

Geometric remarks on the level curves of harmonic functions

L. De Carli and S. M. Hudson

ABSTRACT

Suppose that u is a nonconstant harmonic function on the plane. By the maximum principle, its zero set Z does not contain any simple closed curve. This paper provides bounds on the curvature of Z , and other conditions on Z , which imply that u is of some special type, such as polynomial, linear, or even constant. These theorems resemble unique continuation results, which often apply to various classes of nonharmonic functions. For such functions, Z may contain a closed curve, but a lower bound is established for the area of the enclosed region.

1. Introduction

Many interesting open problems in analysis concern the geometry of level sets of harmonic functions. In particular, we are concerned with the properties of Z , the zero set of a harmonic function u , and with connections to unique continuation problems. There seem to be very few geometric results about these zero sets, even for harmonic functions on R^2 . Of course, if Z contains a simple closed curve, then u vanishes inside, by the maximum principle, and thus it vanishes on R^2 . We usually assume that u is nontrivial, unless stated otherwise.

In [2], the author shows that if a curve in Z is parametrized by $x = p(t)$ and $y = q(t)$, where $p(t)$ and $q(t)$ are polynomials with real coefficients, then the curve is a straight line or a parabola. Therefore the zero set of an harmonic function can contain $y = x^2$ but it cannot contain a cubic, such as $y = x^3$. See also [3]. These results do not apply, for example, to $u(x, y) = xy + 1$, since hyperbolas cannot be parametrized by polynomials.

It is not clear whether any simple geometric properties distinguish curves like $y = x^3$ from those like $y = x^2$. It is certainly possible for Z to have an inflection point. For example, let $u = x^3 - 3xy^2 + x + y$. Then, the origin is an inflection point of Z .

This paper aims to show in various ways that Z cannot ‘wiggle very much’. We investigate variations of the simple remark above, that Z cannot contain a simple closed curve, at least when u is harmonic, and this leads to some results involving area, but our main results concern the curvature κ of Z . It is impossible to bound κ independently of u and D ; an example with large curvature is given in Subsection 2.2, but rescaling of almost any function will also produce a large κ . However, when Z is a *single* connected curve in $D = B_1(0)$, we can prove a fairly sharp bound on κ , at any point $P \in Z$. The bound depends only on the distance from P to the boundary. Our main result assumes for simplicity that P is the origin and that u is continuous on the boundary. It is proved in Section 2.

THEOREM 1.1. *Suppose that $\Delta u = 0$ on $B_r(0) = \{|x| \leq r\} \subset R^2$ with continuous boundary values $f(\theta)$. Assume that $f(\theta) > 0$ on an interval $I = (\alpha, \beta) \subset [-\pi, \pi]$, $f(\alpha) = f(\beta) = 0$ and $f(\theta) < 0$ otherwise. Suppose that $u(0, 0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) d\theta = 0$. Then Z is a curve and its curvature at $(0, 0)$ is bounded in absolute value by C/r , where $C \leq 24$ does not depend on u .*

If Z is a single curve in the entire plane (the case $r = \infty$ of this theorem), we might heuristically conclude that $\kappa = 0$ at every point of Z , so that Z and u must be linear. This claim is true, but we can prove only a special case of it directly from Theorem 1.1. Instead, we prove the general version in Section 3, using complex variables. Likewise, if Z consists of m curves, then u must be a polynomial of degree m ; this is a slightly improved version of a theorem in [9].

The paper [9] is a good introduction to this one, since it clarifies the local structure of Z . Near each point $P \in Z$, u behaves like a rotation of $\operatorname{Re}(z - P)^m$, for some $m \geq 1$. Thus, locally, Z consists of m analytic curves passing through P . We can conclude that on any compact set, Z consists of finitely many maximal analytic curves, for example, ones not contained in a larger analytic curve. In an open set D , Z may split into infinitely many such curves. In fact, our Theorems 3.2 and 3.3 imply that this *must* happen when $D = R^2$ and u is not a polynomial. It can also happen on $D = B_1(0)$, if u does not extend analytically to the boundary. For example, consider the harmonic function with boundary values $\theta \sin(1/\theta)$, with $-\pi < \theta \leq \pi$, which vanishes infinitely often on the boundary of D .

Section 3 also contains some other linearity results. For example, when Z contains a line segment, it contains a full line. Several of these theorems can be viewed as uniqueness results for harmonic functions, based on the properties of Z . They partially answer a very general problem; suppose that u is harmonic on a domain $D \subseteq R^n$ and $u \equiv 0$ on a set $Z \subseteq D$. What properties of Z imply that $u \equiv 0$ on all of D ? For example, if Z contains an open ball, then u vanishes on D ; this is often called the *unique continuation property* (UCP). Or, if Z is a single point, and u vanishes with all its derivatives there (we also say that u vanishes to infinite order at this point), then u vanishes on D ; this is the *strong unique continuation property* (SUCP). These properties extend to certain classes of nonharmonic functions discussed in some parts of this paper. Suppose that

$$\Delta u(x) \leq V(x)|u(x)|, \quad (1.1)$$

where V is called the potential of u . If $V \in L_{\text{loc}}^{n/2}(D)$ (where $D \subseteq R^n$), then u has the UCP and SUCP[†] [4]. There is extensive literature on these problems; see, for example, the survey paper of Wolff [10] and a recent article of Koch and Tataru [6] and the literature cited there.

Typically, unique continuation proofs depend on Carleman estimates. The authors attempted a new approach in [1], and became curious about the intermediate case, in which Z is a set of positive dimension $0 < d < n$, rather than an open ball, or a single point.

Any harmonic function in R^2 that vanishes on a closed curve must also vanish inside, by the maximum principle, but for general solutions of (1.1) there is no maximum principle, and Z may contain closed curves. However, we prove that the area of a region enclosed by Z cannot be arbitrarily small. We prove a lower bound that depends on $\|V\|_\infty$. We also include variations of this result for R^n , with $n > 2$.

In Section 5, we discuss a related conjecture: if the curve is ‘almost closed’, then u must vanish. However, this idea seems very unlikely unless we impose fairly severe conditions on u , for example, that u be a harmonic polynomial. Section 5 contains several other open questions about extending subsets of Z , related to Section 3, and about the local arc length of Z .

