DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation	Mumford, David Bryant, and Jayant Shah. 1989. Optimal approximations by piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics 42(5): 577-685.
Published Version	doi:10.1002/cpa.3160420503
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:3637121
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems[1]

DAVID MUMFORD
Harvard University

AND

JAYANT SHAH
Northeastern University

1. Introduction and Outline

The purpose of this paper is to introduce and study the most basic properties of three new variational problems which are suggested by applications to computer vision. In computer vision, a fundamental problem is to appropriately decompose the domain R of a function $g(x, y)$ of two variables. To explain this problem, we have to start by describing the physical situation which produces images: assume that a three-dimensional world is observed by an eye or camera from some point P and that $g_1(\rho)$ represents the intensity of the light in this world approaching the point P from a direction ρ. If one has a lens at P focussing this light on a retina or a film—in both cases a plane domain R in which we may introduce coordinates x, y—then let $g(x, y)$ be the strength of the light signal striking R at a point with coordinates (x, y); $g(x, y)$ is essentially the same as $g_1(\rho)$—possibly after a simple transformation given by the geometry of the imaging system. The function $g(x, y)$ defined on the plane domain R will be called an *image*. What sort of function is g? The light reflected off the surfaces S_i of various solid objects O_i visible from P will strike the domain R in various open subsets R_i. When one object O_1 is partially in front of another object O_2 as seen from P, but some of object O_2 appears as the background to the sides of O_1, then the open sets R_1 and R_2 will have a common boundary (the 'edge' of object O_1 in the image defined on R) and one usually expects the image $g(x, y)$ to be discontinuous along this boundary: see Figure 1 for an illustration of the geometry.

Other discontinuities in g will be caused by discontinuities in the surface orientation of visible objects (e.g., the 'edges' of a cube), discontinuities in the objects albedo (i.e., surface markings) and discontinuities in the illumination

[1]A preliminary version of this paper was submitted by invitation in 1986 to "Computer Vision 1988", L. Erlbaum Press, but it has not appeared!

Communications on Pure and Applied Mathematics, Vol. XLII 577–685 (1989)
© 1989 John Wiley & Sons, Inc. CCC 0010-3640/89/050577-107$04.00

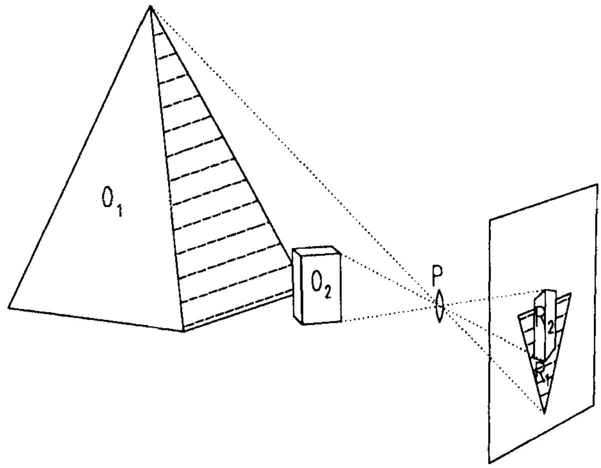

Figure 1. An *image* of a 3D scene.

(e.g., shadows). Including all these effects, one is led to expect that the image $g(x,y)$ is piece-wise smooth to a first approximation, i.e., it is well modelled by a set of smooth functions f_i defined on a set of disjoint regions R_i covering R. This model is, however, far from exact: (i) textured objects such as a rug or fragmented objects such as a canopy of leaves define more complicated images; (ii) shadows are not true discontinuities due to the penumbra; (iii) surface markings come in all sorts of misleading forms; (iv) partially transparent objects (e.g., liquids) and reflecting objects give further complications; (v) the measurement of g always produces a corrupted, noisy approximation of the true image g.

In spite of all this, the piece-wise smooth model is serviceable on certain scales and to a certain approximation. Restating these ideas, the *segmentation problem* in computer vision consists in computing a decomposition

$$R = R_1 \cup \cdots \cup R_n$$

of the domain of the image g such that

(a) the image g varies smoothly and/or slowly *within* each R_i,
(b) the image g varies discontinuously and/or rapidly across most of the boundary Γ between different R_i.

From the point of view of approximation theory, the segmentation problem may be restated as seeking ways to define and compute *optimal approximations* of a general function $g(x, y)$ by piece-wise smooth functions $f(x, y)$, i.e., functions f whose restrictions f_i to the pieces R_i of a decomposition of the domain R are differentiable. Such a problem arises in many other contexts: the perception of speech requires segmenting time, the domain of the speech signal, into intervals during which a single phoneme is being pronounced. Sonar, radar or laser "range" data, in which $g(x, y)$ represents the distance from a fixed point P in direction (x, y) to the nearest solid object, are other computer vision signals whose domains must be segmented. CAT scans are estimates of the density of the body at points (x, y, z) in three-space: segmentation is needed to identify the various organs of the body.

To make mathematics out of this, we must give a precise definition of what constitutes an optimal segmentation. In this paper we shall study three functionals which measure the degree of match between an image $g(x, y)$ and a segmentation. We have a general functional E which depends on two parameters μ and ν and two limiting cases E_0 and E_∞ which depend on only one parameter ν and correspond to the limits of E as the parameter μ tends to 0 and ∞, respectively.

We now define these three functionals and motivate them in terms of the segmentation problem. In all these functionals, we use the following notation: the R_i will be disjoint connected open subsets of a planar domain R each one with a piece-wise smooth boundary and Γ will be the union of the part of the boundaries of the R_i inside R, so that

$$R = R_1 \sqcup \cdots \sqcup R_n \sqcup \Gamma.$$

(Conversely, we could start from a closed set Γ made up of a finite set of singular points joined by a finite set of smooth arcs meeting only at their endpoints, and let the R_i be the connected components of $R - \Gamma$.) For the functional E, let f be a differentiable function on $\cup R_i$, which is allowed to be discontinuous across Γ. Let

$$E(f, \Gamma) = \mu^2 \iint_R (f - g)^2 \, dx\, dy + \iint_{R-\Gamma} \|\nabla f\|^2 \, dx\, dy + \nu|\Gamma|,$$

where $|\Gamma|$ stands for the total length of the arcs making up Γ. The smaller E is, the better (f, Γ) segments g:

(i) the first term asks that f approximates g,
(ii) the second term asks that f – and hence g – does not vary very much on each R_i,
(iii) the third term asks that the boundaries Γ that accomplish this be as short as possible.

Dropping any of these three items, $\inf E = 0$: without the first, take $f = 0$, $\Gamma = \varnothing$; without the second, take $f = g$, $\Gamma = \varnothing$; without the third, take Γ to be a fine grid of N horizontal and vertical lines, $R_i = N^2$ small squares, $f =$ average of g on each R_i. The presence of all three terms makes E interesting.

The pair (f, Γ) has an interesting interpretation in the context of the original computer vision problem: (f, Γ) is simply a *cartoon* of the actual image g; f may be taken as a new image in which the edges are drawn sharply and precisely, the objects surrounded by the edges are drawn smoothly without texture. In other words, (f, Γ) is an idealization of a true-life complicated image by the sort of image created by an artist. The fact that such cartoons are perceived correctly as representing essentially the same scene as g argues that this is a simplification of the scene containing many of its essential features.

We do not know if the problem of minimizing E is well posed, but we conjecture this to be true. For instance we conjecture that for all continuous functions g, E has a minimum in the set of all pairs (f, Γ), f differentiable on each R_i, Γ a finite set of singular points joined by a finite set of C^1-arcs.

A closely related functional defined for functions g and f on a lattice instead of on a plane domain R was first introduced by D. and S. Geman [6], and has been studied by Blake and Zisserman [4], J. Marroquin [10] and others. The variational problem $\delta E = 0$ turns out to be related to a model introduced recently by M. Gurtin [8] in the study of the evolution of freezing/melting contours of a body in three-space.

The second functional E_0 is simply the restriction of E to piecewise constant functions f: i.e., $f =$ constant a_i on each open set R_i. Then multiplying E by μ^{-2} we have

$$\mu^{-2}E(f, \Gamma) = \sum_i \iint_{R_i} (g - a_i)^2 \, dx \, dy + \nu_0|\Gamma|,$$

where $\nu_0 = \nu/\mu^2$. It is immediate that this is minimized in the variables a_i by setting

$$a_i = \text{mean}_{R_i}(g) = \iint_{R_i} g \, dx \, dy \Big/ \text{area}(R_i),$$

so we are minimizing

$$E_0(\Gamma) = \sum_i \iint_{R_i} (g - \text{mean}_{R_i} g)^2 \, dx \, dy + \nu_0 \cdot |\Gamma|.$$

As we shall see, if Γ is fixed and $\mu \to 0$, the f which minimizes E tends to a piecewise constant limit, so one can prove that E_0 is the natural limit functional of E as $\mu \to 0$. E_0 may be viewed as a modification of the usual Plateau problem functional, length (Γ), by an external force term that keeps the regions R_i—soap bubbles in the Plateau problem setting—from collapsing. Whereas the two-dimensional Plateau problem has only rather uninteresting extrema with Γ made up of straight line segments (cf. Allard-Almgren [1]), the addition of the pressure term makes the infimum more interesting. Having the powerful arsenal of geometric measure theory makes it straightforward to prove that the problem of minimizing E_0 is well posed: for any continuous g, there exists a Γ made up of a finite number of singular points joined by a finite set of C^2-arcs on which E_0

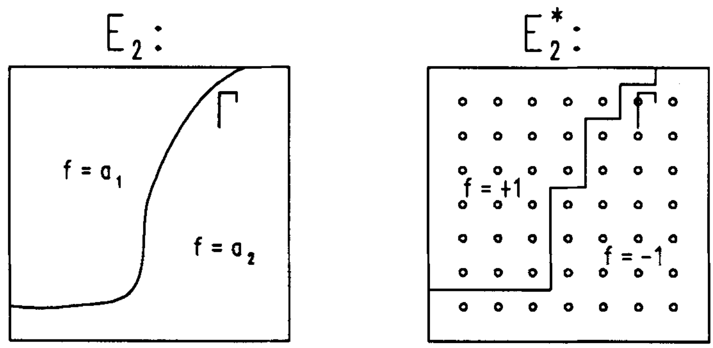

Figure 2. Continuous vs. discrete segmentation.

attains a minimum. E_0 is also closely related to the energy functional in the *Ising model*. For this, we restrict f even further to take on only two values: $+1$ and -1, and we assume g and f are functions on a lattice instead of functions on a two-dimensional region R. In this setting, Γ is the path made up of lines between all pairs of adjacent lattice points on which f changes sign (see Figure 2). Then E_0 reduces to

$$E_0^*(f) = \sum_{\substack{\text{lattice pts} \\ i,j}} (f(i,j) - g(i,j))^2 + \nu_0 \cdot \sum_{\substack{\text{adjacent} \\ \text{lattice pts} \\ (i,j),(i',j')}} (f(i,j) - f(i',j'))^2$$

which is the Ising model energy.

The third functional E_∞ depends only on Γ and is given by

$$E_\infty(\Gamma) = \int_\Gamma \left[\nu_\infty - \left(\frac{\partial g}{\partial n} \right)^2 \right] ds,$$

where ν_∞ is a constant, ds is arc length along Γ and $\partial/\partial n$ is a unit normal to Γ. Using dx, dy as coordinates on the tangent plane to R, so that $ds = \sqrt{dx^2 + dy^2}$, E_∞ may be rewritten as the integral along Γ of a *generalized Finsler metric* $\rho(dx, dy, x, y)$ (a function ρ such that $\rho(t\,dx, t\,dy, x, y) = |t| \cdot \rho(dx, dy, x, y)$), namely:

$$E_\infty(\Gamma) = \int_\Gamma \rho(dx, dy, x, y),$$

$$\rho(dx, dy, x, y) = \frac{\nu_\infty(dx^2 + dy^2) - (g_x\,dy - g_y\,dx)^2}{\sqrt{dx^2 + dy^2}}.$$

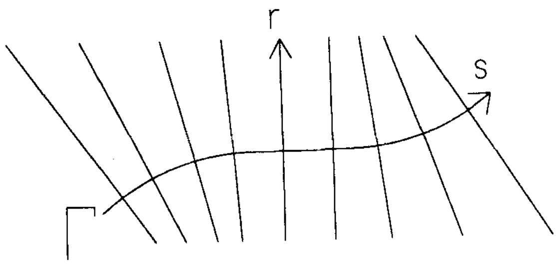

Figure 3. Curvilinear coordinates r, s.

Intuitively, minimizing E_∞ is then a generalized geodesic problem. It asks for paths Γ such that (i) length (Γ) is as short as possible while (ii) normal to Γ, g has the largest possible derivative. Looking at the graph of g as a landscape, Γ is the sort of path preferred by mountain goats—short but clinging to the face of cliffs wherever possible.

At first glance, E_∞ looks completely unrelated to E. In fact, like E_0 it is essentially E with a special choice of f:

We consider only smooth parts of Γ and take $f = g$ outside an infinitesimal neighborhood of Γ. Near Γ, set

$$f(r,s) = g(r,s) + \varepsilon\,\mathrm{sgn}(r) \cdot e^{-|r|/\varepsilon} \frac{\partial g}{\partial r}(0,s),$$

where r, s are curvilinear coordinates defined by normals of Γ (see Figure 3), and ε is infinitesimal. Then if $\mu = 1/\varepsilon$, $\nu = 2\varepsilon\nu_\infty$, it can be checked that

$$E(f,\Gamma) - E(g,\phi) = 2\varepsilon E_\infty(\Gamma) + O(\varepsilon^2).$$

We shall see, moreover, that for fixed Γ, if μ is very large, the f which minimizes E is very close to the above f with $\varepsilon = 1/\mu$, so this E_∞ is a natural limit functional of E as $\mu \to \infty$.

Minimizing E_∞ *over all* Γ is not a well-posed problem in most cases: if $\|\nabla g\|^2 \leqq \nu_\infty$ everywhere, then $E_\infty \geqq 0$ and the simple choice $\Gamma = \varnothing$ minimizes E_∞. But if $\|\nabla g\|^2 > \nu_\infty$ on a non-empty open set U, then consider Γ made up of many pieces of level curves of g within U. On such Γ's, E_∞ tends to $-\infty$. For μ large but not infinite, the pair (f, Γ) which minimizes E itself presumably has a Γ made up of many components in this open set U. Minimizing E_∞ on a suitably restricted class of Γ's can, however, be a well-posed problem.

We finish this introduction by giving an outline of the results of each section. The proofs in the later sections will sometimes be quite technical and hence it seems useful to describe the main results here before plunging into detail.

In Section 2, we analyze the variational equations for the functional E. Fixing Γ, standard calculus of variations shows that E is a positive definite quadratic function in f with a unique minimum. The minimum is the function f which solves the elliptic boundary value problem on *each* R_i,

$$\Delta f = \mu^2(f - g),$$

$$\partial f / \partial n|_{\partial R_i} \equiv 0.$$

Here ∂R_i is the boundary of R_i, and $\partial / \partial n$ is a unit normal vector to ∂R_i. The second condition means that $\partial f / \partial n$ is zero on both sides of Γ and on the inside of ∂R, the boundary of the whole domain R. Writing f_Γ for this solution, E reduces to a function of Γ alone:

$$E(f_\Gamma, \Gamma).$$

Next, we make an infinitesimal variation of Γ by a normal vector field $\mathbf{X} = a(x, y)\, \partial / \partial n$ defined along Γ and zero in the neighborhood of the singular points of Γ. We prove:

$$\frac{\delta}{\delta X} E(f_\Gamma, \Gamma) = \int_\Gamma a(e_+ - e_- + \nu \cdot \operatorname{curv}(\Gamma))\, ds,$$

where

$$e_\pm = \mu^2(f_\Gamma^\pm - g)^2 + \left(\frac{df_\Gamma^\pm}{ds}\right)^2,$$

$$f_\Gamma^\pm = \text{boundary values of } f_\Gamma.$$

Therefore if $E(f_\Gamma, \Gamma)$ is minimized at Γ, Γ must satisfy the variational equation

$$(*) \qquad e_+ - e_- + \nu \cdot \operatorname{curv}(\Gamma) \equiv 0.$$

Finally, we look at possible singular points of Γ. Here the situation is complicated by the fact that although the minimizing f_Γ is pointwise bounded:

$$\min_R g \leq f_\Gamma(P) \leq \max_R g,$$

its gradient may be unbounded in a neighborhood of the singular points of Γ. However, the behavior of the solutions of elliptic boundary value problems in domains with "corners" has been studied by several people, starting with a classic paper of Kondratiev [9] and recently surveyed in a book of Grisvard [7]. Using these bounds, and assuming that the singular points are given by a finite number of C^2-arcs with a common endpoint, it is easy to show, by elementary comparisons of $E(f_\Gamma, \Gamma)$ with E on modified Γ's, that if $E(f_\Gamma, \Gamma)$ attains a minimum at some Γ, then the only possible singularities of Γ are:

(i) "triple points" P where three C^2-arcs meet with 120°-angles,
(ii) "crack-tips" P where a single C^2-arc ends and no other arc meets P.

Moreover, on the boundary of the domain R, another possibility is:

(iii) "boundary points" P where a single C^2-arc of Γ meets perpendicularly a smooth point of ∂R.

There is a wrinkle here though: assuming that E has a minimum when Γ varies over some reasonable set of possible curves, it is not clear that the minimizing Γ will have singular points made up of C^2-arcs. Instead there might well be a larger class of nastier singular points that can arise as minima. In Section 3, we look more closely at crack-tips with this possibility in mind. First of all, assuming the crack-tip is C^2 we calculate the first variation of E with respect to infinitesimal extensions or truncations of Γ at the crack tip. We find a new restriction on minimizing Γ's, which is the analogue of Griffiths' law of cracks in solid mechanics. Secondly, we consider the possibility that the crack-tip might be given by

$$y = \alpha x^{3/2} + \phi(x), \qquad x \geqq 0, \phi \text{ a } C^2\text{-function}.$$

Such singular points are called *cusps* in algebraic geometry. We give several arguments to make it plausible that such a singularity will occur on minimizing Γ's. Moreover, in the natural world, approximate cusps certainly look like they appear at the ends of arcs in other situations which may be modelled by free boundary value problems. For example, consider sand bars that stick out into an area with a strong transverse current. Sand may be washed away or may accrete, hence the boundary is free to shift in both directions, and its tip may likewise erode or grow. Figure 4 shows a chart of Cape Cod and of Monomoy Island, whose outlines are strongly reminiscent of cusps.

In Section 4, we study E for small μ. We use the isoperimetric constant λ_Γ which gets small only if some component W of $R - \Gamma$ has a narrow "neck" (see Figure 5). By definition, λ_Γ is the minimum of $\delta/\min(A_1, A_2)$ over all diagrams as in the figure. We prove that

$$\mu^2\left\{E_0(\Gamma) - \frac{4\mu^2}{\lambda_\Gamma^2 + 4\mu^2}\iint g^2\right\} \leqq E(f_\Gamma, \Gamma) \leqq \mu^2 E_0(\Gamma),$$

where $\nu_0 = \nu/\mu^2$. We then study the first variation of $E_0(\Gamma)$. We show firstly that the first variation of $E(f_\Gamma, \Gamma)$ also tends to the first variation of $E_0(\Gamma)$. Secondly, the equation for the first variation of $E_0(\Gamma)$ being zero turns out to be the second-order differential equation for Γ:

$$(\bar{g}_\Gamma^+ - \bar{g}_\Gamma^-)\cdot(\bar{g}_\Gamma^+ + \bar{g}_\Gamma^- - 2g) + \nu_0 \cdot \mathrm{curv}(\Gamma) \equiv 0,$$

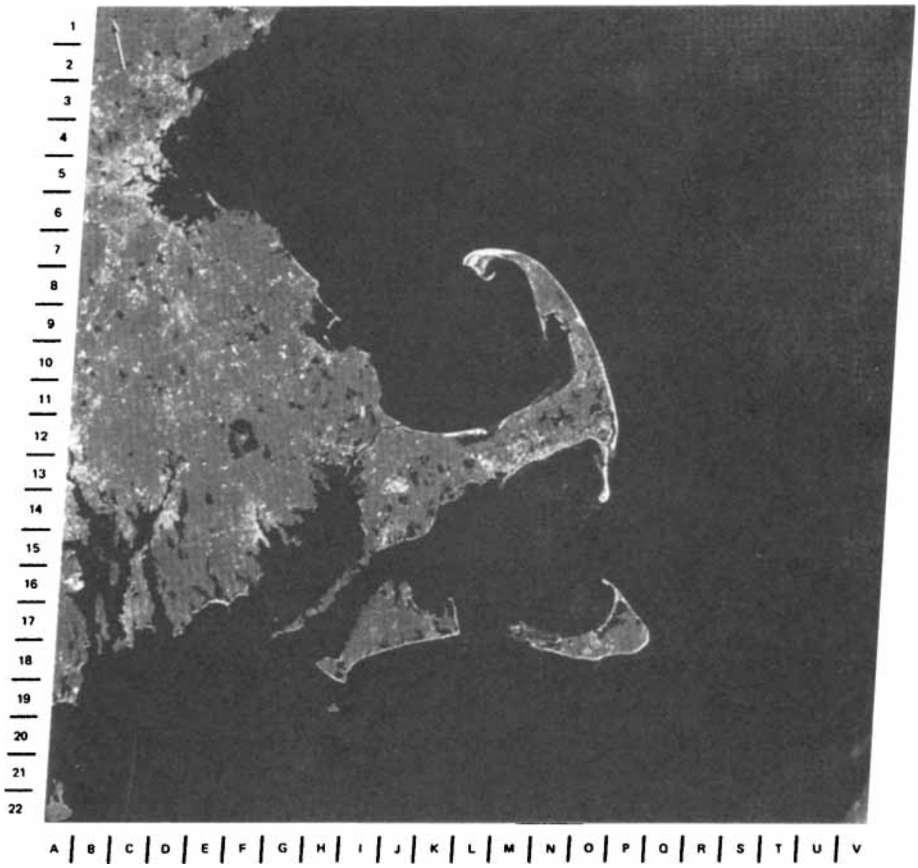

Figure 4. A sandy shore often has cusp-like spits.

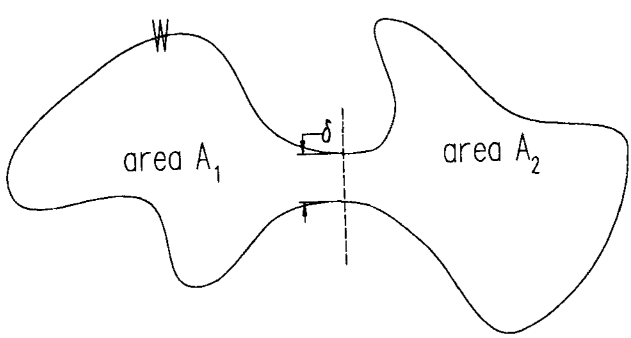

Figure 5. Definition of the isoperimetric constant.

where $\bar{g}_\Gamma^+$ and $\bar{g}_\Gamma^-$ are the means of g on the components of $R - \Gamma$ to each side of Γ.

Section 5 is devoted to the proof that $\min E_0(\Gamma)$ exists for a Γ which is a finite union of C^2-arcs joining a finite number of triple points or border points. We dip deeply into the toolkit of geometric measure theory, using especially the results in Simon [13]. The fundamental idea is to enlarge the set of allowable Γ's until one of the compactness results of geometric measure theory can be applied to show that $\min E_0$ exists for a possibly very wild Γ. And once this Γ is in hand, use the vanishing of its first variation to show that Γ is in fact very nice. More precisely, what we do is to shift the focus from Γ to a decomposition of R into disjoint measurable sets:

$$R = R_1 \sqcup \cdots \sqcup R_n$$

and define

$$E_n(\{R_i\}) = \sum_i \Big(\iint (g - \operatorname{mean}_{R_i} g)^2 \, dx\, dy + \tfrac{1}{2}\nu_0 |\partial R_i| \Big).$$

For this to make sense, R_i must be a so-called *Cacciopoli set*, a measurable set whose boundary as a current ∂R_i has finite length, $|\partial R_i|$. Then, for each n, we show that E_n attains its minimum, and that at this minimum, R_i is an open set (up to a measure zero set) with piecewise C^2 boundary. We show finally that as n increases, the minimum eventually increases if all R_i have positive measure. How should Cacciopoli sets be visualized? Note that the boundaries of Cacciopoli sets are *not* fractals: their Hausdorff dimension is 1 since their length is finite, but they will, in general, be made up of infinitely many rectifiable arcs with finite total length. A good way to imagine what Cacciopoli sets are like is to look at the segmentation of the world into sea and land. Some coastlines are best seen as fractals: Richardson's data on the indefinite growth of the length of rocky coasts with the scale of the yardstick was one of the inspirations of Mandelbrot's theory of fractals. But other coasts are mixtures of smooth and ragged parts, and may serve as models (see Figure 6).

In Section 6 we study E for large μ. The first step is to use Green's theorem and rewrite E as an integral only over Γ: let g_μ and f_Γ be the solutions of

$$\Delta g_\mu = \mu^2 (g_\mu - g), \qquad \Delta f_\Gamma = \mu^2 (f_\Gamma - g),$$

$$\partial g_\mu / \partial n|_{\partial R} \equiv 0, \qquad \partial f_\Gamma / \partial n|_{(\Gamma \cup \partial R)} \equiv 0.$$

Then we show that

$$E(f_\Gamma, \Gamma) = E(g_\mu, \phi) + \int_\Gamma \Big[\nu - \frac{\partial g_\mu}{\partial n} (f_\Gamma^+ - f_\Gamma^-) \Big] \, ds,$$

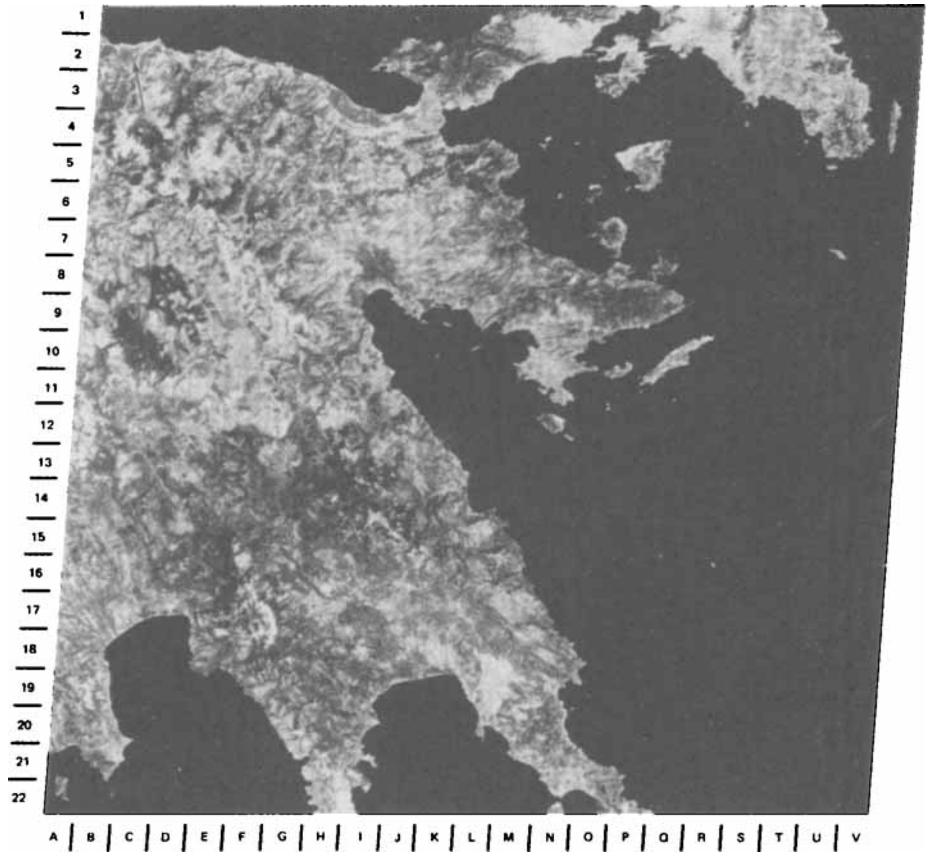

Figure 6. A Cacciopoli-like coastline.

where $f_\Gamma^\pm$ are the boundary values of f_Γ along Γ and $\partial/\partial n$ points from the $-$ side of Γ to the $+$ side. Moreover, if $1/\mu$ is small, we prove that, away from the singularities of Γ,

$$f_\Gamma^+ - f_\Gamma^- = \frac{2}{\mu}\frac{\partial g_\mu}{\partial n} + O\left(\frac{1}{\mu^2}\right),$$

while uniformly on R

$$f_\Gamma - g = O\left(\frac{1}{\mu}\right),$$

from which we deduce that

$$E(f_\Gamma, \Gamma) = E(g_\mu, \phi) + \frac{2}{\mu}E_\infty(\Gamma) + O\left(\frac{\log\mu}{\mu^2}\right),$$

where $\nu_\infty = \frac{1}{2}\mu\nu$. We then study the first variation of $E_\infty(\Gamma)$. We show firstly that the variation of $E(f_\Gamma, \Gamma)$ also tends to the first variation of $E_\infty(\Gamma)$. Secondly, the equation for the first variation of $E_\infty(\Gamma)$ being zero turns out to be a second-order differential equation for Γ. Let H_g be the matrix of second derivatives of g, let t_Γ and n_Γ be the unit tangent vector and unit normal vector for Γ. Then the equation is

$$(n_\Gamma \cdot \nabla g) \cdot \Delta g + (t_\Gamma \cdot \nabla g) \cdot (t_\Gamma \cdot H_g \cdot n_\Gamma)$$
$$+ \operatorname{curv}(\Gamma) \cdot \left[\tfrac{1}{2}\nu_\infty + \tfrac{1}{2}(n_\Gamma \cdot \nabla g)^2 - (t_\Gamma \cdot \nabla g)^2\right] = 0.$$

(To read this equation properly, note that ∇g, Δg and H_g play the role of coefficients, t_Γ and n_Γ are 1-st derivatives of the solution curve and curv(Γ) is its 2-nd derivative.)

Finally, in Section 7, we look briefly at this equation, noting that like the equation for geodesics in a Lorentz metric, it has two types of local solutions: "space-like" solutions which locally minimize E_∞ and "time-like" solutions which locally maximize E_∞. A general solution flips back and forth between the two types, with cusps marking the transition. Discussion of existence theorems for solutions of this differential equation is postponed to a later paper.

We use the following standard notations throughout this paper:

(a) For $k \geqq 0$ an integer and $1 \leqq p \leqq \infty$, $W_p^k(D)$ is the Banach space of functions f on D with norm

$$\|f\|_{k,p,D} = \left[\sum_{\substack{\alpha,\beta \geqq 0 \\ \alpha+\beta \leqq k}} \iint_D \left|\frac{\partial^2 f}{\partial x^\alpha \, \partial y^\beta}\right|^p dx\,dy\right]^{1/p}.$$

(b) C^k is the class of functions with continuous derivatives through order k, and $C^{k,1}$ of those whose k-th derivatives satisfy a Lipshitz condition. The boundary of a domain D is said to be in one of these classes if D is represented locally by $y < f(x)$ (or $y > f(x)$ or $x < f(y)$ or $x > f(y)$), where f belongs to the class.

2. The First Variation

As described in Section 1, we fix a region R in the plane with compact closure and piecewise smooth boundary and we fix a continuous function g on $\bar{R}$. We also fix positive constants μ and ν. We consider next a subset $\Gamma \subset R$ made up of a finite set $\{\gamma_\alpha\}$ of curves. We shall assume for our analysis initially that the γ_α are simple $C^{1,1}$-curves meeting ∂R and meeting each other only at their endpoints. Finally, we consider a function f on $R - \Gamma$ which we take initially to be

C^1 with first derivatives continuous up to all boundary points.[2] For each f and Γ, we have the functional E defined in Section 1:

$$E(f, \Gamma) = \mu^2 \iint_R (f - g)^2 \, dx \, dy + \iint_{R-\Gamma} \|\nabla f\|^2 \, dx \, dy + \nu|\Gamma|.$$

The fundamental problem is to find f and Γ which minimize the value of E. Note that by a scaling in the coordinates x, y in R and by a multiplicative constant in the functions f, g, we can transform E with any constants μ, ν into the E for any other set of constants μ, ν. Put another way, if μ is measured in units of inverse distance in the R- plane, and the constant ν is measured in units of (size of $g)^2$/distance, then the three terms of E all have the same "dimension". So fixing μ, ν is the same as fixing units of distance in R and of size of g. The purpose of this section is to study the effect of small variations of f and Γ on E and determine the condition for the first variation of E to be 0. These are, of course, necessary conditions for f and Γ to minimize E.

The first step is to fix Γ as well as g and vary f. This is a standard variational problem. Let δf represent a function of the same type as f. Then

$$\begin{aligned} &E(f + t\,\delta f, \Gamma) - E(f, \Gamma) \\ &\quad = t\left[\mu^2 \iint 2\,\delta f \cdot (f - g)\, dx\, dy + \iint 2(\nabla(\delta f) \cdot \nabla f)\, dx\, dy\right] \\ &\qquad + t^2\left[\mu^2 \iint (\delta f)^2\, dx\, dy + \iint \|\nabla(\delta f)\|^2\, dx\, dy\right]. \end{aligned}$$

Thus

$$\begin{aligned} \frac{\delta E}{\delta f}(f, \Gamma) &= \lim_{t \to 0} \frac{E(f + t\,\delta f, \Gamma) - E(f, \Gamma)}{t} \\ &= 2\left[\mu^2 \iint \delta f \cdot (f - g)\, dx\, dy + \iint (\nabla(\delta f) \cdot \nabla f)\, dx\, dy\right]. \end{aligned}$$

Integrating by parts and using Green's theorem[3] we obtain

$$\begin{aligned} \frac{1}{2}\frac{\delta E}{\delta f}(f, \Gamma) &= \mu^2 \iint \delta f \cdot (f - g)\, dx\, dy - \iint \delta f \cdot \nabla^2 f\, dx\, dy + \int_B \delta f \frac{\partial f}{\partial n}\, ds \\ &= \iint \delta f\left(\nabla^2 f - \mu^2(f - g)\right) dx\, dy + \int_B \delta f \frac{\partial f}{\partial n}\, ds, \end{aligned}$$

[2] We shall weaken these conditions later to include certain f's with more complicated boundary behavior, but this will not affect the initial arguments. Appendix 1 contains a full discussion of how regular the minimizing functions f will be at different points of R.

[3] For Green's theorem to be applicable, we must assume that $f \in W_p^2$, some $p \geq 1$. Alternatively, by the last equation, a minimum f is a *weak* solution of $\nabla^2 f = \mu^2(f - g)$ and, by the results in Appendix 1, f in fact has L^p second derivatives locally everywhere in $R - \Gamma$.

where B is the entire boundary of $R - \Gamma$, i.e., ∂R and each side of Γ. Now taking δf to be a "test function", non-zero near one point of $R - \Gamma$, zero elsewhere, and taking the limit over such δf, we deduce as usual that f satisfies

$$(*) \qquad \nabla^2 f = \mu^2(f - g)$$

on $R - \Gamma$. Now taking δf to be non-zero near one point of B, zero elsewhere, we also see that

$$(**) \qquad \frac{\partial f}{\partial n} = 0 \text{ on } \partial R, \text{ and on the two sides } \gamma_\alpha^\pm \text{ of each } \gamma_\alpha.$$

As is well known, this determines f uniquely: the operator $\mu^2 - \nabla^2$ is a positive-definite selfadjoint operator, and there is a Green's function $K(x, y; u, v)$ which is C^∞ except on the diagonal $(x, y) = (u, v)$ where it has a singularity like

$$-\frac{1}{2\pi}\log\left(\mu\sqrt{(x-u)^2 + (y-v)^2}\right)$$

such that, for each component R_i of $R - \Gamma$, and for each g,

$$f(x, y) = \mu^2 \cdot \iint_{(u,v)\in R_i} K(x, y; u, v)\, g(u, v)\, du\, dv$$

satisfies $(*)$ and $(**)$ on R_i. Further regularity results for this boundary value problem especially on the behavior of f near Γ are sketched in Appendix 1 and will be quoted as needed.

