Skirting subsets of the plane, with application to marginal stability curves

G. D. McBain*

Abstract

An easily implemented algorithm is described for tracing the margin of a plane region defined by a predicate. Given a point inside and one outside, a sequence of marginal points is produced. The algorithm is a modified specialization of the 'simplicial decomposition' method for n equations in $n+1$ dimensions. The case $n = 1$ has special properties and its importance motivates their present exploitation. It is directly applicable to finding level curves. It does not require differentiability and copes well with cusps. Two questions of accuracy are the proximity of the outputs to the margin and the proximity of the margin to the output set. The first is answered precisely. The second is complicated and predicate-dependent, but is addressed in practical terms by adaptivity, which also improves the scheme's efficiency.

*School of Aerospace, Mechanical & Mechatronic Engineering, The University of Sydney, AUSTRALIA. mailto:geordie.mcbain@aeromech.usyd.edu.au

Contents

1 Introduction

Consider the problem of determining the marginal set of a predicate on a plane convex region; that is, the set of points for which every neighbourhood includes a point for which the predicate is true and one for which it is false.

An important general example is the curve $f(x, y) = 0$, which can be characterized as the marginal set of the predicate $P(x, y) \equiv f(x, y) < 0$.[1]

---

[1]Logical notation: $\equiv$ equivalence; $\neg$ negation.

A specific practical example arises in connection with the Orr–Sommerfeld equation of linear hydrodynamic stability theory [3, p. 157]. This governs the stability of a parallel flow with Reynolds number R with respect to infinitesimal disturbances of wavenumber α. For each pair (R, α), if the Orr–Sommerfeld spectrum is not confined to the lower half complex plane the flow is unstable. Thus, denoting the eigenvalues by c, a stability predicate might be $P(R, \alpha) := \max \Im c < 0$. More details can be found in the accompanying paper [7].

This note describes an algorithm for approximating a margin by a sequence of points; no knowledge of the predicate (for example, the smoothness of an underlying function f) is required, just the ability to evaluate it.

The nonadaptive version of the algorithm described in §2 is essentially a special case with $n = 1$ of the simplicial decomposition method for finding the curve $f(\mathbf{x}) = \mathbf{0}$ for a function $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ described by Allgower & Georg [2, p. 161]. This special case received a similar treatment recently from Mezher & Philippe [8]; the present work goes further by incorporating adaptivity (§3).

Section 4 presents results of the adaptive algorithm for a simple predicate with a cusped margin, and illustrates the huge efficiency gain due to adaptivity. Section 5 compares the algorithm with alternatives, particularly those previously employed for Orr–Sommerfeld stability margins.

2 Simple skirting algorithm

2.1 Preliminaries

2.1.1 The rectangular metric

Although the algorithm is geometric in character, the planes in the envisaged applications are parameter planes rather than physical planes so that a rectangular metric is appropriate. To this end, define the inner product as $\langle \mathbf{a}, \mathbf{b} \rangle := \mathbf{b}^T \mathsf{G} \mathbf{a}$ where $\mathsf{G} := \operatorname{diag} \tau^{-2}$ and the norm as

$$\|\mathbf{a}\| := \langle \mathbf{a}, \mathbf{a} \rangle^{1/2} = \left(\mathbf{a}^T \mathsf{G} \mathbf{a} \right)^{1/2} = \left\{ \left(\frac{a_1}{\tau_1} \right)^2 + \left(\frac{a_2}{\tau_2} \right)^2 \right\}^{-1/2}, \qquad (1)$$

so that a unit disk $\|\mathbf{a} - \mathbf{a}_0\| < 1$ contains the points considered equal to $\mathbf{a}_0$ to within the specified tolerances. For example, in the stability margins in the accompanying paper [7], the horizontal coordinate runs over $[0, 32\,000]$ while the vertical range is $[0, 5]$. Thus we might use tolerance components of $\tau_1 = 1$ and $\tau_2 = 10^{-4}$.

Two operators required later are the matrix to rotate through θ radians

$$\mathsf{R}_{\mathsf{G},\theta} = \mathsf{G}^{-1/2} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \mathsf{G}^{1/2} = \begin{bmatrix} \cos\theta & -\frac{\tau_1}{\tau_2}\sin\theta \\ \frac{\tau_2}{\tau_1}\sin\theta & \cos\theta \end{bmatrix}, \qquad (2)$$

and the matrix to project on to the span of a vector $\mathbf{a}$

$$\operatorname{proj}_{\mathbf{a}} = \frac{\mathbf{a}\mathbf{a}^T \mathsf{G}}{\mathbf{a}^T \mathsf{G} \mathbf{a}}. \qquad (3)$$

2.1.2   Bisection

Given a point each from the truth set and its complement; that is, a *transverse* pair; a marginal point can be located to arbitrary accuracy, as shown in the constructive proof of the following existence