[†]In this context, $u \in W^{2,p}(D)$ vanishes to infinite order at x_0 if for every $N > 0$, we have

$$\lim_{\epsilon \rightarrow 0} \epsilon^{-N} \int_{|x-x_0|<\epsilon} |u(x)|^2 dx = 0.$$

2. Curvature

In this section, we investigate the curvature κ of $Z \subset \mathbf{R}^2$, when u is harmonic. Jerrard and Rubel [5] proved that $\log |\kappa|$ is superharmonic, on a nice enough domain. See [8] for another monotonicity property of curvature. Unfortunately, these do not seem to say much about a given level curve. Before proving Theorem 1.1 and some corollaries, we recall a few useful curvature formulas.

2.1. Curvature in \mathbf{R}^2

The standard formula for the curvature of the graph of a function $y = f(x) \in C^2(\mathbf{R})$ at a point $P = (p, f(p))$ is as follows:

$$\kappa(f)(P) = \frac{f''(p)}{(1 + [f'(p)]^2)^{3/2}}.$$

We use this formula to compute the curvature of the level sets of a $C^2(D)$ function $u(x, y)$ at points where the gradient does not vanish. Here D is an open set of \mathbf{R}^2 .

Let $L = u^{-1}(c)$ be the set of points where $u \equiv c$. Let $P = (p, q) \in L$ be such that $\nabla u(P) \neq 0$. Suppose that L is the graph, in a neighborhood of P , of a function $y = y(x)$. By the implicit function theorem, there exists a neighborhood of P in L in which

$$y'(x) = -\frac{u_x(x, y(x))}{u_y(x, y(x))}.$$

Consequently, we have

$$\begin{aligned} y''(P) &= -\frac{(u_{xx}(P) + y'(p)u_{xy}(P))u_y(P) - u_x(P)(u_{xy}(P) + y'(p)u_{yy}(P))}{u_y(P)^2} \\ &= -\frac{u_y^2(P)u_{xx}(P) - 2u_{xy}(P)u_x(P)u_y(P) + u_x^2(P)u_{yy}(P)}{u_y(P)^3}. \end{aligned}$$

The curvature of L at p is given by

$$\kappa(u)(P) = \frac{y''(p)}{(1 + (y'(p))^2)^{3/2}},$$

and by the previous calculation,

$$\kappa(u)(P) = -\frac{u_y^2(P)u_{xx}(P) - 2u_{xy}(P)u_x(P)u_y(P) + u_x^2(P)u_{yy}(P)}{(u_x^2(P) + u_y^2(P))^{3/2}}. \quad (2.1)$$

Hence, $\kappa(u)(P)$ associates to each $P \in D$ (with $\nabla u(P) \neq (0, 0)$), the curvature of the level set of u that passes through P . We say in short that $\kappa(u)(P)$ is *the curvature of u at P* . If we let $\tilde{H}u(P) = \begin{pmatrix} -u_{yy}(P) & u_{xy}(P) \\ u_{xy}(P) & -u_{xx}(P) \end{pmatrix}$, then

$$\kappa(u)(P) = |\nabla u(P)|^{-3} \langle \nabla u(P), \tilde{H}u(P) \nabla u(P) \rangle. \quad (2.2)$$

We recall one more useful curvature formula. If $u = \operatorname{Re}(w)$, where w is a holomorphic function in D , then the curvature of a level set of u at point P , where $\nabla u(P) \neq 0$, is given by

$$\kappa(u)(P) = |w'(P)| \operatorname{Re} \left(\frac{w''(P)}{(w'(P))^2} \right). \quad (2.3)$$

For example, see [5].

2.2. Proof of Theorem 1.1

Proof of Theorem 1.1. We can assume that $r = 1$ by scaling. Let $u(x, y)$ be harmonic on the unit disk, with $u(0, 0) = 0$. We can assume that $\nabla u(0, 0) \neq \mathbf{0}$; otherwise, Z would intersect

itself there, and therefore intersect the unit circle more than twice (by the maximum principle). By rotation, we can assume that $u_y(0,0) = 0$, and that $D = \pi u_x(0,0) > 0$. Then u can be represented by the Poisson integral in terms of its boundary values, $f(\theta)$, as follows:

$$u(x,y) = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) \frac{(1-x^2-y^2)d\theta}{(x-\cos\theta)^2 + (y-\sin\theta)^2}. \quad (2.4)$$

By hypothesis, $f > 0$ on an interval $I = [a,b] \subset [0,2\pi]$. Without loss of generality, $\int_I f(\theta) d\theta = 1$. We have $\int_0^{2\pi} f(\theta) d\theta = u(0,0) = 0$. Several averages, especially ones involving $\cos\theta$, reappear throughout this proof. We denote by D and C those taken over $[0,2\pi]$ and I , respectively. We use d and c for the corresponding ones with $\cos^2\theta$. Define

$$\begin{aligned} S &= \int_I f(\theta) \sin\theta d\theta, & C &= \int_I f(\theta) \cos\theta d\theta, & c &= \int_I f(\theta) \cos^2\theta d\theta, \\ D &= \pi u_x(0,0) = \int_0^{2\pi} f(\theta) \cos\theta d\theta, & d &= -\frac{\pi}{4} u_{yy}(0,0) = \int_0^{2\pi} f(\theta) \cos^2\theta d\theta. \end{aligned}$$

Furthermore, note

$$0 = u_y(0,0) = \frac{1}{\pi} \int_0^{2\pi} f(\theta) \sin\theta d\theta.$$

From (2.1), we find that $|\kappa| = |u_{yy}/u_x| = |4d/D|$. We can assume $d > 0$ (if not, replace u by $-u$, rotate by π radians, and replace $[a,b]$ by its complement; this operation keeps $D > 0$). Hence, we must prove $d \leq 6D$.

Roughly speaking, this means we need to bound c/C . We will prove a bound like this for the positive part of u , defined from f restricted to $[a,b]$, and for the negative part separately. We define G_0 precisely below, but it is basically the contribution to D from an extreme f , consisting of point masses at a and b . Likewise, g_0 can be thought of as the contribution to c (and d) from the same extreme function. We compare our f to the extreme case by proving Lemma 2.1 below.