If there were no boundaries at all, then Green's function for this operator and the whole plane is obtained from the so-called "modified Bessel's function of the second kind" K_0 (cf. Whittaker and Watson [14], Section 17.71, pp. 373–374):

$$K(x, y; u, v) = \frac{1}{2\pi}K_0\left(\mu\sqrt{(x-a)^2 + (y-v)^2}\right),$$

where $K_0(r)$ (defined for $r > 0$) is the solution of

$$K_0'' + \frac{1}{r}K_0' - K_0 = 0$$

such that

$$K_0(r) \sim \log(1/r) \quad \text{for } r \text{ small},$$

$$K_0(r) \sim \sqrt{\pi/2r}\, e^{-r} \quad \text{for } r \text{ large}.$$

Its graph is depicted in Figure 7.

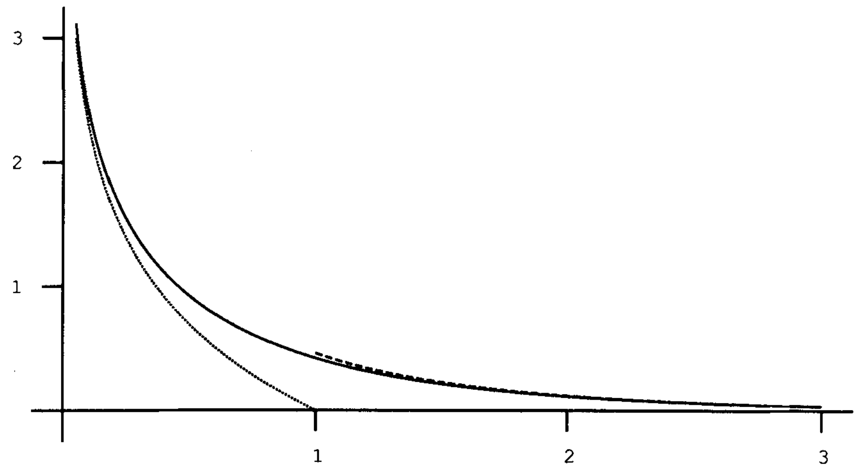

Figure 7. Graph of the Green's function and its asymptotes.

Another way of characterizing K is that it is the Fourier transform of

$$L(\xi, \eta) = 1/(\mu^2 + \xi^2 + \eta^2)$$

evaluated at $(x - u, y - v)$.

Now look at the variation of E with respect to Γ. First consider moving Γ near a simple point P of Γ. Such a P lies on exactly one γ_α and, as γ_α is $C^{1,1}$, we can write it in a small neighborhood U of P either as a graph $y = h(x)$, or as a graph $x = h(y)$. Assume the former; then we can deform γ_α to

$$y = h(x) + t \cdot \delta h(x),$$

where δh is zero outside a small neighborhood of $x(P)$ (see Figure 8). If t is small, the new curve $\gamma_\alpha(t)$ does not meet any other γ_β's except at its endpoints and

$$\Gamma(t) = \gamma_\alpha(t) \cup \bigcup_{\beta \neq \alpha} \gamma_\beta$$

is an allowable deformation of Γ. Now we cannot speak of leaving f *fixed* while Γ moves, because f must be C^1 on $R - \Gamma$ but will usually be discontinuous across Γ. Instead do this: Let f^+ denote the function f in

$$U^+ = \{(x, y) | y > h(x)\} \cap U$$

and let f^- denote f in

$$U^- = \{(x, y) | y < h(x)\} \cap U$$

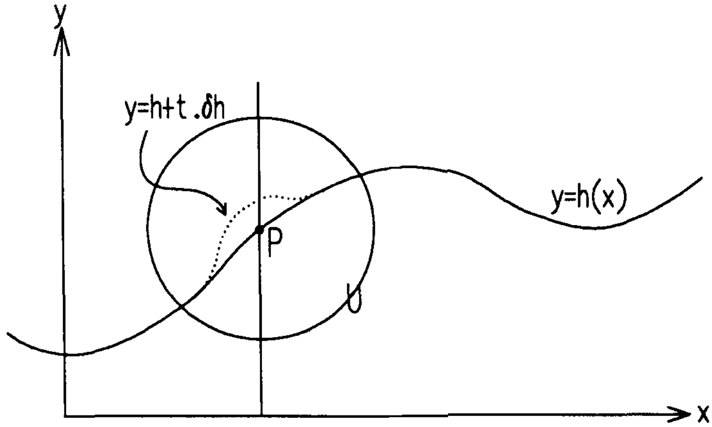

Figure 8. Deformation of Γ.

and choose C^1 extensions of f^+ from U^+ to U and of f^- from U^- to U. By the results in Appendix 1, f is C^1 on both sides of Γ at all simple points of Γ, hence this is possible. Then for any small t, define

$$f^t(P) = \begin{cases} f(P) & \text{if } P \notin U, \\ \text{extension of } f^+ & \text{if } P \in U,\ P \text{ above } \gamma_\alpha(t), \\ \text{extension of } f^- & \text{if } P \in U,\ P \text{ below } \gamma_\alpha(t). \end{cases}$$

Now

$$\begin{aligned} E(f^t, \Gamma(t)) - E(f, \Gamma) &= \mu^2 \iint_U \left[(f^t - g)^2 - (f-g)^2\right] dx\,dy \\ &\quad + \iint_{U - \Gamma(t)} \|\nabla f^t\|^2\, dx\,dy - \iint_{U-\Gamma} \|\nabla f\|^2\, dx\,dy \\ &\quad + \nu\left[|\gamma_\alpha(t)| - |\gamma_\alpha|\right] \\ &= \mu^2 \int\left(\int_{h(x)}^{h(x)+t\,\delta h(x)} \left[(f^- - g)^2 - (f^+ - g)^2\right] dy\right) dx \\ &\quad + \int\left(\int_{h(x)}^{h(x)+t\,\delta h(x)} \left[\|\nabla f^-\|^2 - \|\nabla f^+\|^2\right] dy\right) dx \\ &\quad + \nu \int\left[\sqrt{1 + (h + t\,\delta h)'^2} - \sqrt{1 + h'^2}\right] dx; \end{aligned}$$

hence

$$\frac{\delta E}{\delta \gamma} = \mu^2 \int \left[(f^- - g)^2 - (f^+ - g)^2 \right] \Big|_{y=h(x)} \delta h \, dx$$
$$+ \int \left[\|\nabla f^-\|^2 - \|\nabla f^+\|^2 \right] \Big|_{y=h(x)} \delta h \, dx$$
$$+ \nu \int \frac{h'}{\sqrt{1+h'^2}} (\delta h)' \, dx.$$

Since γ_α is in $C^{1,1}$, it has a curvature almost everywhere which is bounded and is given by

$$\operatorname{curv}(\gamma_\alpha) \quad \text{at} \quad (x, h(x)) = \frac{h''(x)}{\left(1 + h'(x)^2\right)^{3/2}},$$

where curv > 0 means that γ_α is convex upwards (i.e., U_+ is convex, U_- is not) while curv < 0 means γ_α is convex downwards (i.e., U_- is convex, U_+ is not). Integrating by parts, we get

$$\frac{\delta E}{\delta \gamma} = \int_{\gamma_\alpha} \left[\left(\mu^2 (f^- - g)^2 + \|\nabla f^-\|^2 \right) - \left(\mu^2 (f^+ - g)^2 + \|\nabla f^+\|^2 \right) \right] \cdot \frac{\delta h}{\sqrt{1+h'^2}} \, ds$$
$$- \nu \operatorname{curv}(\gamma_\alpha)$$

(ds = infinitesimal arc length on γ_α). Note that the coefficient $\delta h / \sqrt{1+h'^2}$ has the simple interpretation as the amount of the displacement of the deformed Γ along the *normal* lines to the original Γ. Since this formula for $\delta E/\delta\gamma$ holds for any δh, at an extremum of E we deduce that, *along each* γ_α,

$$(* * *) \qquad \left[\mu^2 (f^+ - g)^2 + \|\nabla f^+\|^2 \right] - \left[\mu^2 (f^- - g)^2 + \|\nabla f^-\|^2 \right]$$
$$+ \nu \cdot \operatorname{curv}(\gamma_\alpha) = 0.$$

The terms in square brackets can be simply interpreted as the *energy density* corresponding to the functional E just above and just below the curve. In fact, write the energy density as e:

$$e(f; x, y) = \mu^2 \left(f(x,y) - g(x,y) \right)^2 + \| \nabla f(x,y) \|^2;$$

then $(* * *)$ says

$$(* * *') \qquad e(f^+) - e(f^-) + \nu \cdot \operatorname{curv}(\gamma_\alpha) = 0 \quad \text{on} \quad \gamma_\alpha.$$

(Note that if we fix f^+ and f^-, then what we have here is a 2-nd order ordinary differential equation for γ_α.)

Next, we look at special points P of Γ:

(a) points P where Γ meets the boundary of the region R,
(b) "corners" P where two γ_α's meet,
(c) "vertices" P where three or more γ_α's meet,
(d) "crack-tips" P where a γ_α ends but does not meet any other γ_β or ∂R.

Before analyzing the restrictions at such points implied by the stationarity of E, we should reconsider what assumptions may reasonably be placed on f. We saw above that for each fixed Γ, there was a unique minimizing f: the f satisfying the elliptic boundary value problem ($*$), ($**$). If Γ is $C^{1,1}$ and g is continuous, then the standard theory of elliptic operators sketched in Appendix 1 implies that f is C^1 on the open set $R - \Gamma$ and that, at all simple boundary points of Γ and ∂R, f extends locally to a C^1-function on the region plus its boundary. This shows that ($***$) certainly holds along each γ_α. However, at *corners*, p is not always such a nice function. In fact, the typical behavior of f at a corner point with angle α is shown in the following example:

Let z, w be complex variables and consider the conformal map $w = z^{\pi/\alpha}$ between R_1 and R_2 (see Figure 9). Let $z = re^{i\theta}$ and let

$$\begin{aligned} f(z) &= \mathscr{I}m(w) \\ &= r^{\pi/\alpha}\sin\left(\frac{\pi}{\alpha}\theta\right). \end{aligned}$$

Then f is a real-valued function on R_1 with

$$\nabla^2 f = 0,$$

$$\frac{\partial f}{\partial n} = 0 \quad \text{on} \quad \partial R_1.$$

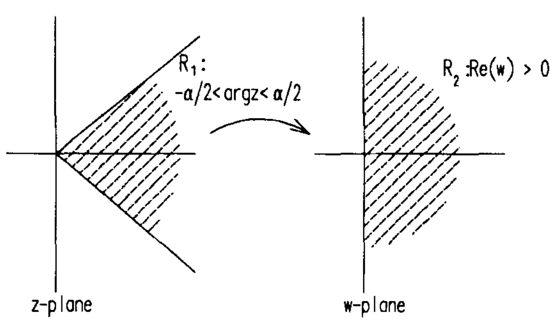

Figure 9. Straightening of a corner by a conformal map.

But note that

$$\frac{\partial f}{\partial r} = \frac{\pi}{\alpha} r^{\pi/\alpha - 1} \sin\left(\frac{\pi}{\alpha}\theta\right)$$

has an infinite limit as $r \to 0$ if $\alpha > \pi$.

V. A. Kondratiev [9] discovered that this type of singularity characterizes what happens to the solutions of elliptic boundary value problems at corners or even in "conical singular points" of n-dimensional domains. We have collected the results we shall use in Appendix 1, where it is shown that

(a) f is bounded everywhere,

(b) f is C^1 at corners P of $R - \Gamma$ whose angle α satisfies $0 < \alpha < \pi$,

(c) at corners P such that $\pi < \alpha \leqq 2\pi$, including the exterior of crack (i.e., P is the endpoint of a $C^{1,1}$-arc which is not continued by any other arc), f has the form

$$f = cr^{\pi/\alpha} \sin\left(\frac{\pi}{\alpha}(\theta - \theta_0)\right) + \tilde{f},$$

where $\tilde{f}$ is C^1, (r, θ) are polar coordinates centered at P and c, θ_0 are suitable constants.

Using these estimates, let us first show that if $E(f, \Gamma)$ is minimum, then Γ has no kinks, i.e., points P where two edges γ_i and γ_j meet at an angle other than π. To do this, let U be a disc around P of radius ε and let $\gamma_i \cup \gamma_j$ divide U into sectors U^- with angle $\alpha^- > \pi$ and U^+ with angle $\alpha^+ < \pi$ (see Figure 10). We assume $0 < \alpha^+ < \pi$ and hence $\pi < \alpha^- < 2\pi$. Fix a C^∞-function $\eta(x, y)$ such

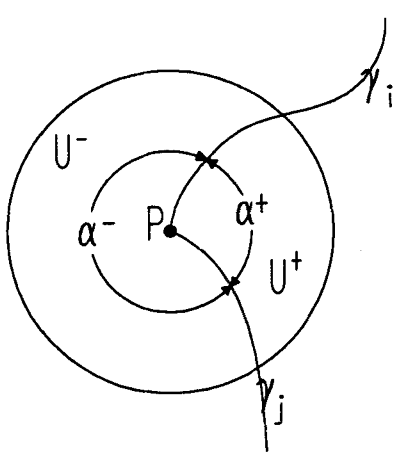

Figure 10. A corner in Γ.

that $0 \leqq \eta \leqq 1$ and

$$\eta(x, y) = \begin{cases} 1 & \text{if } x^2 + y^2 \geqq 1, \\ 0 & \text{if } x^2 + y^2 \leqq \frac{1}{2} \end{cases}$$

and let η_U denote η "adapted" to U, i.e.,

$$\eta_U(x, y) = \eta\left(\frac{x - x(P)}{\varepsilon}, \frac{y - y(P)}{\varepsilon}\right)$$

(so that $\eta_U = 1$ outside U, $\eta_U = 0$ in the concentric disc U' of radius $\frac{1}{2}\varepsilon$).

What we will do is to "cut" the corner at P at a distance $\frac{1}{2}\varepsilon$ shrinking U^+ and expanding U^-; define a new f on the new U^+ by restriction; and define a new f on the new U^- by $f^-(0) + \eta_U(f^- - f^-(0))$. Let us estimate the change in $E(f, \Gamma)$. To make this estimate we can assume that $f^-(0) = 0$ by replacing f by $f - f^-(0)$, g by $g - f^-(0)$, which does not affect E. Then we have:

(a) Change in 1-st term:

$$\iint_{U^-} (\eta_U f^- - g)^2 - \iint_{U^-} (f^- - g)^2$$

$$= \iint_{U^-} \frac{1 - \eta_U}{1 + \eta_U} g^2 - (1 - \eta_U) \cdot \left(\sqrt{1 + \eta_U}\, f^- - \frac{g}{\sqrt{1 + \eta_U}}\right)^2$$

$$\leqq \iint_{U^-} g^2$$

$$\leqq c_1 \varepsilon^2.$$

(b) Change in 2-nd term:

$$\iint_{U^-} \|\nabla(\eta_U f^-)\|^2 - \iint_{U^-} \|\nabla f^-\|^2$$

$$= \iint_{U^-} (f^-)^2 \|\nabla \eta_U\|^2 + 2\eta_U f^- (\nabla \eta_U \cdot \nabla f^-).$$

But $\|\nabla \eta_U\| \leqq c/\varepsilon$. Moreover, the estimates just discussed show that

$$f^- = O(r^{\pi/\alpha^-}),$$

$$\|\nabla f^-\| = O(r^{(\pi/\alpha^-)-1}).$$

Taking into account that $\nabla \eta_U = 0$ outside the annulus between ε, $\frac{1}{2}\varepsilon$, we can

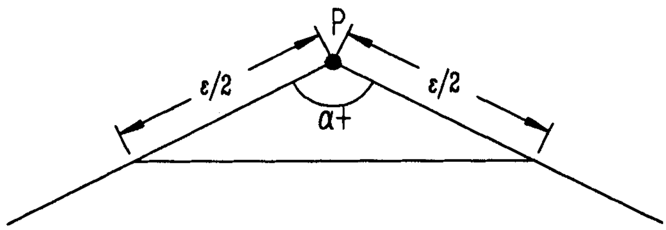

Figure 11. "Cutting" a corner on Γ.

bound the above expression:

$$\leqq \frac{c^2}{\varepsilon^2} \sup_{\text{Annulus}} (f^-)^2\left(\tfrac{3}{4}\pi\varepsilon^2\right) + \frac{2c}{\varepsilon} \sup_{\text{Annulus}} \left(|f^-|\,\|\nabla f^-\|\right)\left(\tfrac{3}{4}\pi\varepsilon^2\right)$$

$$\leqq c_1 \frac{1}{\varepsilon^2} \cdot \varepsilon^{2\pi/\alpha^-} \cdot \varepsilon^2 + c_2 \frac{1}{\varepsilon} \cdot \varepsilon^{(\pi/\alpha^- + \pi/\alpha^- - 1)} \cdot \varepsilon^2$$

$$= c_3 \varepsilon^{2\pi/\alpha^-}.$$

(c) Change in 3-rd term:

Asymptotically, we are replacing the equal sides of an isosceles triangle with angle α^+ at apex by the 3-rd side (see Figure 11) which decreases the length term in E from ε to $\varepsilon \sin(\frac{1}{2}\alpha^+)$.

Note that terms 1 and 2 increase by order ε^2 and $\varepsilon^{2\pi/\alpha^-}$, respectively (and $2\pi/\alpha^- > 1$), while term 3 decreases by order ε which shows that if ε is small, E decreases.

The same argument gives us more results. Consider next points P where an arc γ_i of Γ meets the boundary of the region R where the problem is posed. At a minimum of E, we claim that γ_i must meet ∂R perpendicularly. Let U be the intersection of a disc around P of radius ε and R, and suppose instead that γ_i divides U into a sector U^+ with angle $\alpha^+ < \frac{1}{2}\pi$ and a sector U^- with angle $\alpha^- > \frac{1}{2}\pi$ (see Figure 12). We modify Γ and f as follows: replace γ_i by a new curve which, when it hits U, has a corner and follows a straight line meeting ∂R perpendicularly; replace f^+ by its restriction, and f^- by $f^-(P) + \eta_U \cdot (f - f^-(P))$, extended by $f^-(P)$. The above estimates show that the first two terms go up by order at most ε^2, while the 3-rd decreases by order ε, hence E decreases.

Look next at triple points P where three arcs γ_i, γ_j, γ_k meet with angles α_{ij}, α_{jk}, α_{ki} (see Figure 13). We claim that, at a minimum of E, $\alpha_{ij} = \alpha_{jh} = \alpha_{ki} = \frac{2}{3}\pi$. If not, then one of the three angles is smaller than $\frac{2}{3}\pi$: say $\alpha_{jk} < \frac{2}{3}\pi$ as in the diagram. Then modify Γ by extending γ_i along the bisector of angle α_{jk} and joining it by straight lines to the points where γ_j and γ_k first meet U (see Figure 13). The new triple point is chosen so that its three angles are each $\frac{2}{3}\pi$. A little trigonometry will show that the old segments of γ_j and γ_k of total length

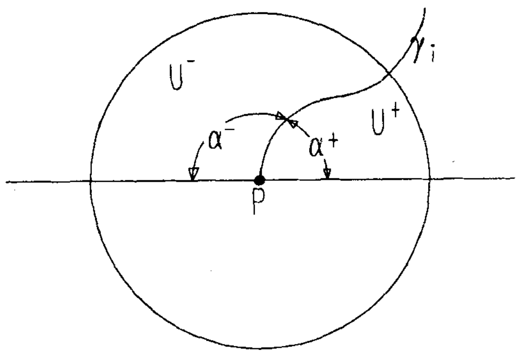

Figure 12. Γ meets the boundary of R.

asymptotically equal to 2ε are then replaced by three new straight segments of total length exactly

$$2\sin\left(\tfrac{1}{2}\alpha_{jk} + \tfrac{1}{6}\pi\right) \cdot \varepsilon$$

which is clearly less than 2ε if $\alpha_{jk} < \frac{2}{3}\pi$. This linear decrease in length is greater than the increase in the 1-st and 2-nd terms of E by the same argument as before.

Finally, we can apply this argument to show that at a minimum of E there are no points where four or more γ_i meet at positive angles; simply choose the smallest angle α at the intersection. Then $\alpha \leqq \frac{1}{2}\pi$ and replace the 4-fold intersection by two 3-fold intersections as in Figure 14.

In all of this, we have ignored "cuspidal corners", i.e., corners where the two arcs γ_i and γ_j are tangent. Our argument breaks down for cusps, but we can show that these do not occur by another argument. We cut off the end of one of

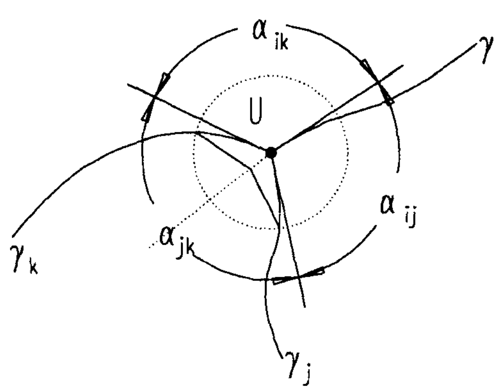

Figure 13. Modifying a triple point in Γ.

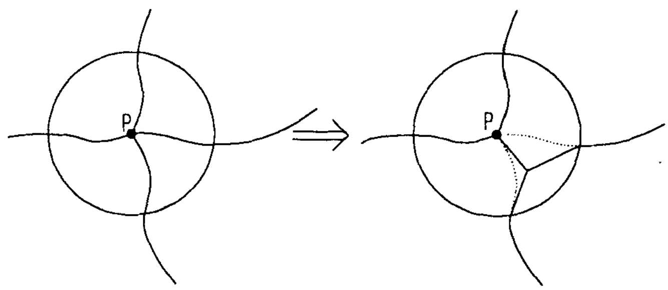

Figure 14. Splitting a 4-fold point into 2 triple points.

the arcs γ_i, join it to the nearest point of γ_j and leave the end of γ_j as a "crack-tip" (see Figure 15). To extend f^- from the old U^- to the new bigger $U^- = U -$ crack γ_j, we simply extend it from γ_i by asking that f^- be constant on circles with center P. Again the length decreases by order ε. To check the increase in the 1-st and 2-nd terms, note the estimate on f^-:

$$f^- = f^{-1}(0) + O(\sqrt{r}),$$
$$\|\nabla f^-\| = O\left(\frac{1}{\sqrt{r}}\right).$$

Write γ_i in polar coordinates as $\theta = \theta_i(r)$ and γ_j by $\theta = \theta_j(r)$, so that $\theta_i - \theta_j = O(r)$. Thus the new part of the second term is estimated like this:

$$\int_0^\varepsilon \int_{\theta_i}^{\theta_j} \|\nabla f^-\|^2 r\, d\theta\, dr \leqq \int_0^\varepsilon \frac{c}{r} \cdot r \cdot (\theta_i(r) - \theta_j(r))\, dr$$
$$\leqq c_1 \varepsilon^2.$$

This proves:

THEOREM 2.1. *If (f, Γ) is a minimum of E such that Γ is a finite union of simple $C^{1,1}$-curves meeting ∂R and meeting each other only at their endpoints, then*

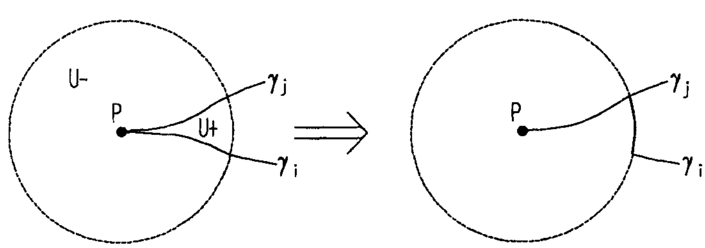

Figure 15. Eliminating cuspidal corners in Γ.

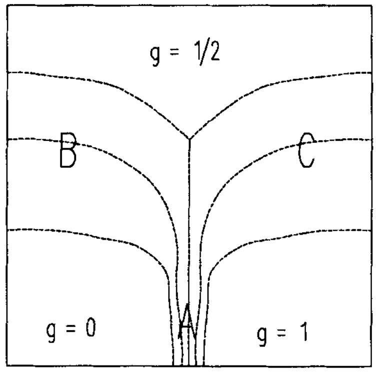

Figure 16. An image leading to a Γ with a crack.

the only vertices of Γ *are*:

(a) *points* P *on the boundary of* R *where one* γ_i *meets* ∂R *perpendicularly*,
(b) *triple points* P *where three* γ_i *meet with angles* $\frac{2}{3}\pi$,
(c) *crack-tips* P *where a* γ_i *ends and meets nothing.*

3. Cuspidal Crack Tips

Considering the original variational problem, we would first like to argue heuristically that "cracks", i.e., arcs in Γ that end at a point P without any continuation, are likely to occur in minima of E. Take g to be a function of the type shown in Figure 16. If ν is set carefully, it will "pay" in terms of decreasing E to have an arc of Γ in strip A separating the $g \equiv 0$ and $g \equiv 1$ regions but it will not pay to have arcs of Γ in strips B and C: this is because g has been concocted so that its gradient is much larger in strip A than in B or C; so putting Γ along A saves a big penalty in $\iint \|\nabla f\|^2$ but putting Γ along B or C does not. We expect that in this case the optimal Γ will run along A to a crack-tip somewhere in the middle.

However, if $\|\nabla g\|$ is somewhat larger in B than C, the crack will be expected to bend to the left into B at its tip. Consider the rule which determines when Γ is in balance between bending left or right:

$$(*) \qquad \left[(f_+ - g)^2 + \|\nabla f_+\|^2\right] - \left[\mu^2(f_- - g)^2 + \|\nabla f_-\|^2\right] = \nu \cdot \operatorname{curv}(\Gamma).$$

At the crack-tip itself, $(f_\pm - g)^2$ is bounded, but $\|\nabla f_\pm\|^2$ is not. $\|\nabla f_\pm\|^2$ grows like $1/r$, where r is the distance to the crack-tip. Some cancellation takes place in $\|\nabla f_+\|^2 - \|\nabla f_-\|^2$. Suppose for instance that the crack-tip is the positive x-axis and f is approximated by

$$f_{\pm} = \pm c\,\mathcal{R}e\sqrt{z} + b \cdot x + \cdots$$

(which satisfies $(\partial f/\partial n)_{\pm} = \partial f/\partial y = 0$). Then

$$\|\nabla f_{+}\|^2 - \|\nabla f_{-}\|^2 = 2bc/\sqrt{r} + \cdots$$

which still grows unboundedly near the crack-tip. Thinking of $(*)$ as an O.D.E., one would expect that this growth would, in general, force curv(Γ) to grow like $1/\sqrt{r}$ also. Curves that do this are curves with *cusps* at their endpoints:

$$\Gamma\colon\ y = c \cdot x^{3/2} + \cdots, \qquad x \geqq 0,$$

where

$$\operatorname{curv}(\Gamma) = \frac{y''}{(1 + y'^2)^{3/2}} = \frac{3c}{4\sqrt{x}} + \cdots.$$

For this reason, it seems logical to expect that crack-tips have cusps at their end. To confirm that such Γ were *consistent*, we have made a careful calculation to produce functions f and g and a cuspidal crack Γ satisfying the variational equations $(*)$ and $\Delta f = \mu^2(f - g)$. In outline this goes as follows.

We work backwards:

(i) Start with Γ given by

$$y = x^{3/2} + \lambda_1 x^2 + \lambda_2 x^{5/2}, \qquad x \geqq 0,$$

where λ_1, λ_2 will be chosen later.

(ii) Choose f of the form

$$f = f_{\text{sing}} + f_{\text{smooth}},$$

where

$$f_{\text{sing}} = 2\sqrt{\nu}\,\frac{\sinh(\mu r)}{\mu\sqrt{r\pi}}\cos\left(\tfrac{1}{2}\theta\right) + \alpha\sqrt{\nu}\,\frac{3\mu r\cosh(\mu r) - 3\sinh(\mu r)}{\mu^2 (r)^{3/2}} \cdot \sin\left(\tfrac{3}{2}\theta\right)$$

α will be chosen later, and f_{smooth} is a C^∞-function. It is easy to verify that

$$\nabla^2 f_{\text{sing}} = \mu^2 \cdot f_{\text{sing}}.$$

(iii) Choose

$$g = f_{\text{smooth}} - \mu^{-2} \cdot \nabla^2 f_{\text{smooth}}.$$

Then by the choice of f and g, it is clear that

$$\Delta f = \mu^2 (f - g).$$

Also, the leading term of f is

$$\frac{2\sqrt{\nu}}{\sqrt{\pi}} \sqrt{r} \cos\left(\tfrac{1}{2}\theta\right);$$

hence the extra variational condition, Griffith's law of cracks, derived later in this section is satisfied. We must check that for suitable λ_1, λ_2, α and suitable f smooth, we also get

(a) $\partial f / \partial n|_\Gamma \equiv 0$,

(b) $e_+ - e_- = \nu \cdot \operatorname{curv}(\Gamma)$.

Taking into account that

$$f_+ = f_{\text{smooth}} + f_{\text{sing}},$$

$$f_- = f_{\text{smooth}} - f_{\text{sing}},$$

and the definition of g, (b) simplifies to

$$\text{(b}'\text{)} \qquad 4\mu^{-2} \nabla^2 f_{\text{smooth}} \cdot f_{\text{sing}} + 4 \frac{\partial f_{\text{smooth}}}{\partial t} \cdot \frac{\partial f_{\text{sing}}}{\partial t} = \nu \cdot \operatorname{curv}(\Gamma)$$

(where $\partial / \partial t$ is the directional derivative along Γ). The construction of f_{smooth} is rather tedious and the values of the constants messy, but in outline, we may see that a solution is *possible* like this: Parametrize Γ by x. Then $(1/\sqrt{\nu})(\partial/\partial n) f_{\text{sing}}$ is a power series in $\sqrt{x}$ beginning with

$$-\sqrt{\frac{1}{\pi}} - \tfrac{3}{2}\lambda_1 \sqrt{\frac{x}{\pi}} + \cdots$$

and $\operatorname{curv}(\Gamma)$ is a Laurent series in $\sqrt{x}$ beginning with

$$\frac{3}{4\sqrt{x}} + 2\lambda_1 + \left(\tfrac{15}{4}\lambda_2 - \tfrac{81}{32}\right)\sqrt{x} + \cdots$$

and $(1/\sqrt{\nu})(\partial/\partial t) f_{\text{sing}}$ is a Laurent series in $\sqrt{x}$ beginning with

$$\sqrt{\frac{1}{\pi x}} - \frac{3}{4}\sqrt{\frac{x}{\pi}} + \cdots.$$

So we must solve the partial differential equations

$$\text{(a)} \qquad \frac{1}{\sqrt{\nu}} \frac{\partial}{\partial n} f_{\text{smooth}} = \sqrt{\frac{1}{\pi}} + \frac{3}{2}\lambda_1 \sqrt{\frac{x}{\pi}} + \cdots ,$$

$$\text{(b}'\text{)} \qquad \frac{4\mu^{-2}}{\nu} \nabla^2 f_{\text{smooth}} \left(\frac{2\sqrt{\nu}}{\sqrt{\pi}} x + \cdots \right) + \frac{4}{\nu} \frac{\partial f_{\text{smooth}}}{\partial t} \left(\sqrt{\frac{\nu}{\pi}} + \cdots \right)$$
$$= \left(\tfrac{3}{4} + 2\lambda_1 \sqrt{x} + \cdots \right).$$

Looking at low-order coefficients, it turns out that f_{smooth} has three too few coefficients: in fact, if

$$f_{\text{smooth}} = \frac{3\sqrt{\pi\nu}}{16} x + \sqrt{\frac{\mu\nu}{\pi}}\, y + \text{higher order terms},$$

then the constant terms in (a) and (b′) are OK, but to make the $\sqrt{x}$-coefficients cancel we must set $\lambda_1 = 3/\pi$ and $\alpha = 3(1 + \frac{1}{16}\pi^2)/\mu \cdot \pi^{3/2}$. Likewise the three quadratic coefficients in f_{smooth} and λ_2 are needed to make the x- and $x^{3/2}$-coefficients cancel. Thereafter the four terms x^n, $x^{n-1}y$, $x^{n-2}y^2$ and $x^{n-3}y^3$ in f_{smooth} are sufficient to satisfy (a) and (b′) mod x^n. To see that convergence is not a problem, note that as soon as $f^{(0)}_{\text{smooth}}$ satisfies (a) and (b) mod x^3, we can set

$$f_{\text{smooth}} = f^{(0)}_{\text{smooth}} + \left(y^2 - x^3\right) \cdot \left(f_1(x) + y f_2(x)\right)$$
$$+ \left(y^2 - x^3\right)^2 \cdot \left(f_3(x) + y f_4(x)\right).$$

Substituting in (a), we solve for unique f_1, f_2 by dividing:

$$f_1(x) + \left(x^{3/2} + \lambda_1 x^2 + \lambda_2 x^{5/2}\right) f_2(x) = \frac{\left[\sqrt{1/\pi} + \cdots - (\partial/\partial n) f^{(0)}_{\text{smooth}}\right]}{(\partial/\partial n)(y^2 - x^3)\big|_\Gamma}$$

(noting that $(\partial/\partial n)(y^2 - x^3)|_\Gamma = 2x^{3/2} + \cdots$, so this is OK so long as $x^{3/2}$ divides the numerator). Substituting in (b), we solve similarly for f_3 and f_4, dividing this time by

$$\nabla^2\!\left(\left(y^2 - x^3\right)^2\right)\Big|_\Gamma = 8x^3 + \cdots .$$

In Section 1, we derived extra variational equations at triple points and boundary points of Γ. There is also a new variational equation at crack-tips. It is the analogue of Griffiths' law of cracks well known in solid mechanics; see [12]. The difference is that in our case, cracks can be "sewn" back together as well as extended, hence we get an equality at critical points of E, not just an inequality. The result is this:

THEOREM 3.1. *Let f, Γ be a critical point of E, and assume Γ includes an arc ending at P of the form*

$$y = ax^{3/2} + g(x), \qquad g(x) \in C^2, \quad g(0) = g'(0) = 0$$

in suitable coordinates x, y where a may or may not be zero. Then if, near P,

$$f(x, y) = a_0 + a_1\, \mathcal{R}e\left(\sqrt{x + iy}\right) + \cdots$$

we must have

$$\pi a_1^2 = 2\nu.$$

To derive Griffiths' law of cracks and at the same time check for possible further conditions on a solution (f, Γ) related to sideways perturbation of crack-tips, we use the general technique for deriving the first variation which is employed in geometric measure theory for highly singular Γ's. Let X be a C^∞-vector field on R, tangent to ∂R and let Φ_t be the one-parameter group of diffeomorphisms from R to R obtained by integrating X. We want to compare $E(f \circ \Phi_t, \Phi_t^{-1}(\Gamma))$ with $E(f, \Gamma)$ and, especially, compute

$$\delta_X E(f, \Gamma) = \lim_{t \to 0} \frac{1}{t}\left[E\left(f \circ \Phi_t, \Phi_t^{-1}(\Gamma)\right) - E(f, \Gamma)\right].$$

For simplicity, we shall assume that X is locally constant near crack-tips in Γ.

