Rn$. The migration velocity depends on the difference in $f_{\text{coherent}}$ between 1 and 2:

$$v(n, R) = M' Y(n) [Y(n)\eta^2(C^2 - C_0, 1)^2 - Y(-Rn)\eta^2(C^2 - C_0, 2)^2]$$

where $M'$ is a combination of parameters that can be assumed to be isotropic. The migration rate and interface orientation, interface shapes can be predicted as a function of the elastic anisotropy for various conditions of stress in adjacent grains.

Often the stress in the growing grain becomes completely relaxed; the grain tends to become uniform in composition during LFM and in equilibrium with the liquid. Effectively $C_0$ for that grain approaches $C^*$, and the strain energy tends to zero in that grain (e.g. grain 2). The migration velocity becomes:

$$v(n) = M Y(n)$$

where $M' = M\eta^2(C^2 - C_0)^2$ is also assumed to be isotropic. The symmetry of the limiting shape of a stress-free growing grain 2 embedded in a self-stress-soluble matrix is that of the dissolving grain 1. Observations on DIR (Handwerker 1988) show that the new growing grains are almost dislocation-free while the dissolving matrix grains contain a high density of defects, including dislocations. The shape of the growing grains is cuboidal but oriented with the axes of the dissolving grains.

For an elastically isotropic material, that is, for $A - 1$ and for all B, $Y(n) = \text{constant}$ and $W_\infty$ is a sphere. Computing $v(n)$ leads to a prediction of how $W_\infty$ varies with $A$ and $B$. The limiting grain shapes $W_\infty$ obtained from $v(n)$ for KCl, Mo and Fe (figure 4), represent two forms for $W_\infty$ for cubic materials. Sections of $v(n)$ and $v_0$ for both types of shapes are shown in figure 5. For Fe, where $A < 1$, the shape is cuboidal with corners at

Figure 4  Limiting grain shapes $W_\infty$ determined by the elastic anisotropy of $Y(n)$ for (a) Mo, (b) KCl and (c) Fe for $n_1 \geq 0$.

3.6 Example of $v(n)$ for discontinuous coarsening models

Discontinuous coarsening is a solid state process in which a region with a fine lamellar structure dissolves and is replaced with a coarser one. The process is driven by capillary forces (reduction of surface energy between the lamellae). In a linear macroscopic theory where we consider only the capillary forces due to the fine lamellae, averaged over distances large compared to the lamellar spacing, the velocity of the process should be proportional to the product of the density of the lamellae ending along the interface and to the resolved 'surface tension' forces of each lamella along the smoothed interface normal. The density is $|\sin \phi |/\lambda$, where $\lambda$ is the lamellar spacing and $\phi$ is the angle between the normal to the interface $n$.

$\langle 111 \rangle$ with curved edges and faces. For KCl and Mo, where $A > 1$, the shape is an octahedron with corners at $\langle 100 \rangle$ and curved edges and faces. The creases in the edges join the edges for the KCl, but for the Mo the middle of the edges around the $\langle 110 \rangle$ are rounded. There are critical points along the edges where the creases terminate that are part of the description of a swallowtail in catastrophe theory (Thom 1975). Such morphologies have been observed (Handwerker 1988; see figures 17 and 19).
the microscope resolution. Furthermore, equilibrium may be attained without rotation, if these lamellar surfaces are so highly anisotropic that they resist rotating over a range of \( n \), the sum of shrinkage and 'torque' terms add to a force equal to the surface energy \( \gamma \), independent of \( \sin \phi \) over that range. We conclude that the velocity of growth \( v \) of the coarser region into the lamellar region is proportional to \( (2\gamma/\lambda)\sin^4\phi \) where \( k \) is probably 1, but could be as large as 2. Therefore we should contrast two cases, \( \sin \phi \) and \( \sin^2 \phi \). A section through a radial plot of these two functions is shown in figure 6. These plots are like doughnuts without a complete hole. In either case \( v \) is zero when the surface is parallel to the lamella. This obviously suggests that growth is confined to the caliper diameter of the initial form \( C_0 \), the range between the highest and lowest point of the initial form transverse to the lamellae. The orientation dependence of \( v \) has not been measured, and \( C_0 \) was not controlled. However, evolving growth forms suggested that \( v \) was indeed zero perpendicular to the lamella, and an average \( v \) measurement confirmed the dependence on \( \lambda \). Let us examine the predictions for these two cases.