We now define g_0 and G_0 . Given two angles, such that $\sin(\theta) > S > \sin(\gamma)$, there is a unique $\alpha = \alpha(\theta, \gamma) \in (0,1)$, such that $\alpha \sin(\theta) + (1-\alpha) \sin(\gamma) = S$, namely $\alpha = (S - \sin(\gamma))/(\sin(\theta) - \sin(\gamma))$. Set $G(\theta, \gamma) = \alpha \cos(\theta) + (1-\alpha) \cos(\gamma)$, which is just the x coordinate of the intersection of the line $y = S$ with the line through $e^{i\theta}$ and $e^{i\gamma}$. According to Lemma 2.3 below, $\sin(b) > S > \sin(a)$, and hence we can define $G_0 = G(b, a)$. Likewise, set $g(\theta, \gamma) = \alpha \cos^2(\theta) + (1-\alpha) \cos^2(\gamma)$ and let $g_0 = g(b, a)$. The proof of Theorem 1.1 is a consequence of the following lemma.

LEMMA 2.1. *We have*

$$|c - g_0| \leq 6(C - G_0). \quad (2.5)$$

To see the plan of the proof, suppose that $f|_I$ is concentrated at the endpoints of I . Then both sides of (2.5) are zero. The claim is that if we smooth out this f , we increase C as fast as c . To make this plan work, we need to define and relate some parameters: an angle θ near b , a lower angle γ near a , and $t \in [0,1]$.

Since $\sin(a) < \sin(b)$, it follows that there is a $p \in (a,b)$ such that $\sin(p) = S$. Likewise, there is a $p_2 \in (b, a + 2\pi)$ such that $\sin(p_2) = S$. Since $\sin(x) = S$ has only two solutions with $x \in (a, a + 2\pi)$, it follows that p and p_2 are unique. Hence, for $a < \gamma < p < \theta < b$, we can define

$$T(\theta, \gamma) = \frac{f(\theta) \sin\theta - S}{f(\gamma)(S - \sin\gamma)} > 0.$$

Define $\gamma = \gamma(\theta)$ by $d\gamma/d\theta = -T(\theta, \gamma) < 0$, with $\gamma(b) = a$.

The domain and range of γ are not yet clear, since the construction requires $\gamma < p$; but we claim that it defines a bijection $\gamma : (p, b] \rightarrow [a, p)$. As long as $\gamma < p < \theta$, we have

$$\int_a^\gamma f(\gamma)(S - \sin \gamma) d\gamma = \int_\theta^b f(\theta)(\sin \theta - S) d\theta,$$

and from the definition of S , we obtain

$$\int_a^b f(\theta)(\sin \theta - S) d\theta = 0.$$

If either $\gamma = p$ or $\theta = p$, these equations imply that $\gamma = \theta = p$, which proves our claim about γ . Also, we can define $\gamma(p) = p$, with continuity.

We now regard α , G , and g as functions of θ alone. Yet another parameter will be useful.

Define $t = t(\theta)$ by $dt/d\theta = -(1/\alpha)f(\theta)$, with $t(b) = 0$, which implies that $dt/d\gamma = f(\gamma)/(1 - \alpha)$. Then we have

$$\int_p^b f d\theta = \int_0^{t(p)} \alpha dt \quad \text{and} \quad \int_a^p f d\gamma = \int_0^{t(p)} (1 - \alpha) dt.$$

Now, we add and get $1 = \int_0^{t(p)} dt = t(p)$, and thus $t : [p, b] \rightarrow [0, 1]$. Likewise, we can multiply the integrands by $\cos(\theta)$ in the first equation and by $\cos(\gamma)$ in the second, to get

$$C = \int_p^b f(\theta) \cos(\theta) d\theta + \int_a^p f(\gamma) \cos(\gamma) d\gamma = \int_0^1 G(\theta(t)) dt.$$

The same procedure with \cos^2 leads to

$$c = \int_0^1 g(\theta(t)) dt.$$

Recall that $G_0 = G(b, a) = G(\theta(0))$ and that $g_0 = g(b, a) = g(\theta(0))$. We will prove that $\partial G/\partial \theta < 0$ and hence $\partial G/\partial \gamma > 0$, and hence $dG/d\theta < 0$ and $G(\theta(t))$ increases with t . Thus, $C > G_0$. To prove Lemma 2.1, we must bound

$$\frac{|c - g_0|}{C - G_0} = \frac{|\int_0^1 g dt - g_0|}{\int_0^1 G dt - G_0} = \frac{|g(\theta) - g_0|}{G(\theta) - G_0}$$

for some $\theta \in (p, b)$, by the mean value theorem. We must show that

$$|g(\theta) - g_0| \leq 6(G(\theta) - G_0). \tag{2.6}$$

LEMMA 2.2. *We have*

$$\left| \frac{\partial g}{\partial \theta} \right| \leq -6 \frac{\partial G}{\partial \theta} \quad \text{and} \quad \left| \frac{\partial g}{\partial \gamma} \right| \leq 6 \frac{\partial G}{\partial \gamma}. \tag{2.7}$$

Assuming this for the moment, we can prove (2.6) and Lemma 2.1. Indeed

$$\left| \frac{dg}{d\theta} \right| \leq |\partial g \partial \theta| + \left| \partial g \partial \gamma \frac{d\gamma}{d\theta} \right| \leq -6 \frac{\partial G}{\partial \theta} - 6 \frac{\partial G}{\partial \gamma} \frac{d\gamma}{d\theta} = -6 \frac{dG}{d\theta}.$$

Proof of Lemma 2.2. From the definition of α , it is clear that $\partial \alpha / \partial \theta = -\alpha(\cos(\theta)/\Delta_s)$, where $\Delta_s = \sin(\theta) - \sin(\gamma) > 0$. Likewise, set $\Delta_c = \cos(\theta) - \cos(\gamma)$, and $\Delta_{c^2} = \cos^2(\theta) - \cos^2(\gamma)$. Thus, we have

$$\frac{\partial G}{\partial \theta} = \frac{\partial \alpha}{\partial \theta} \Delta_c - \alpha \sin(\theta)$$

and

$$-\frac{\partial G}{\partial \theta} \frac{\Delta_s}{\alpha} = \cos(\theta)\Delta_c + \sin(\theta)\Delta_s = \mathbf{w}_\theta \cdot (\mathbf{w}_\theta - \mathbf{w}_\gamma) > \frac{\Delta_c^2}{4}, \quad (2.8)$$

where \mathbf{w}_θ is the vector $(\cos \theta, \sin \theta) \in \mathbf{R}^2$ and the dot is the usual scalar product in \mathbf{R}^2 . For the last step above, see Lemma 2.4. Note that $\partial G/\partial \theta < 0$. Likewise, we have

$$\begin{aligned} -\frac{\partial g}{\partial \theta} \frac{\Delta_s}{\alpha} &= \cos \theta \Delta_{c^2} + 2 \sin \theta \cos \theta \Delta_s \\ &= 2 \cos \theta [\cos \theta \Delta_c + \sin \theta \Delta_s] - \cos \theta \Delta_c^2 \\ &= 2 \cos \theta \left[-\frac{\partial G}{\partial \theta} \frac{\Delta_s}{\alpha} \right] - \cos \theta \Delta_c^2. \end{aligned}$$

With (2.8) and the triangle inequality, this proves the first part of Lemma 2.2.