First, write

$$\begin{aligned}\delta_X E = {} & \mu^2 \lim_{t \to 0} \iint_R \left(\frac{f \circ \Phi_t - f}{t}\right) \cdot (f \circ \Phi_t + f - 2g) \\ & + \lim_{t \to 0} \frac{1}{t}\left[\iint_{R - \Phi_t^{-1}(\Gamma)} \|\nabla f \circ \Phi_t\|^2 - \iint_{R - \Gamma} \|\nabla f\|^2\right] \\ & + \nu \cdot \lim_{t \to 0}\left[\frac{\text{length}\, \Phi_t^{-1}\Gamma - \text{length}\, \Gamma}{t}\right].\end{aligned}$$

Call these terms T_1, T_2 and T_3. To evaluate the first integral T_1, let

$$B_t = \bigcup_{0 \leqq s \leqq t} \Phi_s^{-1}(\Gamma)$$

be the band swept out by Γ while moving from Γ to $\Phi_t^{-1}(\Gamma)$ and let

$$h_t = \begin{cases} (f \circ \Phi_t - f)/t & \text{if } (x, y) \notin B_t, \\ 0 & \text{if } (x, y) \in B_t. \end{cases}$$

We claim

$$\lim_{t \to 0} \iint_R h_t \cdot (f \circ \Phi_t + f - 2g)$$

$$= \iint_R \lim_{t \to 0} h_t \cdot (f \circ \Phi_t + f - 2g)$$

$$= \iint_R Xf \cdot (2f - 2g).$$

This follows from the Lebesgue bounded convergence theorem, plus the estimate:

LEMMA 3.2. *Near a crack-tip P,*

$$|h_t(x, y)| \leqq C/\sqrt{d},$$

where d is distance from (x, y) to the line $\{\Phi_t(P) | t \in \mathbb{R}\}$.

Proof: The leading term in $f \circ \Phi_t$ is

$$c\,\mathcal{R}e\sqrt{z - at},$$

where z is a suitable complex coordinate centered at P, $a \in \mathbb{C}$, $c \in \mathbb{R}$. The corresponding term in h_t is

$$c\,\mathcal{R}e\left(\frac{\sqrt{z - at} - \sqrt{z}}{t}\right) = c\,\mathcal{R}e\frac{1}{\sqrt{z - at} + \sqrt{z}}.$$

But in $R - B_t$,

$$|\sqrt{z - at} + \sqrt{z}| \geqq |\sqrt{z}|$$

(let $z = r_1 \exp\{i\theta_1\}$, $z - at = r_2 \exp\{i\theta_2\}$. Then, if $|\theta_1 - \theta_2| \leqq \pi$,

$$|\sqrt{z - at} + \sqrt{z}|^2 = r_1 + r_2 + \sqrt{r_1 r_2}\cos\left(\tfrac{1}{2}(\theta_1 - \theta_2)\right) \geqq r_1.$$

In $R - B_t$, $|\theta_1 - \theta_2| \leqq \pi$ (see Figure 17). Therefore the leading term in h_t is bounded by $c/\sqrt{|z|}$, hence $c/\sqrt{d}$. The remainder S in f satisfies $S \in o(r^{1-\varepsilon})$,

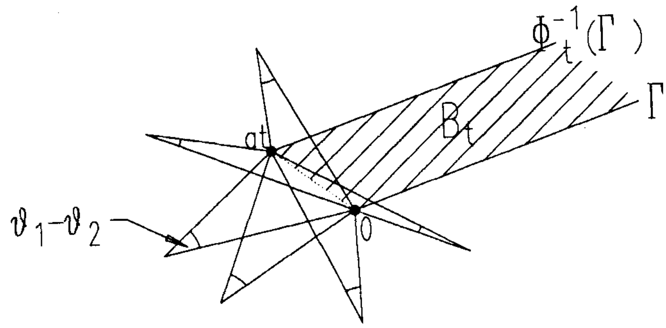

Figure 17. The function $\theta_1 - \theta_2$.

$\|\nabla S\| \in o(r^{-\varepsilon})$ for all $\varepsilon > 0$, hence the remainder S_t in h_t satisfies

$$|S_t(z)| \leqq \sup_{0 \leqq s \leqq t} \|\nabla R\|(z - as) \leqq Cd^{-\varepsilon}.$$

This proves the lemma.

The rest of the first integral is

$$\lim_{t \to 0} \frac{1}{t} \iint_{B_t} (f \circ \Phi_t - f)(f \circ \Phi_t + f - 2g)$$

which equals

$$\int_\Gamma (f_+ - f_-)(f_+ + f_- - 2g)\, ds$$

($f_\pm$ = boundary values of f along Γ).

To evaluate the second integral T_2, let D be a union of discs D_P around all the crack-tips P of Γ of small radius δ. Since X is constant in D, Φ_t is a translation on D and

$$\iint_{\Phi_t^{-1}(D-\Gamma)} \|\nabla f \circ \Phi_t\|^2 = \iint_{D-\Gamma} \|\nabla f\|^2.$$

Therefore the second integral equals:

$$\lim_{t \to 0} \frac{1}{t} \left[\iint_{R - \Phi_t^{-1} D - \Phi_t^{-1}(\Gamma)} \|\nabla f \circ \Phi_t\|^2 - \iint_{R - D - \Gamma} \|\nabla f\|^2 \right]$$

$$= \lim_{t \to 0} \iint_{R - D - \Gamma - \Phi_t^{-1}(\Gamma)} \frac{1}{t} \left(\|\nabla f \circ \Phi_t\|^2 - \|\nabla f\|^2 \right.$$

$$+ \lim_{t \to 0} \frac{1}{t} \left[\iint_{\Phi_t D - D} \|\nabla f\|^2 - \iint_{D - \Phi_t D} \|\nabla f\|^2 \right].$$

As above, define

$$h_t(x, y) = \begin{cases} \frac{1}{t}\left(\|\nabla f \circ \Phi_t\|^2 - \|\nabla f\|^2\right) & \text{if } \ (x, y) \notin D \cup B_t, \\ 0 & \text{if } \ (x, y) \in D \cup B_t. \end{cases}$$

Then it is easy to check that

$$\begin{aligned} \lim_{t \to 0} \iint_{R-D} h_t &= \iint_{R-D} \lim_{t \to 0} h_t \\ &= 2\iint_{R-D} (\nabla(Xf) \cdot \nabla f) \\ &= 2\int_{\partial D} Xf \frac{\partial f}{\partial n}\, ds - 2\iint_{R-D} Xf \cdot \Delta f, \end{aligned}$$

while

$$\lim_{t \to 0} \iint_{B_t - B_t \cap D} \frac{1}{t}\|\nabla f \circ \Phi_t\|^2 - \|\nabla f\|^2 = \int_{\Gamma - \Gamma \cap D} \left(\|\nabla f_+\|^2 - \|\nabla f_-\|^2\right) ds.$$

But

$$\begin{aligned} &\lim_{t \to 0} \frac{1}{t}\left(\iint_{\Phi_t D - D} \|\nabla f\|^2 - \iint_{D - \Phi_t D} \|\nabla f\|^2\right) \\ &\quad = \sum_P \int_{\partial D_P} \|\nabla f\|^2 \cos(\theta - \theta_0) \cdot r\, d\theta, \end{aligned}$$

where (r, θ) are polar coordinates at P and θ_0 is the direction of $X(P)$. At each crack-tip P, write

$$f = c_P\, \mathcal{R}e\sqrt{z} + S_P.$$

Then

$$\|\nabla f\|^2 = \frac{1}{4} c_P^2 \frac{1}{r} + S'_P,$$

where $S'_P \in o(r^{-1/2-\varepsilon})$. Therefore,

$$\int_{\partial D_P} \|\nabla f\|^2 \cos(\theta - \theta_0) \cdot r\, d\theta = \int_{\partial D_P} S'_P \cos(\theta - \theta_0) \cdot r\, d\theta \leqq C\,\delta^{1/2-\varepsilon}.$$

Finally the third term T_3 is easily evaluated as

$$\nu\int_{\Gamma}\kappa(P)\big(X(P)\cdot\vec{n}_P\big)\,ds+\sum_{\substack{\text{singular}\\ \text{points}\\ Q\text{ of }\Gamma}}\left[\sum_{\substack{\text{branches}\\ \beta\text{ of }\Gamma\\ \text{ending}\\ \text{at }Q}}\big(X(Q)\cdot\vec{t}_\beta\big)\right].$$

where $\kappa(P)$ is the curvature of Γ at P, $\vec{n}_P$, $\vec{t}_P$ are the unit normal and tangent vectors to Γ. Putting this together, we get

$$\begin{aligned}\delta_X E &= 2\mu^2\iint_R Xf\cdot(f-g)-2\iint_{R-D}Xf\cdot\Delta f\\ &\quad+\mu^2\int_\Gamma(f_+-f_-)(f_++f_--2g)+\int_{\Gamma-D}\|\nabla f_+\|^2-\|\nabla f_-\|^2\\ &\quad+\nu\int_\Gamma\kappa\,ds+\sum_{\substack{\text{crack-}\\ \text{tips }P}}2\int_{\partial D_P}Xf\frac{\partial f}{\partial n}\,ds\\ &\quad+\sum_{\substack{\text{singular}\\ \text{points}\\ Q\text{ of }\Gamma}}\;\sum_{\substack{\text{branches}\\ \beta\text{ of }\Gamma}}\nu\big(X_Q\cdot\vec{t}_\beta\big)+o(\delta).\end{aligned}$$

Now let $\delta\to 0$. Use the fact that $\Delta f=\mu^2(f-g)$ on $R-\Gamma$, and that $e_+-e_-+\nu\kappa\equiv 0$ along Γ, as proved in Section 1. Moreover, at all singular points Q other than crack-tips, it follows as before that either Q is a triple point with 120° angles, hence

$$\sum_{\substack{\text{branches}\\ \beta\text{ at }Q}}\vec{t}_\beta=\vec{0},$$

or Q is a point where Γ meets ∂R perpendicularly, hence $(X_Q\cdot\vec{t}_\beta)=0$. Thus all terms drop out as expected except for those at crack-tips, where we have

$$\delta_X E=\sum_{\substack{\text{crack-}\\ \text{tips }P}}2\lim_{\delta\to 0}\int_{\partial D_P(\delta)}X(f)\frac{\partial f}{\partial n}\,ds+\nu\big(X_P\cdot\vec{t}_P\big).$$

At each P, take coordinates such that Γ is tangent to the positive x-axis, and write $f = C_P\,\mathcal{R}e\sqrt{z} + S_P$. Suppose $X = \alpha\,\partial/\partial x + \beta\,\partial/\partial y$. Then

$$X(f) = \alpha\frac{C_P}{2\sqrt{r}}\cos(\tfrac{1}{2}\theta) + \beta\frac{C_P}{2\sqrt{r}}\sin(\tfrac{1}{2}\theta) + o(r^{-\varepsilon}),$$

$$\frac{\partial f}{\partial n} = \frac{C_P}{2\sqrt{r}}\cos(\tfrac{1}{2}\theta) + o(r^{-\varepsilon}),$$

from which it follows easily that

$$\lim_{\delta\to 0}\int_{\partial D_P} X(f)\frac{\partial f}{\partial n}\,ds = \tfrac{1}{4}\pi c_P^2\alpha.$$

4. Approximation when μ is Small

We derive here limiting forms of the energy and its first variation when μ is small. As before, let f_Γ minimize $E(f, \Gamma)$ for fixed Γ. Let R_i be the components of $R - \Gamma$:

$$R - \Gamma = R_1 \sqcup \cdots \sqcup R_n$$

and let

$$\bar{g}_\Gamma = \begin{cases} \text{function constant on each } R_i \\ \text{with value } \operatorname{mean}_{R_i}(g). \end{cases}$$

We shall prove that f_Γ is very close to $\bar{g}_\Gamma$ when μ is small. Throughout this section, we assume that Γ is a finite union of $C^{1,1}$-arcs meeting at corners with angles α, $0 < \alpha < 2\pi$, or ending at crack-tips.

The error term depends on the smallest "necks" of each component R_i. For any region W, define the isoperimetric constant $h(W)$ by

$$h(W) = \inf_\gamma\left\{\frac{|\gamma|}{\min(|W_1|, |W_2|)} : \begin{array}{c}\gamma \text{ is a curve dividing } W\\ \text{into 2 disjoint open}\\ \text{sets } W_1 \text{ and } W_2\end{array}\right\},$$

where $|\gamma|$ = length of γ and $|W_i|$ = area of W_i (compare Figure 5). Note that we have excluded cuspidal corners on the components R_i of $R - \Gamma$ so $h(R_i)$ is positive. A bad case is shown in Figure 18. Let

$$\lambda_\Gamma = \min_i h(R_i).$$

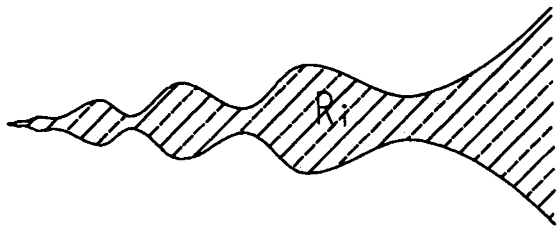

Figure 18. A domain with zero isoperimetric constant.

As in the introduction, we write

$$E_0(\Gamma) = \iint_R (g - \bar{g}_\Gamma)^2 \, dx \, dy + \nu_0 |\Gamma|.$$

We now set $\nu_0 = \nu/\mu^2$.

THEOREM 4.1.

$$\mu^2 \left\{ E_0(\Gamma) - \frac{4\mu^2}{\lambda_\Gamma^2 + 4\mu^2} \|g\|_{0,2,R}^2 \right\} \leqq E(f_\Gamma, \Gamma) \leqq \mu^2 E_0(\Gamma).$$

Proof: Let $f_1 = f_\Gamma - \bar{g}_\Gamma$ and $g_1 = g - \bar{g}_\Gamma$. Then f_1 satisfies the equation $\nabla^2 f_1 = \mu^2(f_1 - g_1)$ in $R - \Gamma$ and homogeneous Neumann boundary condition along $\Gamma \cup \partial R$. By Green's identity for each component W of $R - \Gamma$,

$$\mu^2 \int_W f_1 g_1 = \int_W (-\nabla^2 f_1 + \mu^2 f_1) f_1 = \int_W |\nabla f_1|^2 + \mu^2 \int_W f_1^2 - \int_{\partial W} \frac{\partial f_1}{\partial n} f_1.$$

Since the last integral is zero,

$$\text{(a)} \quad \|\nabla f_1\|_{0,2,W}^2 + \mu^2 \|f_1\|_{0,2,W}^2 = \mu^2 \int_W f_1 g_1 \leqq \mu^2 \|f_1\|_{0,2,W} \cdot \|g_1\|_{0,2,W}$$

by Cauchy–Schwarz inequality. By Cheeger's inequality (see [3]),

$$\inf_{\substack{u \neq 0 \\ \int_W u = 0}} \frac{\|\nabla u\|_{0,2,W}^2}{\|u\|_{0,2,W}^2} \geq \tfrac{1}{4} h^2(W).$$

Since $\int_W g_1 = 0$ and, by Green's identity,

$$\mu^2 \int_W (f_1 - g_1) = \int_W \nabla^2 f_1 = \int_{\partial W} \frac{\partial f_1}{\partial n} = 0,$$

we have

$$\int_W f_1 = 0.$$

Therefore,

$$\text{(b)} \qquad \|\nabla f_1\|_{0,2,W}^2 \geqq \tfrac{1}{4} h^2(W) \|f_1\|_{0,2,W}^2 \geqq \tfrac{1}{4} \lambda_\Gamma^2 \|f_1\|_{0,2,W}^2.$$

Combining (a) and (b), we have

$$\|f_1\|_{0,2,W} \leqq \frac{4\mu^2}{\lambda_\Gamma^2 + 4\mu^2} \|g_1\|_{0,2,W}$$

and

$$\|\nabla f_1\|_{0,2,W} \leqq \frac{2\mu^2}{\sqrt{\lambda_\Gamma^2 + 4\mu^2}} \|g_1\|_{0,2,W}.$$

Squaring these inequalities and summing over the components of $R - \Gamma$, we obtain the same inequalities with W replaced by $R - \Gamma$. Finally,

$$0 \leqq E(\bar{g}_\Gamma, \Gamma) - E(f_\Gamma, \Gamma) = \mu^2 \int_R g_1^2 - \mu^2 \int_R (g_1 - f_1)^2 - \int_{R-\Gamma} |\nabla f_1|^2$$

$$= 2\mu^2 \int_R f_1 g_1 - \mu^2 \int f_1^2 - \int_{R-\Gamma} |\nabla f_1|^2$$

$$= \mu^2 \int_R f_1 g_1, \qquad \text{by Green's identity as above,}$$

$$\leqq \mu^2 \|f_1\|_{0,2,R} \|g_1\|_{0,2,R}$$

$$\leqq \frac{4\mu^4}{\lambda_\Gamma^2 + 4\mu^2} \|g_1\|_{0,2,R}^2$$

$$\leqq \frac{4\mu^4}{\lambda_\Gamma^2 + 4\mu^2} \|g\|_{0,2,R}^2;$$

so the theorem follows.

We consider now the first variation of $E(f_\Gamma, \Gamma)$. Let $C_0^1(\Gamma)$ be the space of continuously differentiable functions on γ which vanish in a neighborhood of the singular points of Γ. For each $\varphi \in C_0^1(\Gamma)$ and all sufficiently small $\tau > 0$ define nearby curves $\Gamma_{\tau\varphi}$ as follows: let $(x(s), y(s))$ be a local parametrization of Γ by arc length. Then $\Gamma_{\tau\varphi}$ is the curve

$$s \mapsto (x(s), y(s)) + \tau\varphi(s)(-y'(s), x'(s)).$$

THEOREM 4.2.

(i)

$$\left.\frac{dE_0(\Gamma_{\tau\varphi})}{d\tau}\right|_{\tau=0} = \int_\Gamma (\psi_0 + \nu_0\kappa)\varphi\, ds,$$

where $\psi_0 = (\bar{g}_\Gamma^+ - \bar{g}_\Gamma^-)(\bar{g}_\Gamma^+ + \bar{g}_\Gamma^- - 2g)$, κ = *curvature of* Γ.

(ii) *There exists a constant* C_Γ *independent of* μ *(but which depends on* Γ*) such that*

$$\left| \left.\frac{dE(f_{\Gamma_{\tau\varphi}}, \Gamma_{\tau\varphi})}{d\tau}\right|_{\tau=0} - \mu^2 \left.\frac{dE_0(\Gamma_{\tau\varphi})}{d\tau}\right|_{\tau=0} \right| \leqq C_\Gamma \mu^4 \|\varphi\|_{0,\infty,\Gamma} \cdot \|g\|_{0,\infty,R}^2.$$

Proof: Since $\bar{g}_\Gamma$ minimizes $E(f, \Gamma)$ over the space of locally constant functions over $R - \Gamma$, the first variation formula in part (i) follows in the same way as in the general case, derived in Section 1.

Define f_1 and g_1 as in the proof of Theorem 4.1. Let

$$\begin{aligned} e_1(\Gamma) &\stackrel{\text{def}}{=} e(f_\Gamma) - e(\bar{g}_\Gamma) \\ &= \mu^2\{(g - f_\Gamma)^2 - (g - \bar{g}_\Gamma)^2\} + |\nabla f_\Gamma|^2 \\ &= \mu^2\{(g_1 - f_1)^2 - g_1^2\} + |\nabla f_1|^2 \\ &= \mu^2(f_1^2 - 2f_1 g_1) + |\nabla f_1|^2. \end{aligned}$$

The left-hand side in the inequality in part (ii) equals

$$\left| \int_\Gamma (e_1^+(\Gamma) - e_1^-(\Gamma))\varphi\, ds \right| \leqq \sum_{\substack{\text{components} \\ W \text{ of } R - \Gamma}} \|\varphi\|_{0,\infty,\Gamma} \int_{\partial W} |e_1(\Gamma)|\, ds.$$

We need to estimate $\|f_1\|_{1,2,\partial W}$. The boundary value problem

$$\nabla^2 u = v, \qquad \left.\frac{\partial u}{\partial n}\right|_{\Gamma \cup \partial R} = 0$$

has a solution provided that $\int_W v = 0$ for all components W of $R - \Gamma$. The solution is unique if we require that $\int_W u = 0$ for every W. In this case, since Γ has no cuspidal corners, $u \in W_p^2(R - \Gamma)$ for every p, $1 \leqq p < \frac{4}{3}$, by Appendix 1, **H**, and

$$\|u\|_{2,p,R-\Gamma} \leqq C_1\|v\|_{0,p,R-\Gamma},$$

where C_1 is a constant which depends on Γ. Applying this to the equation $\nabla^2 f_1 = \mu^2(f_1 - g_1)$, we get

$$\|f_1\|_{2,p,W} \leqq \mu^2 C_1\big(\|f_1\|_{0,p,W} + \|g_1\|_{0,p,W}\big).$$

Since $L_2(W)$ embeds in $L_p(W)$ and we know from the proof of Theorem 4.1 that

$$\|f_1\|_{0,2,W} \leqq \frac{4\mu^2}{\lambda_\Gamma^2 + 4\mu^2}\|g_1\|_{0,2,W},$$

we have

$$\|f_1\|_{2,p,W} \leqq \mu^2 C_2\|g_1\|_{0,2,W}.$$

By the trace theorem (see [7]), the restriction of $u \in W_p^2(W)$ to ∂W defines a continuous linear map

$$W_p^2(W) \to W_p^1(\partial W) \subset W_2^1(\partial W).$$

Therefore there exists a constant C_3 depending on Γ such that

$$\|f_1\|_{1,2,\partial W} \leqq C_3\mu^2\|g_1\|_{0,2,W}.$$

Going back to the energy density $e_1(\Gamma)$, we have

$$\begin{aligned}\int_{\partial W}|e_1(\Gamma)|\,ds &\leqq \mu^2\int_{\partial W} f_1^2 + 2\mu^2\left(\int_{\partial W} f_1^2\right)^{1/2}\left(\int_{\partial W} g_1^2\right)^{1/2} + \int_{\partial W}|\nabla f_1|^2\\ &\leqq \left(\mu^2+1\right)\|f_1\|_{1,2,\partial W}^2 + 2\mu^2\|f_1\|_{0,2,\partial W}\cdot\|g_1\|_{0,2,\partial W}\\ &\leqq C_\Gamma\,\mu^4\|g\|_{0,\infty,R}^2.\end{aligned}$$

5. Existence of Solutions when $\mu = 0$

In case $\mu = 0$, our free boundary value problem is not much more complicated than minimal—but singular—soap bubble problems. This is an especially easy case since we are dealing with singular sets Γ which have dimension as well as codimension equal to one. The main result of this section is

THEOREM 5.1. *Let R be an open rectangle*[4] *in the plane and let g be a continuous function on $R \cup \partial R$. For all one-dimensional sets $\Gamma \subset R$ such that $\Gamma \cup \partial R$ is made up of a finite number of $C^{1,1}$-arcs, meeting each other only at their end-points, and, for all locally constant functions f on $R - \Gamma$, let*

$$E_0(f, \Gamma) = \iint_R (f - g)^2 + \nu_0 \cdot \text{length}(\Gamma).$$

Then, there exist an f and a Γ which minimize E_0.

Geometric measure theory approaches problems of this sort by embedding them in larger minimizing problems in which extremely singular Γ's are allowed, then showing that in this larger world, a "weak solution" Γ exists and finally arguing that any weak solution Γ must be of the restricted type envisioned in the original formulation. As a weak version of our problem, we consider segmentations of R by Cacciopoli sets which are measurable subsets of R with finite "perimeter". Our standard reference for geometric measure theory is the book by L. Simon [13]. We let L^2 denote the Lebesgue measure on $\mathbb{R}^2$ and H^1 denote the one-dimensional Hausdorff measure on $\mathbb{R}^2$.

We begin by recalling De Giorgi's theory of Cacciopoli sets (see Section 14 in [13]). A bounded subset F of $\mathbb{R}^2$ is called a Cacciopoli set if it is L^2-measurable and has finite "perimeter"; that is, the characteristic function χ_F of F has bounded variation. For such a set F, there exists a Radon measure μ_F on $\mathbb{R}^2$ and a μ_F-measurable function η_F: $\mathbb{R}^2 \to \mathbb{R}^2$ with $|\eta_F| = 1$ μ_F-a.e. such that

$$\int_{\mathbb{R}^2} \chi_F \operatorname{div}(\vec{g})\, dL^2 = -\int_{\mathbb{R}^2} (\vec{g} \cdot \eta_F)\, d\mu_F$$

for all C^1-functions $\vec{g}$: $\mathbb{R}^2 \to \mathbb{R}^2$ with compact support. We call μ_F the "generalized boundary measure" and η_F the "generalized inward unit normal". Let $D\chi_F$ denote the gradient of χ_F in the sense of distributions. Then, $D\chi_F = \eta_F\, d\mu_F$ and η_F, μ_F can be recovered from $D\chi_F$ by

$$\mu_F(U) = \sup_{\substack{|\vec{g}| \leq 1 \\ \operatorname{supp}(\vec{g}) \subset U \\ \vec{g} \text{ smooth}}} \int (D\chi_F \cdot \vec{g})\, dL^2, \qquad U \text{ open},$$

and

$$\eta_F(x) = \lim_{\rho \downarrow 0} \frac{\int_{B_\rho(x)} D\chi_F\, dL^2}{\mu_F(B_\rho(x))},$$

[4] The restriction to a rectangle R is not essential, but simplifies some technical aspects of the proof.

where $B_\rho(x)$ denotes the ball of radius ρ, centered at x. The "perimeter" of F is defined to be equal to $\mu_F(\mathbb{R}^2)$. Notice that if $F_1 = F_2$ L^2-a.e., then $\mu_{F_1}(\mathbb{R}^2) = \mu_{F_2}(\mathbb{R}^2)$.

We may restate these results in the language of currents. Let $\vec{F}$ be the 2-current defined by F:

$$\vec{F}(\varphi) = \int\int_F \varphi,$$

for all C^∞ 2-forms φ with compact support. Let $\partial\vec{F}$ denote the current boundary of $\vec{F}$:

$$\partial\vec{F}(\varphi) = \vec{F}(d\varphi)$$

for all C^∞ 1-forms φ with compact support. Then

$$\partial\vec{F}(\varphi) = \int\langle\eta_F, \varphi\rangle\, d\mu_F.$$

For any set S, let ∂S be the topological boundary $\bar{S} - \mathrm{Int}(S)$ of S. Then the topological boundary ∂F of a Cacciopoli set F may have positive L^2-measure and hence, infinite H^1-measure, even though F still will have finite perimeter. Fortunately, it is possible to define the *reduced boundary* $\partial^* F$ so that the perimeter of F equals the H^1-measure of $\partial^* F$:

$$\partial^* F = \left\{ x \in \mathbb{R}^2 \colon \begin{array}{c} \eta_F(x) \text{ as defined above exists} \\ \text{and has length } 1 \end{array} \right\}.$$

By De Giorgi's theorem (cf. [13], Section 14):

(i) $\partial^* F$ is 1-rectifiable.
(ii) $\mu_F = H^1 \llcorner \partial^* F$ (i.e., H^1 restricted to $\partial^* F$).
(iii) For any set $S \subset \mathbb{R}^2$, $x \in \mathbb{R}^2$, $\rho > 0$, let

$$S_{x,\rho} = \{\rho^{-1}(y - x) \mid y \in S\}.$$

Then for every point $x \in \partial^* F$, the approximate tangent space $T_{x,\partial^* F}$ exists and is given by $\{\vec{y} \in \mathbb{R}^2 \mid \vec{y} \cdot \eta_F(x) = 0\}$, i.e.,

$$\lim_{\rho \to 0} \int_{(\partial^* F)_{x,\rho}} f\, dH^1 = \int_{\{y \mid y \cdot \eta_F(x) = 0\}} f\, dH^1$$

for all continuous f with compact support. Moreover,

$$\lim_{\rho \to 0} \int_{F_{x,\rho}} f\,dL^2 = \int_{\{y | y \cdot \eta_F(x) > 0\}} f\,dL^2$$

for all $f \in L^1(\mathbb{R}^2)$.

(iv) Rotating η_F by 90° defines a unit tangent vector t_F to $\partial^* F$. Using t_F to orient the 1-rectifiable *set* $\partial^* F$, we obtain a 1-current $\overrightarrow{\partial^* F}$ such that

$$\partial(\vec{F}) = \overrightarrow{(\partial^* F)}.$$

In particular, for all bounded open subsets $U \subset \mathbb{R}^2$, the mass $\mathbf{M}_U(\partial \vec{F})$ equals $H^1(U \cap \partial^* F)$.

We next reformulate E_0 using Cacciopoli sets. Note that if Γ has an arc γ which is surrounded by a *single* component of $R - \Gamma$, we can reduce the energy E_0 simply by removing γ. Hence, we might as well assume in our original formulation that the boundary of each component of $R - \Gamma$ consists of piecewise $C^{1,1}$-loops which mutually intersect in only finitely many points (see Figure 19). Therefore, for each component F of $R - \Gamma$, and each arc γ in ∂F, γ is the boundary of F from only one side; hence using the notion of the mass of a current:

$$\text{length}(\partial F) = \mathbf{M}(\partial F).$$

Thus

$$2\,\text{length}(\Gamma) + \text{length}(\partial R) = \sum_{\left(\substack{\text{conn. comp.} \\ F \text{ of } R - \Gamma}\right)} \mathbf{M}(\partial F)$$

and

$$\text{length}(\Gamma) = \frac{1}{2} \sum_{\left(\substack{\text{conn. comp.} \\ F \text{ of } R - \Gamma}\right)} \mathbf{M}_R(\partial F).$$

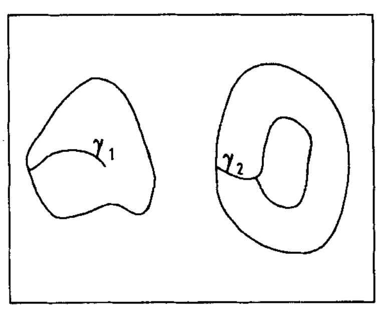

Figure 19. Removable arcs γ_i.

This motivates the following new functional: fix some positive integer n, and consider sets of n Cacciopoli sets, $U_1, U_2, \cdots, U_n$, such that

$$U_i \subset R \quad \text{for} \quad 1 \leqq i \leqq n,$$

and

$$\sum_{i=1}^{n} \vec{U_i} = \vec{R}.$$

For such $\{U_i\}$ and for all constants $a_1, \cdots, a_n$, define

$$E_n(\{U_i, a_i\}) = \sum_i \left[\tfrac{1}{2}\nu_0 \mathbf{M}_R(\partial \vec{U_i}) + \int_{U_i} (g - a_i)^2 \, dL^2 \right].$$

This is our "weak" formulation. Note that to make a reasonable "weak" problem, we have to fix n. However, the U_i are required to be neither connected nor nonempty. Hence,

$$\inf_{\{U_i, a_i\}} E_{n'} \leqq \inf_{\{U_i, a_i\}} E_n \quad \text{if} \quad n' \geqq n.$$

We shall show:

THEOREM 5.2. (a) *For each $n \geqq 1$, E_n takes on a minimum value for some $\{U_i, a_i\}$.*

(b) *If $\{U_i, a_i\}$ minimizes E_n, then each U_i is an open set with a finite number of components and piecewise C^2 boundary.*

(c) *There exists an integer n_0 such that*

$$\min_{\{U_i, a_i\}} E_n = \min_{\{U_i, a_i\}} E_{n_0} \quad \textit{for all} \quad n \geqq n_0.$$

In fact,

$$E_n(\{U_i, a_i\}) > \min E_{n_0}$$

if $n > n_0$ and each U_i is nonempty.

Clearly, this will prove Theorem 5.1. Now consider Theorem 5.2. Part (a) is an immediate consequence of the compactness theorem for functions of bounded variation (see [13], Section 6). In fact, if $\{U_i^\alpha, a_i^\alpha\}$ is a minimizing sequence so that

$$\lim_\alpha E_n(\{U_i^\alpha, a_i^\alpha\}) = \inf E_n,$$

then, by the compactness theorem, for each i, there exists a subsequence of $\{U_i^\alpha\}$

which converges to a Cacciopoli set U_i (i.e., the integral of any $L^1(\mathbb{R}^2)$ function on U_i^α converges to its integral on U_i) such that

$$H^1(\partial^* U_i) \leqq \liminf_\alpha H^1(\partial^* U_i^\alpha).$$

Moreover, minimizing with respect to a_i is seen immediately to mean

$$a_i = \frac{1}{L^2(U_i)} \int_{U_i} g \, dL^2.$$

Hence, we may assume that $0 \leqq a_i^\alpha \leqq \max|g|$ for all i and α. Therefore, on a suitable subsequence, all U_i^α and a_i^α converge, and

$$\inf E_n \leqq E_n(\{U_i, a_i\}) \leqq \liminf_\alpha E_n(\{U_i^\alpha, a_i^\alpha\}) = \inf E_n.$$

The identity $\sum_{i=1}^n \vec{U}_i^\alpha = \vec{R}$ passes to the limit so that $\sum_{i=1}^n \vec{U}_i = \vec{R}$.

Part (b) is the hardest to prove. Our method is essentially a generalization of the theory of minimizing currents of codimension 1. In outline, the proof is as follows: Let

$$\Gamma^* = \bigcup_{i=1}^n (R \cap \partial^* U_i).$$

Γ^* is the weak version of the set of curves Γ specified in Theorem 5.1. We rewrite E_n in terms of Γ^* and then consider the first variation of E_n with respect to Γ^*. This shows that Γ^* has generalized mean curvature and hence the monotonicity formula of geometric measure theory applies. We conclude that if $\{U_i, a_i\}$ minimizes E_n, then Γ^* equals its closure Γ H^1-a.e. in R. Hence, we may assume that the U_i are open and Γ is the union of their *topological* boundaries in R. We next study the singularities of Γ. To do this, we first show that tangent cones exist everywhere on Γ and that they have multiplicity away from the origin. We conclude that the singularity set of Γ is discrete. Allard's theorem implies that Γ is C^2 away from singularities. The rest of the proof now follows easily.

Before giving the details of the proof, we introduce a construction which will be used several times in the proof to handle the behavior of Γ^* along ∂R. This is to consider a larger region $R^\#$ built out of four copies of R centered around one of the corners P of R; see Figure 20. The function g is extended to a continuous function $g^\#$ on $R^\#$ by *reflection*. Thus suppose $P = (0,0)$ after a translation. Then in the situation of the figure

$$g^\#(x, y) = \begin{cases} g(x, y) & \text{in } R_1, \\ g(-x, y) & \text{in } R_3, \\ g(x, -y) & \text{in } R_2, \\ g(-x, -y) & \text{in } R_4. \end{cases}$$

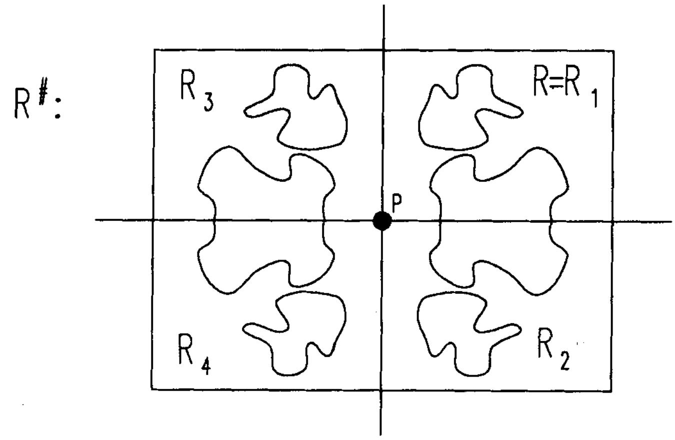

Figure 20. A segmentation reflected around a corner P.