![Figure 5](image1)

**Figure 5** Sections of (a) polar plots of \( v(n)(nv(n)) \) and (b) \( v(n) \) for Fe and KCl for \( n_1 = 0 \), showing ears at \( \langle 110 \rangle \) and \( \langle 100 \rangle \), respectively.

and the normal to the lamellae, which we may take to be \((0, 0, 1)\), so that \( \cos(\phi) = n_3 \). The second factor depends on whether or not the lamellar ends can rotate, so that the trijunctions, where lamellar surfaces meet the grain boundary, can attain equilibrium angles. If rotation occurs until equilibrium is reached, the second factor is simply twice the energy of the surface of the lamellae, \( \gamma \). If not, and we assume isotropic surface energies, then the surface tension force resolved in the direction of motion is reduced by another factor of \( \sin \phi \). Since lamellar rotations were not observed, the \( \sin \phi \) in the second factor is a possibility. However, the steady state motion required to rotate enough of the tips to eliminate this factor was well below

![Figure 6](image2)

**Figure 6** Sections through radial plots of \( \sin \phi \) and \( \sin^2 \phi \); two possibilities for \( v(n) \) for discontinuous coarsening. \( v(n) \) for these two cases are respectively vectors of constant magnitude in the plane \( n_1 = 0 \), and the double swallowtail in figure 2(b) with \( b/a = -1 \).

Given \( p \) and \( n = p/|p| \), we set

\[
pr = (p_1^2 + p_2^2)^{1/2}
\]

and

\[
n_r = (pr|p_3|pr).
\]

Note that \( n_r \) is a unit vector in the direction of the projection of \( n \) into the plane of the lamellae.
For the first case of \( k = 1 \),
\[
v(n) = a(1 - ((0, 0, 1) \cdot n)^2)^{1/2} = a(1 - n_3^2)^{1/2}
\]
\[
v(p) = a(\|p\|^2 - (p_3)^2)^{1/2} = ap
\]
\[
v_0(p) = 2a[p_1p_1, p_2p_2, p_3p_3, 0] = 2am
\]
Note that \( v_0 \) is a radial vector of constant magnitude in this plane. For the case of \( k = 2 \),
\[
v(n) = a - a(n \cdot (0, 0, 1))^2 = a(1 - n_3^2)
\]
\[
v(p) = (a|p|)(|p|^2 - p_3^2) = (a|p|)p^2
\]
\[
v_0(p) = a(2(1 - n_3^2)n_3, -(1 - n_3^2)n_3)
\]
Note that \( v_0 \) is composed of two terms, an outward radial term along the lamellae, that is dependent on the inclination, plus a smaller inward term that is along \( n \). In component form the significance of the second term is more apparent:
\[
v_0(p) = a(1 - n_3^2)((2 - (1 - n_3^2)n_3, -(2 - (1 - n_3^2)n_3, -(1 - n_3^2)n_3)
\]

The first case, \( k = 1 \), is a good physical example of constant radial characteristics. The evolving forms tend to circular discs with thickness equal to initial caliper diameter; all the elements of \( C_0 \) translate radially in the direction of the projection of their \( n \) with constant velocity. A vertical cross section of the evolving growth forms would look blocky; this was similar to the shapes observed in the Co–Si alloys (Livingston and Cahn 1974, see their figure 1c). In the second case, the characteristics have a vertical component that is opposite in sign to \( n_3 \). With \( k = 2 \) the stronger orientation dependence of \( v \) is such that, in a convex front, the fastest growth direction is grown out and the slower growth directions take over. A convex front (as viewed from the coarser region) sticks out further and develops shocks and in vertical cross section will look pointed; a concave front flattens. Such features were seen in the Ni–In and Cu–In alloys (Livingston and Cahn 1974, see their figures 2 and 3). Thus the two cases give quite different predictions for the evolving forms; hints of both predictions were seen.

4 CONNECTIONS AND EXTENSIONS OF THIS THEORY

4.1 Analogy with surface energy
The gradient of (the extension of) \( v \) turns out to be of central importance in this model. It is similarly of central importance (and called \( \xi \) there) in determining equilibrium shapes of curved surfaces with anisotropic surface free energy \( \gamma(n) \) (Hoffman and Cahn 1972, Cahn and Hoffman 1974). We can apply some of the known relations between \( \gamma \), its gradient, and its

\[
W = \{x: x \cdot n \leq \gamma(n) \text{ for every unit vector } n\}
\]
to the case of \( v \), its gradient, and the shape
\[
W_0 = \{x: x \cdot n \leq v(n) \text{ for every unit vector } n\}
\]
since they are simply results about certain functions defined on unit normal vectors; this is done in the appendix. If \( v \) depends on position and time as well as the normal vectors, then those variables must be held fixed for the analogy to hold, and the dependence of \( v \) on those variables will be temporarily suppressed. Similarly, \( v \) ought to be non-negative.