A similar calculation proves $|\partial g/\partial \gamma| \leq 6(\partial G/\partial \gamma)$. Set $\beta = 1 - \alpha$. Then $\beta' = \partial \beta/\partial \gamma = \beta \cos \gamma/\Delta_s$ and

$$\frac{\Delta_s}{\beta} \frac{\partial G}{\partial \gamma} = -(\cos \gamma \Delta_c + \sin \gamma \Delta_s) = \mathbf{w}_\theta \cdot (\mathbf{w}_\theta - \mathbf{w}_\gamma) > \frac{\Delta_c^2}{4}$$

much as before. Furthermore, we have

$$\begin{aligned} \frac{\partial g}{\partial \theta} \frac{\Delta_s}{\beta} &= 2 \cos \gamma [\cos \gamma \Delta_c + \sin \gamma \Delta_s] + \cos \gamma [\Delta_c]^2 \\ &= 2 \cos \gamma \left[-\frac{\partial G}{\partial \gamma} \frac{\Delta_s}{\beta} \right] + \cos \gamma \Delta_c^2, \end{aligned}$$

which shows $|\partial g/\partial \gamma| \leq 6(\partial G/\partial \gamma)$, as claimed. \square

This settles Lemma 2.1, which was about the derivatives of the positive harmonic function u_+ with boundary values $f \cdot \chi_I$. Set $u_- = u_+ - u \geq 0$, which has boundary values $-f \cdot \chi_{I^c} \geq 0$. Set $v(x, y) = u_-(-x, y)$. Its boundary values are simply a reflection of those of u_- , non-negative on (a_v, b_v) , where $\sin(a_v) = \sin(a)$ and $\cos(a_v) = -\cos(a)$, with similar relations for b_v and b . We can apply Lemma 2.1 to v instead of u_+ , with no change to the values of S or $\alpha(b, a)$ or g_0 . However, we must replace G_0 by $-G_0$, c by $c - d$, and C by $D - C$. Lemmas 2.3 and 2.4 still hold. This gives

$$|(c - d) - g_0| \leq 6(D - C + G_0). \quad (2.9)$$

The triangle inequality, with Lemma 2.1 and (2.9), shows $d \leq |c - g_0| + |(c - d) - g_0| \leq 6D$, which proves the theorem. \square

LEMMA 2.3. *With S , a and b as in Theorem 1.1, we have $\sin(b) > S > \sin(a)$.*

Proof. Let m be the midpoint of (a, b) and let \mathbf{w}_m be the unit vector $(\cos(m), \sin(m)) \in \mathbf{R}^2$. For every $\theta \in (a, b)$, we have $\mathbf{w}_\theta \cdot \mathbf{w}_m > \mathbf{w}_b \cdot \mathbf{w}_m = \cos(m - b) = \nu$. Recall that $u = u_+ - u_-$ is the difference of two positive harmonic functions, which satisfy formulas similar to the ones for u . Thus, $\pi \nabla u_+(0, 0) = (C, S)$ is an average of vectors of the form \mathbf{w}_θ , and hence $\pi \nabla u_+(0, 0) \cdot \mathbf{w}_m > \nu$. Likewise, $\pi \nabla u_-(0, 0) \cdot \mathbf{w}_m < \nu$. Thus $0 < \pi \nabla u(0, 0) \cdot \mathbf{w}_m = D \cos(m)$, and $\cos(m) > 0$. Hence, $\sin(b) - \sin(a) = \sin(m + (b - m)) - \sin(m - (b - m)) = 2 \cos(m) \sin(b - m) > 0$, since $0 < b - m < \pi$. Thus, $\sin(b) > \sin(a)$.

Recall that S is a weighted average of $\sin(\theta)$ over $a < \theta < b$, and hence it is less than the maximum value of this function. The maximum is 1 if and only if $\pi/2 \in (a, b)$; otherwise it is $\sin(b)$. Since $u_y(0, 0) = 0$, it follows that S is an average of $\sin(\theta)$ over $(b, a + 2\pi)$ as well, with

similar conclusions about the max. However, $\pi/2$ can not belong to both intervals, and thus in one case the maximum is $\sin(b)$. Hence, $S < \sin(b)$. By similar reasoning, $S > \sin(a)$. \square

LEMMA 2.4. *For any two angles $\gamma \neq \theta$, we have $\Delta_c^2 < 4\mathbf{w}_\theta \cdot (\mathbf{w}_\theta - \mathbf{w}_\gamma)$.*

Proof. Letting $\mathbf{v} = \mathbf{w}_\theta - \mathbf{w}_\gamma = (\Delta_c, \Delta_s)$, we see $\Delta_c^2 \leq \|\mathbf{v}\|^2$. Now, suppose that β is the angle between \mathbf{w}_θ and \mathbf{v} , and that $\nu \in (0, \pi)$ is the angle between \mathbf{w}_θ and \mathbf{w}_γ . Thus, $2\beta + \nu = \pi$. Also, $\|\mathbf{v}\| < \nu < 4\sin(\nu/2)$ (since $\nu < \pi$). Furthermore, $\mathbf{w}_\theta \cdot (\mathbf{w}_\theta - \mathbf{w}_\gamma) = \|\mathbf{v}\| \cos(\beta) = \|\mathbf{v}\| \sin(\nu/2) > \|\mathbf{v}\|^2/4$, as desired. \square

2.3. Corollaries of Theorem 1.1

The previous theorem controls curvature only at the center of the ball. Next, we shall study the curvature of Z at a general point $a \in Z \cap B_1(0)$ via conformal mapping; but the bound must grow as $|a| \rightarrow 1$. To see this, recall the example u given at the beginning of Subsection 2.2. There, Z is a hyperbola centered at (ϵ, ϵ) with very large curvature, $\kappa(u)(p) = 1/\epsilon\sqrt{2}$ at the origin. We can translate this function so that the new hyperbola is centered on the circle, at $(\sqrt{1/2}, \sqrt{1/2})$. This new Z crosses the circle only twice, thus satisfying the assumptions of the theorem. It has the same large curvature, but at a new point a , with $1 - |a| = \epsilon/\sqrt{2} = 1/2\kappa$. The following corollary gives the best possible bound on the curvature at a general point a in the ball, in terms of the best constant C in Theorem 1.1 (but C is not currently known).