Theorem 1 *If $\mathcal{D}$ is convex and contains a pair $\mathbf{a}$ and $\mathbf{b}$ traverse withe respect to a predicate P, then the segment $[\mathbf{a}, \mathbf{b}] = \{(1 - k)\mathbf{a} + k\mathbf{b} : 0 \leq k \leq 1\}$ contains a marginal point of P.*

Proof: Set $\mathbf{a_0} := \mathbf{a}$ and $\mathbf{b_0} := \mathbf{b}$, then for $n = 0, 1, 2, \ldots$ define $\mathbf{m}_n := (\mathbf{a}_n + \mathbf{b}_n)/2$, which is contained in $\mathcal{D}$ by convexity. Define $[\mathbf{a}_{n+1}, \mathbf{b}_{n+1}]$ as $[\mathbf{m}_n, \mathbf{b}_n]$ if $P(\mathbf{m}_n)$ and as $[\mathbf{a}_n, \mathbf{m}_n]$ otherwise. Then $P(\mathbf{a}_n) \equiv P(\mathbf{a}) \equiv \neg P(\mathbf{b}_n) \equiv \neg P(\mathbf{b})$ for all n, and $\lim_{n\to\infty} \mathbf{m}_n = \lim_{n\to\infty} \mathbf{a}_n = \lim_{n\to\infty} \mathbf{b}_n$ is a point in $\mathcal{D}$ and all $[\mathbf{a}_n, \mathbf{b}_n]$. Any neighbourhood of $\mathbf{m}_\infty$ contains an infinite number of terms from the sequences $\{\mathbf{a}_n\}$ and $\{\mathbf{b}_n\}$ and thus at least one true and one false point. Therefore $\mathbf{m}_\infty$ is marginal point for P.   ♠

This is, of course, an adaptation of the bisection method for finding zeros of continuous univariate functions. The procedure is described in Algorithm 1 in terms of the stepping procedure in Algorithm 2 which is also required later

---

Algorithm 1: *Bisect*($\mathbf{a} \in \mathcal{D}, \mathbf{b} \in \mathcal{D}, P : \mathcal{D} \to \{T, F\}, \tau \in \mathbb{R}^2$)

Require: $P(\mathbf{a}) \equiv \neg P(\mathbf{b})$
Ensure: *Bisect* is within tolerance of a marginal point of P
 1: $\mathbf{u} := \mathbf{a}$, $\mathbf{v} := \mathbf{b}$; {# Preserve $\mathbf{a}$ and $\mathbf{b}$ from modification. #}
 2: **do**
 3:    $Bisect := Step(\mathbf{u}, \mathbf{v}, P, \frac{1}{2})$
 4: **until** $\|\mathbf{v} - \mathbf{u}\| < \frac{1}{2}$

---

for the skirting algorithm.

Algorithm 2: $Step(\mathbf{a} \in \mathcal{D}, \mathbf{b} \in \mathcal{D}, P : \mathcal{D} \to \{T, F\}, \mathsf{L} : \mathbb{R}^2 \to \mathbb{R}^2)$

Require: $P(\mathbf{a}) \equiv \neg P(\mathbf{b})$
Ensure: $P(\mathbf{a}) \equiv \neg P(\mathbf{b})$ and one of $\mathbf{a}$ and $\mathbf{b}$ has been changed.
1: $Step := \mathbf{a} + \mathsf{L}(\mathbf{b} - \mathbf{a})$;
2: **if** $P(Step) \equiv P(\mathbf{a})$ **then**
3:     $\mathbf{a} := Step$ **else** $\mathbf{b} := Step$ **fi**

2.2   Accessing the second dimension

Given a transverse pair $\mathbf{a}, \mathbf{b}$, another can be found by taking any point $\mathbf{c} \in D$ not collinear with $\mathbf{a}$ and $\mathbf{b}$. Then if $P(\mathbf{c}) \equiv P(\mathbf{a})$, $\mathbf{c}$ and $\mathbf{b}$ are transverse, otherwise $\mathbf{a}$ and $\mathbf{c}$ are. One choice for $\mathbf{c}$ is that point forming an anticlockwise equilateral triangle with $\mathbf{a}$ and $\mathbf{b}$ [8]; using (2)

$$\mathbf{c} := \mathbf{a} + \mathsf{R}_{\mathsf{G}, \pi/3}(\mathbf{b} - \mathbf{a}) \ . \tag{4}$$

The state is now equivalent to the initial condition, so stepping can repeat indefinitely, as in Algorithm 3. Figure 1 illustrates Algorithm 3 for the pred-

Algorithm 3: $Skirt(\mathbf{a}, \mathbf{b} \in \mathcal{D}, P : \mathcal{D} \to \{T, F\}, \tau : \mathbb{R}^2)$