If \( \gamma(n) \) is replaced by \( \gamma_0(n) = \gamma(n) + n \cdot b \), then the Wulff shapes for \( \gamma \) and for \( \gamma_0 \) differ only by a translation by \( b \), since \( \nabla\gamma_0(n) = \nabla\gamma(n) + b \); this was first observed by Lee et al. (1975) for isotropic \( \gamma \), and is a counterexample to Chernov's conjecture that the area of a face decreases as \( \gamma \) increases. Also, by Stokes' theorem, the integral of \( \gamma_0 \) over any piece of surface differs from the integral of \( \gamma \) over that piece of surface only by the integral of \( b \cdot m \) over the two-dimensional boundary of that piece of surface, where \( m \) is the exterior normal of the surface, so that \( \gamma_0 \) and \( \gamma \) have the same surface-energy-minimising surfaces. This was first noted by Arbel and Cahn (1975), and theorems showing that two surface energy functions have the same minimal surfaces if and only if they differ by the addition of \( n \cdot b(x) \) for \( b(x) \) a divergence-free vector field were proved by Taylor (1982). For crystal growth, the parallel results are that the growth shapes using \( u(t, n) \) and \( v(t, n) + n \cdot b(t) \) are indistinguishable by shape alone, as at each time one shape is just a translation of the other if they start with the same initial shape.

An important feature about the \( W \) corresponding to a given surface free energy function \( \gamma \) is that it has less surface energy than any other shape having the same volume. Since the surface energy is the integral of \( \gamma \) over the surface, and the integral of \( v \) over a surface is the rate at which volume is swept out during growth, we conclude that volume is consumed at a slower rate by the growth of \( W_0 \) than by the growth of any other shape of the same volume, assuming that the growth rate is given by \( v \). Similarly, results about the structure of surface-energy-minimising surfaces under other constraints can be interpreted as results about the structure of surfaces which instantaneously consume volume at the slowest rate under those constraints; such surface-energy-minimising shapes (for example, with the constraint of having a fixed boundary curve) have been extensively studied theoretically (Taylor 1983, Taylor and Cahn 1986a, b, Zia et al. 1988). The direct utility of these other results is not clear, however, since such surfaces rarely remain of the same shape as they grow.

However, many results for \( \gamma \) do not carry over to growth with \( v \) because they explicitly involve minimisation in ways that growth does not. For example, if a planar surface has direction \( n \) such that \( \nabla\gamma(n) \) is part of an
interesting features of crystal growth with a prescribed normal velocity do not appear in the solution of the anisotropic wave equation.

4.3 Alternative numerical techniques for solution
The method of characteristics to construct the family of growing crystal surfaces is nice for existence theorems and for constructing the solution ‘by hand’, but is relatively difficult to implement numerically when shocks are present. Sethian (1985, 1990) has investigated alternative numerical techniques for solving the isotropic case \( u(n) = 1 \) (which he interprets as a simple model for the case of flames burning at a constant speed normal to themselves). The condition at shocks is the same as the one we use, but he formulates it as ‘once a particle is burnt, it remains burnt’. It is described as an entropy condition because of the irreversibility of the motion at shocks. By replacing \( v \) by \( v_e = 1 - eK \), where \( K \) is the curvature of the interface, he ‘regularises’ the equation to a second-order partial differential equation. Now the method of characteristics no longer applies, because the PDE is no longer first order. But the advantage of doing so is that curvature plays a smoothing role, preventing the occurrence of shocks. He proves (at least in the two-dimensional case (Sethian 1985)) that, as \( e \) approaches zero, the limit of these smooth solutions is the solution to the problem with \( e = 0 \) (which may of course no longer be smooth but have shocks). Viewing, as we do, the \((N - 1)\)-dimensional moving surface as a level set of a time-dependent function of \( N \) dimensions, he observes that ‘the equation of motion for this function resembles an initial value Hamilton–Jacobi equation with parabolic right-hand-side and is closely related to a viscous hyperbolic conservation law’ and ‘numerical schemes designed to approximate hyperbolic conservation laws may be used to approximate the motion of the propagating surface’. (The model of having \( v \) be \( 1 - eK \) is certainly also of interest in itself for crystal growth.)