COROLLARY 2.5. *Let u be a harmonic function in the unit disk. Suppose that Z is a single connected curve, as in Theorem 1.1. Then, for every $a \in Z$, the curvature $\kappa(u)(a)$ of Z at a satisfies the following inequality:*

$$|\kappa(u)(a)| \leq \frac{C + 2|a|}{1 - |a|^2} \leq \frac{C_2}{1 - |a|}, \quad (2.10)$$

where C is the constant in Theorem 1.1 and C_2 is some constant independent of both a and u .

Proof. Let $a \in Z$ be such that $u(a) = 0$. Assume that $u = \operatorname{Re}(w)$, where w is a holomorphic function in the unit disk. Consider the conformal transformation $g_a(z) = (z + a)/(1 - \bar{a}z)$, with $z = x + iy$. This transformation maps the unit disk and the unit circle onto themselves. The function $w_a(x, y) = w \circ g_a(x, y)$ is holomorphic and its real part, $u_a(x, y) = u \circ g_a(x, y)$, is harmonic. By the formula (2.3), the curvature of the zero set of u_a at the origin is given by

$$\begin{aligned} \kappa(u_a)(0) &= |w'_a(0)| \operatorname{Re} \left(\frac{w''_a(0)}{(w'_a(0))^2} \right) \\ &= |w'(a)|(1 - |a|^2) \operatorname{Re} \left(\frac{(g'_a(0))^2 w''(a) + w'(a) g''_a(0)}{(g'_a(0) w'(a))^2} \right) \\ &= |w'(a)|(1 - |a|^2) \left(\operatorname{Re} \left(\frac{w''(a)}{(w'(a))^2} \right) + \operatorname{Re} \left(\frac{g''_a(0)}{w'(a) (g'_a(0))^2} \right) \right) \\ &= (1 - |a|^2) \kappa(u)(a) - 2|w'(a)| \operatorname{Re} \left(\frac{\bar{a}}{w'(a)} \right) \\ &= (1 - |a|^2) \kappa(u)(a) - 2 \operatorname{Re} \left(\frac{\bar{a} |w'(a)|}{w'(a)} \right). \end{aligned}$$

We have used the fact that

$$g'_a(z) = \frac{1 - |a|^2}{(1 + z\bar{a})^2} \quad \text{and} \quad g''_a(z) = \frac{-2\bar{a}(1 - |a|^2)}{(1 + z\bar{a})^3}.$$

Since $|\operatorname{Re}(\bar{a}|w'(a)|/w'(a))| < |a|$, we have the following inequalities:

$$\kappa(u)(a) = (1 - |a|^2)^{-1} \left(\kappa(u_a)(0) + 2\operatorname{Re} \left(\frac{\bar{a}|w'(a)|}{w'(a)} \right) \right) \leq (1 - |a|^2)^{-1} (\kappa(u_a)(0) + 2|a|),$$

and

$$\kappa(u)(a) \geq (1 - |a|^2)^{-1} (\kappa(u_a)(0) - 2|a|).$$

Since by Theorem 1.1, $|\kappa(u_a)(0)| < C$, we have proved (2.10). \square

3. Some special classes of level curves

We now consider some conditions on Z , which imply that u must be linear. This section also includes results about polynomials, and on the linearity of Z . Recall that Theorem 1.1 controls the curvature of Z inside a ball in the plane, if Z cuts the boundary only twice. That proof does not seem to extend directly to the entire plane, but a slightly stronger result can be proved in that case using complex analytic methods.

THEOREM 3.1. *Let u be harmonic on the plane, and assume that $U_+ = \{z : u(z) > 0\}$ and $U_- = \{z : u(z) < 0\}$ are both connected. Then u is linear.*

The theorem implies that every analytic bijection of the complex plane must be linear (this is also part of the proof).

Proof of Theorem 3.1. The authors' original proof, using mostly real analysis, resembles that of Theorem 5 in [9]. We present here an elegant complex analytic version, found independently by Bao Qin Li (Personal communication). Let $f = v + iu$ be analytic on the plane. All the zeroes of f lie in Z . Suppose that n zeroes lie inside a simple closed curve Γ , which intersects Z just twice. Thus, $f(\Gamma)$ winds around the origin at most once. By the argument principle, $n \leq 1$. Thus, f has at most one zero in the plane. The same is true for $f - a$, for any $a \in \mathbf{R}$. Hence, f must be a polynomial (otherwise, $f(z) = a$ would occur infinitely often, for most values of a). A polynomial with only one zero (counting multiplicity) must be linear. \square

The following two theorems generalize Theorem 3.1 and are roughly equivalent to each other.