Define $U_i^{\#}$ to be the union of U_i and its three mirror images in $R^{\#}$. Note that $\mathbf{M}_{R^{\#}}(\partial \vec{U}_i^{\#}) = 4\mathbf{M}_R(\partial \vec{U}_i)$. In fact,

$$R^{\#} = R_1 \cup R_2 \cup R_3 \cup R_4 \cup \Delta,$$

where Δ is the part of the horizontal and vertical axes through P which is in $R^{\#}$. Then

$$\mathbf{M}_{R^{\#}}\left(\partial \vec{U}_i^{\#}\right) = \sum_{\alpha=1}^{4} \mathbf{M}_{R_\alpha}\left(\partial \vec{U}_i^{\#}\right) + H^1\left(\partial^* U_i^{\#} \cap \Delta\right).$$

But

$$\partial^* \vec{U}_i^{\#} = \partial \vec{U}_i^{(1)} + \partial \vec{U}_i^{(2)} + \partial \vec{U}_i^{(3)} + \partial \vec{U}_i^{(4)}$$

and if $\partial^* U_i^{\#}$ were to include a portion of Δ of positive H^1-measure, it would occur in *two* boundaries $\partial^* U_i^{(\alpha)}$, but with the opposite normal vectors $\eta_i^{(\alpha)}$. Thus it would have to cancel out in $\partial(\vec{U}_i^{\#})$. This shows that in fact

$$E_n\left(\left\{U_i^{\#}, a_i\right\}\right) = 4E_n(\{U_i, a_i\}).$$

If $\{U_i, a_i\}$ minimizes E_n and if $\{V_i, b_i\}$ is any other decomposition of $R^{\#}$, then

$$\begin{aligned} E_n(\{V_i, b_i\}) &\geqq \sum_{\alpha=1}^{4} E_n(\{V_i \cap R_\alpha, b_i\}) \\ &\geqq 4E_n(\{U_i, a_i\}) \\ &= E_n\left(\left\{U_i^{\#}, a_i\right\}\right); \end{aligned}$$

hence $\{U_i^{\#}, a_i\}$ minimizes E_n for $g^{\#}$ on $R^{\#}$. Finally, the curve $\Gamma^{\#,*}$ defined by $\{U_i^{\#}, a_i\}$ is obtained from Γ^* and its three reflections and a set of H^1-measure zero in Δ.

We now begin the detailed proof.

LEMMA 5.3. *Let F_1, F_2 be Cacciopoli sets. Let $F = F_1 \cup F_2$ and assume that $\vec{F} = \vec{F_1} + \vec{F_2}$. Then,*

$$\partial^* F = \partial^* F_1 \cup \partial^* F_2 - \partial^* F_1 \cap \partial^* F_2 \qquad H^1\text{-a.e.}$$

Hence, for all open bounded subsets $U \subset \mathbb{R}^2$,

$$H^1(U \cap \partial^* F) = \left[H^1(U \cap \partial^* F_1) - H^1(U \cap \partial^* F_1 \cap \partial^* F_2)\right] + \left[H^1(U \cap \partial^* F_2) - H^1(U \cap \partial^* F_1 \cap \partial^* F_2)\right].$$

Proof: Let M denote $\partial^* F_1 \cup \partial^* F_2$ and N denote $\partial^* F_1 \cap \partial^* F_2$. We shall show

(i) $N \cap \partial^* F$ is empty,
(ii) $M - N \subset \partial^* F \quad H^1$-a.e.,
(iii) $\partial^* F \subset M - N \quad H^1$-a.e.

(i) Let $x \in N$. For $\rho > 0$, let

$$(F_i)_{x,\rho} = \{\rho^{-1}(y - x) : y \in F_i\} \quad \text{for} \quad i = 1, 2,$$

and

$$F_{x,\rho} = \{\rho^{-1}(y - x) : y \in F\}.$$

By De Giorgi's theorem, $\lim_{\rho \downarrow 0} (F_i)_{x,\rho} = H_i$, where H_i is an open half-space in $\mathbb{R}^2$. H_1 and H_2 must be disjoint. For, if $H_1 \cap H_2$ is non-empty, there exists a ball $B \subset H_1 \cap H_2$ and $f \in C_c^\infty(B)$ such that, for $i = 1, 2$,

$$\lim_{\rho \downarrow 0} \int_{(F_i)_{x,\rho}} f \, dL^2 = \int_{H_i} f \, dL^2 > 0.$$

Since $\vec{F} = \vec{F_1} + \vec{F_2}$,

$$\lim_{\rho \downarrow 0} \int_{F_{x,\rho}} f \, dL^2 = \lim_{\rho \downarrow 0} \int_{(F_1)_{x,\rho}} f \, dL^2 + \lim_{\rho \downarrow 0} \int_{(F_1)_{x,\rho}} f \, dL^2$$

$$= 2 \int_{\mathbb{R}^2} f \, dL^2.$$

Therefore, there exists $\rho > 0$ such that

$$\int_{F_{x,\rho}} f\,dL^2 > \int_{\mathbb{R}^2} f\,dL^2$$

which is absurd. It follows that H_1 and H_2 are complementary half-spaces; hence $\lim_{\rho\downarrow 0} F_{x,\rho} = \mathbb{R}^2$ and thus $x \notin \partial^* F$ by De Giorgi's theorem.

(ii) Let $x \in M - N$, say $x \in \partial^* F_1 - N$. Outside a set of H^1-measure zero, we may assume that the upper density $\Theta^*(H^1, \partial^* F_2, x) = 0$. Write out $\eta_F(x)$:

$$\eta_F(x) = \lim_{\rho\downarrow 0} \frac{\int_{B_\rho(x)} D\chi_F\,dL^2}{\mu_F\big(B_\rho(x)\big)} = \frac{\lim_{\rho\downarrow 0}(1/2\rho)\int_{B_\rho(x)} D\chi_F\,dL^2}{\Theta(H^1, \partial^* F, x)}.$$

Since $\Theta(H^1, \partial^* F_1, x) = 1$ and $\Theta^*(H^1, \partial^* F_2, x) = 0$,

$$\Theta(H^1, \partial^* F, x) = \lim_{\rho\downarrow 0} \frac{1}{2\rho}\mu_F\big(B_\rho(x)\big) = 1.$$

Since $D\chi_F = D\chi_{F_1} + D\chi_{F_2}$ it also follows that

$$\lim_{\rho\downarrow 0} \frac{1}{2\rho}\int_{B_\rho(x)} D\chi_F\,dL^2 = \lim_{\rho\downarrow 0} \frac{1}{2\rho}\int_{B_\rho(x)} D\chi_{F_1}\,dL^2 = \eta_{F_1}(x).$$

Hence $\eta_F(x)$ exists and has length 1, i.e., $x \in \partial^* F$.

(iii) If $x \in \partial^* F - M$, we may assume that

$$\Theta^*\big(H^1, \partial^* F_i, x\big) = 0 \quad \text{for} \quad i = 1, 2.$$

Then, $\Theta^*(H^1, \partial^* F, x) \leqq \Theta^*(H^1, \partial^* F_1, x) + \Theta^*(H^1, \partial^* F_2, x) = 0$ and so $H^1(\partial^* F - M) = 0$.

LEMMA 5.4. *Let $U_1, U_2, \cdots, U_n$ be Cacciopoli sets such that $\Sigma_{i=1}^n \vec{U}_i = \vec{R}$. Then,*

(a) *for H^1-a.e., $x \in \Gamma^*$ belongs to precisely two $\partial^* U_i$'s,*

(b) $H^1 \llcorner \Gamma^* = \frac{1}{2}\Sigma_{i=1}^n H^1 \llcorner (R \cap \partial^* U_i)$.

In particular,

$$H^1(\Gamma^*) = \tfrac{1}{2}\sum_{i=1}^{n} \mathbf{M}_R\big(\partial \vec{U}_i\big).$$

Proof: (a) Suppose $x \in R \cap \partial^* U_i$ for some i. Let $U = \bigcup_{j\neq i} U_j$. Since $\vec{U}_i = \vec{R} - \vec{U}$, $\partial \vec{U}_i = \partial \vec{R} - \partial \vec{U}$ and hence

$$R \cap \partial^* U_i = R \cap \partial^* U \qquad H^1\text{-a.e.}$$

By Lemma 5.3,

$$R \cap \partial^* U = \bigcup_{j \neq i} (R \cap \partial^* U_j)$$

$$- \{ y : y \in (\partial^* U_j) \cap (\partial^* U_k) \cap R \text{ for } j, k \neq i \} \quad H^1\text{-a.e.}$$

$$= \{ y : y \in R \cap \partial^* U_j \text{ for some unique } j \neq i \} \quad H^1\text{-a.e.}$$

(b) Let $\mu = H^1 \llcorner \Gamma^*$ and $\mu_i = H^1 \llcorner (R \cap \partial^* U_i)$:

$$D_\mu \sum_{i=1}^n \mu_i(x) \underset{\text{def}}{=} \lim_{\rho \downarrow 0} \frac{\sum_{i=1}^n \mu_i(B_\rho(x))}{\mu(B_\rho(x))}$$

$$= \frac{\sum_{i=1}^n \Theta(\mu_i, x)}{\Theta(\mu, x)}$$

$$= 2 \quad H^1\text{-a.e.} \qquad \text{by part (a).}$$

Part (b) now follows from Theorem 4.7 in [13].

LEMMA 5.5. *Let $\{U_i, a_i\}$ minimize E_n. Let Γ = closure of Γ^* in $\overline{R}$. Let η_i denote the generalized inward unit normal in R corresponding to the Cacciopoli set U_i. Letting η_i be zero on $\Gamma^* - \partial^* U_i$, define $\kappa : \Gamma^* \to \mathbb{R}^2$ by*

$$\kappa = -\frac{1}{\nu_0} \sum_{i=1}^n (g - a_i)^2 \eta_i.$$

Then,

(a) *Γ^* has the generalized curvature κ. That is, for all C^1-vector fields $\mathbf{X}$ on $\overline{R}$ tangent to ∂R along ∂R, if $D\mathbf{X}$ is the 2×2 matrix of derivatives of the components of $\mathbf{X}$ and if $t : \Gamma^* \to \mathbb{R}^2$ is a unit tangent vector (with any choice of signs),*

$$\int_{\Gamma^*} (t' \cdot D\mathbf{X} \cdot t)\, dH^1 = -\int_{\Gamma^*} (\kappa \cdot \mathbf{X})\, dH^1.$$

(b) *Monotonicity: Define a function ε on $\overline{R}$ by*

$$\varepsilon(x) = \begin{cases} 1 & \text{if } x \in R, \\ 2 & \text{if } x \in \partial R, \text{ but } x \text{ is not a corner}, \\ 4 & \text{if } x \text{ is a corner of } \overline{R}. \end{cases}$$

Then the density $\Theta(H^1, \Gamma^*, x)$ exists for all $x \in \Gamma$, $\varepsilon\Theta$ is greater than or equal to 1, *and is upper semi-continuous.*

(c) $\Gamma = \Gamma^*$ *H^1-a.e. so that* $\frac{1}{2}\Sigma_{i=1}^n \mathbf{M}'_R(\partial \vec{U_i}) = H^1(\Gamma)$.

(d) *We may assume that the U_i are open and that Γ equals the union of their topological boundaries in R.*

Proof: (a) By Lemma 5.4,

$$E_n = \nu_0 H^1(\Gamma^*) + \sum_{i=1}^n \int_{U_i} (g - a_i)^2 \, dL^2.$$

Consider the first variation $\delta_{\mathbf{X}} E_n$ of E_n with respect to a C^1-vector field $\mathbf{X}$ on $\overline{R}$, tangent to ∂R along ∂R. The formula

$$\delta_{\mathbf{X}} H^1(\Gamma^*) = \int_{\Gamma^*} (t' \cdot D\mathbf{X} \cdot t) \, dH^1$$

is standard, and

$$\delta_{\mathbf{X}} \int_{U_i} (g - a_i)^2 \, dL^2 = \int_{U_i} \operatorname{div}\big((g - a_i)^2 \mathbf{X}\big) \, dL^2 \qquad \text{by direct computation,}$$

$$= -\int_{\partial^* U_i} (g - a_i)^2 (\eta_i \cdot \mathbf{X}) \, dH^1 \quad \text{by De Giorgi's theorem.}$$

Since $\{U_i, a_i\}$ minimizes E_n, $\delta_{\mathbf{x}} E_n = 0$; hence

$$0 = \nu_0 \int_{\Gamma^*} (t' \cdot D\mathbf{X} \cdot t) \, dH^1 - \sum_{i=1}^n \int_{\partial^* U_i} (g - a_i)^2 (\eta_i \cdot \mathbf{X}) \, dH^1$$

$$= \nu_0 \left[\int_{\Gamma^*} (t' \cdot D\mathbf{X} \cdot t) \, dH^1 + \int_{\Gamma^*} \kappa \cdot \mathbf{X} \, dH^1 \right].$$

(b) The fact that $\Theta(H^1, \Gamma^*, x)$ exists at every point of R and that it is upper semicontinuous follows from the monotonicity formula (see Corollary 17.8 in [13] or Almgren–Allard [1]). Since Γ^* is rectifiable, $\Theta(H^1, \Gamma^*, x) = 1$ H^1-a.e. on Γ^*. (This follows from Lemma 5.4.) Hence, $\Theta(H^1, \Gamma, x) \geq 1$ for all $x \in \Gamma \cap R$ by upper semicontinuity. To extend this argument to $\overline{R}$, we use the reflection technique explained in the beginning of this proof. Then $\Theta(H^1, \Gamma^{\#,*}, x)$ exists, is at least 1 and is upper semi-continuous on $R^{\#}$. But it is easy to see that $\Theta(H^1, \Gamma^{\#,*}, x) = \varepsilon(x)\Theta(H^1, \Gamma^*, x)$ for $x \in R^{\#} \cap \overline{R}$, i.e., along edges of R, Γ^* is half of $\Gamma^{\#,*}$, and at corners, Γ^* is a quarter of $\Gamma^{\#,*}$.

(c) This follows from part (b).

(d) Let $U = R - \Gamma$. Since $L^2(\Gamma) = 0$, we may replace U_i by $U_i \cap U$ without altering $E_n(\{U_i, a_i\})$. Let $\{V_\alpha\}$ be the set of components of U. Then $\vec{V}_\alpha = \Sigma_i \overrightarrow{(U_i \cap V_\alpha)}$; hence, by the constancy theorem, $\vec{V}_\alpha = \overrightarrow{U_{i(\alpha)} \cap V_\alpha}$ for exactly one $i(\alpha)$. Therefore we may replace U_i by $\cup_\alpha \{V_\alpha | i(\alpha) = i\}$.

We now consider the tangent cones of Γ. Let

$$\Gamma_{\text{reg}} = \{x \in \Gamma : \varepsilon\Theta(H^1, \Gamma, x) = 1\}$$

and

$$\Gamma_{\text{sing}} = \{x \in \Gamma : \varepsilon\Theta(H^1, \Gamma, x) > 1\}.$$

By Theorem 6.3 in [13], for any sequence $\rho_k \downarrow 0$, there exists a subsequence $\rho_{k(j)} \downarrow 0$ and Cacciopoli sets$\{V_i\}$ such that, for $1 \leqq i \leqq n$, $(U_i)_{x, \rho_{k(j)}} \to V_i$ in the $L^1(\mathbb{R}^2)$ sense.

LEMMA 5.6. *Suppose* $\{U_i, a_i\}$ *minimizes* E_n. *Let* $x \in \Gamma \cap R$. *Let* $\rho_k \downarrow 0$ *be a sequence such that, for* $1 \leqq i \leqq n$, $(U_i)_{x, \rho_k} \to V_i$. *Let* $N^* = \cup_{i=1}^n \partial^* V_i$ *and let* $N =$ *(closure of* N^* *in* $\mathbb{R}^2$*). Then*:

(a) $\Gamma_{x, \rho_k} \to N^*$; *that is, for all* $f \in L^1(\mathbb{R}^2)$ *with compact support,*

$$\int_{(\Gamma_{x, \rho_k})} f\, dH^1 \to \int_{N^*} f\, dH^1.$$

(b) N^* *is stationary in* $\mathbb{R}^2$; *that is, for all* C^1*-vector fields* $\mathbf{X}$ *on* $\mathbb{R}^2$ *with compact support,*

$$\int_{N^*} (t'_{N^*} \cdot D\mathbf{X} \cdot t_{N^*})\, dH^1 = 0.$$

Hence, we may assume that the V_i *are open and* N *is the union of their topological boundaries.*

(c) N *is a finite union of rays and each* V_i *is a finite union of sectors of* $\mathbb{R}^2$.

(d) $\Theta(H^1, \Gamma, x) \in \frac{1}{2}\mathbb{Z}$.

(e) *If* $\{t_i\}$ *is the set of unit tangent vectors along the rays of* N, *pointing away from the origin, then* $\Sigma t_i = 0$.

Finally, if $x \in \Gamma \cap \partial R$, (a) *and* (b) *hold for the extension* $\{U_i^\#, \Gamma^\#\}$ *described above*; *hence* (c) *holds,* $\varepsilon(x)\Theta(H^1, \Gamma, x) \in \frac{1}{2}\mathbb{Z}$ *and, along edges of* R, Σt_i *is normal to the edge.*

Proof: To simplify notation in this proof, let

$$f_k(y) = \rho_k^{-1}(y - x)$$

so that, for any set S, S_{x, ρ_k} is the same as $f_k(S)$: the set S expanded around x by a large factor ρ_k^{-1}.

To prove (a), it is enough to show that

(i) for all $W \subset \mathbb{R}^2$ open, $H^1(N^* \cap W) \leqq \liminf_k H^1(f_k(\Gamma) \cap W)$;

(ii) for all $K \subset \mathbb{R}^2$ compact, $H^1(N^* \cap K) \geqq \limsup_k H^1(f_k(\Gamma) \cap K)$.

Since $\overrightarrow{\partial f_k(U_i)} \to \partial \vec{V}_i, \mathbf{M}_W(\partial \vec{V}_i) \leqq \liminf \mathbf{M}_W(\overrightarrow{f_k(U_i)})$ for all open W. Hence (i) follows. Now fix $K \subset \mathbb{R}^2$ compact and $\varepsilon > 0$. Choose a smooth function $\varphi : \mathbb{R}^2 \to [0,1]$ such that

$$\begin{aligned} &\varphi \equiv 1 \quad \text{on} \quad K, \\ &\operatorname{supp}(\varphi) \subset \{y \mid \operatorname{dist}(y, K) < \varepsilon\}. \end{aligned}$$

Let k_0 be an integer such that

$$\operatorname{supp}(\varphi) \subset f_k(R) \quad \text{for} \quad k \geqq k_0.$$

Let

$$W_\alpha = \{y \mid \varphi(y) > \alpha\},$$

$$\vec{P}_i^{(k)} = \vec{V}_i - f_k\left(\vec{U}_i\right) \qquad \text{(difference of 2- currents)}.$$

Then $\mathbf{M}_W(\vec{P}_i^{(k)}) \to 0$ for all open $W \subset \mathbb{R}^2$. By the slicing theorem, we may choose $0 < \alpha < 1$ such that for $1 \leqq k \leqq n$ and $k \geqq k_0$:

$$\partial\left(\vec{P}_i^{(k)} \llcorner W_\alpha\right) = \left(\partial \vec{P}_i^{(k)}\right) \llcorner W_\alpha + \vec{Q}_i^{(k)},$$

$$\vec{Q}_i^{(k)} = \text{“slice” of } \vec{P}_i^{(k)} \text{ by } \partial W_\alpha,$$

so that

$$\operatorname{supp}\left(\vec{Q}_i^{(k)}\right) \subset \partial W_\alpha,$$

$$\lim_k \mathbf{M}\left(\vec{Q}_i^{(k)}\right) = 0.$$

Moreover, for suitable α, we may assume

$$\mathbf{M}\left(\overrightarrow{\partial f_k(U_i)} \llcorner \partial W_\alpha\right) = \mathbf{M}\left(\partial \vec{V}_i \llcorner \partial W_\alpha\right) = 0.$$

The main idea is to define, for all k, a modified decomposition of R into Caccioppoli sets $\{U_i^{(k)}\}$, namely:

$$U_i^{(k)} \cap f_k^{-1}(W_\alpha) = f_k^{-1}(V_i),$$

$$U_i^{(k)} \cap \left(R - f_k^{-1}(W_\alpha)\right) = U_i,$$

or alternately, define $U_i^{(k)}$ as a current by

$$\vec{U}_i^{(k)} = \vec{U}_i + f_k^{-1}\left(\vec{P}_i^{(k)} \llcorner W_\alpha\right);$$

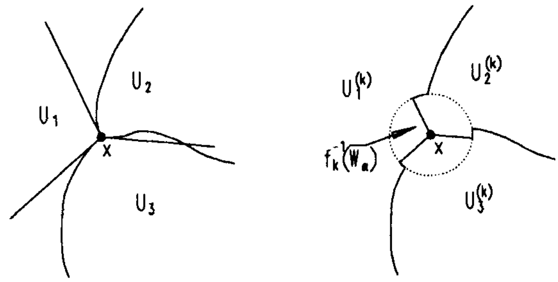

Figure 21. Modifying a decomposition via its tangent cone.

(see Figure 21). It is easy to check, using the second description, that

$$\mathbf{M}\left(\partial \vec{U}_i^{(k)}\right) = \mathbf{M}\left(\partial \vec{U}_i\right) + \rho_k\left[\mathbf{M}_{W_\alpha}\left(\partial \vec{V}_i\right) - \mathbf{M}_{W_\alpha}\left(\partial f_k\left(\vec{U}_i\right)\right) + \mathbf{M}\left(\vec{Q}_i^{(k)}\right)\right].$$

Letting $\omega_{i,k}(y) = (g(f_k^{-1}(y)) - \alpha_i)^2\, dx\, dy$, we find

$$\begin{aligned} E_n\left(\left\{U_i^{(k)}, a_i\right\}\right) &= E_n(\{U_i, a_i\}) \\ &\quad + \frac{1}{2}\nu_0\rho_k \sum_{i=1}^{n} \left[\mathbf{M}_{W_\alpha}\left(\partial \vec{V}_i\right) - \mathbf{M}_{W_\alpha}\left(\overrightarrow{\partial f_k(U_i)}\right) + \mathbf{M}\left(\vec{Q}_i^{(k)}\right)\right] \\ &\quad + \rho_k^2 \sum_{i=1}^{n} \left(\vec{P}_i^{(k)} \llcorner W_\alpha\right)(\omega_{i,k}). \end{aligned}$$

But now $E_N(\{U_i^{(k)}, a_i\}) \geqq E_n(\{U_i, a_i\})$ by hypothesis, so

$$\begin{aligned} 0 &\leqq H^1(N^* \cap W_\alpha) - H^1(f_k(\Gamma) \cap W_\alpha) \\ &\quad + \sum_{i=1}^{n} \left[\frac{1}{2}\mathbf{M}\left(\vec{Q}_i^{(k)}\right) + \frac{\rho_k}{\nu_0}\left(\vec{P}_i^{(k)} \llcorner W_\alpha\right)(\omega_{i,k})\right]. \end{aligned}$$

It follows that

$$H^1(N^* \cap W_\alpha) \geqq \limsup H^1(f_k(\Gamma) \cap W_\alpha)$$

and hence

$$H^1(N^* \cap \{y : \operatorname{dist}(y, K) < \varepsilon\}) \geqq \limsup H^1(f_k(\Gamma) \cap K).$$

Letting $\varepsilon \downarrow 0$, we get

$$H^1(N^* \cap K) \geqq \limsup H^1(f_k(\Gamma) \cap K).$$

It follows that $f_k(\Gamma) \to N^*$.

To prove (b), note that $f_k(\Gamma)$ has generalized curvature κ_k such that $|\kappa_k(y)| = \rho_k|\kappa(\rho_k y + x)|$. Therefore, N is stationary by lower semicontinuity of variation measures (Theorem 40.6 in [13]). The rest of part (b) now follows as in Lemma 5.5. Part (c) follows from Theorem 19.3 in [13]. (d) holds because $\Theta(H^1, \Gamma, x) = \Theta(H^1, N, \{0\})$. Part (e) follows from part (b). The extension to ∂R follows immediately by reflection.

LEMMA 5.7. *If $\{U_i, a_i\}$ minimizes E_n, then Γ_{sing} is finite.*

Proof: Suppose that Γ_{sing} is not finite. Let $x_1, x_2, \cdots$ be a sequence of points in $\Gamma_{\text{sing}} - \{x\}$ converging to x. If $x \in \partial R$, replace $\{U_i, a_i\}$ by $\{U_i^{\#}, a_i\}$. Let $\rho_k = |x_k - x|$. If we replace $\{\rho_k\}$ by a suitable subsequence, then Γ_{x,ρ_k} converges to a cone N as in Lemma 5.6. The points $\xi_k = \rho_k^{-1}(x_k - x)$ are on $\partial B_1(0)$ and hence converge to a point $\xi \in \partial B_1(0)$. By monotonicity and Lemma 5.6, $\Theta(H^1, N, \xi) \geqq \limsup \Theta(H^1, \Gamma_{x,\rho_k}, \xi_k) \geqq 1\frac{1}{2}$. Therefore, $\xi \in N - \{0\}$. But, $\Theta(H^1, N, y) = 1$ for all $y \in N - \{0\}$ by Lemma 5.6. Contradiction!

LEMMA 5.8. *If $\{U_i, a_i\}$ minimizes E_n, then Γ_{reg} is C^2.*

Proof: By Allard's regularly theorem (see Theorem 24.2 in [13]), Γ_{reg} is C^1. Its curvature κ is C^0 along Γ_{reg} since g is. It follows from the standard regularity theory of elliptic differential equations that Γ_{reg} must be C^2.

Proof of Theorem 5.2: Parts (a) and (b) have already been proven. Note that the theory of Section 1 now applies and we have a classification of the singularities of Γ. The key point in proving (c) is to show that

$$|R \cap \partial W| \geqq \min\left\{\frac{\pi v_0}{12|g|^2_{\max}}, \quad \text{width}(R)\right\}$$

for all components W of $R - \Gamma$. Notice that the inequality does not depend on n. To prove the inequality, suppose that there is a W for which the inequality is false. First consider the case when one of the components γ of ∂W is wholly contained in R, i.e., does not meet ∂R. Let m be the number of singular points on γ. We claim that $m \leqq 6$. To see this, orient γ and let $\{\theta_i\}$ be the set of angle changes in the tangent directions at the singular points of γ (see Figure 22).

By the Gauss–Bonnet formula,

$$2\pi = \sum \theta_i + \int_\gamma \kappa \cdot n,$$

where κ is the curvature vector and n is the inward unit normal. By the results of

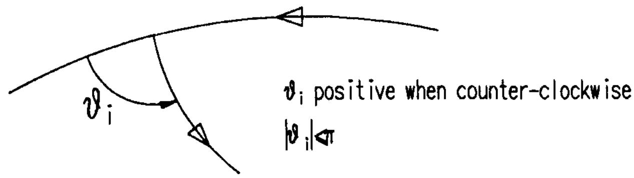

Figure 22. Conventions for the exterior angle θ_i.

Section 1, $\theta_i = \frac{1}{3}\pi$ for all i. By Lemmas 5.4 and 5.5,

$$|\kappa| \leqq \frac{4|g|^2_{\max}}{\nu_0}.$$

It follows that

$$m \leqq 6 + \frac{12}{\pi\nu_0}|g|^2_{\max}|\gamma| < 7.$$

Therefore γ contains a C^2-arc l of length greater than or equal to $\frac{1}{6}|\gamma|$. Suppose $l \subset U_i \cap U_j$, where U_i meets l along the side interior to γ. Let W' be the component of U_i meeting l. Consider a new segmentation $\{U_i'\}$ of R so that

$$\vec{U}_i' = \vec{U}_i - \vec{W}',$$

$$\vec{U}_j' = \vec{U}_j + \vec{W}',$$

$$U_k' = U_k \quad \text{if} \quad k \neq i, j.$$

$$\Delta E_n = E_n(\{U_i, a_i\}) - E_n(\{U_i', a_i\}) \geqq \nu_0 l - 2|g|^2_{\max} \cdot \operatorname{area}(W').$$

By the isoperimetric inequality,

$$\operatorname{area}(W') \leqq \frac{|\gamma|^2}{4\pi}.$$

Therefore,

$$\begin{aligned}
\Delta E_n &\geqq \tfrac{1}{6}\nu_0|\gamma| - \frac{|g|^2_{\max}}{2\pi}|\gamma|^2 \\
&\geqq \tfrac{1}{6}\nu_0|\gamma|\left[1 - \frac{3|g|^2_{\max}}{\pi\nu_0}|g|\right] \\
&> \tfrac{1}{6}\nu_0|\gamma|\left(\tfrac{3}{4}\right) > 0
\end{aligned}$$

which is a contradiction since $\{U_i, a_i\}$ minimizes E_n.

If $R \cap \partial W$ does not contain a loop, then it must have a component γ meeting ∂R twice. Moreover, since $|R \cap \partial W| < \text{width}(R)$, γ cannot meet the opposite sides of R. Now we apply the reflection construction where $R^{\#}$ is obtained by reflecting R across the two sides met by γ. Let $W^{\#}$ be the set in $R^{\#}$ corresponding to W, obtained by reflecting W. Then $\partial W^{\#}$ must contain a loop $\gamma^{\#}$. We argue now as before to get a contradiction again.

Observe that

$$E_n(\{U_i, a_i\}) = \tfrac{1}{2}\nu_0 \sum_{\substack{\text{components} \\ W \text{ of } R-\Gamma}} |R \cap \partial W| + \sum_{i=1}^{n} \int_{U_i} (g - a_i)^2$$

$$\leqq E_n(\{V_i, b_i\}), \quad \text{where } V_1 = R,\ b_1 = 0 \text{ and } V_i = \varnothing \text{ for } i > 1,$$

$$\leqq |g|_{\max}^2 \cdot \text{area}(R).$$

Therefore,

$$[\text{number of components of } R - \Gamma] \leqq \frac{(2/\nu_0)|g|_{\max}^2 \cdot \text{area}(R)}{\min\{\pi\nu/12|g|_{\max}^2, \text{width}(R)\}}.$$

Part (c) of Theorem 5.2 now follows.

6. Approximation when μ is Large

We derive here limiting forms of the energy and its first variation (along smooth portions of Γ) as $\mu \to \infty$. When μ is large, the effect of Γ on the energy is essentially confined to a narrow strip along Γ. We can even express the contribution of Γ to the energy as an integral along Γ and analyze the first variation in the form of variation of this line integral. The whole approach is based on the following:

LEMMA. *Suppose g_μ satisfies the equation $\nabla^2 g_\mu = \mu^2(g_\mu - g)$ everywhere in R and $\partial g_\mu/\partial n|_{\partial R} = 0$. Let f_Γ minimize $E(f, \Gamma)$ with Γ fixed. Let $h_\Gamma = f_\Gamma - g_\mu$. Then*

$$E(f_\Gamma, \Gamma) = E(g_\mu, \varnothing) - \int_\Gamma \left[\frac{\partial g_\mu}{\partial n}(h_\Gamma^+ - h_\Gamma^-) - \nu\right] ds,$$

where the superscripts + and − distinguish between the values of a variable on the two sides of Γ and $\partial/\partial n$ is the normal derivative in the direction from the − side to the + side.

Proof:

$$
\begin{aligned}
&E(f_\Gamma, \Gamma) - E(g_\mu, \phi) \\
&\quad = \iint_{R-\Gamma} \left[\mu^2(f_\Gamma - g)^2 + \|\nabla f_\Gamma\|^2\right] + \nu|\Gamma| \\
&\qquad - \iint_R \left[\mu^2(g_\mu - g)^2 + \|\nabla g_\mu\|^2\right] \\
&\quad = \iint_{R-\Gamma} \left[2\left\{\mu^2(g_\mu - g)h_\Gamma + \nabla g_\mu \cdot \nabla h_\Gamma\right\}\right. \\
&\qquad\qquad \left. + \left\{\mu^2 h_\Gamma^2 + \|\nabla h_\Gamma\|^2\right\}\right] + \nu|\Gamma| \\
&\quad = \iint_{R-\Gamma} \left[2h_\Gamma\left\{\mu^2(g_\mu - g) - \nabla^2 g_\mu\right\} + h_\Gamma\left(\mu^2 h_\Gamma - \nabla^2 h_\Gamma\right)\right] \\
&\qquad - \int_\Gamma \left[2\frac{\partial g_\mu}{\partial n}(h_\Gamma^+ - h_\Gamma^-) + \frac{\partial h_\Gamma^+}{\partial n}h_\Gamma^+ - \frac{\partial h_\Gamma^-}{\partial n}h_\Gamma^- - \nu\right] ds,
\end{aligned}
$$

by Green's identity.

The lemma follows since the first integrand in the last step is zero and $\partial h_\Gamma^+/\partial n = \partial h_\Gamma^-/\partial n = -\partial g_\mu/\partial n$ along Γ.

Thus Γ minimizes $E(f_\Gamma, \Gamma)$ if and only if it maximizes the line integral

$$\int_\Gamma \left[\frac{\partial g_\mu}{\partial n}(h^+ - h^-) - \nu\right] ds,$$

where h is the solution to the boundary value problem

$$\nabla^2 h = \mu^2 h \quad \text{on} \quad R - \Gamma,$$

$$\frac{\partial h}{\partial n} = -\frac{\partial g_\mu}{\partial n} \quad \text{along} \quad \Gamma \cup \partial R.$$

To understand the limiting behavior of this integral as $\mu \to \infty$, we need to describe the asymptotic behavior of h_Γ as $\mu \to \infty$. This involves considerable technical details, which we have put in Appendices 2 and 3. Appendix 2 is devoted to proving that

$$\sup_{P \in R} |h_\Gamma(P)| = O(1/\mu) \quad \text{as} \quad \mu \to \infty.$$

This estimate is very simple away from Γ and near smooth points of Γ, but to prove this near singularities of Γ seems harder (in fact, we had to exclude cusps on Γ). Appendix 3 is devoted to studying h_Γ near smooth points of Γ and deriving the precise asymptotic form. Introduce coordinates r and s along Γ, where $r(P)$ is the distance from P to the nearest point $\bar{P}$ on Γ and $s(P)$ is arc length on Γ from some origin to $\bar{P}$ (see Figure 3, Section 1). We can prove that if Γ is sufficiently smooth, e.g. $C^{3,1}$, then

$$h_\Gamma(r, s) = \frac{e^{-\mu r}}{\mu\sqrt{1-\kappa(s)r}}\left(1 + \frac{\kappa(s)}{2\mu}\right)\frac{\partial g}{\partial n}(0, s) + O\left(\frac{1}{\mu^3}\right),$$

where $\kappa(s)$ is the curvature of Γ at s. The proof is rather long, however, but the essence of it, a careful application of Green's theorem, is given in Appendix 3 for the points on Γ, where we shall use it. It is easy to derive the form of this expansion for h_Γ by examining the case $\Gamma = \textit{circle}$, and using the explicit expression for h_Γ in terms of Bessel functions of the 2-nd kind (cf. [14], Ch. 17).

An interesting question is to find the asymptotic expansion of h_Γ for large μ near the singularities of Γ. We were quite puzzled looking for appropriate "elementary" functions from which to construct this expansion. In the case where *Γ is the positive x-axis*, John Myers found a beautiful explicit formula for the h_Γ satisfying

(a) $\nabla^2 h_\Gamma = \mu^2 h_\Gamma$,
(b) $\partial h_\Gamma/\partial n = 1$ along Γ,
(c) $h_\Gamma(x, y) \approx \operatorname{sgn}(y)\cdot e^{-\mu y}/\mu$ if $x \gg 0$, $|y| \leqq C$,
(d) $h_\Gamma = O(e^{-\mu r})$ if $0 < \theta_0 \leqq \theta \leqq 2\pi - \theta_0$.

Using the error function, he introduces

$$g(x, y) = \frac{e^{-\mu y}}{\mu\sqrt{\pi}}\int_{-\infty}^{\mathcal{R}e(1+i)\sqrt{\mu z}} \exp\{-s^2\}\, ds.$$

One can check that (a) and (d) hold for g and that (b) and (c) also hold for

$$h_\Gamma(x, y) = g(x, y) - g(x, -y).$$

Using this special h_Γ, one should be able to construct good asymptotic approximations to the general h_Γ near crack-tips, as $\mu \to \infty$.

Applying the estimates for h_Γ in Appendices 2 and 3, we can estimate the behavior of $E(f_\Gamma, \Gamma)$ *when $\mu \to \infty$ and Γ is fixed.* As in the introduction, we write

$$E_\infty(\Gamma) = \int_\Gamma\left[\nu_\infty - \left(\frac{\partial g}{\partial n}\right)^2\right] ds.$$

We now set $\nu_\infty = \frac{1}{2}\mu\nu$. Then we have

THEOREM 6.1. *Suppose that g is $C^{1,1}$ and that Γ is the union of finitely many $C^{1,1}$-arcs. Assume that $\Gamma \cup \partial R$ has no cusps. Then, as $\mu \to \infty$,*

$$E(f_\Gamma, \Gamma) = E(g_\mu, \varnothing) + \frac{2}{\mu}E_\infty(\Gamma) + \|g\|^2_{2,\infty,R}\, O\left(\frac{\log\mu}{\mu^2}\right).$$

Proof: We first construct g_μ and derive estimates for it. Pave $\mathbb{R}^2$ by successive reflections of R about edges and thus extend g to all of $\mathbb{R}^2$. The solution g_μ can then be expressed as

$$g_\mu(P) = \frac{\mu^2}{2\pi}\int\int_{\mathbb{R}^2_Q} K_0(\mu|PQ|)g(Q)$$

for all $P \in \mathbb{R}^2$. Here K_0 is the zeroth-order modified Bessel function of the second kind.