4.4 Limitations
Our formulation, which assumes that the normal velocity is a function only of \( n, x \) and \( t \), is an idealisation. It excludes all factors that alter the type of equation, such as a dependence on surface curvature which would make the equation second order and render the method of characteristics inapplicable. In particular, the formulation is inapplicable where two surfaces with the same normal \( n \) have different velocities \( v \). This may arise, for example, from the presence or absence of dislocations with screw components normal to the surface, or from differences in large-scale geometry; a facet on a convex surface encounters a different physical problem to a nucleate and/or grow a new layer than a facet on a concave surface.

A different formulation is also needed to consider cases that arise from diffusion control, in which the velocity does change with time, but only implicitly because the crystal significantly alters its environment during
growth. An attempt by Chernov (1963) to treat this case as part of a unified formulation in terms of orientation and supersaturation seems of limited usefulness because it requires explicit knowledge of how the supersaturation or temperature varies with orientation.

ACKNOWLEDGMENTS

We are most grateful to Craig Carter for generous and tireless help with the symbolic calculations and with the graphics, and for critical readings of the manuscript, and to Robert Parker for help with some of the early crystal growth literature. We gratefully acknowledge support of the ONR, DARPA, Air Force and NSF.

APPENDIX. SUMMARY OF RESULTS RELATING \( v(n), v(p) \), \( \nabla v \) and \( W_\infty \)

First, we extend \( v \) to a function on all vectors \( p \), rather than just unit vectors \( n \), by defining \( v(p) = |p|v(n) \) for all \( p \) (when we use \( p \) and \( n \) in the same formula or description in this paper, we use implicitly assuming that \( p = |p|n \)). One can easily verify that \( \nabla v \) is thereby constant in the radial direction, i.e. \( \nabla v(p) = \nabla v(n) \); whenever we write \( \nabla v \), we mean \( \nabla v = (\partial u/\partial P_1, \partial u/\partial P_2, \partial u/\partial P_3) \) (whether or not \( v \) depends additionally on \( x \) and \( t \)) and \( \nabla v(n) \) means that gradient evaluated at the unit normal \( n \), not any type of surface gradient. Note that \( p \cdot \nabla v = v(p) = 1 \), the first equation holding because \( v(p) = |p|v(n) \) and the last equation holding because of the PDE.

Secondly, for any \( n_0 \) such that \( \nabla v(n_0) \) is undefined (which happens where \( \nabla v \) is discontinuous), we use in place of \( \nabla v(n_0) \) the set of all convex combinations of the limits of \( \nabla v(n) \) as \( n \) approaches \( n_0 \). For example, if we define \( v_{\text{cube}}(n) = |n_1| + |n_2| + |n_3| \), then for every \( n \) in the open first octant, \( v_{\text{cube}} = (1, 1, 1) \). Because of the absolute values in the orientation dependence of \( v_{\text{cube}} \), there are discontinuities in the gradient at the octant boundaries. If \( p = c(1, 0, 0) \) for some \( c > 0 \), then wherever \( v_{\text{cube}}(p) \) is called for, one uses instead all values \((1, t, s), |s|, |t| \leq 1, \) and if \( p = c(1, 1, s) \), one uses all values \((t, (1 - t^2)^{1/2}, 0), -1 \leq t \leq 1 \), for \( v_{\text{cube}}(p) \) (Figure 3). The plot of \( v_{\text{cube}} \), when understood as above, is a cube; each corner is the value of \( v_{\text{cube}}(n) \) for all \( n \) in its octant. This is the same as using the 'subgradient' of \( v \) (Rockafellar 1981).

The plot of \( n|v(n) \) is a level set of the function \( v(p) \), namely, the set of \( p \) such that \( v(p) = 1 \). Therefore \( v(n) \) is normal to this plot. \( 1/v(n) \) has been called the 'slowness' not only by Frank in crystal growth but also in crystal optics and elastic waves.