THEOREM 3.2. *If U_+ and U_- have a total of n components, then u is a polynomial of degree m , with $n/2 \leq m \leq n - 1$.*

THEOREM 3.3. *If Z is a finite union of m maximal analytic curves, then u is a polynomial of degree m .*

Proof. Theorem 5 of [9] shows that u is a polynomial of degree $d < \infty$, but it does not bound d ; nor does the proof of Theorem 3.2, without additional work. Nevertheless, outside some large ball $B_R(0, 0)$, the highest order term of u dominates, and it behaves like $\operatorname{Re}(z^d)$. Thus, outside B_R , Z consists of exactly $2d$ curves extending to infinity. However, [9] shows that each maximal analytic curve in Z extends to infinity exactly twice. So, $2m = 2d$ and $d = m$. \square

Proof of Theorem 3.2. Euler's theorem for planar graphs states that in a planar graph the numbers of vertices v , faces f and edges e satisfy $v - e + f = 2$. This implies that $n =$

$m + v + 1$, where m is the number of intersections of the analytic curves, by multiplicity. The intersections can occur only at the roots of $f'(z)$ on Z ; thus $0 \leq v \leq m - 1$ and the theorem follows. \square

It is interesting to speculate whether this kind of result generalizes to other domains. Suppose that u is harmonic on the sphere S_2 (the boundary of $B_1(0) \subset \mathbf{R}^3$), except possibly at the point $(0,0,1)$. Assume that U_+ and U_- have only one component each. Then we can map the sphere to the plane conformally (for example by the stereographic projection $f(x, y, z) = (x/(1-z), y/(1-z))$, so that $(0,0,1)$ maps to infinity). Then $u \circ f^{-1}$ is harmonic on the plane, and $f(Z)$ splits the plane in two. Thus, we can conclude that $f(Z)$ is a straight line, and Z is a circle in S_2 through $(0,0,1)$ (and so is every other level curve of u).

There seems to be no obvious generalization of these three theorems if u is harmonic on \mathbf{R}^3 instead of \mathbf{R}^2 . Here is a simple counterexample to a generalized Theorem 3.1. Let $u(x, y, z) = xy + z$, which is harmonic, but not linear, on \mathbf{R}^3 . We claim that every point in U_+ is connected to $P = (0, 0, 1)$ by a path in U_+ , and hence this set is connected. To see this, let $U_a = U_+ \cap \{z = a\}$. If $a > 0$, then U_a is connected, and the union of these is also connected (it contains the positive z -axis). If $a \leq 0$, then each U_a has two components. One contains the point $(1 - a, 1 - a, a)$ (since $(1 - a)^2 + a > 0$ for $a \leq 1$) and the other contains $(a - 1, a - 1, a)$. The union of these points form two lines, contained in U_+ (at least while $z \leq 1$), which intersect at P . Thus, U_+ is path-connected via $(0,0,1)$; we have that U_- is the image of U_+ under the rotation $(x, y, z) \rightarrow (-x, y, -z)$, and hence it is also connected. Essentially the same example works in \mathbf{R}^4 , etc.

THEOREM 3.4. *Suppose that L is an analytic curve parametrized by $\phi : \mathbf{R}^1 \rightarrow \mathbf{C}$, that v is harmonic on some neighborhood of the curve, and that $v(\phi(t)) = 0$ on some interval, $a < t < b$. Then $v = 0$ on L .*

This result may follow from the theory of maximal analytic curves in [9], but we provide an independent proof. It is enough to show that $v(\phi(t))$ is real analytic at an arbitrary point, such as $t = 0$. We may assume that $\phi(0) = (0, 0)$. We can express $\phi(t)$ as a complex-valued series that converges uniformly for $|t| < \epsilon$ as follows:

$$\phi(t) = \sum_{j=1}^{\infty} (b_j + ic_j)t^j.$$

Now replace t by a complex variable z . The radius of convergence of the series is at least ϵ , and hence this formula extends ϕ to an analytic function Φ defined on a small ball centered at the origin. Thus, $v \circ \Phi$ is harmonic, and therefore real analytic, with a power series expansion in s and t . When restricted to $s = 0$, it is still a real analytic function of t , namely $v \circ \phi$.

This has some simple corollaries, probably not new. The first could also be proved from the Schwartz reflection principle. The second is similar to a UCP result. It shows that a nontrivial harmonic function on the plane cannot vanish on an arc of a circle. It follows from the theorem and the maximum principle.

COROLLARY 3.5. *Suppose that $\Delta v = 0$ on the plane \mathbf{C} and $v = 0$ on some line segment $[a, a + \epsilon] \subset \mathbf{R}^1 \subset \mathbf{C} = \mathbf{R}^2$. Then $v = 0$ on \mathbf{R}^1 .*

COROLLARY 3.6. *Suppose that $\Delta v = 0$ on $\mathbf{C} = \mathbf{R}^2$ and $v = 0$ on the semicircle $S = \{z \in \mathbf{C} : |\mathbf{2z} - \mathbf{i}| = 1 \text{ and } \text{Im}(\mathbf{z}) \geq 1/2\}$. Then $v = 0$ on \mathbf{C} .*

Of course, S could be replaced by any arc of the circle, with the same result. However, if v fails to be harmonic at some point on the circle, the result does not hold (consider the Poisson kernel for the disk). If v is harmonic only on a neighborhood of the circle, then v must vanish on the circle, but not necessarily at other points. For example, $v = \ln(|2z - i|)$ vanishes on the circle, but it has a pole at the center. Likewise, 3.5 implies that $u = 0$ on R^1 but not necessarily on R^2 (consider $u(x, y) = xy$). We mention some related open questions in Section 5.

4. Unique continuation

In this section we investigate the properties of the solutions[†] of the differential inequality as follows:

$$\Delta u(x) \leq V(x)|u(x)|, \quad x \in D \subset \mathbf{R}^n, \quad (4.1)$$

here $V(x) \in L^1_{\text{loc}}(D)$. We assume that u vanishes on the boundary of D . Of course harmonic functions satisfy (4.1), and if they vanish on the boundary of D , then they vanish everywhere. This property is not generally true for nonharmonic solutions, but we can prove it under suitable assumptions on the size of V and $|D|$. Major tools in our proofs are sharp Sobolev inequalities. The following standard Sobolev inequality on a domain $D \subset \mathbf{R}^n$:

$$\|u\|_{np/(n-p)} \leq c(p, D) \|\nabla u\|_p \quad u \in W_0^{1,p}(D) \quad (4.2)$$

is valid for every $p < n$. It is easy to prove that this inequality is translation and scaling invariant. When $p = 1$ and D is a ball in \mathbf{R}^n , the best possible constant $c(p, D)$ is achieved by the characteristic function of the ball. It is given by

$$c(1, B_1(0)) = (n|\omega_n|^{1/n})^{-1}, \quad (4.3)$$

where ω_n is the volume of the unit ball in \mathbf{R}^n . In particular, when $n = 2$, we have $c(1, B_1(0)) = (2\pi^{1/2})^{-1}$. Using rearrangements, it is possible to prove that $c(p, D) \leq c(p, B_1(0))$ for every $D \subset \mathbf{R}^n$ such that $|D| = |B_1(0)|$, and for all $1 \leq p < n$. When $p \neq 1$, we find that

$$c(p, B_1(0)) = p \frac{p-1}{n-p}. \quad (4.4)$$

For the proof of these inequalities see, for example, [7].