We need estimates on g_μ and its first and second derivatives:

LEMMA. *For all points P in R and all μ such that $\mu \cdot \operatorname{dist}(P, \partial R) \geqq 3\log(\mu|\partial R|)$,*

$$g_\mu(P) = g(P) + \|g\|_{2,\infty,R}\, O\left(\frac{1}{\mu^2}\right),$$

$$\nabla g_\mu(P) = \nabla g(P) + \|g\|_{2,\infty,R}\, O\left(\frac{1}{\mu}\right),$$

$$\left|\frac{\partial^2 g_\mu}{\partial x_i\, \partial x_j}\right|_P \leqq \text{const.}\|g\|_{2,\infty,R} \quad \textit{for} \quad 1 \leqq i, j \leqq 2.$$

Moreover,

$$\|g_\mu\|_{1,\infty,R} \leqq \text{const.}\|g\|_{0,\infty,R}.$$

Proof: To see this, we first note that, for $n \geqq 0$ and $i = 0$ or 1,

$$\int\int_{\mathbb{R}^2} r^n K_i(\mu r) = 2\pi\int_0^\infty r^{n+1}K_i(\mu r)\, dr = \frac{2\pi}{\mu^{n+2}}\int_0^\infty z^{n+1}K_i(z)\, dz = O\left(\frac{1}{\mu^{n+2}}\right)$$

and

$$\int\int_{r \geqq d} K_i(\mu r) = 2\pi\int_d^\infty rK_i(\mu r)\, dr = \frac{2\pi}{\mu^2}\int_{\mu d}^\infty zK_i(z)\, dz$$

$$\leqq (\mu d)^{1/2} e^{-\mu d}\, O\left(\frac{1}{\mu^2}\right),$$

by Lemma 3 in Appendix 3,

$$\leqq O\left(\frac{1}{\mu^4}\right) \quad \text{if} \quad \mu d \geqq 3\log(\mu|\partial R|).$$

Let $d = \text{dist}(P, \partial R)$ and let U_d be the disc of radius d with center at P. Since

$$\frac{\mu^2}{2\pi}\iint_{\mathbf{R}^2} K_0(\mu r) = 1,$$

$$g_\mu(P) - g(P) = \frac{\mu^2}{2\pi}\iint_{\mathbf{R}^2_Q} K_0(\mu r)[g(Q) - g(P)]$$

we find

$$\begin{aligned}
&\left|\iint_{U_d} K_0(\mu r)[g(Q) - g(P)]\right| \\
&\quad \leqq \left|\iint_{U_d} K_0(\mu r)\left[\left.\frac{\partial g}{\partial x}\right|_P x + \left.\frac{\partial g}{\partial y}\right|_P y + \|g\|_{2,\infty,R}\, O(r^2)\right]\right| \\
&\quad \leqq \|g\|_{2,\infty,R}\, O\left(\frac{1}{\mu^4}\right)
\end{aligned}$$

and

$$\left|\iint_{\mathbf{R}^2_Q - U_d} K_0(\mu r)[g(Q) - g(P)]\right| \leqq \|g\|_{0,\infty,R}\, O\left(\frac{1}{\mu^4}\right).$$

Likewise,

$$\begin{aligned}
\nabla g_\mu(P) - \nabla g(P) &= \frac{\mu^2}{2\pi}\iint_{\mathbf{R}^2_Q}[\nabla_P K_0(\mu r) g(Q) - K_0(\mu r)\nabla g(P)] \\
&= \frac{\mu^2}{2\pi}\iint_{\mathbf{R}^2_Q}[-\nabla_Q K_0(\mu r) g(Q) - K_0(\mu r)\nabla g(P)],
\end{aligned}$$

$$\left|\int_{U_d}[\nabla_Q K_0(\mu r) g(Q) + K_0(\mu r)\nabla g(Q)]\right| \leqq 2\pi K_0(\mu d)\|g\|_{0,\infty,R},$$

by Green's identity,

$$\leqq \text{const.}(\mu d)^{-1/2} e^{-\mu d}\|g\|_{0,\infty,R},$$

by Lemma 3, Appendix 3,

$$\leqq \|g\|_{0,\infty,R}\, O\left(\frac{1}{\mu^3}\right).$$

Therefore,

$$\left| \iint_{U_d} \nabla_Q K_0(\mu r) g(Q) + K_0(\mu r) \nabla g(P) \right|$$

$$\leqq \left| \iint_{U_d} K_0(\mu r)[\nabla g(Q) - \nabla g(P)] \right| + \|g\|_{0,\infty,R}\, O\left(\frac{1}{\mu^3}\right)$$

$$\leqq \left| \iint_{U_d} K_0(\mu r) O(r) \|g\|_{2,\infty,R} \right| + \|g\|_{0,\infty,R}\, O\left(\frac{1}{\mu^3}\right)$$

$$\leqq \|g\|_{2,\infty,R}\, O\left(\frac{1}{\mu^3}\right).$$

Also

$$\left| \iint_{\mathbf{R}^2_Q - U_d} [\nabla_Q K_0(\mu r) g(Q) - K_0(\mu r) \nabla g(P)] \right|$$

$$\leqq \iint_{\mathbf{R}^2_Q - U_d} [\mu K_1(\mu r)|g(Q)| + K_0(\mu r)|\nabla g(P)|]$$

$$\leqq \|g\|_{1,\infty,R}\, O\left(\frac{1}{\mu^3}\right).$$

The remaining estimates may be proved in the same way.

We proceed now with the proof of the theorem. If Γ contains arcs which terminate in free ends (i.e., crack-tips), extend these arcs in some $C^{1,1}$ way until they meet $\Gamma \cup \partial R$ (without creating cusps, however). Let $\overline{\Gamma}$ denote the extended Γ. In order to apply the theorem in Appendix 3, we need to define several constants which depend only on Γ. If $\gamma \subset \gamma^* \subset \Gamma$ are curves, let

$$d(\gamma, \gamma^*) = \begin{cases} \infty, \text{ if } \gamma^* \text{ is a connected closed curve;} \\ \text{arc length between } \gamma \text{ and the end points of } \gamma^*, \\ \quad \text{if } \gamma \text{ and } \gamma^* \text{ are connected and open;} \\ \min_i d(\gamma_i, \gamma_i^*), \text{ in general, where } \gamma_i \text{ are the components of } \gamma \text{ and} \\ \quad \gamma_i^* \text{ is the component of } \gamma^* \text{ containing } \gamma_i. \end{cases}$$

For some $\varepsilon > 0$, we introduce constants:

$$\Gamma_{\text{sing}} = \text{singular points of } \Gamma \cup \partial R \text{ (including the crack-tips)},$$

$$\Lambda_P = \text{the diameter of the largest circle } \textit{through } P$$
$$\text{contained in a single component of } R - \overline{\Gamma},$$

$$\gamma_\varepsilon = \left\{ P \in \Gamma | \operatorname{dist}(P, \Gamma_{\text{sing}}) \geqq \varepsilon \right\},$$

$$\Lambda_\varepsilon = \min_{P \in \gamma_\varepsilon} \Lambda_P.$$

Since there are no cusps and since each arc of Γ is $C^{1,1}$, there exist constants C_1, C_2 and C_3 such that

$$\Lambda_\varepsilon \geqq C_1 \varepsilon,$$

$$\operatorname{dist}(\gamma_\varepsilon, \partial R) \geqq C_2 \varepsilon,$$

and

$$\text{arc length}(\Gamma - \gamma_\varepsilon) \leqq C_3 \varepsilon.$$

Let $C = \min\{\frac{1}{32}, \frac{1}{8}C_1, \frac{1}{2}C_2\}$, α_i be the C^2-norm of the i-th arc of Γ and $\alpha = \max_i \alpha_i$. Let β be the universal constant as defined in Lemma 2 of Appendix 3. Let

$$\mu_0 = \frac{6\alpha}{\beta} \log(\mu|\partial R|).$$

Now fix $\mu \geqq \mu_0$ and let $\varepsilon = (6/\mu C)\log(\mu|\partial R|)$. Let $\gamma = \gamma_\varepsilon$ and $\gamma^* = \gamma_{\varepsilon/2}$. Then $d(\gamma, \gamma^*) \geqq \frac{1}{2}\varepsilon$ and if

$$\delta = \min\left\{ \frac{\beta}{\alpha}, \tfrac{1}{16} d(\gamma, \gamma^*), \tfrac{1}{4}\Lambda_{\varepsilon/2} \right\},$$

then

$$\mu\delta \geqq 6 \log(\mu|\partial R|).$$

Moreover,

$$\mu \operatorname{dist}(\gamma^*, \partial R) \geqq 6 \log(\mu|\partial R|)$$

and

$$\text{arc length}(\Gamma - \gamma) \leqq \frac{6C_3}{C} \frac{\log(\mu|\partial R|)}{\mu}.$$

Let $h_\Gamma = f_\Gamma - g_\mu$ as above. We now apply Part A of the theorem in Appendix 3. This shows that, for each component Ω of $R - \bar{\Gamma}$,

$$h_\Gamma|_\gamma = \frac{1}{\mu}\frac{\partial g_\mu}{\partial n} + \left\|\frac{\partial g_\mu}{\partial n}\right\|_{1,\infty,\gamma^*} O\left(\frac{1}{\mu^2}\right) + \|h_\Gamma\|_{0,\infty,\partial\Omega} O\left(\frac{1}{\mu^4}\right) + \left\|\frac{\partial h_\Gamma}{\partial n}\right\|_{0,1,\partial\Omega} O\left(\frac{1}{\mu^6}\right).$$

Substitute the following estimates in this formula:

$$\frac{\partial g_\mu}{\partial n} = \frac{\partial g}{\partial n} + \|g\|_{2,\infty,R}\, O\left(\frac{1}{\mu}\right),$$

$$\|h_\Gamma\|_{0,\infty,\partial\Omega} \leqq \|g_\mu\|_{0,\infty,\partial\Omega} + \|f_\Gamma\|_{0,\infty,\partial\Omega} \leqq 2\|g\|_{0,\infty,R}.$$

Note that $\partial h_\Gamma/\partial n$ must be estimated on $\bar{\Gamma} - \Gamma$ as well as on Γ and it will certainly have singularities at crack-tips. To deal with the last term, recall that $h_\Gamma \in W_p^2(R - \Gamma)$ for $1 \leqq p < \frac{4}{3}$ by the results in Appendix 1. Therefore, by Theorem 2.3.3.6 in Grisvard [7] (applied to g_μ and f_Γ separately),

$$\|h_\Gamma\|_{2,p,R-\Gamma} \leqq \|g\|_{0,p,R-\Gamma}\, O(\mu^2) \leqq \|g\|_{0,\infty,R}\, O(\mu^2);$$

hence

$$\left\|\frac{\partial h_\Gamma}{\partial n}\right\|_{0,1,\partial\Omega} \leqq \|g\|_{0,\infty,R}\, O(\mu^2),$$

by the trace formula in Sobolev spaces (see [7], Section 1.5). We get

$$h_\Gamma|_\gamma = \frac{1}{\mu}\frac{\partial g}{\partial n} + \|g\|_{2,\infty,R}\, O\left(\frac{1}{\mu^2}\right)$$

and

$$\int_\gamma \frac{\partial g_\mu}{\partial n}(h_\Gamma^+ - h_\Gamma^-)\, ds = \frac{2}{\mu}\int_\gamma \left(\frac{\partial g}{\partial n}\right)^2 ds + \|g\|_{2,\infty,R}^2\, O\left(\frac{1}{\mu^2}\right).$$

By the theorem in Appendix 2,

$$\|g_\mu - g\|_{0,\infty,R} \leqq \|g\|_{1,\infty,R}\, O\left(\frac{1}{\mu}\right)$$

and

$$\|f_\Gamma - g\|_{0,\infty,R} \leqq \|g\|_{1,\infty,R}\, O\left(\frac{1}{\mu}\right).$$

Therefore,

$$\begin{aligned}\|h_\Gamma\|_{0,\infty,R} = \|f_\Gamma - g_\mu\|_{0,\infty,R} &\leqq \|f_\Gamma - g\|_{0,\infty,R} + \|g - g_\mu\|_{0,\infty,R}\\ &\leqq \|g\|_{1,\infty,R}\, O\left(\frac{1}{\mu}\right).\end{aligned}$$

It follows that

$$\left|\int_{\Gamma-\gamma} \frac{\partial g_\mu}{\partial n}(h_\Gamma^+ - h_\Gamma^-)\, ds\right| \leqq \|g\|_{1,\infty,R}^2 |\Gamma - \gamma| O\left(\frac{1}{\mu}\right)$$

$$\leqq \|g\|_{1,\infty,R}^2\, O\left(\frac{\log \mu}{\mu^2}\right).$$

This finishes the proof of Theorem 6.1. Note that, by a similar argument, we can also prove

$$E(g_\mu, \varnothing) = E(g, \varnothing) - \frac{1}{\mu}\int_{\partial R}\left(\frac{\partial g}{\partial n}\right)^2 ds + \|g\|_{2,\infty,R}^2\, O\left(\frac{\log \mu}{\mu^2}\right).$$

We derive now the asymptotic form of the equation of first variation. We show that the first variation of $E_\infty(\Gamma)$ agrees with that of $E(f_\Gamma, \Gamma)$, at least when Γ_{sing} is kept fixed. For simplicity, we assume that Γ consists of $C^{2,1}$ arcs only. We shall work with an *oriented* piece $\gamma \subset \Gamma - \Gamma_{\text{sing}}$ and then label the left side of γ "$-$", the right side "$+$". This fixes the sign of the unit tangent vector t to γ, the unit normal n (let it point from the $-$ side to the $+$ side) and the curvature κ of γ via its definition:

$$\frac{dt}{ds} = \kappa \cdot n.$$

THEOREM 6.2. *Assume that g is $C^{2,1}$. Let γ be an oriented $C^{2,1}$-curve contained in Γ, not meeting the singularities of $\Gamma \cup \partial R$. Then for all C^1-fields $\mathbf{X}$ of normal vectors along Γ with support contained in γ:*

(i) *The first variation is given by*

$$\delta_{\mathbf{X}} E_\infty(\gamma) = -\int_\gamma [\mathscr{D}(g) + \tfrac{1}{2}\mu\nu\kappa](\mathbf{X}\cdot n)\, ds,$$

where

$$\mathscr{D}(g) = 2\left(\frac{\partial g}{\partial n}\nabla^2 g + \frac{\partial g}{\partial s}\cdot\frac{\partial}{\partial s}\frac{\partial g}{\partial n}\right) + \kappa\left(\frac{\partial g}{\partial n}\right)^2.$$

(ii) *There exists a constant ρ_γ such that for all μ satisfying $\mu\rho_\gamma \geqq 6\log(\mu|\partial R|) \geqq 1$:*

(a) $e^+(f_\Gamma) - e^-(f_\Gamma) = \dfrac{2}{\mu}\mathscr{D}(g) + \|g\|_{3,\infty,R}^2\, O(1/\mu^2)$,

where $e(f_\Gamma)$ is the energy density corresponding to f_Γ.

Hence,

(b) $\delta_{\mathbf{X}} E(f_\Gamma, \Gamma) = (2/\mu)\,\delta_{\mathbf{X}} E_\infty(\gamma) + \|\mathbf{X}\cdot n\|_{0,\infty,\gamma}\|g\|_{3,\infty,R}^2\, O(1/\mu^2)$.

Proof: Let $\varphi = (\mathbf{X} \cdot n)$. Let $\vec{r}(s)$ denote the parametrization of γ by its arc length. Let $\vec{r}_\lambda(s) = \vec{r}(s) + \lambda \mathbf{X}(s)$. For small $|\lambda|$, $\vec{r}_\lambda$ parametrizes curves γ_λ near γ. Let s_λ, t_λ and n_λ denote, respectively, the arc length, unit tangent vectors and unit normal vectors along γ_λ:

$$\frac{\partial \vec{r}_\lambda}{\partial s} = t_0 + \lambda \frac{d\varphi}{ds} n_0 - \lambda \varphi \kappa t_0.$$

Hence, up to the first-order terms in λ, we have

$$ds_\lambda = (1 - \lambda \varphi \kappa)\, ds,$$

$$t_\lambda = t_0 + \lambda \frac{d\varphi}{ds} n_0,$$

$$n_\lambda = n_0 - \lambda \frac{d\varphi}{ds} t_0.$$

Therefore,

$$\begin{aligned}\delta_{\mathbf{X}}\left[\left(\frac{\partial g}{\partial n}\right)^2 ds - \frac{1}{2}\mu\nu\, ds\right] &= 2\frac{\partial g}{\partial n}\left[(\delta_{\mathbf{X}} \nabla g) \cdot n + (\nabla g) \cdot \delta_{\mathbf{X}} n\right] ds \\ &\quad + \left[\left(\frac{\partial g}{\partial n}\right)^2 - \frac{1}{2}\mu\nu\right] \delta_{\mathbf{X}}(ds) \\ &= \left\{2\frac{\partial g}{\partial n}\left[\frac{\partial^2 g}{\partial n\, \partial n}\varphi - \frac{\partial g}{\partial s}\frac{d\varphi}{ds}\right] - \kappa\left[\left(\frac{\partial g}{\partial n}\right)^2 - \tfrac{1}{2}\mu\nu\right]\varphi\right\} ds\end{aligned}$$

(where $\partial^2 g/\partial n\, \partial n$ means that, at each point of γ, we take the 2-nd derivative of g along the straight line normal to γ at that point). But

$$-\int_\gamma \frac{\partial g}{\partial n}\frac{\partial g}{\partial s}\frac{d\varphi}{ds}\, ds = \int_\gamma \left[\frac{\partial g}{\partial n}\frac{\partial^2 g}{\partial s^2} + \frac{\partial g}{\partial s}\frac{\partial}{\partial s}\left(\frac{\partial g}{\partial n}\right)\right]\varphi\, ds$$

(here $\partial^2 g/\partial s^2$ means the 2-nd derivative of g restricted to γ) and

$$\frac{\partial^2 g}{\partial n\, \partial n} + \frac{\partial^2 g}{\partial s^2} = \nabla^2 g + \kappa \frac{\partial g}{\partial n}$$

so we get part (i) by substitution.

To prove part (ii), we now apply Part B of the theorem in Appendix 3. If Γ has arcs terminating in free ends, extend these arcs to meet $\Gamma \cup \partial R$ without creating cusps. Pick γ^* such that $\gamma \subset \gamma^* \subset \Gamma$, $\operatorname{dist}(\gamma, \gamma^*) > 0$ and γ^* does not meet the singularities of $\Gamma \cup \partial R$. Let δ be the constant specified in the theorem in Appendix 3. (We may choose δ so that it serves all the components of $R - \bar{\Gamma}$).

Let $\varepsilon = \operatorname{dist}(\gamma^*, \partial R)$. Let $\rho_\gamma = \min\{\delta, \varepsilon\}$. Let $R_\varepsilon = \{P \in R | \operatorname{dist}(P, \partial R) \geqq \varepsilon\}$. Following the method used in the proof of the lemma in Theorem 4.1, we can extend the estimates for g_μ as follows. For all $P \in R_\varepsilon$ and for all $\mu \geqq 3\log(\mu|\partial R|)/\rho_\gamma$,

(i) for $0 \leqq k \leqq 3$, $\|g_\mu\|_{k,\infty,R_\varepsilon} \leqq (\text{const.})\|g\|_{k,\infty,R}$,

(ii) $\|g_\mu\|_{4,\infty,R_\varepsilon} \leqq \|g\|_{3,\infty,R}\, O(\mu)$,

(iii) Since

$$g_\mu - g = \frac{\nabla^2 g_\mu}{\mu^2} = \frac{\nabla^2 g + \nabla^2(g_\mu - g)}{\mu^2} = \frac{\nabla^2 g}{\mu^2} + \frac{\nabla^4 g_\mu}{\mu^4},$$

$$g_\mu = g + \frac{\nabla^2 g}{\mu^2} + \|g\|_{3,\infty,R}\, O\left(\frac{1}{\mu^3}\right)$$

and

$$\|g_\mu - g\|_{k,\infty,R_\varepsilon} \leqq \|g\|_{3,\infty,R}\, O\left(\frac{1}{\mu^{3-k}}\right) \quad \text{for} \quad k = 1,2.$$

Let $h_\Gamma = f_\Gamma - g_\mu$ as before. From Appendix 3 we get, for each component Ω of $R - \overline{\Gamma}$,

$$h_\Gamma|_{\gamma \cap \partial\Omega} = \frac{1}{\mu}\frac{\partial g_\mu}{\partial n}\left(1 + \frac{\kappa}{2\mu}\right) + C_{h,\gamma^*} O\left(\frac{1}{\mu^3}\right)$$

and

$$\left.\frac{\partial h_\Gamma}{\partial s}\right|_{\gamma \cap \partial\Omega} = \frac{1}{\mu}\frac{\partial}{\partial s}\frac{\partial g}{\partial n} + C_{h,\gamma^*} O\left(\frac{1}{\mu^2}\right),$$

where

$$C_{h,\gamma^*} = \left\|\frac{\partial g_\mu}{\partial n}\right\|_{2,\infty,\gamma^*} + \|h_\Gamma\|_{0,\infty,\partial\Omega}\, O\left(\frac{1}{\mu}\right) + \left\|\frac{\partial h_\Gamma}{\partial n}\right\|_{0,1,\partial\Omega} O\left(\frac{1}{\mu^3}\right).$$

We have

$$\frac{\partial g_\mu}{\partial n} = \frac{\partial g}{\partial n} + \|g\|_{3,\infty,R}\, O\left(\frac{1}{\mu^2}\right),$$

$$\left\|\frac{\partial g_\mu}{\partial n}\right\|_{2,\infty,\gamma^*} \leqq \text{const.}\|g\|_{3,\infty,R},$$

$$\|h_\Gamma\|_{0,\infty,\partial\Omega} \leqq 2\|g\|_{0,\infty,R},$$

$$\left\|\frac{\partial h_\Gamma}{\partial n}\right\|_{0,1,\partial\Omega} = \|g\|_{0,\infty,R}\, O(\mu^2).$$

Therefore,

$$h^+ - h^- = \frac{2}{\mu}\frac{\partial g}{\partial n} + \|g\|_{3,\infty,R}\, O\left(\frac{1}{\mu^3}\right),$$

$$h^+ + h^- = \frac{\kappa}{\mu^2}\frac{\partial g}{\partial n} + \|g\|_{3,\infty,R}\, O\left(\frac{1}{\mu^3}\right),$$

$$\frac{\partial h^\pm}{\partial s} = \frac{\pm 1}{\mu}\frac{\partial}{\partial s}\left(\frac{\partial g}{\partial n}\right) + \|g\|_{3,\infty,R}\, O\left(\frac{1}{\mu^2}\right).$$

Now

$$\begin{aligned} e^+(f_\Gamma) - e^-(f_\Gamma) &= \mu^2\left[(f_\Gamma^+ - g)^2 - (f_\Gamma^- - g)^2\right] + \left[\left(\frac{\partial f_\Gamma^+}{\partial s}\right)^2 - \left(\frac{\partial f_\Gamma^-}{\partial s}\right)^2\right] \\ &= \mu^2(f_\Gamma^+ - f_\Gamma^-)(f_\Gamma^+ + f_\Gamma^- - 2g) \\ &\quad + \left(\frac{\partial f_\Gamma^+}{\partial s} - \frac{\partial f_\Gamma^-}{\partial s}\right)\left(\frac{\partial f_\Gamma^+}{\partial s} + \frac{\partial f_\Gamma^-}{\partial s}\right) \\ &= \mu^2(h^+ - h^-)(2g_\mu - 2g + h^+ + h^-) \\ &\quad + \left(\frac{\partial h^+}{\partial s} - \frac{\partial h^-}{\partial s}\right)\left(2\frac{\partial g_\mu}{\partial s} + \frac{\partial h^+}{\partial s} + \frac{\partial h^-}{\partial s}\right). \end{aligned}$$

Part (ii) of the theorem follows by substitution.

7. The Case $\mu = \infty$

In the last section, we have argued that, in a certain sense, E_∞ is the limiting functional of E when $\mu \to \infty$ and $\frac{1}{2}\mu\nu$ has a finite limit ν_∞. However, the sense in question involves *fixing* Γ *while* μ *increases*. To see if this is reasonable, consider the problem of minimizing the limit functional

$$E_\infty(\Gamma) = \int_\Gamma \left(\nu_\infty - \left(\frac{\partial g}{\partial n}\right)^2\right) dx$$

over all Γ. There are two cases:

(a) $\sup_R \|\nabla g\|^2 \leqq \nu_\infty$.
(b) For some $P \in R$, $\|\nabla g(P)\|^2 > \nu_\infty$.

In the first case, E_∞ is clearly minimized by $\Gamma = \varnothing$. In the second case, we may make Γ's with more and more components each of which is a short arc from a

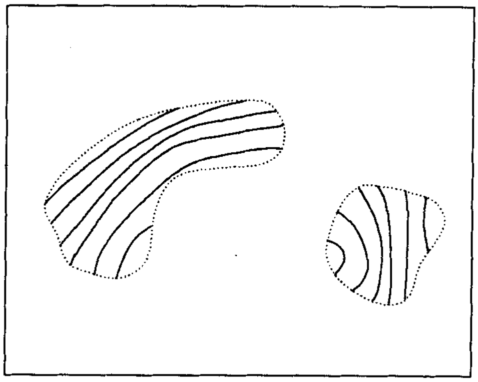

Figure 23. Conjectural form of Γ, μ large.

level curve of g in the region R_0 where $\|\nabla g\|^2 > \nu_\infty$. On such Γ's, $\inf E_\infty = -\infty$, so no minimum exists!

On the other hand, we conjecture that minimizing E itself is a well-posed problem. The explanation of this is that for $\mu \gg 0$ and $\mu\nu$ fixed, the minimum of $E(f_\Gamma, \Gamma)$ will presumably be taken on for a Γ supported in R_0 and made up of a curve which locally has more and more components, these being smooth nearly parallel curves with a separation of about c/μ (see Figure 23). The limit of such Γ should be taken as a *current*, not a curve.

Instead, the reason we are interested in the limit E_∞ is the hope that an approximate solution to minimizing E can sometimes be obtained by a very different procedure, involving two steps:

(a) smoothing g by convolution with a suitable kernel of size c/μ, so as to create a modified problem in which $1/\mu$ is already small compared to fluctuations in f,
(b) solving for a curve Γ which maximizes $E_\infty(\Gamma)$ for this smoothed g and with respect to all small deformations of Γ (as in Section 3).

Unfortunately, it is still unclear whether there is a natural class of g's for which Γ's satisfying (b) can be found. We hope to clarify this in a subsequent paper, and will only make a few preliminary remarks here. As we have seen in Section 4, requiring that the first variation of E_∞ vanishes on a smooth curve Γ is equivalent to asking that, *along* Γ,

$$(*) \qquad \frac{\partial g}{\partial n}\nabla^2 g + \frac{\partial g}{\partial s}\frac{\partial}{\partial s}\left(\frac{\partial g}{\partial n}\right) + \tfrac{1}{2}\kappa\left[\left(\frac{\partial g}{\partial n}\right)^2 + \nu_\infty\right] = 0.$$

Now g is a given function and this is *not* a PDE for g: instead it is a 2-nd order ODE for Γ! To see this, let $(x(t), y(t))$ be a parametrization of Γ and let $\theta(t)$ be the angle between the tangent line to Γ and the positive x-axis. Then

$$\frac{\partial}{\partial s} = \dot{x}(t)\frac{\partial}{\partial x} + \dot{y}(t)\frac{\partial}{\partial y},$$

$$\frac{\partial}{\partial n} = -\dot{y}(t)\frac{\partial}{\partial x} + \dot{x}(t)\frac{\partial}{\partial y}.$$

If we let t be the arc length s, then

$$(\dot{x}, \dot{y}) = (\cos\theta, \sin\theta),$$

$$\dot{\theta} = \kappa.$$

Thus

$$\frac{\partial g}{\partial n} = -\sin\theta \cdot g_x + \cos\theta \cdot g_y,$$

$$\frac{\partial g}{\partial s} = \cos\theta \cdot g_x + \sin\theta \cdot g_y,$$

and

$$\begin{aligned}\frac{\partial}{\partial s}\left(\frac{\partial g}{\partial n}\right) &= \frac{d}{dt}\left(-\sin\theta \cdot g_x + \cos\theta \cdot g_y\right)\\ &= \cos\theta\sin\theta(g_{yy} - g_{xx}) + (\cos^2\theta - \sin^2\theta)g_{xy}\\ &\quad -\left(\cos\theta \cdot g_x + \sin\theta \cdot g_y\right)\kappa.\end{aligned}$$

With these formulae we can rewrite the requirement on Γ as the system of first-order ODE's for (x, y, θ):

$$\begin{aligned}\dot{x} &= \cos\theta,\\ \dot{y} &= \sin\theta,\\ \dot{\theta} &= \frac{d_g^*(\theta)\,\Delta g + d_g(\theta)\cdot\left(\frac{1}{2}\sin 2\theta(g_{yy} - g_{xx}) + \cos 2\theta \cdot g_{xy}\right)}{\tau(\theta)}.\end{aligned}$$

where

$$\begin{aligned}d_g(\theta) &= g_x\cos\theta + g_y\sin\theta,\\ d_g^*(\theta) &= -g_x\sin\theta + g_y\cos\theta,\\ \tau(\theta) &= d_g(\theta)^2 - \tfrac{1}{2}d_g^*(\theta)^2 - \tfrac{1}{2}\nu_\infty.\end{aligned}$$

However, we see immediately that these equations may be singular if the denominator $\tau(\theta)$ in the expression for $\dot{\theta}$ is 0. It is better not to use arc-length parametrization, but to rewrite the equations by multiplying each equation by $\tau(\theta)$ and then choose a new "time" t so that

$$\dot{x} = \tau(\theta)\cos\theta,$$

$$\dot{y} = \tau(\theta)\sin\theta,$$

$$\dot{\theta} = (\Delta g)\, d_g^*(\theta) + \left((g_{yy} - g_{xx})\tfrac{1}{2}\sin 2\theta + g_{xy}\cos 2\theta\right) \cdot d_g(\theta)$$

This form shows that the solutions are the integral curves of a smooth vector field on the unit tangent bundle T to R. (Note that $d_g : T \to \mathbb{R}$ is the function on T defined by the differential dg and d_g^* is the function on T defined by the differential $(^*dg)$.)

However, we may divide the unit tangent bundle T into two zones:

"space-like" zone: $\tau(\theta) < 0$,

"time-like" zone: $\tau(\theta) > 0$.

On the boundary between these two zones, $\dot{x} = \dot{y} = 0$, i.e., the vector field is "vertical". Integral curves of this field will, in general, be smooth curves on T, crossing from space-like to time-like and back again. But they project to curves on R with smooth space-like segments that end in cusps, at which they turn into smooth time-like segments. A computer-generated example of such a curve is shown in Figure 24. It corresponds to the case $g(x, y) = e^y$ and $\nu_\infty = 0$, for which all solutions are identical up to translation.

It can be shown that along its time-like segments, an integral curve locally *maximizes* E_∞ whereas along its space-like segments, it locally *minimizes* E_∞. For instance, if $\nu_\infty = 0$, the gradient curves of g are time-like integral curves (on gradient curves, $\partial g/\partial n \equiv 0$, hence $(*)$ is satisfied) and these give the absolute maximum, namely 0, of the functional

$$E_\infty = -\int \left(\frac{\partial g}{\partial n}\right)^2 ds.$$

Therefore, the existence problem for $\nu = \infty$ becomes:

> Find criteria for the existence of closed everywhere space-like integral curves of $(**)$, or singular curves, as in Section 2, whose smooth pieces are everywhere space-like integral curves of $(**)$.

Unfortunately, the requirement of being everywhere space-like is rather unstable.

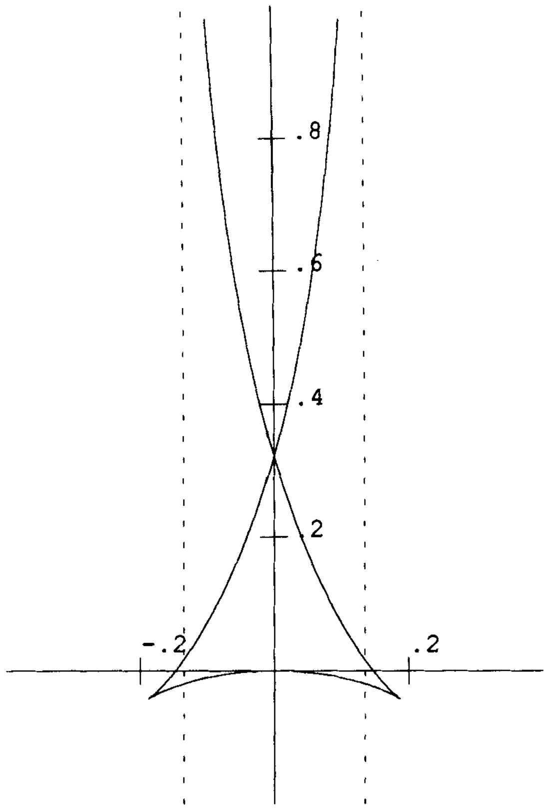

Figure 24. A solution of the ODE with "space"-and "time"-like segments.

Appendix 1

Boundary Behavior of the Solutions of Neumann Problems in Non-smooth Domains

A. We want to study the problem: given a bounded open set Ω in $\mathbb{R}^2$ and a function $g(x, y)$, continuous on $\overline{\Omega}$, solve for a function $f(x, y)$ such that

$$(*) \qquad \begin{aligned} \Delta f &= \mu^2(f - g), \\ \left.\frac{\partial f}{\partial n}\right|_{\partial\Omega} &\equiv 0. \end{aligned}$$

If Ω has a smooth, e.g., $C^{1,1}$-boundary, it is well known that $(*)$ has a unique solution f which is C^1 and has second derivatives which are in L_p for all $p < \infty$. If g and $\partial\Omega$ are a bit smoother, then so is f. Unfortunately we need the solution f for domains Ω with corners and slits and even with cusps at the end of the slits. We can find no standard reference for the existence and regularity of the solutions f in this case, hence we must summarize "standard" theory here. Grisvard [7] contains all the basic ideas that we need, but they must be adapted to our particular case.

B. First of all, we use the Hilbert space $W_2^1(\Omega)$ defined in the introduction and consider the positive definite continuous quadratic function on $W_2^1(\Omega)$:

$$E(f) = \int\int_\Omega \|\nabla f\|^2 \, dx\, dy + \mu^2 \int\int_\Omega (f-g)^2 \, dx\, dy.$$

By the Lax-Milgram lemma, E has a unique minimum, i.e., the function f such that

$$(**) \qquad \int\int_\Omega (\nabla\phi \cdot \nabla f)\, dx\, dy + \mu^2 \iint \phi \cdot f\, dx\, dy = \mu^2 \iint \phi \cdot g\, dx\, dy$$

for all functions $\phi \in W_2^1(\Omega)$. This f exists because $\phi \mapsto \mu^2 \int\int \phi \cdot g$ is a bounded linear functional on $W_2^1(\Omega)$, hence is given by dot product with some f.

C. Secondly, as a distribution on Ω, f satisfies $\Delta f = \mu^2(f-g)$. Putting Ω inside a large domain R with $C^{1,1}$-boundary, let f_R be the solution of $(*)$ for this domain R. f_R is C^1 with L^p-second derivatives for all p. But then, on Ω, $\Delta(f - f_R) = \mu^2(f - f_R)$, i.e., $f - f_R$ is in the kernel of the elliptic operator $\Delta - \mu^2$. Thus $f - f_R$ is C^∞. Therefore f is also C^1 on Ω with L^p-second derivatives. Thus f is a strong solution of $\Delta f = \mu^2(f-g)$ on Ω, i.e., we have equality a.e. of Δf with $\mu^2(f-g)$.