The relationship between \( v \) and \( W_\infty \) has been well studied, especially in the case where \( v \) is a convex function (which is equivalent to that plot being convex) (see, for example, Rockafellar (1970, 1981); Cahn and Hoffman (1974) rediscovered many of these results for the case when \( v \) is interpreted as a surface-energy function \( \gamma \). When \( v \) is a convex function, \( v \) is the support function of its \( W_\infty \), and \( W_\infty \) is the plot of \( v \). Note that if \( x \) is a boundary point of \( W_\infty \) and the plane with normal \( n \) through \( x \) is a support plane of \( W_\infty \), then \( v(x) = v(n) \) when \( v \) is convex. One can regard \( W_\infty \) itself as the plot of \( m|w(m) \) for some function \( w \) defined on unit vectors and then extend \( w \) by \( w(m) = m\cdot w(m) \). This can be written as

\[ w(x) = v^*(x) = \sup_{\|p\|} (p \cdot x - v(p)) \]

The mapping \( v \rightarrow w = v^* \) is called the Fenchel transform (or Legendre transform when \( v \) is differentiable); the Fenchel transform of \( w = v^* \) takes one back to \( v \). Thus the 'Wulff shape' for the function \( w \) is the plot of \( n|v(n) \); Frank (1963) refers to this plot as the pedal of \( W_\infty \).

When \( v \) is not a convex function, the plot of \( v \) (interpreted as above) coincides with the surface of its \( W_\infty \) except for additional 'ears' at corners and/or edges of \( W_\infty \). As is the case for a surface energy function \( \gamma \) and its Wulff shape \( W \).

Corners and sharp edges occur in \( W_\infty \) where the plot of \( v \) is self-intersecting, i.e. where \( v(n_1) = v(n_2) \) for some \( n_1 \neq n_2 \); there are no 'ears' if and only if \( v(n) \) is constant on all directions \( n \) omitted at that corner or edge. Facets and straight edges on \( W_\infty \) correspond to discontinuities in \( v \).

As a simple example, the \( W_\infty \) for \( v_{\text{cube}}(n) = |n_1| + |n_2| + |n_3| \) is a cube, and so is the plot of \( v_{\text{cube}} \). This is the smallest it can be and still have its \( W_\infty \) be that cube; there are no 'ears' on its \( W_\infty \) plot (with \( v \) being interpreted as described above). In some situations, theories predict that the surface energy \( \gamma \) should be a constant times this \( v \); in those theories, the plot of \( v_{\text{cube}}(n) \) is sometimes called a 'raspberry' figure. Small changes in this \( v \) can either round the edges or corners of its \( W_\infty \) or add swallowtails there.

The most important property of \( W_\infty \) for our purposes is that it is the limiting growth shape for any finite seed.

REFERENCES

Burton W K, Cabrera N and Frank F C 1951 Phil. Trans. R. Soc. A 243 299–358
Chernov A A 1963 Sov. Phys.—Cryst. 7 278–30
—— 1984 Modern Crystallography III, Crystal Growth (Berlin: Springer) p 219ff
Interfacial Dislocations in Epitaxy

J H van der Merwe

1 INTRODUCTION

Interfacial dislocations occur in a variety of systems, for example, in epitaxial interfaces, in grain boundaries and in precipitate–matrix interfaces. This paper shall be mainly concerned with epitaxial interfaces where the dislocations have commonly become known as misfit dislocations.

In 1949 Frank and van der Merwe (F&M) published two papers, based on what has become known as the Frenkel–Kontorowa (FK) model (1938). N F Mott suggested that van der Merwe, a PhD student at the time, exploit this model with the view of using it as a first step in developing a theoretical description of intercrystalline boundaries. With the ingenious leadership, insight and collaboration of his supervisor F C Frank this culminated in the abovementioned papers addressing the role of interfacial dislocations in epitaxial interfaces.

The objectives of this paper are to show: (i) that the abovementioned work has laid the foundations for the description of the role of interfacial dislocations in the accommodation of misfit at submonolayer to multilayer epitaxial film interfaces; foundations which are still valid after 40 years—Hibma writes for example in 1988 ‘... the most successful theory is the original theory of Frank and van der Merwe ...’; (ii) that the description lends itself ideally for extensions to more complicated systems and (iii) that there are further important aspects which the authors did not deal with at the time.

In relation to item (i) above it is of interest that several subsequent
Sir Charles Frank, OBE, FRS:
An eightieth birthday tribute

Edited by
R G Chambers, J E Enderby, A Keller, A R Lang and J. W. Steeds

H H Wills Physics Laboratory, University of Bristol

Adam Hilger
Bristol, Philadelphia and New York