Since (4.2) is scaling-invariant, it follows that $c(p, \mathbf{R}^n)$ is still the best constant when the functions are restricted to have their support in a fixed open set. We use the inequalities (4.2) and (4.4) to estimate the size of the support of the solutions of the differential inequality (4.1).

THEOREM 4.1. *Let $u(x) \in W_0^{2,p}(D)$, with D a domain of \mathbf{R}^n , be a nontrivial solution of the differential inequality (4.1). Assume that $u(x) > 0$ in D and $u \equiv 0$ on the boundary of D . Then we have the following.*

(i) *If $n = 2$ and $|V(x)|$ is bounded in D , then*

$$|D| \geq 4\pi \|V\|_{\infty}^{-1}. \quad (4.5)$$

(ii) *If $n > 2$ and $V \in L^r(D)$ with $r > n/2$, then*

$$|D| \geq \left(\frac{n-2}{2\|V\|_r^{1/2}} \right)^{2nr/(2r-n)}. \quad (4.6)$$

[†]By solution we mean: $u \in W^{2,p}(D)$ that satisfy (4.1) almost everywhere.

(iii) If $V \in L^{n/2}(D)$, with $n > 2$, then necessarily

$$\|V\|_{n/2} \geq \frac{(n-2)^2}{4}. \quad (4.7)$$

REMARK. We noted in Section 1 that the differential inequality (4.1) has the SUCP (unique continuation from a point) and hence also the UCP (unique continuation from an open set) whenever $V \in L_{\text{loc}}^{n/2}(D)$. Theorem 4.1 can be interpreted as a unique continuation theorem from the boundary of the domain of u . That is, if u vanishes on the boundary of its domain of definition and the $L^{n/2}$ norm of V is sufficiently small, then $u \equiv 0$.

Proof of Theorem 4.1. (i) ($n = 2$). By the standard Sobolev inequality (4.3), with $p = 1$ and $n = 2$, Hölder's inequality, (4.1), and the identity

$$\int_D |\nabla u(x)|^2 dx = \int_D u(x) \Delta u(x) dx, \quad u(x) \in W_0^{2,p}(D),$$

we get the following chain of inequalities:

$$\begin{aligned} \|u\|_2 &\leq (2\pi^{1/2})^{-1} \|\nabla u\|_1 \\ &\leq (2\pi^{1/2})^{-1} |D|^{1/2} \|\nabla u\|_2 \\ &\leq (2\pi^{1/2})^{-1} |D|^{1/2} \left(\int_D u(x) \Delta u(x) dx \right)^{1/2} \\ &\leq (2\pi^{1/2})^{-1} |D|^{1/2} \left(\int_D u^2(x) V(x) dx \right)^{1/2} \\ &\leq (2\pi^{1/2})^{-1} \|V\|_{\infty}^{1/2} |D|^{1/2} \|u\|_2. \end{aligned}$$

Since $u \not\equiv 0$, it follows that $(2\pi^{1/2})^{-1} \|V\|_{\infty}^{1/2} |D|^{1/2} \geq 1$, which implies that $|D| \|V\|_{\infty} \geq 4\pi$.

(ii) and (iii), ($n > 2$). Assume that $V \in L^{n/2}(D)$. We apply the sharp Sobolev inequality (4.4) with $p = 2$ and the tools that we have used in the first part of the proof. We obtain

$$\begin{aligned} \|u\|_{2n/(n-2)} &\leq \frac{2}{n-2} \|\nabla u\|_2 \\ &= \frac{2}{n-2} \left(\int_D u(x) \Delta u(x) dx \right)^{1/2} \\ &\leq \frac{2}{n-2} \left(\int_D u^2(x) V(x) dx \right)^{1/2} \\ &\leq \frac{2}{n-2} \|u\|_{\frac{2n}{n-2}} \|V^{1/2}\|_n \\ &= \frac{2}{n-2} \|u\|_{2n/(n-2)} \|V\|_{n/2}^{1/2} \end{aligned}$$

In the step before the last we have applied Hölder's inequality with $1/2 = (n-2)/2n + 1/n$. To summarize, we have

$$\|u\|_{2n/(n-2)} \leq 2/(n-2) \|u\|_{2n/(n-2)} \|V\|_{n/2}^{1/2}. \quad (4.8)$$

This implies (iii). For (ii), if $V(x) \in L^r(D)$ for some $r > n/2$, then Hölder's inequality (with $2/n = 1/r + (2/n - 1/r)$), and (4.8) yield the following inequality, which implies (ii):

$$\|u\|_{2n/(n-2)} \leq \frac{2}{n-2} \|u\|_{2n/(n-2)} |D|^{1/n-1/2r} \|V\|_r^{1/2}. \quad \square$$

The authors originally proved Theorem 4.1 in a less direct manner, but obtained exactly the same lower bound for $\|V\|_{\infty}$. We do not know if it is sharp, but the following example

shows it is not extremely far from that. If $D = B_1(0)$, then (4.5) implies that $\|V\|_\infty \geq 4$. Let $u_t(x) = r^2 e^{-tr^2}$, with $t = \frac{1}{2}(3 - \sqrt{5})$ and $r = \sqrt{x^2 + y^2}$. Then $u_t(1) - u_t(r)$ vanishes on the unit circle.

In polar coordinates, its Laplacian is $-(\partial_{rr}^2 u_t(r) + \frac{1}{r} \partial_r u_t(r))$, or

$$4e^{-r^2 t}(t(3 - r^2 t)r^2 - 1).$$

We chose $t = \frac{1}{2}(3 - \sqrt{5})$ so that this Laplacian vanishes for $r = 1$, and

$$V(r) = -\frac{\Delta u_t(r)}{u_t(1) - u_t(r)} = \frac{4e^t(t(3 - r^2 t)r^2 - 1)}{e^t r^2 - e^{r^2 t}}$$

is bounded. It is not too difficult to see that V is decreasing in $[0, 1]$. Hence, the maximum of V is $V(0) = 4e^{(1/2)(3-\sqrt{5})} \sim 5.86$.