Require: $P(\mathbf{a}) \equiv \neg P(\mathbf{b})$
Ensure: An infinite sequence of points within tolerance of P's margin.
1: **do** {# until interrupted #}
2:    $Step(\mathbf{a}, \mathbf{b}, P, \mathsf{R}_{\mathsf{G}, \pi/3})$;
3:    $Print(Bisect(\mathbf{a}, \mathbf{b}, P, \tau))$
4: **od**

icate

$$P(x, y) := y < \sqrt{|x| + x^2} \ . \tag{5}$$

for which the true margin has a cusp at the origin but away from that approaches $y = |x|$. Such a curve would trouble or foil a predictor–corrector continuation method [2, pp. 13–15] since the tangent flips at $(0, 0)$.
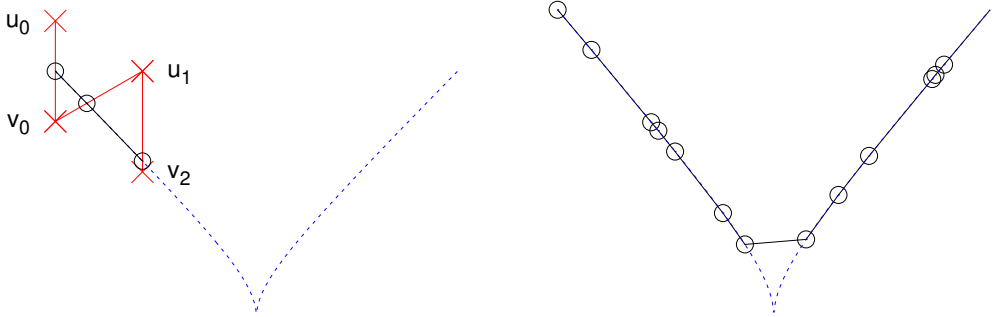
FIGURE 1: Simple skirting (Algorithm 3) on predicate (5); true margin dashed and outputs marked '∘'. *Left*: The first three transverse segments are $[\mathbf{u}_0, \mathbf{v}_0]$, $[\mathbf{u}_1, \mathbf{v}_1 = \mathbf{v}_0]$, and $[\mathbf{u}_2 = \mathbf{u}_1, \mathbf{v}_2]$. *Right*: first 15 steps.

This method of selecting the new point differs from the Coxeter flipping scheme proposed by [2, p. 154] although they both lead to affine transformations of Freundenthal's triangulation of the plane. The present method leads to simpler bookkeeping in two dimensions, though not for higher dimensions.

Mezher & Philippe's [8] integer coordinates for a triangular grid also involve considerably more bookkeeping; however, they do have a distinct advantage when the margin is a closed curve (see §2.3). Aside from its simplicity, the present approach is preferred because it allows adaptive control of the step size (§3).

2.3 Termination criteria

The two possible outcomes are that the sequences of triangles and marginal points will either leave D or (at least in exact arithmetic) return to the original triangle (and therefore the first found marginal point) and thereafter cycle [2, p. 162]. This is the counterpart of the fact that connected level curves

of continuous functions are homeomorphic to either the line or the circle [2, p. 2]. Although the former possibility is more typical for Orr–Sommerfeld marginal curves, the latter also occur in related problems [6, Figure 16].

The unbounded case is handled by testing whether we remain in D.

The cyclic possibility is best treated by considering the triangular grid implied by the images of a line segment under the transformation. This was the choice of Mezher & Philippe [8], since their application was level curves of resolvent norms ('pseudospectra') which they proved bounded. This approach was rejected here, since it renders adaptive step sizing much more difficult, although the sequential quaternary addressing system used in the analysis of Sierpinski's triangle [9, p. 93] might be a possibility.

The crude catchall of a limit to the number of iterations is also useful.

3 Adaptive skirting algorithm

3.1 Measures of accuracy

Two measures of accuracy for the algorithm are the proximity of output points to the margin, and the proximity of the margin to the output points. These are not the same! See Figure 1: all outputs lie near the margin, but some points on the margin are far from any output points.

The first measure is covered by Theorem 1.

The second is a problem of interpolation, so that little can be said about it without restricting the predicate. Analysis texts are replete with pathological examples such as $P(\mathbf{x}) \equiv (x_i \text{ is rational})$ for which the marginal set is $\mathbb{R}^2$ [11, p. 10], but predicates arising in applications are usually better behaved; for example, the Orr–Sommerfeld stability predicate is continuous and piecewise

differentiable. Numerical experience suggests that at least when the marginal set forms a piecewise smooth curve satisfactory results are obtained as the resolution of the triangulation is decreased. At any rate, the present method is quite analogous to representing functions $\mathbb{R} \to \mathbb{R}$ by tables or plots of values at equally spaced abscissas, and similarly, the margin is most obviously reconstructed by linear interpolation. This suggests the local departure of the margin from a straight line as a measure of accuracy. Using the norm (1) and projector (3) of the rectangular metric, the perpendicular distance from a point $\mathbf{c}$ to the line containing points $\mathbf{a}$ and $\mathbf{b}$ is

$$\left\| \mathbf{c} - \left\{ \mathbf{a} + \mathrm{proj}_{(\mathbf{b}-\mathbf{a})}(\mathbf{c} - \mathbf{a}) \right\} \right\| \ . \tag{6}$$

This is used in §§ 3.2–3.3 to develop an adaptive step size skirting algorithm.