D. Thirdly, note that f is bounded on Ω by the max and min of g:

$$\min_\Omega g \leqq f(x) \leqq \max_\Omega g \quad \text{for all} \quad x \in \Omega.$$

To see this, let

$$f_1(x) = \max\Big(f(x), \min_\Omega g\Big),$$

$$f_2(x) = \min\Big(f_1(x), \max_\Omega g\Big).$$

Then at almost all points of Ω either $f_2 = f$ or f_2 is locally constant, hence, at almost all $x \in \Omega$,

$$\|\nabla f_2\|^2(x) \leqq \|\nabla f\|^2(x).$$

Moreover, at every $x \in \Omega$,

$$(f_2(x) - g(x))^2 \leqq (f(x) - g(x))^2.$$

Therefore,

$$E(f_2) \leqq E(f);$$

hence $f = f_2$ since f is the unique minimum of E.

E. Fourthly, look at points $P \in \partial\Omega$ such that $\partial\Omega \cap U$ is $C^{1,1}$ for some neighborhood U of P. We claim f is C^1 on $\overline{\Omega} \cap U$ and that $\partial f/\partial n|_{\partial\Omega \cap U} = 0$. To see this, introduce a subdomain U_P of Ω which is $C^{1,1}$ and whose boundary coincides with that of Ω near P (see Figure 25). Let η be a $C^{1,1}$-function which is 1 near P, which is zero on $\partial U_P - (\partial\Omega \cap \partial U_P)$ and such that $\partial\eta/\partial n|_{\partial U_P} = 0$. Then $\tilde{f} = \eta f$ satisfies

$$\Delta\tilde{f} - \mu^2\tilde{f} = \eta\mu^2 g + 2\nabla\eta \cdot \nabla f + f\Delta\eta.$$

Since $f \in W_2^1(U_P)$, $\nabla f \in L_2(U_P)$; hence

$$(\Delta\tilde{f} - \mu^2\tilde{f}) \in L_2(U_P).$$

Moreover, $\tilde{f}$ satisfies Neumann boundary conditions on U_P weakly, in the sense

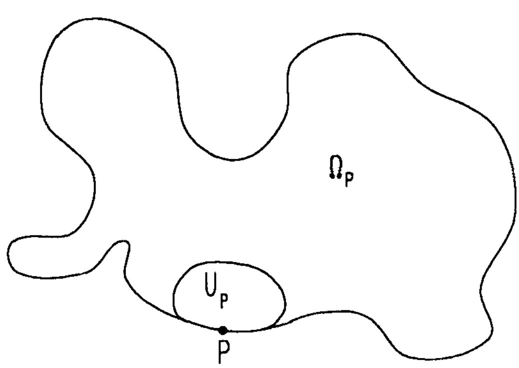

Figure 25. A smooth subdomain with same boundary near P.

of Lions and Magenes, i.e.,

$$\iint_{U_P} \psi \Delta \tilde{f} = \iint_{U_P} \tilde{f} \Delta \psi,$$

for all $\psi \in W_2^2(U_P)$ such that $\partial\psi/\partial\nu|_{\partial U_P} \equiv 0$.

This follows from applying the basic property $(**)$ of f with $\varphi = \eta\psi$ and some juggling with Green's theorem. We apply the weak existence theorem, Proposition 2.5.2.3 in Grisvard [7], to deduce that $\tilde{f} \in W_2^2(U_P)$. We can soup this up a bit, because now we know that $\partial\tilde{f}/\partial x$ and $\partial\tilde{f}/\partial y$ are in the space $W_2^1(U_P)$, hence they are in $L_p(U_P)$ for all p, by the Sobolev inequalities. Now repeat the argument with $\tilde{f} = \eta^2 f$, but starting with

$$\Delta\tilde{f} - \mu^2\tilde{f} = \eta^2\mu^2 g + 4\eta(\nabla\eta \cdot \nabla f) + f\Delta(\eta^2) \in L_p(U_P).$$

By the same result in Grisvard, $\tilde{f} \in W_p^2(U_P)$ for all p. This tells us that f is C^1 on $\bar{\Omega}$ near P, by the Sobolev inequalities, and now the weak Neumann condition becomes a strong Neumann condition:

$$\partial f/\partial n|_{\partial\Omega}(P) = 0.$$

F. Before studying the behavior of u in corners of Ω, we must make a digression on $C^{1,1}$-coordinate changes that straighten corners. We claim that for any two $C^{1,1}$-arcs meeting at $(0,0)$ with distinct tangent lines, in a neighborhood of $(0,0)$, there is a $C^{1,1}$-diffeomorphism Φ such that

(a) Φ maps the two arcs to straight rays through $(0,0)$,

(b) $d\Phi$ is the identity at $(0,0)$,

(c) $d\Phi$ carries normals to the two arcs to normals to the corresponding rays.

We prove this in two steps. In the first, we satisfy (a) and (b) but not (c). In the second, we satisfy (c) while preserving conditions (a) and (b).

Let the two arcs be

$$y = f(x), \qquad y = g(x), \qquad f, g \in C^{1,1}([O, c]),$$

and

$$\lambda = f'(0), \quad \mu = g'(0) \quad \text{where} \quad f(x) > g(x) \quad \text{if} \quad x \neq 0.$$

Define $\Phi_1 : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\Phi_1(x, y) = \begin{cases} (x, y) & \text{if } x \leqq 0, \\ \left(\dfrac{f(x) - g(x)}{\lambda - \mu}, y + \dfrac{\mu f(x) - \lambda g(x)}{\lambda - \mu} \right) & \text{if } x \geqq 0. \end{cases}$$

Φ_1 carries vertical lines to vertical lines. It carries $y = f(x)$ to the line $y = \lambda x$

and $y = g(x)$ to the line $y = \mu x$ and $d\Phi_1(0,0) = I_2$. Therefore Φ_1 is a $C^{1,1}$-diffeomorphism in some neighborhood of $(0,0)$. Next for any $C^{0,1}$-function $h(x)$ such that $h(x) = 0$ if $x \leqq 0$ and for any $\nu > 0$, define $\Psi_{h,\nu}$ by

$$\Psi_{h,\nu}(x, y) = \left(x + \int_{\nu x - 2y}^{\nu x - y} h(t)\, dt,\ y \right).$$

Then

$$\Psi_{h,\nu}(x, 0) = (x, 0),$$

$$\Psi_{h,\nu}(x, \nu x) = (x, \nu x),$$

$$d\Psi_{h,\nu}(x, y) = \begin{pmatrix} 1 + \nu(h(\nu x - y) - h(\nu x - 2y)) & 0 \\ 2h(\nu x - 2\eta) - h(\nu x - y) & 1 \end{pmatrix},$$

especially

$$d\Psi_{h,\nu}(x, \nu x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$d\Psi_{h,\nu}(x, 0) = \begin{pmatrix} 1 & 0 \\ h(\nu x) & 1 \end{pmatrix}.$$

Thus $\Psi_{h,\nu}$ is a $C^{1,1}$ diffeomorphism near $(0,0)$ which maps the sector $0 < y < \nu x$ to itself, preserves perpendiculars along $y = \nu x$, but maps the perpendiculars $\partial/\partial y$ to $y = 0$ to the vectors $\partial/\partial y + h(\nu x)\, \partial/\partial x$. $\Psi_{h,\nu}$ may be extended to any sector bounded by $y = 0$, $x > 0$ at one edge and including $0 < y < \nu x$ by setting it equal to the identity at the new points: it is still $C^{1,1}$.

We may then use a conjugate

$$\Phi_2 = A^{-1} \circ \Psi_{h,\nu} \circ A,$$

where A is a rotation or reflection, to modify the field of normal vectors to any sector at one edge without changing those at the other. Two such maps then will modify the two fields of normal vectors by arbitrary $C^{0,1}$-maps, so long as the normals of the apex do not move. The resulting $\Phi_2 \circ \Phi_1$ can be made to have property (c).

Finally note that this argument can be extended to map the exterior sector of two $C^{1,1}$ arcs *tangent* at $(0,0)$ or the complement of one $C^{1,1}$-arc to the complement of a single ray by a diffeomorphism Φ with properties (a), (b) and (c) above. Let the arcs be given by

$$y = f(x) \quad \text{and} \quad y = g(x),$$

where

$$\lambda = f'(0) = g'(0), \qquad f(x) \geqq g(x).$$

The exterior sector is given by $(y > f(x)) \cup (y < g(x)) \cup (x < 0)$. The complement of the $C^{1,1}$-arc is just the special case $f(x) \equiv g(x)$. To construct Φ, we first use the map

$$\Phi_1(x, y) = \begin{cases} (x, y) & \text{if } y \leqq 0, \\ (x, y - f(x) + \lambda x) & \text{if } y > f(x), \\ (x, y - g(x) + \lambda x) & \text{if } y < g(x), \end{cases}$$

and then "correct" the fields of normal vectors as before.

G. We next look at singular points $P \in \partial\Omega$ which are corners with convex angles α, $0 < \alpha < \pi$. At such points P, we claim that f is still C^1 in $\bar{\Omega} \cap U$ for a suitable neighborhood U of P, and since the two arcs of $\partial\Omega$ at P have distinct tangent lines, $(\nabla f)(P) = (0,0)$. To prove this we use the techniques in Grisvard, Section 5.2, together with a $C^{1,1}$-map Φ of the type just constructed. More precisely, we assume Φ is a mapping of a neighborhood W of P in $\mathbb{R}^2$ to a neighborhood of $(0,0)$ in $\mathbb{R}^2$ carrying $W \cap \Omega$ to an acute triangle $\tilde{\Omega}$ in $\mathbb{R}^2$ and such that $d\Phi(P) =$ identity, and Φ carries normals to $W \cap \partial\Omega$ to normals to $\partial\tilde{\Omega}$ (see Figure 26). Let Γ_1 be the two sides of $\tilde{\Omega}$ through $(0,0)$, and Γ_2 the third side. Let η be a C^∞-function of $\|z\|$ on $\tilde{\Omega}$ which is 1 near $(0,0)$ and 0 on Γ_2. Consider the function

$$\tilde{f}(z) = \eta(z) f(\Phi^{-1}(z))$$

on $\tilde{\Omega}$; $\tilde{f}$ satisfies a variable coefficient elliptic problem on $\tilde{\Omega}$:

$$A\tilde{f} = \Delta\tilde{f} + \sum b_{ij} \frac{\partial^2 \tilde{f}}{\partial x_i\, \partial x_j} + \sum b_i \frac{\partial \tilde{f}}{\partial x_i} - \mu^2 \tilde{f} = \tilde{g},$$

$$\partial\tilde{f}/\partial n|_{\Gamma_1} \equiv 0, \qquad \tilde{f}|_{\Gamma_2} \equiv 0,$$

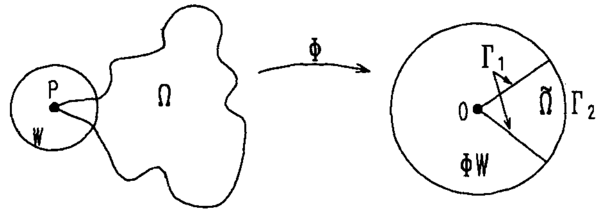

Figure 26. Straightening a corner of Ω.

for suitable bounded functions b_{ij}, b_i, c and $\tilde{g}$ on $\tilde{\Omega}$, b_{ij} continuous with $b_{ij}(0,0) = 0$, $\tilde{g}$ continuous. Next choose p such that

$$2 < p < \min_{\substack{\text{angles} \\ \beta \text{ of } \tilde{\Omega} \\ \text{with } \beta > \pi/2}} \left(\frac{2\beta}{2\beta - \pi} \right)$$

and consider the family of operators

$$B_t : W_p^2(\tilde{\Omega}) \to L_p(\tilde{\Omega}) \oplus W_p^{1-1/p}(\Gamma_1) \oplus W_p^{2-1/p}(\Gamma_2),$$

$$B_t(\tilde{f}) = \left(((1-t)A + t\Delta)\tilde{f} \partial\tilde{f}/\partial n|_{\Gamma_1}, \tilde{f}|_{\Gamma_2} \right).$$

As in Grisvard, B_t are bounded operators, B_1 is an isomorphism by Theorem 4.4.3.7. Therefore B_0 is an isomorphism. This shows $\tilde{f} \in W_p^2(\tilde{\Omega}) \subset C^1(\bar{\tilde{\Omega}})$ using the Sobolev inequalities and the assumption $p > 2$.

H. We may extend this last argument to cover corners P of $\partial\Omega$ with angles α such that $\pi < \alpha \leqq 2\pi$. Here $\alpha = 2\pi$ includes both the exterior sector to "cuspidal corners", pairs of $C^{1,1}$-arcs tangent at P and the exterior of a single $C^{1,1}$-arc ending at P. However, the result is different: f is *not* generally C^1 on $\bar{\Omega}$ near P. Instead f has a singular leading term.

Let $z = x + iy$ be a complex coordinate on $\mathbb{R}^2$ and let θ_1, θ_2 be the angles with the positive x-axis of the tangent lines to the arcs of $\partial\Omega$ at P ($\theta_1 = \theta_2$ allowed); see Figure 27. Let

$$f_{\text{sing}}(x, y) = \mathcal{R}e\left[e^{-i\theta_1}(z - z(P))\right]^{\pi/\alpha}.$$

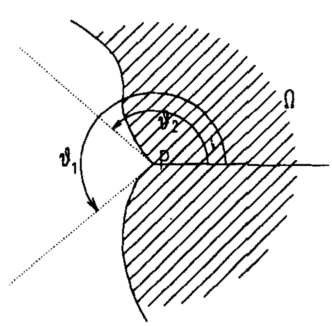

Figure 27. Orienting the sides at a corner of Ω.

Then we claim that, near P,

$$f = cf_{\text{sing}} + f_{\text{reg}}, \qquad f_{\text{reg}} \text{ a } C^1\text{-function}.$$

To prove this, use a Φ carrying a neighborhood W of P to a polygon $\tilde{\Omega}$, an η zero on $\partial W \cap \Omega$ and consider $\tilde{f} = \eta f \circ \Phi^{-1}$ as a function on $\tilde{\Omega}$ satisfying an elliptic boundary value problem with mixed Dirichlet-Neumann boundary conditions. Take p slightly larger than 2 and use Theorem 4.4.3.7 plus the methods of Section 5.2 to deduce that, on W_p^2, this boundary value problem has index 1, i.e., the map B_t defined in subsection **G** has index 1.

What we have to verify is that f_{sing} fills in the required range, i.e.,

(i) $(\Delta - \mu^2) f_{\text{sing}} \in L_p(\Omega)$,
(ii) $\partial f_{\text{sing}} / \partial n|_{\Gamma_1} \in W_p^{1-1/p}(\Gamma_1)$.

But f_{sing} is harmonic and continuous, so (i) follows.

To show (ii), let r, θ be polar coordinates at P and use r as a coordinate on the arcs of Γ_1. Take one of these and, assuming for simplicity that it is tangent to the x-axis, write it as

$$y = f(r), \qquad f \in C^{1,1},\ f(0) = f'(0) = 0,$$

or

$$\theta = \sin^{-1}\left(\frac{f(r)}{r}\right).$$

Let $g(r) = \sin^{-1}(f(r)/r)$. Now if $f(r)$ is $C^{1,1}$ and $f(0) = 0$, then $f(r)/r$ is $C^{0,1}$, hence $g(r) \in C^{0,1}$. Moreover,

$$rg'(r) = \frac{f'(r) - f(r)/r}{\sqrt{1 - (f(r)/r)^2}}$$

which is in $C^{0,1}$ also. Now, in these coordinates,

$$f_{\text{sing}} = r^{\pi/\alpha} \cos\left(\frac{\pi}{\alpha}\theta\right)$$

so

$$\begin{aligned}
\left.\frac{\partial f_{\text{sing}}}{\partial n}\right|_{\Gamma_1} &= \left(\frac{1}{r}\frac{\partial}{\partial \theta} - rg'(r)\frac{\partial}{\partial r}\right)\left(r^{\pi/\alpha}\cos\left(\frac{\pi}{\alpha}\theta\right)\right)\Big|_{\theta = g(r)} \\
&= -\frac{\pi}{\alpha}\sin\left(\frac{\pi}{\alpha}g(r)\right) \cdot r^{\pi/\alpha - 1} - \frac{\pi}{\alpha}\cos\left(\frac{\pi}{\alpha}g(r)\right)(rg'(r))r^{\pi/\alpha - 1} \\
&= (C^{0,1}\text{-function})\ r^{\pi/\alpha - 1}.
\end{aligned}$$

We use the following easy calculation:

LEMMA. *If* $g(x)$ *is a* $C^{0,1}$*-function on* $[0,1]$, *then* $g(x)\,x^k \in W_p^{1-1/p}([0,1])$ *if* $k \geqq -\frac{1}{2}$ *and* $p < 4$.

Differentiating f_{sing}, we also note:

LEMMA. *If* $1 \leqq p < \frac{4}{3}$, *then* f_{sing} *and hence* f *belongs to* $W_p^2(\Omega)$.

I. Another case that we need is a point $P \in \partial\Omega$ such that Ω is the complement of an arc ending at P with a *cusp* at P, i.e., an arc Γ given by

$$y = ax^{3/2} + g(x), \quad g \in C^{1,1}, \; g(0) = g'(0) = 0, \; a \in \mathbb{R}, \; a \neq 0,$$

after a rotation. This is a $C^{1,1/2}$-boundary point and cannot be straightened by a $C^{1,1}$-map Φ. Nevertheless we claim that, near P, the solution $f(x, y)$ has the form

$$f = a_0 + a_1\sqrt{r}\cos\left(\tfrac{1}{2}\theta\right) + R(x, y),$$

where

$$R \in O(r^{1-\varepsilon}),$$

$$\|\nabla R\| \in O(r^{-\varepsilon}) \quad \text{for all} \quad \varepsilon > 0.$$

In particular, f has a continuous extension to $\overline{\Omega}$ near P (where the two sides of Γ are considered as distinct points of $\overline{\Omega}$).

This can be shown by considering

$$\tilde{f}(u, v) = f(u^2 - v^2, 2uv)$$

defined on $\tilde{\Omega}$, one of the two components of

$$\left\{(u, v) | (u^2 - v^2, 2uv) \in \Omega\right\}.$$

In complex coordinates, we are letting $x + y = (u + iv)^2$. This opens up the cusp into $C^{1,1}$-boundary (with a discontinuity in the curvature at $(0, 0)$); see Figure 28. On $\tilde{\Omega}$, $\tilde{f}$ satisfies

$$\Delta\tilde{f} = 4\mu^2(u^2 + v^2)(\tilde{f} - \tilde{g}),$$

$$\partial\tilde{f}/\partial n|_{\tilde{\Gamma}} = 0.$$

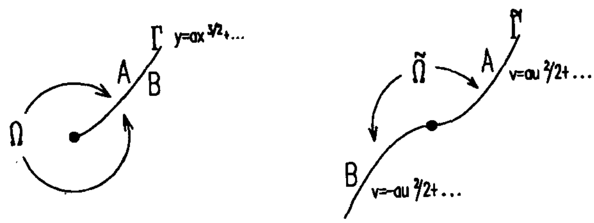

Figure 28. Smoothing a cuspidal crack by complex square root.

Therefore, as in subsection **G**, $\tilde{f} \in W_p^2(\tilde{\Omega})$, but for all $p < \infty$. Thus $\tilde{f}$ can be written as

$$\tilde{f} = a_0 + a_1 u + \tilde{R},$$

$$\tilde{R}(0,0) = \|\nabla \tilde{R}\|(0,0) = 0.$$

By the Sobolev inequalities, $\tilde{R} \in C^{1,\alpha}$ and $\nabla \tilde{R} \in C^{0,\alpha}$ for all $\alpha < 1$; hence

$$\tilde{R} = O\left((u^2 + v^2)^{(1+\alpha)/2}\right),$$

$$\|\nabla \tilde{R}\| = O\left((u^2 + v^2)^{\alpha/2}\right).$$

Going back to f, this gives us the expansion above.

J. One estimate is missing from our analysis: it would be nice to state it *inside* a cuspidal corner, i.e., in a domain

$$g(x) < y < f(x), \qquad x > 0,$$

where

$$f, g \in C^{1,1}, \quad f(0) = g(0) = 0, \quad f'(0) = g'(0), \quad g(x) < f(x) \quad \text{if} \quad x > 0,$$

the solution to the Neumann problem is C^1 at the corner $(0,0)$, or at least C^0. We do not know how to prove this. Fortunately we do not need it either.

Appendix 2

Estimation of Neumann Problems via Brownian Motion

Let D be a domain in $\mathbb{R}^2$ bounded by a finite number of $C^{1,1}$-curves, meeting at corners with various angles α_i or ending at crack-tips. The case $\alpha_i = 0$, i.e., cusps sticking out from D, is excluded, but $\alpha_i = 2\pi$, i.e., cusps sticking in, is allowed. Let $g \in C^1(\overline{D})$. We are interested in solving

$$\Delta f_\mu = \mu^2(f_\mu - g),$$

$$\partial f_\mu / \partial n|_{\partial D} \equiv 0,$$

for various μ's. The main result of this appendix is

THEOREM 1. *Given a domain D as above, there is a constant C_D such that, for all g and all μ,*

$$\sup_{P \in D} |f_\mu(P) - g(P)| \leq \sup_{P \in D} \|\nabla g(P)\| \cdot \frac{C_D}{\mu}.$$

We know two proofs of this result. One that we found and that seems very intuitive to us which uses Brownian motion, but does require some painful estimates. The other was found by S.-T. Yau after we posed the problem to him, which is based on iterating the Sobolev inequality infinitely often to estimate $\|f_\mu - g\|_p$ for higher and higher p in terms of $\|f_\mu - g\|_2$. Yau's proof demonstrates once again that things that can be proven with Brownian motion can generally be done directly. We give here our original proof because it seems so natural however.

We need Brownian motion $\vec{P}(t)$ with reflection on the boundary of D. The simplest way to define this process is to use the Riemann mapping theorem: Let $\Phi : D \to$ unit disk Δ be the conformal map. Then Φ extends homeomorphically from $\overline{D}$ to $\overline{\Delta}$, provided that boundary curves of D which meet D on both sides (e.g., edges ending in crack-tips) are counted twice as boundary curves. By a theorem of ordinary Brownian motion, $\vec{P}(t)$ on D becomes Brownian motion $\vec{P}_\Delta(t)$ on Δ run with a new clock $\tau(t)$:

$$\Phi\left(\vec{P}(t)\right) = \vec{P}_\Delta(\tau(t)),$$

$$\tau(t) = \int_0^t \left|\Phi'\left(\vec{P}(s)\right)\right|^2 ds.$$

(See e.g. [11].) On Δ, define Brownian motion with reflection in the usual way and carry it back to D by

$$\vec{P}(\tau^*(t)) = \Phi^{-1}\left(\vec{P}_\Delta(t)\right),$$

$$\tau^*(t) = \int_0^t \left|(\Phi^{-1})'\left(\vec{P}_\Delta(s)\right)\right|^2 ds.$$

Note that although $\Phi' = 0$ or ∞ at the corners of D, $\vec{P}_\Delta(t)$ hits the image of these corners with probability zero; hence there is no difficulty defining τ^* on almost all paths, and $\vec{P}(t)$ also hits a corner with probability zero.

The solution to the boundary value problem is now given by

$$f(\vec{x}) = \mathbf{E}\left[\int_0^\infty \mu^2 g\left(\vec{P}(t)\right)\exp\{-\mu^2 t\}\, dt \mid \vec{P}(0) = \vec{x}\right].$$

For details on the construction of $\vec{P}(t)$, see [11], especially Sections 3.9, 4.3 and 4.6, or [5], Section 1.6. For a proof of the formula, see [5], Section 2.5, Theorem 5.2.

We need a version of Ito's lemma to calculate the expected position of $\vec{P}(t)$:

Ito's Lemma. *Let f be a $C^{1,1}$-function on a subdomain $\overline{D}_0$ in $\overline{D}$ such that*

$$\partial f/\partial n|_{\overline{D}_0 \cap \partial D} \equiv 0.$$

Then

$$f\left(\vec{P}(t)\right) = b\left(\int_0^t \|\nabla f\left(\vec{P}(s)\right)\|^2\, ds\right) + \int_0^t \Delta f\left(\vec{P}(s)\right)\, ds,$$

where b is a one-dimensional Brownian motion.

Proof: This is Ito's lemma applied to $f(\vec{P}(t))$: cf. [11], Section 2.9 and [5], p. 96 (note: formula **C** there has a misprint.)

Using Ito's lemma, we now prove

Theorem 2. *Fix the domain D. Then there is an $r_0 > 0$ and a constant C_1 such that, for all $r \leqq r_0$, all $t \geqq 0$ and all $P \in D$,*

$$\mathbf{P}\left\{\begin{array}{l} \|\vec{P}(s) - P\| > r \text{ for} \\ \text{some } 0 \leqq s \leqq t, \text{ given} \\ \text{that } \vec{P}(0) = P \end{array}\right\} \leqq C_1 \exp\{-r^2/1600t\}.$$

Proof: First of all, we need coordinates adapted to D. Near the λ-th edge Γ_λ of D, let $d_\lambda(P)$ equal distance from P to Γ_λ and let $s_\lambda(P)$ equal arc length along Γ_λ from the initial point of Γ_λ to the nearest point on Γ_λ to P (cf. Figure 3, Section 1). Note that, since Γ_λ is $C^{1,1}$, s_λ and d_λ are $C^{1,1}$.

Secondly, near each corner Q_α of D, let Φ_α be a diffeomorphism of a neighborhood of Q_α to a neighborhood of $(0,0)$ as in Appendix 1: Φ_α carries the two edges of D abutting at Q_α to two rays in $\mathbb{R}^2$, it carries vectors perpendicular to these edges to vectors perpendicular to the rays and its differential at Q_α is the identity. Choose $r_0 \leqq r_1$ as follows:

(i) (s_λ, d_λ) are coordinates on a neighborhood U_λ of Γ_λ defined by $d_\lambda \leqq r_0$,
(ii) the (x, y)-metric and the (s_λ, d_λ)-metric differ by at most 1.1 on this neighborhood, i.e.,

$$dx^2 + dy^2 \leqq 1.1\left(ds_\lambda^2 + dd_\lambda^2\right) \leqq 1.21\left(dx^2 + dy^2\right),$$

(iii) $U_\lambda \cap U_\mu \neq \varnothing$ only if $\Gamma_\lambda, \Gamma_\mu$ meet at a vertex Q_α and then $U_\lambda \cap U_\mu \subset B_{r_1}(Q_\alpha)$,
(iv) Φ_α is defined in $B_{r_1}(Q_\alpha)$ and the (x, y)-metric differs from the $(x_\alpha = x \circ \Phi_\alpha, y_\alpha = y \circ \Phi_\alpha)$-metric by at most 1.1.

Define γ as the max of

$$|\Delta s_\lambda|, \quad |\Delta d_\lambda|, \quad |\Delta d_\lambda^2| \quad \text{in} \quad U_\lambda$$

and of

$$|\Delta \rho_\alpha^2| \quad \text{in} \quad B_{r_1}(Q_\alpha),$$

where $\rho_\alpha^2 \equiv x_\alpha^2 + y_\alpha^2$.

Now to prove the theorem, we distinguish three cases:

Case (i). $B_{2r/5}(P) \subset \operatorname{Int} D$.

Case (ii). The connected component of $(B_{2r/5} \cap D)$ containing P meets exactly one edge Γ_λ of D and no corner.

Case (iii). The above component meets two edges Γ_λ, Γ_μ of D.

In case (i), we have a small square S:

$$x(P) - \tfrac{1}{5}r \leqq x \leqq x(P) + \tfrac{1}{5}r,$$

$$y(P) - \tfrac{1}{5}r \leqq x \leqq y(P) + \tfrac{1}{5}r,$$

with $S \subset B_{2r/5}(P) \subset B_r(P)$ and $S \subset \operatorname{Int}(D)$. We estimate the probability that $\vec{P}(t)$ leaves S in time t by adding the probabilities of it hitting the four edges of S:

$$\begin{aligned}
\mathbf{P}\left\{\begin{matrix}\vec{P}(s) \notin S \text{ for} \\ \text{some } 0 \leqq s \leqq t \\ \text{given } \vec{P}(0) = P\end{matrix}\right\} &\leqq 4\mathbf{P}\left\{\begin{matrix}x\left(\vec{P}(s)\right) = x(P) + \tfrac{1}{5}r \\ \text{for some } 0 \leqq s \leqq t \\ \text{given } \vec{P}(0) = P\end{matrix}\right\} \\
&= 8\int_{r/5}^{\infty} \frac{\exp\{-x^2/2t\}}{\sqrt{2\pi t}}\,dx \\
&= 8\int_{r/5\sqrt{t}}^{\infty} \frac{\exp\{-\frac{1}{2}y^2\}}{\sqrt{2\pi}}\,dy \\
&\leqq 4\exp\{-r^2/50t\}.
\end{aligned}$$

We use here and below the simple bound (for any $a \geqq 0$):

$$\int_a^\infty \frac{\exp\{-\frac{1}{2}y^2\}}{\sqrt{2\pi}}\,dy = \exp\{-a^2h\}\int_0^\infty \frac{e^{-ay}\exp\{-\frac{1}{2}y^2\}}{\sqrt{2\pi}}\,dy \leqq \tfrac{1}{2}\exp\{-\tfrac{1}{2}a^2\}.$$

In case (ii), we have a curvilinear rectangle S around P:

$$s_\lambda(P) - \tfrac{1}{5}r \leqq s_\lambda \leqq s_\lambda(P) + \tfrac{1}{5}r,$$

$$\max\left(0, d_\lambda(P) - \tfrac{1}{5}r\right) \leqq d_\lambda \leqq d_\lambda(P) + \tfrac{1}{5}r.$$

Again, in this case, $S \subset B_{2r/5}(P) \subset B_r(P)$, because the furthest point of S from P in d_λ, s_λ-coordinates has distance $\sqrt{2r}/5$, and $\|\nabla d_\lambda\|$, $\|\nabla s_\lambda\| \leqq 1.1$ in S, hence these points have distance at most $\frac{2}{5}r$ from P in x, y-coordinates. Now there are really two subcases:

Case (iia). $d_\lambda(P) > \frac{1}{5}r$.

Case (iib). $d_\lambda(P) \leqq \frac{1}{5}r$.

In case (iia), S is in the interior of D and we estimate the probability of a large excursion of d_λ or s_λ using Ito's lemma. Since

$$|\Delta d_\lambda| \leqq \gamma \quad \text{in} \quad S,$$

and since $\|\nabla d_\lambda\| \leqq 1.1$, the lemma tells us that

$$\mathbf{P}\begin{pmatrix} d_\lambda(\vec{P}(s)) = d_\lambda(P) + \frac{1}{5}r \\ \text{for some } 0 \leqq s \leqq t, \\ \text{given } \vec{P}(0) = P \end{pmatrix} \leqq 2\int_{r/5-\gamma t}^\infty \frac{\exp\{-x^2/2.42t\}}{\sqrt{2.42\pi t}}\,dx.$$

(Note that we are adjusting for a worst case of Brownian clock speed 1.21 times normal and a maximum possible drift γt—so that the Brownian factor must have an excursion of at least $\frac{1}{5}r - \gamma t$.) A similar estimate holds for $d_\lambda(\vec{P}(s)) = d_\lambda(P) - \frac{1}{5}r$ and for $s_\lambda(P) \pm \frac{1}{5}r$. Thus

$$\begin{aligned} \mathbf{P}\begin{pmatrix} \vec{P}(s) \notin S \text{ for} \\ \text{some } 0 \leqq s \leqq t \\ \text{given } \vec{P}(0) = P \end{pmatrix} &\leqq 8\int_{r/5-\gamma t}^\infty \frac{\exp\{-x^2/2.42t\}}{\sqrt{2.42\pi t}}\,dx \\ &\leqq 8\int_{(1/1.1\sqrt{t})(\frac{1}{5}r-\gamma t)}^\infty \frac{\exp\{-\frac{1}{2}y^2\}}{\sqrt{2\pi}}\,dy \\ &\leqq 4\exp\left\{-(r-5\gamma t)^2/60.5t\right\}. \end{aligned}$$

Now either $r \geqq 25\gamma t$ or $r \leqq 25\gamma t$. If $f \geqq 25\gamma t$, then $r - 5\gamma t \geqq 0.8r$; hence

$$\mathbf{P}(\ ) \leqq 4\exp\{-r^2/100t\}.$$

If $r \leqq 25\gamma t$, then $r^2/100t \leqq \frac{1}{4}\gamma r_0$ and hence

$$\mathbf{P}(\ \ ) \leqq 1 \leqq \exp\{\tfrac{1}{4}\gamma_0 r_0\}\exp\{-r^2/100t\};$$

thus in all cases

$$\mathbf{P}(\ ) \leqq \max(4, \exp\{\tfrac{1}{4}\gamma r_0\})\exp\{-r^2/100t\}.$$

In case (iib), $\vec{P}(t)$ can leave S in only 3 out of its 4 sides, and bounces off the 4-th. We must use d_λ^2 instead of d_λ in order to satisfy $\partial/\partial n|_{\partial D} = 0$. The modified calculation shows that

$$\mathbf{P}\begin{pmatrix} d_\lambda^2(\vec{P}(s)) = (d_\lambda(P) + \frac{1}{5}r)^2 \\ \text{for some } 0 \leqq s \leqq t \\ \text{given } \vec{P}(0) = P \end{pmatrix}$$

$$\leqq 2\int_{(d_\lambda(P)+\frac{1}{5}r)^2 - d_\lambda(P)^2 - \gamma t}^{\infty} \frac{\exp\{-x^2/4.84(d_\lambda(P) + \frac{1}{5}r)^2 2t\}}{\sqrt{2\pi t}\cdot 2.2(d_\lambda(P) + \frac{1}{5}r)}\,dx,$$

which follows from the lemma, using $|\Delta(d_\lambda^2)| \leqq \gamma$ and

$$\max_S \|\nabla(d_\lambda^2)\|^2 \leqq 4(d_\lambda(P) + \tfrac{1}{5}r)^2\cdot 1.21.$$

The right side is estimated by

$$2\exp\left\{-\left(\frac{(d_\lambda(P) + \frac{1}{5}r)^2 - d_\lambda(P)^2 - \gamma t}{d_\lambda(P) + \frac{1}{5}r}\right)^2 \Big/ 9.68t\right\}$$

$$\leqq 2\exp\left\{-(\tfrac{1}{5}r - 5\gamma t/r)^2/9.68t\right\}.$$

As in case (iia), separating the cases $r^2 \geqq 125\gamma t$ and $r^2 \leqq 125\gamma t$, we deduce that

$$\mathbf{P}(\ ) \leqq \text{const.}\exp\{-r^2/400t\}.$$

In case (iii), by the choice of r_0, the two edges Γ_λ and Γ_μ meet at a corner Q_α of D. Again we distinguish two subcases:

Case (iiia). $Q_\alpha \notin B_{2r/5}(P)$.

Case (iiib). $Q_\alpha \in B_{2r/5}(P)$.

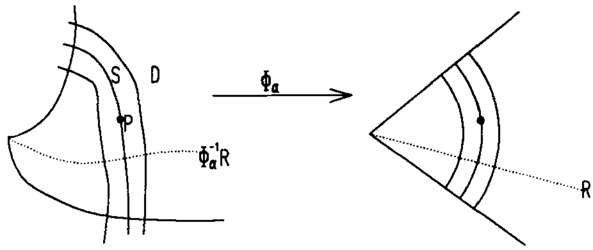

Figure 29. Neighborhoods of a point P near a corner.