The argument used to prove (iii) of Theorem 4.1 can be generalized to prove the following unique continuation theorem.

PROPOSITION 4.2. *Let $c = c(p, q, D)$ denote the best constant in the following Sobolev inequality:*

$$\|u\|_{L^q(D)} \leq c(p, q, D) \|\Delta u\|_{L^p(D)}, \quad (4.9)$$

where $u \in W_0^{2,p}(D)$ and $1/q \geq 1/p - 2/n$. Then the differential inequality $|\Delta u| \leq V(x)|u(x)|$ has only the trivial solution $u \equiv 0$ if

$$c(p, q, D) \|V\|_{L^{n/2}(D)} < 1.$$

Proof. By (4.9) and Hölder's inequality, we have

$$\|u\|_q \leq c(p, q, D) \|\Delta u\|_p \leq c(p, q, D) \|V\|_{L^{n/2}} \|u\|_q$$

and if $c(p, q, D) \|V\|_{n/2} < 1$, then $u \equiv 0$. □

5. Open problems and conjectures

Our first conjecture is somewhat related to Theorem 4.1(i), and the isoperimetric inequality, about the possible area inside a curve. Recall that if $\Delta u = 0$, then Z cannot contain a closed loop. Can Z contain a C-shaped curve that is 'almost' closed?

CONJECTURE 5.1 (Isoperimetric conjecture). Suppose that $\Delta u = 0$ on \mathbf{C} and that $u > 0$ on $B_1(0)$. Suppose that L is a line segment in the plane of length ϵ such that $L \cup Z$ encloses $B_1(0)$ (for example it does not intersect $B_1(0)$, but it disconnects the origin from infinity). Then $\epsilon > c > 0$ for some absolute constant c .

Unfortunately, we have heuristic evidence that this is false, as stated. Consider the Poisson kernel $P(z)$ for $B_1(0)$, which is positive on the open unit ball, and vanishes on the boundary, except at $z = 1$. It does not directly disprove the conjecture, since it is not harmonic at $z = 1$. However, if u is a Taylor polynomial of P (roughly $u = \operatorname{Re}(1 + z + z^2 + \dots + z^n)$), then u is harmonic on the plane. Computer experiments indicate that u has a level curve Z that almost encloses $B_{1/4}(0)$, with a small gap near $z = 1$, with length $\epsilon \rightarrow 0$, as $n \rightarrow \infty$. Thus, the conjecture seems to fail, but it still seems plausible for harmonic polynomials of fixed degree.

CONJECTURE 5.2. Suppose that in Corollary 3.5, the hypothesis $\Delta v = 0$ is weakened to $V \in L^\infty$ (see (1.1)). Is it still true that v must vanish on R^1 ?

This conjecture seems plausible at first glance, since Corollary 3.5 is local in nature, and $V \in L^\infty$ implies that $\Delta v = 0$ on Z . At this writing, the authors expect a negative answer; but our work is not quite conclusive. The next conjecture, which aims to extend Theorem 3.4 to higher dimensions, seems more likely.

CONJECTURE 5.3. Suppose that $\Delta v = 0$ on R^n and $v = 0$ on $B_1^{n-1}(0) \times \{0\} \subset R^n$. Then $v = 0$ on $R^{n-1} \times \{0\}$.

The thrust of this paper is that Z cannot wiggle arbitrarily. We have expressed this in terms of curvature and area, and would like some result about the local arc length of Z . However, it seems very difficult to say much. Examples such as $u = \operatorname{Re}(z^k + \epsilon)$ produce a set Z with very large total arc length inside the unit ball. Examples such as spirals show that curvature does not directly control local arc length.

Nevertheless, in the spirit of unique continuation, we conjecture that a large concentration of arc length cannot occur in only one small region of the plane.

CONJECTURE 5.4 (Arc length conjecture). Suppose that u is harmonic on the unit ball B and that Z is a finite union of rectifiable curves there, of total length $|Z|$. Then

$$|Z \cap B_{1/2}(0)| < 0.99|Z|.$$

Acknowledgements. The authors wish to thank Bao Qin Li for several insightful conversations.

References

1. L. DE CARLI and S. HUDSON, ‘Unique continuation for nonnegative solutions of Schrödinger type inequalities’, *J. Math. Anal. Appl.* 318 (2006) 467–471.
2. L. FLATTO, ‘A theorem on level curves of harmonic functions’, *J. London Math. Soc.* (2) 1 (1969) 470–472.
3. L. FLATTO, D. J. NEWMAN, and H. S. SHAPIRO, ‘The level curves of harmonic functions’, *Trans. Amer. Math. Soc.* 123 (1966) 425–436.
4. D. JERISON and C. E. KENIG, ‘Unique continuation and absence of positive eigenvalues for Schrödinger operators’, *Ann. of Math.* (2) 121 (1985) 463–494, With an appendix by E. M. Stein.
5. R. P. JERRARD and L. A. RUBEL, ‘On the curvature of the level lines of a harmonic function’, *Proc. Amer. Math. Soc.* 14 (1963) 29–32.
6. H. KOCH and D. TATARU, ‘Carleman estimates and unique continuation for second order elliptic equations with nonsmooth coefficients’, *Comm. Pure Appl. Math.* 54 (2001) 339–360.
7. E. H. LIEB and M. LOSS, *Analysis*, 2nd edn, Graduate Studies in Mathematics 14 (American Mathematical Society, Providence, RI, 2001) ISBN: 0-8218-2783-9.
8. J. L. WALSH, ‘Note on the curvature of level curves of Green’s function’, *Proc. Nat. Acad. Sci. USA* 23 (1937) 84–89.
9. Z. Y. WEN, L. M. WU, and Y. ZHANG, ‘Set of zeros of harmonic functions of two variables’, *Harmonic analysis*, Tianjin, 1988, Lecture Notes in Mathematics 1494 (Springer, Berlin, 1991) 196–203.
10. T. H. WOLFF, ‘Recent work on sharp estimates in second order elliptic unique continuation problems’, *Fourier analysis and partial differential equations*, Miraflores de la Sierra, 1992, Studies of Advanced Mathematics (CRC, Boca Raton, FL, 1995) 99–128.

L. De Carli and S. M. Hudson
Department of Mathematics
Florida International University
University Park
Miami, FL 33199
USA

decarlil@fiu.edu
 hudsons@fiu.edu