If the angle at the corner Q_α of D is greater than π, we must be in case (iiib). We use the diffeomorphism Φ_α defined near Q_α in both cases. Let $\rho_\alpha = \|\Phi_\alpha\|$. In case (iiia), let S be the domain defined by

$$\rho_\alpha(P) - \tfrac{1}{5}r \leqq \rho_\alpha \leqq \rho_\alpha(P) + \tfrac{1}{5}r;$$

see Figure 29. We claim that

$$(*) \qquad S \subset B_r(P),$$

so that if $\vec{P}(t)$ leaves $B_r(P)$, it must also leave S and therefore $\rho_\alpha(\vec{P}(t))$ must equal $\rho_\alpha(P) \pm \frac{1}{5}r$ at some point. To check this, use the fact that Φ_α changes distances by at most a factor of 1.1 and some elementary geometry for the sector $\Phi_\alpha(S)$, depicted in Figure 30. In the figure, it is easy to check that the length of the dotted line is at most $\frac{4}{5}r$, so that

$$\Phi_\alpha(S) \subset B_{4r/5}(\Phi_\alpha(P)).$$

Going back, $S \subset B_r(P)$ follows.

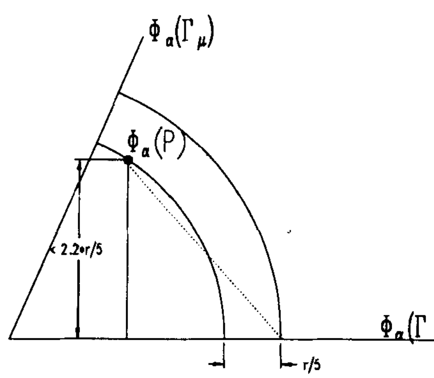

Figure 30. Estimate of certain distances.

We now apply the lemma to the function ρ_α^2. (Since $\Delta\rho_\alpha \to \infty$ at Q_α, we use ρ_α^2 instead of ρ_α.) Note that

$$\left|\Delta\left(\rho_\alpha^2\right)\right| \leqq \gamma,$$

$$\left\|\nabla\left(\rho_\alpha^2\right)\right\| = 2\rho_\alpha\|\nabla\rho_\alpha\| \leqq 2.2\left(\rho_\alpha(P) + \tfrac{1}{5}r\right).$$

Thus

$$\mathbf{P}\begin{pmatrix} \rho_\alpha\left(\vec{P}(s)\right) = \rho_\alpha(P) \pm \frac{1}{5}r \\ \text{for some } 0 \leqq s \leqq t, \\ \text{given } \vec{P}(0) = P \end{pmatrix} \leqq 2\int_*^\infty \frac{\exp\left\{-x^2/9.68\left(\rho_\alpha(P) + \frac{1}{5}r\right)^2 t\right\}}{\sqrt{9.68\pi t}\left(\rho_\alpha(P) + \frac{1}{5}r\right)}\, dx,$$

where

$$* = \begin{cases} \left(\rho_\alpha(P) + \frac{1}{5}r\right)^2 - \rho_\alpha(P)^2 - \gamma t & \text{for the outer boundary,} \\ \rho_\alpha(P)^2 - \left(\rho_\alpha(P) - \frac{1}{5}r\right)^2 - \gamma t & \text{for the inner boundary.} \end{cases}$$

We get as above

$$\mathbf{P}(\quad) \leqq 2\exp\left\{-\left[*/\left(\rho_\alpha(P) + \tfrac{1}{5}r\right)\right]^2\Big/9.68t\right\}.$$

But one checks that

$$\frac{*}{\rho_\alpha(P)} \geqq \frac{1}{5}r - \frac{5\gamma t}{r} \quad \text{for both boundaries,}$$

so, as in case (iib)

$$\mathbf{P}(\quad) \leqq \text{const.}\, e^{-r^2/400t}.$$

Finally, in case (iiib), let S be the domain

$$\rho_\alpha \leqq \rho_\alpha(P) + \tfrac{1}{10}r.$$

Note that $\rho_\alpha(Q) \leqq \rho_\alpha(P) + \frac{1}{10}r$ implies

$$\begin{aligned} \|P - Q\| &\leqq \|P - Q_\alpha\| + \|Q_\alpha - Q\| \\ &\leqq \tfrac{2}{5}r + 1.1\left(\rho_\alpha(P) + \tfrac{1}{10}r\right) \\ &\leqq \left(\tfrac{2}{5} + 1.1\cdot\left(1.1\cdot\tfrac{2}{5} + \tfrac{1}{10}\right)\right)r < r; \end{aligned}$$

hence $S \subset B_r(P)$. There is only one way to leave S now. Use ρ_α^2 as in case (iiia), and get the estimate

$$\mathbf{P}(\text{leaving } S) \leqq \text{const.} \exp\{-r^2/1600t\}.$$

The main result of this appendix can now be proven using Theorem 2 and the formula for the solution f of the Neumann Boundary Value Problem:

Proof: Note that

$$g(P) = E_{\vec{P}(0)=P}\left[\int_0^\infty \mu^2 g(\vec{P}(0)) e^{-\mu^2 t}\, dt\right]$$

hence, by the formula,

$$|f(P) - g(P)| = \left|E_{\vec{P}(0)=P}\left[\int_0^\infty \mu^2\left(g(\vec{P}(t)) - g(\vec{P}(0))\right) e^{-\mu^2 t}\, dt\right]\right|$$

$$\leqq \sup_{Q \in D} \|\nabla g(0)\| \, E_{\vec{P}(0)=P}\left[\int_0^\infty \mu^2 \|\vec{P}(t) - \vec{P}(0)\| e^{-\mu^2 t}\, dt\right].$$

For fixed P and t, let

$$h(r) = \mathbf{P}\left[\|\vec{P}(t) - \vec{P}(0)\| \geqq r \text{ if } \vec{P}(0) = P\right].$$

Then

$$E_{\vec{P}(0)=P}\left[\|\vec{P}(t) - \vec{P}(0)\|\right] = -\int_0^\infty r\, dh(r)$$

$$= \int_0^\infty h(r)\, dr$$

$$\leqq \int_0^{r_0} C_1 \exp\{-r^2/1600t\}\, dr$$

$$+ \int_{r_0}^{\operatorname{diam}(D)} C_1 \exp\{-r_0^2/1600t\}\, dr,$$

by Theorem 2,

$$\leqq 10 C_1 \sqrt{\pi t} + C_1 \operatorname{diam}(D) \exp\{-r_0^2/1600t\}.$$

Therefore,

$$|f(P)-g(P)| \leqq \sup\|\nabla g\| C_1 \int_0^\infty \left(10\sqrt{\pi t} + (\operatorname{diam} D)\exp\{-r_0^2/1600t\}\right)$$
$$\cdot \mu^2 \exp\{-\mu^2 t\}\, dt.$$

Now

$$\int_0^\infty \sqrt{t}\,\mu^2 \exp\{-\mu^2 t\}\, dt = \frac{1}{\mu}\left(\int_0^\infty \sqrt{s}\, e^{-s}\, ds\right) = \frac{1}{\mu}\left(\tfrac{1}{2}\sqrt{\pi}\right)$$

and

$$\int_0^\infty e^{-a/t} + \mu^2 \exp\{-\mu^2 t\}\, dt \leqq \int_0^{\sqrt{a}/\mu} \mu^2 e^{-a/t}\, dt + \int_{\sqrt{a}/\mu}^\infty \mu^2 \exp\{-\mu^2 t\}\, dt$$
$$= \int_{\mu/\sqrt{a}}^\infty a e^{-as}\, ds + e^{-\sqrt{a}\mu}$$
$$= 2\, e^{-\sqrt{a}\mu}$$
$$\leqq \frac{2}{\sqrt{a}\, e\mu}.$$

Thus

$$|f(P)-g(P)| \leqq \sup\|\nabla g\| C_1 \left(5\pi + \frac{80 \operatorname{diam} D}{r_0 e}\right)\frac{1}{\mu}.$$

Appendix 3

Boundary Estimates with μ Large

Let Ω be a bounded domain in $\mathbb{R}^2$ with Lipschitz boundary, $\partial\Omega$. In particular, Ω has no cracks. This is not a serious restriction since we can split cracked domains along the cracks and apply the results derived here to each piece separately.

For a curve $\gamma \subset \partial\Omega$, let $\partial\gamma$ be the set of end points of γ and $|\gamma|$ be arc length of γ. For any point P on γ, let

$\Lambda_P =$ the diameter of the largest circle through P, contained in a single component of $\mathbb{R}^2 - \partial\Omega$.

Let

$$\Lambda_\gamma = \min_{P\in\gamma} \Lambda_P.$$

If γ is connected and contained in another connected curve γ^*, let

$$d(\gamma, \gamma^*) = \begin{cases} \infty \text{ if } \gamma^* \text{ is a closed curve,} \\ \text{arc distance between } \gamma \text{ and } \partial_\gamma{}^* \text{ otherwise.} \end{cases}$$

If γ is $C^{1,1}$, define curvilinear coordinates in a tubular neighborhood of γ in Ω as shown in Figure 3, Section 1: r is the normal distance from γ and s is the arc length along γ. Let n and t denote the unit (inward) normal and tangent vectors along γ. Let θ be the angle that t makes with the positive x-axis. If γ is C^2, the curvature κ along γ is defined by the equation

$$\frac{dn}{ds} = -\kappa t.$$

Let

$$k_\gamma = \begin{cases} \dfrac{1}{|\partial\Omega|} + \mathrm{Lip}_\gamma(\theta) & \text{if } \gamma \text{ is } C^{1,1}, \text{ but not necessarily } C^{2,1}, \\ 1/|\partial\Omega| + \|\kappa\|_{0,\infty,\gamma} + \sqrt{\mathrm{Lip}_\gamma(\kappa)} & \text{if } \gamma \text{ is } C^{2,1}, \end{cases}$$

where $\mathrm{Lip}_\gamma(\theta)$ and $\mathrm{Lip}_\gamma(\kappa)$ are the Lipschitz constants. For $p > 1$, let $V_{\gamma,p,\mu} \subset H_p^2(\Omega)$ denote the subspace of all functions v such that

(a) $\nabla^2 v - \mu^2 v = 0$ in Ω (in the weak sense),

(b) v is bounded on $\partial\Omega$,

(c) $\partial v/\partial r$ is $C^{0,1}$ along γ.

Let

$$|||v|||_{1,\gamma} = \mathrm{Lip}_\gamma\left(\frac{\partial v}{\partial r}\right) + k_\gamma\left\|\frac{\partial v}{\partial r}\right\|_{0,\infty,\gamma} + \frac{k_\gamma^3}{\mu}\|v\|_{0,\infty,\partial\Omega} + \frac{\|\partial v/\partial r\|_{0,1,\partial\Omega}}{\mu^4|\partial\Omega|^6}.$$

If γ is $C^{2,1}$ and $\partial v/\partial r$ is $C^{1,1}$ along γ, let

$$|||v|||_{2,\gamma} = \mathrm{Lip}_\gamma\left(\frac{\partial}{\partial s}\frac{\partial v}{\partial r}\right) + k_\gamma\left\|\frac{\partial}{\partial s}\frac{\partial v}{\partial r}\right\|_{0,\infty,\gamma} + k_\gamma^2\left\|\frac{\partial v}{\partial r}\right\|_{0,\infty,\gamma}$$

$$+ \frac{k_\gamma^4}{\mu}\|v\|_{0,\infty,\partial\Omega} + \frac{\|\partial v/\partial r\|_{0,1,\partial\Omega}}{\mu^3|\partial\Omega|^6}.$$

THEOREM. *Let $\gamma \subset \gamma^*$ be connected $C^{1,1}$-curves $\subset \partial\Omega$ such that $\partial\gamma \cap \partial\gamma^*$ is empty. Then, there exists a universal constant β such that with*

$$\delta = \min\left\{\beta/k_{\gamma^*}, \tfrac{1}{16}d(\gamma,\gamma^*), \tfrac{1}{4}\Lambda_{\gamma^*}\right\},$$

for all μ *satisfying* $\mu\delta \geqq 6\log(\mu|\partial\Omega|) \geqq 1$ *and for all* $v \in V_{\gamma^*, p, \mu}$, *we have along* γ:

A.

$$v(o, s) = -\frac{1}{\mu}\frac{\partial v}{\partial r}(o, s) + \frac{\varepsilon_0(s)}{\mu^2}|||v|||_{1,\gamma^*},$$

where $|\varepsilon_0(s)|$ *is bounded by a universal constant.*

B. *If* γ^* *is* $C^{2,1}$ *and* $\partial v/\partial r$ *is* $C^{1,1}$ *along* γ^*, *then*

$$v(o, s) = -\frac{1}{\mu}\left(1 + \frac{\kappa(s)}{2\mu}\right)\frac{\partial v}{\partial r}(o, s) + \frac{\varepsilon_1(s)}{\mu^3}|||v|||_{2,\gamma^*},$$

$$\frac{\partial v}{\partial s}(o, s) = -\frac{1}{\mu}\left(\frac{\partial}{\partial s}\frac{\partial v}{\partial r}(o, s)\right) + \frac{\varepsilon_2(s)}{\mu^2}|||v|||_{2,\gamma^*},$$

where $|\varepsilon_1(s)|$ *and* $|\varepsilon_2(s)|$ *are bounded by universal constants.*

C. *Suppose that* γ^* *is* $C^{3,1}$ *and* $\partial v/\partial r$ *is* $C^{1,1}$ *along* γ^*. *Then there exists a constant* A *depending only on* δ, $|\partial\Omega|$ *and* $\|\kappa\|_{2,\infty,\gamma^*}$ *such that, for all* $r < \delta$,

$$v(r, s) = -\frac{e^{-\mu r}}{\mu\sqrt{1 - \kappa(s)r}}\left(1 + \frac{\kappa(s)}{2\mu}\right)\frac{\partial v}{\partial r}(o, s) + \frac{\varepsilon_1'(r, s)}{\mu^3},$$

$$\frac{\partial v}{\partial r}(r, s) = \frac{e^{-\mu r}}{\mu\sqrt{1 - \kappa(s)r}}\frac{\partial v}{\partial r}(o, s) + \frac{\varepsilon_2'(r, s)}{\mu^2},$$

$$\frac{\partial v}{\partial s}(r, s) = -\frac{e^{-\mu r}}{\mu\sqrt{1 - \kappa(s)r}}\left(\frac{\partial}{\partial s}\frac{\partial v}{\partial r}(o, s)\right) + \frac{\varepsilon_3'(r, s)}{\mu^2},$$

where, for $1 \leqq i \leqq 3$,

$$|\varepsilon_i'(r, s)| \leqq A(1 + \mu r)e^{-\mu r}\left(\left\|\frac{\partial v}{\partial r}\right\|_{2,\infty,\gamma^*} + \frac{\|v\|_{0,\infty,\partial\Omega} + \|\partial v/\partial r\|_{0,1,\partial\Omega}}{\mu}\right).$$

In order to make the proof more readable we give a proof only for Part B of the theorem. The proof of Part A is essentially contained in the proof of Part B except that we have to replace all the $C^{2,1}$-estimates of γ^* by $C^{1,1}$-estimates and develop the error terms up to one order of $1/\mu$ less. Part C is not needed in our application of the theorem in this paper. Its proof is analogous to the proof of Part B involving a few more calculations of the same type.

The theorem is proved by estimating the integrals in Green's integral representation formula which we now recall. By Sobolev imbedding theorem,

$$v \in H_q^1(\Omega) \quad \text{where} \quad q > 2,$$

and by the trace theorems (see [7]),

$$\frac{\partial v}{\partial n} \in L_q(\partial\Omega).$$

The assumption that $(-\nabla^2 + \mu^2)v = 0$ in the weak sense means that

$$\int_\Omega (\nabla\phi \cdot \nabla v + \mu^2 \phi v) + \int_{\partial\Omega} \phi \frac{\partial v}{\partial n} = 0 \quad \text{for all} \quad \phi \in H_{q'}^1(\Omega),$$

$$\text{where} \quad \frac{1}{q} + \frac{1}{q'} = 1.$$

Fix a point $P \in \mathbb{R}^2$ and let $\phi = K_0(\mu|PQ|)$, where K_0 is the zeroth-order modified Bessel function of the second kind and $Q \in \mathbb{R}^2$. We get

$$\int_{\Omega_Q} \left(\nabla_Q v \cdot \nabla_Q K_0 + \mu^2 v_Q K_0\right) + \int_{\partial\Omega_Q} K_0 \frac{\partial v}{\partial n_Q} = 0. \tag{1}$$

K_0 satisfies the equation

$$\left(-\nabla_Q^2 + \mu^2\right) K_0(\mu|PQ|) = 2\pi\omega_P$$

in the weak sense, where ω_P is the delta function concentrated at P. Therefore, for all $v \in V_{\gamma, p, \mu}$,

$$\int_{\Omega_Q} \left(-\nabla_Q^2 + \mu^2\right) K_0(\mu|PQ|) v(Q) = 2\pi \int_{\Omega_Q} \omega_P v(Q).$$

If $P \notin \partial\Omega$ or if $P \in \partial\Omega$ where it is at least $C^{1,1}$, then $\partial K_0/\partial n_Q$ is bounded in a neighborhood of P and hence, by applying the Green's formula to $\Omega -$ (ε-neighborhood of P) and then taking the limit as $\varepsilon \to 0$, we get

$$\int_{\Omega_Q} \left(\nabla_Q v \cdot \nabla_Q K_0 + \mu^2 v_Q K_0\right) + \int_{\partial\Omega_Q} \frac{\partial K_0}{\partial n_Q} v_Q = 2\pi \int_{\Omega_Q} \omega_P v_Q \tag{2}$$

for all $v \in V_{\varnothing, p, \mu}$ and all P in whose neighborhood $\partial\Omega$ is $C^{1,1}$. Subtracting (1) from (2), we get

$$\int_{\Omega_Q} \omega_P v_Q = \frac{1}{2\pi} \int_{\partial\Omega_Q} \left(-K_0 \frac{\partial v}{\partial n_Q} + \frac{\partial K_0}{\partial n_Q} v_Q \right),$$

$$\int_{\Omega_Q} v_Q \omega_P = m_P v(P),$$

where

$$m_P = \begin{cases} 1 & \text{if } P \in \bar{\Omega} - \partial\Omega, \\ \frac{1}{2} & \text{if } P \in \partial\Omega \text{ and } \partial\Omega \text{ is } C^{1,1} \text{ at } P, \end{cases}$$

since v is continuous at P.

We have

$$(3) \quad m_P v(P) = \frac{1}{2\pi} \int_{\partial\Omega_Q} \left(-K_0 \frac{\partial v}{\partial n_Q} + \frac{\partial K_0}{\partial n_Q} v(Q) \right) \quad \text{where} \quad m_P = \int_{\Omega_Q} \omega_P.$$

Our task is to estimate the integrals in the integral equation above and solve explicitly for v in terms of $\partial v / \partial n$. The key point is that K_0 and its derivatives decrease exponentially with increasing $\mu|PQ|$ and hence the values of the integrals when μ is large depend essentially on a small segment of $\partial\Omega$ near P which may be approximated by a circle.

We prove a series of lemmas first. In Lemma 1, we derive geometric estimates which depend only on Ω. We use these estimates in Lemma 2 to specify δ. In Lemma 3, we derive some basic formulae for K_0, Lemma 4 lists estimates for the basic integrals that arise when we try to estimate the integrals in the Green's formula and their derivatives.

Define $R : \mathbb{R}^2 \times \partial\Omega \to \mathbb{R}_{\geqq 0}$ by setting $R(P, Q) =$ distance between $P \in \Omega$ and $Q \in \partial\Omega$. If P and Q lie in a tubular neighborhood of a connected curve $\gamma \subset \partial\Omega$ so that P and Q have coordinates (r, s_P) and $(0, s_Q)$, respectively, then R becomes a function of r, s_P and s_Q. Define $\sigma : \gamma \times \gamma \to \mathbb{R}_{\geqq 0}$ by setting $\sigma(P, Q) = |s_Q - s_P|$. Let ∂_σ denote the symmetric operator $\partial/\partial s_P + \partial/\partial s_Q$. Let n_P, t_P and κ_P denote the normal vector, tangent vector and the curvature at the point $(0, s_P)$. Let n_Q, t_Q and κ_Q be the corresponding quantities at Q. In what follows, if an estimate contains an order term $O(\varepsilon^m)$ in variable ε, then $|O(\varepsilon^m)/\varepsilon^m|$ is bounded by a universal constant.

LEMMA 1. I. *Let x_1, x_2 be the coordinates in $\mathbb{R}^2$ and let D_i denote the operator $\partial/\partial x_i$. Let $q \in \partial\Omega$ be such that n_Q is defined. Then for all $P \in \mathbb{R}^2 - \{Q\}$,*

$i = 1, 2$ *and* $j = 1, 2$,

$$|D_{i_P} R| \leqq 1,$$

$$\left|\frac{\partial R}{\partial n_Q}\right| \leqq 1,$$

$$|D_{i_P}\, \partial R/\partial n_Q| \leqq \frac{1}{R}.$$

II. *Suppose that P belongs to a tubular neighborhood of a connected $C^{2,1}$-curve $\gamma \subset \partial\Omega$ and $Q \in \gamma$. Then, for all $r, \sigma \leqq 1/k_\gamma$, we have the following estimates*:

(i) $$R^2 = r^2 + \sigma^2(1 - \kappa_P r) + k_\gamma^2[r O(\sigma^3) + O(\sigma^4)],$$

(ii) $$R\frac{\partial R}{\partial r} = r - \tfrac{1}{2}\kappa_P \sigma^2 + k_\gamma^2 O(\sigma^3),$$

(iii) $$R\frac{\partial R}{\partial s_P} = -(1 - \kappa_P r)(s_Q - s_P) + k_\gamma^2 O(\sigma^3),$$

(iv) $$R\frac{\partial R}{\partial s_Q} = (1 - \kappa_P r)(s_Q - s_P) + k_\gamma^2[r O(\sigma^2) + O(\sigma^3)],$$

(v) $$\partial_\sigma R^2 = k_\gamma^2[r O(\sigma^2) + O(\sigma^3)],$$

(vi) $$R\frac{\partial R}{\partial n_Q} = -r - \tfrac{1}{2}\kappa_P \sigma^2(1 - \kappa_P r) + k_\gamma^2 O(\sigma^3),$$

(vii) $$\frac{\partial}{\partial r}\left(R\frac{\partial R}{\partial n_Q}\right) = -1 + \tfrac{1}{2}\kappa_P^2 \sigma^2 + k_\gamma^2 O(\sigma^3),$$

(viii) $$\frac{\partial}{\partial s_P}\left(R\frac{\partial R}{\partial n_Q}\right) = \kappa_P(1 - \kappa_P r)(s_Q - s_P) + k_\gamma^2 O(\sigma^2),$$

(ix) $$\partial_\sigma\left(R\frac{\partial R}{\partial n_Q}\right) = k_\gamma^2 O(\sigma^2).$$

(If γ is only $C^{1,1}$, the corresponding estimates may be obtained by dropping the terms containing κ_P and replacing $k_\gamma^2 O(\sigma^m)$ by $k_\gamma O(\sigma^{m-1})$. We leave it to the reader to verify this.)

Proof: Part I may be proved directly by differentiating R. (It is convenient to let the x_1-axis coincide with n_Q.)

Consider now Part II. Let P_0 be the point with coordinates $(0, s_P)$ and let m be the point on γ with coordinates $(0, \frac{1}{2}(s_P + s_Q))$. From the error formulae for divided differences, we obtain the following estimates in which $k_\gamma^2 O(\sigma^m)$ is abbreviated as $O_1(\sigma^m)$:

$$\overrightarrow{P_0Q} = (s_Q - s_P)\left[t_m + \overrightarrow{O_1(\sigma^2)}\right],$$

where t_m is the tangent vector at m and $\overrightarrow{O_1(\sigma^2)}$ is a vector of magnitude $O_1(\sigma^2)$,

$$\begin{aligned}
n_Q - n_P &= (s_Q - s_P)\left[-\kappa_m t_m + \overrightarrow{O_1(\sigma)}\right] \\
&= (s_Q - s_P)\left[-\kappa_P t_P + \overrightarrow{O_1(\sigma)}\right] = (s_Q - s_P)\left[-\kappa_Q t_Q + \overrightarrow{O_1(\sigma)}\right], \\
t_Q - t_P &= (s_Q - s_P)\left[\kappa_m n_m + \overrightarrow{O_1(\sigma)}\right] \\
&= (s_Q - s_P)\left[\kappa_P n_P + \overrightarrow{O_1(\sigma)}\right] = (s_Q - s_P)\left[\kappa_Q n_Q + \overrightarrow{O_1(\sigma)}\right], \\
n_Q + n_P &= 2n_m + \overrightarrow{O_1(\sigma^2)}, \\
t_Q + t_P &= 2t_m + \overrightarrow{O_1(\sigma^2)}.
\end{aligned}$$

From these estimates, we obtain, in turn, the following:

$$\begin{aligned}
\overrightarrow{P_0Q} \cdot \overrightarrow{P_0Q} &= \sigma^2 + O_1(\sigma^4), \\
\overrightarrow{P_0Q} \cdot n_P &= \tfrac{1}{2}\kappa_m \sigma^2 + O_1(\sigma^3), \\
\overrightarrow{P_0Q} \cdot n_Q &= -\tfrac{1}{2}\kappa_m \sigma^2 + O_1(\sigma^3), \\
\overrightarrow{P_0Q} \cdot t_P &= (s_Q - s_P) + O_1(\sigma^3), \\
\overrightarrow{P_0Q} \cdot t_Q &= (s_Q - s_P) + O_1(\sigma^3), \\
\overrightarrow{P_0Q} \cdot (t_Q - t_P) &= O_1(\sigma^3), \\
n_P \cdot t_Q = -n_Q \cdot t_P &= \kappa_P(s_Q - s_P) + O_1(\sigma^2), \\
1 - n_P \cdot n_Q &= \tfrac{1}{2}\kappa_P^2 \sigma^2 + k_\gamma O_1(\sigma^3).
\end{aligned}$$

The estimates in Part II now follow by substituting the above estimates into the following formulae for R^2 and its derivatives:

$$R^2 = \left(-rn_P + \overrightarrow{P_0Q}\right) \cdot \left(-rn_P + \overrightarrow{P_0Q}\right),$$

$$R\frac{\partial R}{\partial r} = r - \overrightarrow{P_0Q} \cdot n_P,$$

$$R\frac{\partial R}{\partial s_P} = -(1 - \kappa_P r)\overrightarrow{P_0Q} \cdot t_P,$$

$$R\frac{\partial R}{\partial s_Q} = -rn_P \cdot t_Q + \overrightarrow{P_0Q} \cdot t_Q,$$

$$R\frac{\partial R}{\partial n_Q} = -rn_P \cdot n_Q + \overrightarrow{P_0Q} \cdot n_Q,$$

$$\frac{\partial}{\partial r}\left(R\frac{\partial R}{\partial n_Q}\right) = -n_P \cdot n_Q,$$

$$\frac{\partial}{\partial s_P}\left(R\frac{\partial R}{\partial n_Q}\right) = -(1 - \kappa_P r)t_P \cdot n_Q,$$

$$\frac{\partial}{\partial s_Q}\left(R\frac{\partial R}{\partial n_Q}\right) = -\kappa_Q\left(-rn_P + \overrightarrow{P_0Q}\right) \cdot t_Q.$$

LEMMA 2. *There exists a universal constant $\beta \leqq \frac{1}{2}$ such that if γ is a connected $C^{1,1}$-curve $\subset \partial\Omega$ and if we choose $\delta \leqq \beta/k_\gamma$, then*

(a) *for all points P in a tubular neighborhood of γ for which $r \leqq \delta$ and $\{Q \in \partial\Omega | \sigma \leqq 4\delta\} \subset \gamma$, we have*

$$r^2 + 4\sigma^2 \geqq R^2 \geqq r^2 + \tfrac{1}{4}\sigma^2,$$

$$2\sigma \geqq \left|R\frac{\partial R}{\partial s_Q}\right| \geqq \tfrac{1}{2}\sigma.$$

(b) *If, furthermore, $\delta \leqq \frac{1}{4}\Lambda_\gamma$, then $R \geqq 2\delta$ whenever $\sigma \geqq 4\delta$ or $Q \notin \gamma$.*

Proof: The estimates for R^2 and $R\,\partial R/\partial s_Q$ in Lemma 1 imply the existence of β such that the inequalities in part (a) are satisfied. Suppose now that $\delta \leqq \frac{1}{4}\Lambda_\gamma$. Let S be the circle with radius 2δ and center at P. Let Q_1, Q_2, q_1, q_2 be the points as shown in Figure 31. Since $R \geqq 2\delta$ when $\sigma = 4\delta$, $|s_{Q_i} - s_P| \leqq 4\delta$ for $i = 1, 2$. Therefore, Q_1 and Q_2 lie on γ. For $i = 1, 2$, the arc $\widehat{Q_i q_i}$ must lie inside the circle of diameter Λ_γ, tangent to γ at Q_i, since $4\delta < \Lambda_\gamma$. The rest of

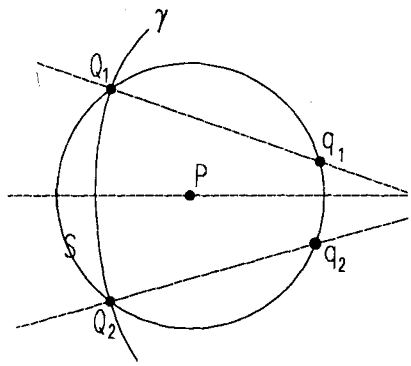

Figure 31. Definition of Q_i and q_i, $i = 1$ and 2.

the circle S must cut any normal to a point on γ between Q_1 and Q_2 at a distance at most $4\delta < \Lambda_\gamma$. It follows that $\partial\Omega \cap S = \{Q_1, Q_2\}$. Any ray from P to a point on $\partial\Omega$ with $\sigma \geqq 4\delta$ must intersect S first and hence, has length at least 2δ.

LEMMA 3. *Let $K_0(z)$ be the zeroth-order modified Bessel function of the second kind. Let L denote the operator $-(1/z)\, d/dz$. Then*:

(i) $Lz^2LK_0(z) = K_0(z)$.

(ii) There exists a constant c such that, for $0 \leqq z < \infty$,

$$0 \leqq z^{1/2}K_0(z) \leqq ce^{-z},$$

$$0 \leqq z^2LK_0(z) \leqq c\max(1, \sqrt{z})e^{-z}.$$

(iii) For any integers $m, n \geqq 0$, there exist constants c_n and $d_{m,n}$ such that, for all $z \geqq 1$,

$$(-1)^n\frac{d^n}{dz^n}K_0(z) \leqq c_n z^{-1/2}e^{-z}, \qquad L^nK_0(z) \leqq c_n z^{-n-1/2}e^{-z},$$

$$\int_z^\infty z^mL^nK_0(z)\,dz \leqq d_{m,n}z^{m-n-1/2}e^{-z}.$$

Proof: Part (i) follows from the identities

$$\frac{d}{dz}K_0(z) = -K_1(z),$$

$$\frac{d}{dz}(zK_1(z)) = -zK_0(z),$$

where $K_1(z)$ is the first-order modified Bessel function of the second kind. (For a collection of formulae involving Bessel functions, see the Handbook of Mathematical Functions, [2], vol. 55.) For proving part (ii), use the polynomial approximations for $z^{1/2}K_0(z)$ and $zK_1(z)$ in the range $0 \leqq z \leqq 2$ and $2 \leqq z < \infty$. The last part follows from (i) and (ii) by differentiation and integration by parts.

LEMMA 4. *Let γ be a connected $C^{2,1}$-curve $\subset \partial\Omega$. Let j be an integer, $0 \leqq j \leqq 2$. Choose $\delta \leqq \beta/k_\gamma$ where β is the constant as in Lemma 2. Let P with coordinates (r, s_P) be a point in a tubular neighborhood of γ with $r < \delta$. Let $N_P = \{Q \in \partial\Omega | \sigma < 4\delta\}$ and assume that $N_P \subset \gamma$. Let $\Phi \in L_1(\partial\Omega)$ be a function on $\partial\Omega$ which is $C^{j-1,1}$ in γ. Let*

$$\Phi_P = \Phi(s_P),$$

$$C^{\Phi}_{0,\gamma} = \|\Phi\|_{0,\infty,\gamma},$$

$$C^{\Phi}_{1,\gamma} = k_\gamma\|\Phi\|_{0,\infty,\gamma} + \mathrm{Lip}_\gamma(\Phi),$$

and

$$C^{\Phi}_{2,\gamma} = k_\gamma^2\|\Phi\|_{0,\infty,\gamma} + k_\gamma\left\|\frac{d\Phi}{ds}\right\|_{0,\infty,\gamma} + \mathrm{Lip}_\gamma\left(\frac{d\Phi}{ds}\right).$$

Let $\mathscr{L}$ denote the operator $-(1/R)\,d/dR$. Then, there exist universal constants $C_1, C_2, \cdots, C_1', C_2', \cdots, C'', C_1'', C_2'', \cdots, \alpha_0$, and polynomials $\alpha_1(\mu r)$ and $\alpha_3(\mu r)$ of degrees 1 and 3 in variable $\sqrt{\mu r}$, respectively, with universal coefficients such that the following estimates hold for all μ satisfying $\mu\delta \geqq 6\log(\mu|\partial\Omega|) \geqq 1$:

$$\frac{1}{2\pi}\int_{N_P} K_0(\mu R)\Phi\, ds_Q \approx \frac{\Phi_P e^{-\mu r}}{2\mu\sqrt{1-\kappa_P r}}$$

$$\text{(4.1)} \qquad \textit{with absolute error} \leqq \begin{cases} \dfrac{C^{\Phi}_{1,\gamma}}{\mu^2}\alpha_1(\mu r)e^{-\mu r} & \textit{if } j = 1, \\ \dfrac{C^{\Phi}_{2,\gamma}}{\mu^3}\alpha_3(\mu r)e^{-\mu r} & \textit{if } j = 2, \end{cases}$$

$$\frac{1}{2\pi}\int_{N_P}\mathscr{L}K_0(\mu R)\Phi\,ds_Q \approx \frac{\Phi_P e^{-\mu r}}{2r\sqrt{1-\kappa_P r}}$$

$$\text{(4.2)}\qquad \textit{with absolute error} \leqq \begin{cases} \dfrac{C^{\Phi}_{1,\gamma}}{\sqrt{\mu r}}\alpha_1(\mu r)e^{-\mu r} & \textit{if } j=1,\\[2mm] \dfrac{C^{\Phi}_{2,\gamma}}{\mu}\alpha_3(\mu r)e^{-\mu r} & \textit{if } j=2,\end{cases}$$

$$\frac{1}{2\pi}\int_{N_P}\mathscr{L}K_0(\mu R)\sigma^2\Phi\,ds_Q \approx \frac{\Phi_P e^{-\mu r}}{2\mu(1-\kappa_P r)^{3/2}}$$

$$\text{(4.3)}\qquad \textit{with absolute error} \leqq \begin{cases} \dfrac{C^{\Phi}_{1,\gamma}}{\mu^2}\alpha_1(\mu r)e^{-\mu r} & \textit{if } j=1,\\[2mm] \dfrac{C^{\Phi}_{2,\gamma}}{\mu^3}\alpha_3(\mu r)e^{-\mu r} & \textit{if } j=2,\end{cases}$$

$$\frac{1}{2\pi}\int_{N_P}\mathscr{L}^2K_0(\mu R)\Phi\,ds_Q \approx \frac{\Phi_P(1+\mu r)e^{-\mu r}}{2r^3\sqrt{1-\kappa_P r}}$$

$$\text{(4.4)}\qquad \textit{with absolute error} \leqq \begin{cases} \dfrac{C^{\Phi}_{1,\gamma}}{r^2}\alpha_1(\mu r)e^{-\mu r} & \textit{if } j=1,\\[2mm] \dfrac{C^{\Phi}_{2,\gamma}}{r}\alpha_3(\mu r)e^{-\mu r} & \textit{if } j=2,\end{cases}$$

$$\text{(4.5)}\qquad \begin{gathered}\frac{1}{2\pi}\int_{N_P}\mathscr{L}^2K_0(\mu R)\sigma^2\Phi\,ds_Q \approx \frac{\Phi_P}{2r}e^{-\mu r}\\ \textit{with absolute error} \leqq \frac{C^{\Phi}_{1,\gamma}}{\sqrt{\mu r}}\alpha_1(\mu r)e^{-\mu r} \quad \textit{if } j=1,\end{gathered}$$

$$\text{(4.6)}\qquad \left|\int_{N_P}K_0(\mu R)\sigma^2\Phi\,ds_Q\right| \leqq \frac{C^{\Phi}_{0,\gamma}}{\mu^3}(1+\mu r)\alpha_0 e^{-\mu r}, \qquad j=0,$$

$$\text{(4.7)}\qquad \left|\int_{N_P}\mathscr{L}K_0(\mu R)\sigma^4\Phi\,ds_Q\right| \leqq \frac{C^{\Phi}_{0,\gamma}}{\mu^3}(1+\mu r)\alpha_0 e^{-\mu r}, \qquad j=0,$$

$$\text{(4.8)}\qquad \left|\int_{N_P}\mathscr{L}^2K_0(\mu R)\sigma^4\Phi\,ds_Q\right| \leqq \frac{C^{\Phi}_{0,\gamma}}{\mu}\alpha_0 e^{-\mu r}, \qquad j=0.$$

If $r=0$, *then for* $m, n \geqq 0$, $j=0$,

$$\text{(4.9)}\qquad \left|\int_{N_P}\mathscr{L}^nK_0(\mu R)\sigma^{2n+m}\Phi\,ds_Q\right| \leqq \frac{C_n}{\mu^{m+1}}C^{\Phi}_{0,\gamma}.$$

For $m, n \geqq 0$ *and* $j = 1$,

$$\lim_{r\to 0}\left|\int_{N_P} \mathscr{L}^n K_0(\mu R)\sigma^{2n+m-2}(s_Q - s_P)\Phi \, ds_Q\right| \leqq \frac{C_n'}{\mu^{m+1}} C_{1,\gamma}^{\Phi}. \tag{4.10}$$

If, additionally, $\delta \leqq \frac{1}{4}\Lambda_\gamma$, *then for* $n \geqq 0$ *and* $j = 0$,

$$\left|\int_{\partial\Omega - N_P} \frac{d^n}{dR^n} K_0(\mu r)\Phi \, ds_Q\right| \leqq \frac{C_n''\mu^n}{(\mu|\partial\Omega|)^6}\|\Phi\|_{0,1,\partial\Omega} e^{-\mu r}. \tag{4.11}$$

Proof: Estimate (4.11) follows from the estimate for $(d^n/dz^n)K_0(z)$ and $L^n K_0(z)$ in Lemma 3. For the rest, first consider the special case where Ω is the upper half-plane in $\mathbb{R}^2$, $N_P = \partial\Omega$ is the y-axis in $\mathbb{R}^2$, r is measured along the x-axis and $\Phi \equiv 1$.

Let

$$u(r) = \frac{1}{2\pi}\int_{-\infty}^{\infty} K_0(\mu R)\, dy.$$

Then, u satisfies the equation $-\nabla^2 u + \mu^2 u = l$, where l is the line potential of unit density along the y-axis; i.e.,

$$\int_{-\infty}^{\infty} l\, dx = 1.$$

The equation is solved easily by reducing to the one-dimensional case:

$$u(r) = \frac{1}{2\mu} e^{-\mu r}.$$

Therefore,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} K_0(\mu R)\, dy_Q = \frac{1}{2\mu} e^{-\mu r}.$$

Differentiation of this formula and integration by parts show that estimates (4.1) through (4.8) hold such that the error terms in the first five estimates are actually zero. To obtain (4.9), just note that $R = \sigma = |y_Q - y_P|$ and

$$\mu^{m+1}\int_{-\infty}^{\infty} \mathscr{L}^n K_0(\mu R)\sigma^{2n+m}\, dy_Q = \int_{-\infty}^{\infty} \frac{1}{\mu\sigma}\left(\frac{d}{d(\mu\sigma)}\right)^n K_0(\mu\sigma)(\mu\sigma)^{2n+m}\, d(\mu y_Q)$$

which does not depend on μ and the integrand, at worst, has only logarithmic pole at $y_Q = y_P$. The integrand in (4.10) has a pole of the order $1/(y_Q - y_P)$ when $m = r = 0$ and hence the integral is defined only as a limit as indicated.

The numerator is an even function of y_Q while the denominator is odd. Therefore, the integral is zero for all $r > 0$.

Consider now the general case. The estimates in the lemma follow by performing a change of variables as follows. We let

$$\eta = (s_Q - s_P)\sqrt{\frac{R^2 - r^2}{\sigma^2}}\,.$$

As before, we abbreviate $k_\gamma^2 O(\sigma^m)$ and $k_\gamma^2 O(\eta^m)$ as $O_1(\sigma^m)$ and $O_1(\eta^m)$, respectively. By Lemma 1,

$$\frac{\eta}{s_Q - s_P}\frac{\partial \eta}{\partial s_Q} = (1 - \kappa_P r) + r\,O_1(\sigma) + O_1(\sigma^2)$$

and

$$\frac{\eta}{s_Q - s_P} = \sqrt{(1 - \kappa_P r) + r\,O_1(\sigma) + O_1(\sigma^2)}\,.$$

By Lemma 2, $\frac{1}{4} \leqq \eta^2/\sigma^2 \leqq 4$ and $\frac{1}{2} \leqq \eta/(s_Q - s_P) \leqq 2$. Hence, η_Q is an invertible function of s_Q:

$$\begin{aligned}
\frac{ds_Q}{d\eta_Q} &= \frac{\sqrt{(1 - \kappa_P r) + r\,O_1(\sigma) + O_1(\sigma^2)}}{(1 - \kappa_P r) + r\,O_1(\sigma) + O_1(\sigma^2)} \\
&= \frac{1}{\sqrt{1 - \kappa_P r}} + r\,O_1(\eta) + O_1(\eta^2) \\
&= 1 + rk_\gamma O(1) + k_\gamma O(\eta),
\end{aligned}$$

$$s_Q - s_P = \frac{\eta}{\sqrt{1 - \kappa_P r}} + r\,O_1(\eta^2) + O_1(\eta^3) = \eta + rk_\gamma O(\eta) + k_\gamma O(\eta^2),$$

and

$$\sigma^2 = \frac{\eta^2}{\sqrt{1 - \kappa_P r}} + r\,O_1(\eta^3) + O_1(\eta^4) = \eta^2 + rk_\gamma O(\eta^2) + k_\gamma O(\eta^3).$$

Note also that, if $|\eta| \geqq 2\delta$,

$$1 \leqq \left|\frac{R}{\eta}\right| = \sqrt{\frac{\eta^2 + r^2}{\eta^2}} \leqq 2$$

and

$$1 \leqq \left|\frac{d\eta_Q}{dR_Q}\right| = \left|\frac{R}{\eta}\right| \leqq 2.$$

We also need three more integral estimates:

(i)
$$\int_0^\infty K_0(\mu R)\eta\, d\eta$$
$$= -\frac{1}{\mu}\int_r^\infty \frac{d}{dR}(RK_1(\mu R))\, dR$$
$$= \frac{rK_1(\mu r)}{\mu} \leqq \max(1, \sqrt{\mu r})e^{-\mu r}O\left(\frac{1}{\mu^2}\right),$$

(ii)
$$\int_0^\infty \mathscr{L}K_0(\mu R)\eta\, d\eta$$
$$= -\int_r^\infty \frac{dK_0(\mu R)}{dR}\, dR$$
$$= K_0(\mu r) \leqq \frac{e^{-\mu r}}{\sqrt{\mu r}}O(1).$$

Integration by parts gives us

(iii) $$\int_0^\infty \mathscr{L}K_0(\mu R)\eta^3\, d\eta = \frac{2rK_1(\mu r)}{\mu} \leqq \max(1, \sqrt{\mu r})e^{-\mu r}O\left(\frac{1}{\mu^2}\right).$$

Finally, note that if $j = 1$,

$$\Phi = \Phi_P + C_{1,\gamma}^{\Phi}O(\eta)$$

and if $j = 2$,

$$\Phi = \Phi_P + \left|\frac{d\Phi}{ds}\right|_P \eta + C_{2,\gamma}^{\Phi}[rO(\eta) + O(\eta^2)].$$

Consider now the first estimate in the lemma with $j = 2$. Perform the change of variables and substitute the above estimates for $ds_Q/d\eta_Q$ and Φ. Let η_- and η_+ be the limits of η_Q corresponding to the limits of s_Q in N_P. We get

$$\frac{1}{2\pi}\int_{N_P} K_0(\mu R)\Phi\, ds_Q$$
$$= \frac{1}{2\pi}\int_{\eta_-}^{\eta_+} K_0(\mu R)\left\{\frac{\Phi_P}{\sqrt{1-\kappa_P r}} + \frac{d\Phi}{ds}\bigg|_P \eta + C_{2,\gamma}^{\Phi}[rO(\eta) + O(\eta^2)]\right\} d\eta_Q.$$

Let I be the interval $[\eta_-, \eta_+]$ and $J = (-\infty, \infty) - (\eta_-, \eta_+)$,

$$\frac{1}{2\pi}\int_I K_0(\mu R)\, d\eta_Q = \frac{1}{2\pi}\int_{-\infty}^{\infty} K_0(\mu R)\, d\eta_Q - \frac{1}{2\pi}\int_J K_0(\mu R)\, d\eta_Q.$$

By the special case above,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} K_0(\mu R)\, d\eta_Q = \frac{e^{-\mu r}}{2\mu}.$$

Since $|\eta_\pm| \geqq 2\delta$ by Lemma 2,

$$\begin{aligned}
\left|\int_J K_0(\mu R)\, d\eta_Q\right| &\leqq \int_{2\delta}^{\infty} K_0(\mu R)\, d\eta_Q \\
&\leqq 2\int_{2\delta}^{\infty} K_0(\mu R)\frac{d\eta_Q}{dR_Q}\, dR_Q \\
&\leqq 4\int_{2\delta}^{\infty} K_0(\mu R)\, dR \\
&\leqq e^{-2\mu\delta}\, O\left(\frac{1}{\mu}\right), \qquad \text{by Lemma 3,} \\
&\leqq \frac{e^{-\mu r}}{\mu^6|\partial\Omega|^6} O\left(\frac{1}{\mu}\right) \\
&\leqq k_\gamma^2 e^{-\mu r} O\left(\frac{1}{\mu^3}\right),
\end{aligned}$$

$$\int_I K_0(\mu R)\eta\, d\eta_Q = \int_{-\infty}^{\infty} K_0(\mu R)\eta\, d\eta_Q - \int_J K_0(\mu R)\eta\, d\eta_Q.$$

The first integral is zero because η is an odd function while $K_0(\mu R)$ is even. Thus

$$\begin{aligned}
\left|\int_J K_0(\mu R)\eta\, d\eta_Q\right| &\leqq 2\int_{2\delta}^{\infty} K_0(\mu R)\eta\, d\eta_Q \\
&\leqq 2\int_{2\delta}^{\infty} K_0(\mu R) R\, dR \\
&\leqq (\mu\delta)^{1/2} e^{-2\mu\delta} O\left(\frac{1}{\mu^2}\right) \\
&\leqq k_\gamma e^{-\mu r}\, O\left(\frac{1}{\mu^3}\right),
\end{aligned}$$

$$\int_I K_0(\mu R)|\eta|\, d\eta \leqq 2\int_0^{\infty} K_0(\mu R)\eta\, d\eta \leqq \max(1, \sqrt{\mu r})\, e^{-\mu r}\, O\left(\frac{1}{\mu^2}\right),$$

$$\begin{aligned}
\int_I K_0(\mu R)\eta^2\, d\eta &\leqq \int_{-\infty}^{\infty} K_0(\mu R)\eta^2\, d\eta_Q \\
&\leqq (1+\mu r)\, e^{-\mu r}\, O\left(\frac{1}{\mu^3}\right), \qquad \text{by the special case.}
\end{aligned}$$

Putting all of the above estimates together, we get the first estimate in the lemma. The other estimates are obtained in the same way.

Proof of the Theorem (Part B): For $1 \leqq k \leqq 4$, define

$$\gamma_k = \{ P \in \gamma^* | \text{arc distance}(P, \gamma) \leqq 4k\delta \} \subset \gamma^*.$$

For a point $P \in \gamma_3$, define its 4δ-neighborhood N_P in $\partial\Omega$ as in Lemma 4. We keep N_P fixed when we differentiate with respect to s_P.

We begin with the Green's formula:

$$v(P) = -\frac{1}{\pi}\int_{\partial\Omega_Q} K_0(\mu R)\frac{\partial v}{\partial n_Q}\, ds_Q + \frac{1}{\pi}\int_{\partial\Omega_Q} \frac{\partial K_0(\mu R)}{\partial n_Q} v(Q)\, ds_Q.$$

First, consider these integrals over $\partial\Omega - N_P$, where P is a point in γ_3. R^2 and $R\,\partial R/\partial n_Q$ are $C^{2,1}$-functions of s_P. Moreover, by Lemma 2, $R \geqq 2\delta$ over $\partial\Omega - N_P$. Therefore, R and $\partial R/\partial n_Q$ are $C^{2,1}$-functions of s_P. $K_0(\mu R)$ is an analytic function of R if $R \neq 0$. Therefore, $K_0(\mu R)$ and $\partial K_0(\mu R)/\partial n_Q$ are $C^{2,1}$-functions of s_P over $\partial\Omega - N_P$. We may compute their derivatives by differentiating inside the integral sign. Let ψ denote $\partial v/\partial n$. Let

$$w_1 = -\frac{1}{\pi}\int_{\partial\Omega - N_P} K_0(\mu R)\psi\, ds_Q.$$

By estimate (4.11) in Lemma 4,

$$|w_1| = \frac{\|\psi\|_{0,1,\partial\Omega}}{\mu^3|\partial\Omega|^6}\, O\left(\frac{1}{\mu^3}\right),$$

$$\frac{dw_1}{ds} = -\frac{1}{\pi}\int_{\partial\Omega - N_P} \frac{d}{dR}K_0(\mu R)\frac{\partial R}{\partial s_P}\psi\, ds_Q,$$

where $|\partial R/\partial s_P| \leqq 1$ by Lemma 1. Therefore,

$$\left|\frac{dw_1}{ds_P}\right| = \frac{\|\psi\|_{0,1,\partial\Omega}}{\mu^3|\partial\Omega|^6}\, O\left(\frac{1}{\mu^2}\right).$$

Let

$$w_2 = \frac{1}{\pi}\int_{\partial\Omega - N_P} \frac{\partial K_0(\mu R)}{\partial \eta_Q} v(Q)\, ds_Q,$$

$$\frac{\partial K_0(\mu R)}{\partial \eta_Q} = \frac{d}{dR}K_0(\mu R)\frac{\partial R}{\partial n_Q} \quad\text{and}\quad \left|\frac{\partial R}{\partial n_Q}\right| \leqq 1.$$

Consequently,

$$|w_2| = \frac{\|v\|_{0,1,\partial\Omega}}{\mu^3|\partial\Omega|^6}\, O\left(\frac{1}{\mu^2}\right) = \frac{\|v\|_{0,\infty,\partial\Omega}}{\mu|\partial\Omega|^4}\, O\left(\frac{1}{\mu^3}\right),$$

$$\frac{dw_2}{ds} = \frac{1}{\pi}\int_{\partial\Omega - N_P}\left[\frac{d^2}{dR^2}K_0(\mu R)\frac{\partial R}{\partial s_P}\frac{\partial R}{\partial n_Q} + \frac{d}{dR}K_0(\mu R)\frac{\partial}{\partial s_P}\frac{\partial R}{\partial n_Q}\right] v(Q)\, ds_Q.$$

By Lemma 2,

$$\left|\frac{\partial R}{\partial s_P}\frac{\partial R}{\partial n_Q}\right| \leqq 1 \quad \text{and} \quad \left|\frac{\partial}{\partial s_P}\frac{\partial R}{\partial n_Q}\right| \leqq \frac{1}{R} \leqq \tfrac{1}{2}\mu.$$

Therefore,

$$\left|\frac{dw_2}{ds}\right| = \frac{\|v\|_{0,\infty,\partial\Omega}}{\mu|\partial\Omega|^4} O\left(\frac{1}{\mu^2}\right).$$

Consider now the same integrals over N_P. Since $K_0(\mu R)$ has a logarithmic pole, differentiation inside the integral sign produces divergent integrals in most cases. Hence we proceed as follows. Let $T_{P,\delta} = [0, \delta] \times N_P$. If an integral has the form

$$\int_{N_P} G(P, Q)\Phi(Q)\, ds_Q,$$

where G is a function on $T_{P,\delta} \times N_P$ and Φ is a function on N_P, then

$$\begin{aligned}
\frac{d}{ds}\int_{N_P} G(P,Q)\Phi(Q)\, ds_Q &= \int_{N_P} \frac{\partial}{\partial s_P} G(P,Q)\Phi(Q)\, ds_Q \\
&= \int_{N_P}\left[-\frac{\partial}{\partial s_Q} + \partial_\sigma\right] G(P,Q)\Phi(Q)\, ds_Q, \\
&\qquad \text{where } \partial_\sigma = \partial/\partial s_P + \partial/\partial s_Q \\
&= -[G(P,Q)\Phi(Q)]_{Q_-}^{Q_+} + \int_{N_P} G(P,Q)\frac{d\Phi}{ds_Q}\, ds_Q \\
&\quad + \int_{N_P} \partial_\sigma G(P,Q)\Phi(Q)\, ds_Q,
\end{aligned}$$

where $Q_\pm$ are the end points of the arc N_P. $\partial_\sigma G(P, Q)$ will turn out to have singularity no worse than the singularity of $G(P, Q)$. Therefore, everything in the last step is convergent if the original integral is. If any of the integrals in any of the steps above turn out to be divergent we consider the integral

$$\int_{N_P} G(P^{\#}, Q)\Phi(Q)\, ds_Q,$$

where $P^{\#}$ is a point in $T_{P,\delta}$ with coordinates (r, s_P), carry out the above steps and then take the limit as $r \to 0$.

Let $v_1(P) = -(1/\pi)\int_{N_P} K_0(\mu R)\psi(Q)\,ds_Q$. By (4.1) in Lemma 4,

$$
\begin{aligned}
v_1(P) &= \|\psi\|_{0,\infty,N_P} O\left(\frac{1}{\mu}\right)\\
&= -\frac{\psi_P}{\mu} + C^{\psi}_{1,\gamma^*} O\left(\frac{1}{\mu^2}\right)\\
&= -\frac{\psi_P}{\mu} + C^{\psi}_{2,\gamma^*} O\left(\frac{1}{\mu^3}\right).
\end{aligned}
$$

Consider $v_1(P^{\#}) = -(1/\pi)\int_{N_P} K_0(\mu R^{\#})\psi(Q)\,ds_Q$ where $R^{\#} = |P^{\#}Q|$. Then,

$$
\begin{aligned}
\frac{\partial v_1}{\partial s} &= \frac{1}{\pi}\left[K_0(\mu R^{\#})\psi(Q)\right]_{Q-}^{Q+} - \frac{1}{\pi}\int_{N_P} K_0(\mu R^{\#})\frac{d\psi}{ds_Q}\,ds_Q\\
&\quad + \frac{1}{2\pi}\int \mathscr{L}K_0(\mu R^{\#})\,\partial_\sigma (R^{\#})^2 \psi(Q)\,ds_Q,
\end{aligned}
$$

where $\mathscr{L} = -(1/R)\,d/dR$.

By Lemma 1,

$$
\partial_\sigma (R^{\#})^2 = \left[rO(\sigma^2) + O(\sigma^3)\right] k^2_{\gamma^*}.
$$

Since $\mathscr{L}K_0(\mu R)$ has singularity like $1/\sigma^2$ at $R = 0$, the last integral is convergent when $r = 0$. Therefore,

$$
\begin{aligned}
\frac{dv_1(P)}{ds} &= \frac{1}{\pi}\left[K_0(\mu R)\psi(Q)\right]_{Q-}^{Q+} + \frac{1}{2\pi}\int_{N_P} \mathscr{L}K_0(\mu R)\,\partial_\sigma R^2 \psi(Q)\,ds_Q\\
&\quad - \frac{1}{\pi}\int_{N_P} K_0(\mu R)\frac{d\psi}{ds_Q}\,ds_Q.
\end{aligned}
$$

By Lemma 3, the first term equals

$$
\|\psi\|_{0,\infty,N_P}\frac{e^{-2\mu\delta}}{\sqrt{2\mu\delta}}O(1) = \frac{\|\psi\|_{0,\infty,\gamma^*}}{\mu^4|\partial\Omega|^6} O\left(\frac{1}{\mu^2}\right) = \frac{\|\psi\|_{0,\infty,\gamma^*}}{|\partial\Omega|^2} O\left(\frac{1}{\mu^2}\right).
$$

Consider now the second term. Since $\partial_\sigma R^2 = k^2_{\gamma^*}O(\sigma^3)$, the term equals $k^2_{\gamma^*}\|\psi\|_{0,\infty,N_P} O(1/\mu^2)$ by (4.9) in Lemma 4. By (4.1) again,

$$
-\frac{1}{\pi}\int_{N_P} K_0(\mu R)\frac{d\psi}{ds_Q}\,ds_Q = \|\psi'\|_{0,\infty,N_P} O\left(\frac{1}{\mu}\right) = -\frac{1}{\mu}\psi'_P + C^{\psi'}_{1,\gamma^*} O\left(\frac{1}{\mu^2}\right).
$$

Putting all the pieces together, we get

$$\frac{dv_1}{ds} = C^{\psi}_{1,\gamma^*} O\left(\frac{1}{\mu}\right) = -\frac{1}{\mu}\psi'_P + C^{\psi}_{2,\gamma^*} O\left(\frac{1}{\mu^2}\right).$$

Consider

$$v_2 = \frac{1}{\mu}\int_{N_P} \frac{\partial K_0(\mu R)}{\partial n_Q} v(Q)\, ds_Q = -\frac{1}{\pi}\int_{N_P} \mathscr{L}K_0(\mu R) R \frac{\partial R}{\partial n_Q} v(Q)\, ds_Q.$$

Let

$$|||v|||_{0,\gamma^*} = \left\| \frac{\partial v}{\partial n} \right\|_{0,\infty,\gamma^*} + \frac{k^2_{\gamma^*}}{\mu} \|v\|_{0,\infty,\partial\Omega} + \frac{\|\partial v/\partial n\|_{0,1,\partial\Omega}}{\mu^5 |\partial\Omega|^6}.$$

Since v appears on both sides of the equation, we have to bootstrap the result, successively shrinking the neighborhood of γ, from γ_4 to γ_1, during the process.

Start with $P \in \gamma_3$. Since $R\, \partial R/\partial n_Q = k_{\gamma^*} O(\sigma^2)$,

$$v_2(P) = \|v\|_{0,\infty,\gamma^*} k_{\gamma^*} O\left(\frac{1}{\mu}\right).$$

Hence, for all $P \in \gamma_3$,

$$\begin{aligned} v(P) &= w_1(P) + w_2(P) + v_1(P) + v_2(P) \\ &= \left[\|\psi\|_{0,\infty,\gamma^*} + k_{\gamma^*} \|v\|_{0,\infty,\partial\Omega} + \frac{\|\psi\|_{0,1,\partial\Omega}}{\mu^5|\partial\Omega|^6} \right] O\left(\frac{1}{\mu}\right). \end{aligned}$$

Restrict P now to γ_2 so that $N_P \subset \gamma_3$. Then,

$$v_2(P) = \|v\|_{0,\infty,N_P} K_{\gamma^*} O\left(\frac{1}{\mu}\right) = |||v|||_{0,\gamma^*} O\left(\frac{1}{\mu}\right),$$

$$v(P) = |||v|||_{0,\gamma^*} O\left(\frac{1}{\mu}\right).$$

Next, restrict P to γ_1 so that $N_P \subset \gamma_2$. Then,

$$v_2(P) = k_{\gamma^*} |||v|||_{0,\gamma^*} O\left(\frac{1}{\mu^2}\right),$$

and

$$v(P) = -\frac{\psi(P)}{\mu} + |||v|||_{1,\gamma^*} O\left(\frac{1}{\mu^2}\right),$$

which essentially proves Part A of the theorem. Finally, restrict P to γ. Let $u = v + \psi/\mu$. Then,

$$v_2(P) = -\frac{1}{\pi}\int_{N_P} \mathscr{L}K_0(\mu R) R \frac{\partial R}{\partial n_Q} v(Q)\, ds_Q$$

$$= -\frac{1}{\pi}\int_{N_P} \mathscr{L}K_0(\mu R)\left[-\tfrac{1}{2}\kappa_P \sigma^2 + k_{\gamma^*}^2 O(\sigma^3)\right]\left[-\frac{\psi(Q)}{\mu} + u(Q)\right] ds_Q$$

$$= -\frac{\kappa_P}{2\pi\mu}\int_{N_P} \mathscr{L}K_0(\mu R)\sigma^2\psi(Q)\, ds_Q$$

$$+\frac{1}{\mu}\int_{N_P} \mathscr{L}K_0(\mu R) k_{\gamma^*}^2 O(\sigma^3)\psi(Q)\, ds_Q$$

$$+\int_{N_P} \mathscr{L}K_0(\mu R) k_{\gamma^*} O(\sigma^2) u(Q)\, ds_Q.$$

By (4.3),

$$\text{the first integral} = -\frac{\kappa_P}{2\mu^2}\psi(P) + k_\gamma^* C_{1,\gamma^*}^\psi O\left(\frac{1}{\mu^3}\right),$$

$$\text{the second integral} = k_{\gamma^*}^2 \|\psi\|_{0,\infty,\gamma^*} O\left(\frac{1}{\mu^3}\right),$$

and

$$\text{the third integral} = k_{\gamma^*}\|u\|_{0,\infty,N_P} O\left(\frac{1}{\mu}\right) = k_{\gamma^*} |||v|||_{1,\gamma^*} O\left(\frac{1}{\mu^3}\right).$$

Therefore,

$$v_2(P) = -\frac{\kappa_P}{2\mu^2}\psi_P + |||v|||_{2,\gamma^*} O\left(\frac{1}{\mu^3}\right),$$

and

$$v(P) = -\left(1 + \frac{\kappa_P}{2\mu}\right)\frac{\psi_P}{\mu} + |||v|||_{2,\gamma^*} O\left(\frac{1}{\mu^3}\right).$$

It remains to estimate dv_2/ds. For all $P \in \gamma$,

$$\frac{dv_2}{ds} = -\frac{1}{\pi}\frac{d}{ds}\int_{N_P} \mathscr{L}K_0(\mu R) R \frac{\partial R}{\partial n_Q}\left(-\frac{\psi(Q)}{\mu} + u(Q)\right) ds_Q,$$

$$\frac{1}{\pi\mu}\frac{d}{ds}\int_{N_P} \mathscr{L}K_0(\mu R) R \frac{\partial R}{\partial n_Q}\psi(Q)\, ds_Q$$

$$= -\frac{1}{\pi\mu}\left[\mathscr{L}K_0(\mu R) R \frac{\partial R}{\partial n_Q}\psi(Q)\right]_{Q_-}^{Q_+}$$

$$+ \frac{1}{\pi\mu}\int_{N_P} \partial_\sigma\left[\mathscr{L}K_0(\mu R) R \frac{\partial R}{\partial n_Q}\right]\psi(Q)\, ds_Q$$

$$+ \frac{1}{\pi\mu}\int_{N_P} \mathscr{L}K_0(\mu R) R \frac{\partial R}{\partial n_Q}\frac{d\psi}{ds_Q}\, ds_Q.$$

The first term is easily seen to be of order $k_{\gamma^*}^2\|\psi\|_{0,\infty,\gamma^*}O(1/\mu^2)$:

$$\frac{1}{\pi\mu}\int_{N_P} \partial_\sigma\left[\mathscr{L}K_0(\mu R) R \frac{\partial R}{\partial n_Q}\right]\psi(Q)\, ds_Q$$

$$= \frac{1}{\pi\mu}\int_{N_P:}\left[-\tfrac{1}{2}\mathscr{L}R^2\mathscr{L}K_0(\mu R)(\partial_\sigma R^2)\left(\frac{1}{R}\frac{\partial R}{\partial n_Q}\right)\right.$$

$$\left. + \mathscr{L}K_0(\mu R) R^2 \partial_\sigma\left(\frac{1}{R}\frac{\partial R}{\partial n_Q}\right)\right]\psi(Q)\, ds_Q$$

$$= -\frac{1}{\pi\mu}\int_{N_P}\left[\tfrac{1}{2}\mu^2 K_0(\mu R) + \mathscr{L}K_0(\mu R)\right](\partial_\sigma R^2)\left(\frac{1}{R}\frac{\partial R}{\partial n_Q}\right)\psi(Q)\, ds_Q$$

$$+ \frac{1}{\pi\mu}\int_{N_P} \mathscr{L}K_0(\mu R)\,\partial_\sigma\left(R\frac{\partial R}{\partial n_Q}\right)\psi(Q)\, ds_Q.$$

By Lemmas 1 and 2,

$$\left(\frac{1}{R}\frac{\partial R}{\partial n_Q}\right)(\partial_\sigma R^2) = \frac{k_{\gamma^*}^2 O(\sigma^2) O(\sigma^2)}{R^2} = k_{\gamma^*}^2 O(\sigma^2)$$

and

$$\partial_\sigma\left(R\frac{\partial R}{\partial n_Q}\right) = k_{\gamma^*}^2 O(\sigma^2).$$

Therefore, by (4.3) and (4.6),

$$\frac{1}{\pi\mu}\int_{N_P}\partial_\sigma\left[\mathscr{L}K_0(\mu R)R\frac{\partial R}{\partial n_Q}\right]\psi(Q)\,ds_Q = k_{\gamma^*}^2\|\psi\|_{0,\infty,\gamma^*}O\left(\frac{1}{\mu^2}\right).$$

Since $R\,\partial R/\partial n_Q = k_{\gamma^*}O(\sigma^2)$,

$$\frac{1}{\pi\mu}\int_{N_P}\mathscr{L}K_0(\mu R)R\frac{\partial R}{\partial n_Q}\frac{d\psi}{ds_Q}\,ds_Q = k_{\gamma^*}\|\psi'\|_{0,\infty,\gamma^*}O\left(\frac{1}{\mu^2}\right).$$

The remaining integral is

$$= -\frac{1}{\pi}\frac{d}{ds}\int_{N_P}\mathscr{L}K_0(\mu R)R\frac{\partial R}{\partial n_Q}u(Q)\,ds_Q$$

$$= -\frac{1}{\pi}\int_{N_P}\left[\mathscr{L}R^2\,\mathscr{L}K_0(\mu R)\left(R\frac{\partial R}{\partial s_P}\right)\left(\frac{1}{R}\frac{\partial R}{\partial n_Q}\right)\right.$$

$$\left.-R^2\mathscr{L}K_0(\mu R)\frac{\partial}{\partial s_P}\left(\frac{1}{R}\frac{\partial R}{\partial n_Q}\right)\right]u(Q)\,ds_Q$$

$$= \frac{\mu^2}{\pi}\int_{N_P}K_0(\mu R)\left(R\frac{\partial R}{\partial s_P}\right)\left(\frac{1}{R}\frac{\partial R}{\partial n_Q}\right)u(Q)\,ds_Q$$

$$-\frac{1}{\pi}\int_{N_P}\mathscr{L}K_0(\mu R)\left[\frac{\partial}{\partial s_P}\left(R\frac{\partial R}{\partial n_Q}\right) - 2\left(R\frac{\partial R}{\partial s_P}\right)\left(\frac{1}{R}\frac{\partial R}{\partial n_Q}\right)\right]u(Q)\,ds_Q.$$

Since

$$\left(R\frac{\partial R}{\partial s_P}\right)\left(\frac{1}{R}\frac{\partial R}{\partial n_Q}\right) = \frac{O(\sigma)k_{\gamma^*}O(\sigma^2)}{R^2} = k_{\gamma^*}O(\sigma),$$

$$\text{the first integral} = \mu^2\|u\|_{0,\infty,N_P}k_{\gamma^*}O\left(\frac{1}{\mu^2}\right) = k_{\gamma^*}|||v|||_{1,\gamma^*}O\left(\frac{1}{\mu^2}\right).$$

In the second integral,

$$\frac{1}{R^2}\left[R^2\frac{\partial}{\partial s_P}\left(R\frac{\partial R}{\partial n_Q}\right) - 2\left(R\frac{\partial R}{\partial s_P}\right)\left(R\frac{\partial R}{\partial n_Q}\right)\right]$$

$$= \frac{1}{R^2}\Big\{\left[\sigma^2 + k_{\gamma^*}O(\sigma^4)\right]\left[\kappa_P(s_Q - s_P) + \kappa_{\gamma^*}^2 O(\sigma^2)\right]$$

$$-2\left[(s_Q - s_P) + k_{\gamma^*}^2 O(\sigma^3)\right]\left[\tfrac{1}{2}\kappa_P\sigma^2 + k_{\gamma^*}^2 O(\sigma^3)\right]\Big\}$$

$$= k_{\gamma^*}^2 O(\sigma^2).$$

Therefore,

$$\text{the second integral} = k_{\gamma^*}^2\|u\|_{0,\infty,N_P} O\left(\frac{1}{\mu}\right)$$

$$= k_{\gamma^*}^2 |||v|||_{1,\gamma^*} O\left(\frac{1}{\mu^3}\right)$$

$$= k_{\gamma^*} |||v|||_{1,\gamma^*} O\left(\frac{1}{\mu^2}\right).$$

Adding up all the estimates, we get

$$\frac{dv_2}{ds}(P) = |||v|||_{2,\gamma^*} O\left(\frac{1}{\mu^2}\right)$$

and

$$\frac{dv}{ds}(P) = -\frac{\psi_P'}{\mu} + |||v|||_{2,\gamma^*} O\left(\frac{1}{\mu^2}\right).$$

Acknowledgment. The first author gratefully acknowledges the Army Research Office and the National Science Foundation for support under grants DAAL03-86-K-0171 and IST-85-12736. The second author gratefully acknowledges the support of the National Science Foundation under grant IRI-8704467.

Bibliography

[1] Allard, W. K., and Almgren, F. J., *The structure of stationary one-dimensional varifolds with positive density*, Inventiones Math. 34, 1976, pp. 83–97.

[2] Abramowitz, M., and Stegun, I. A., eds., *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematics Series, 1965.

[3] Buser, P., *On Cheeger's inequality*, Proc. of Symp. in Pure Math. 36, 1980.

[4] Blake, A., and Zisserman, A., *Using Weak Continuity Constraints*, Report CSR-186-85, Dept. of Comp. Sci., Edinburgh University, 1985.

[5] Freidlin, M., *Functional Integration and Partial Differential Equations*, Annals of Math. Studies, Princeton University Press, 1985.

[6] Geman, S., and Geman, D., *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*, IEEE Trans., PAMI 6, 1984, pp. 721–741.

[7] Grisvard, P., *Elliptic Problems in Nonsmooth Domains*, Pitman, 1985.

[8] Gurtin, M. E., *On a theory of phase transitions with interfacial energy*, Archive Rat. Mech. Anal. 87, 1985, pp. 187–212.

[9] Kondratiev, V. A., *Boundary value problems for elliptic equations in domains with conical or angular points*, Trans. Moscow Math. Soc. 16, 1967, pp. 227–313.

[10] Marroquin, J. L., *Surface reconstruction preserving discontinuities*, Artificial Intelligence Lab. Memo 792, M.I.T., 1984.

[11] McKean, H., *Stochastic Integrals*, Academic Press, 1969.

[12] Rice, J. R., *Mathematical analysis in the mechanics of fracture*, in *Fracture, An Advanced Treatise*, H. Liebowitz, ed., Academic Press, 1968.

[13] Simon, L., *Lectures on Geometric Measure Theory*, Proc. Centre for Math. Anal., Australian Nat. Univ., 3, 1983.

[14] Whittaker, E. T., and Watson, G. N., *A Course of Modern Analysis*, 4th ed., Cambridge University Press, 1952.

Received July, 1988.