3.2   Adaptive skirting

Using (6), the perpendicular distance of the latest marginal point from the line containing the last two can be compared against a specified tolerance to assess the adequacy of resolution. If the distance is excessive, the last marginal point and transverse pair can be discarded ($\mathbf{m}_2$, $\mathbf{u}_2$, and $\mathbf{v}_2$ in Figure 2) and the second last transverse pair ($\mathbf{u}_1$ and $\mathbf{v}_1$) shrunk about their marginal point ($\mathbf{m}_1$), checking that the pair remain transverse (§ 3.4). If the distance is acceptable, the latest transverse pair can be expanded about their marginal point, again checking that transverseness is maintained. Quantitatively, for shrinking halving has proven workable in practice, but for expanding the ratio of the tolerance to the distance, capped at some maximum (for example, 1.2), has been used.

There is no need to enforce a minimum step length since once the transverse points become sufficiently close, the next marginal point will be so close to the last that its projection on to the line of the last two points cannot lie outside the acceptable circle; this follows from the triangle inequality as applied to (6).
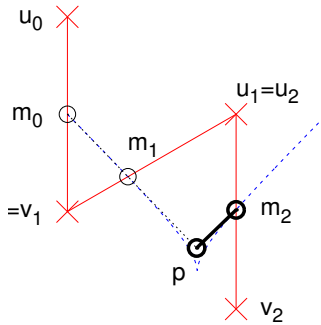
FIGURE 2: The adaptivity criterion: $\|\mathbf{m}_2 - \mathbf{p}\|$.

3.3 Initialization

The main adaptive skirting loop (§ 3.2) requires two previous marginal points, so that an initialization is required before the loop can be entered. This can be done by taking the initial transverse pair ($\mathbf{u}_0$ and $\mathbf{v}_0$ with marginal point $\mathbf{m}_0$ in Figure 3) and forming another ($\mathbf{u}_{0s}$ and $\mathbf{v}_{0s}$) by moving each point halfway toward the marginal point $\mathbf{m}_0$, checking that truth-values are maintained. Then step each forward to form $\mathbf{m}_1$ and $\mathbf{m}_{1s}$. If $\mathbf{m}_1$ is close to the line containing $\mathbf{m}_0$ and $\mathbf{m}_{1s}$, the main loop can be begun using $\mathbf{u}_1$ and $\mathbf{v}_1$ with $\mathbf{m}_1$ and $\mathbf{m}_0$ as the last two marginal points.

Adaptivity complicates the coding, so the algorithm is not listed here. For example, the GNU Octave [4] implementation[2] occupies eighty lines of code compared with less than twenty for the minimal implementation of § 2. Most of the additional code is needed for checking if points change truth-value during shrinking or expansion transverse and finding alternatives when they do. Nevertheless, this is well rewarded since the reduction in the number of predicate evaluations required exceeds the expansion in the number of lines

---

[2]Available from the author's website http://www.aeromech.usyd.edu.au/~mcbain/stability/skirting.html

FIGURE 3: Initialization of adaptive skirting. After copying and shrinking the first triangle, the criterion becomes $\|\mathbf{m}_1 - \mathbf{p}\|$.

of code, as demonstrated in §4.

3.4   Maintaining truth values

With adaptivity, it is often necessary to shrink or expand transverse pairs. It is essential that the end points retain their truth values. As an example, say in Figure 2 the deviation is deemed excessive so that $\mathbf{u}_1$ and $\mathbf{v}_1$ are to be shrunk but that $P(\mathbf{u}_1)$ changes as a result (this would not occur for the margin shown in Figure 2): then the old and new $\mathbf{u}_1$ are transverse and can be used thereafter, discarding $\mathbf{v}_1$. If expansion causes a change of truth value, the expansion can be aborted. If the shortened segment in Figure 3 differs logically from the initial, the initialization should begin again with a new shorter pair such as $\mathbf{u}_0$ and $\mathbf{u}_{0s}$.

FIGURE 4: Adaptive skirting of predicate (5) beginning from $(-2, 3)$ and $(-2, 2)$ with $\tau_1 = \tau_2 = 10^{-3}$ and dogleg tolerance (6) $128, 64, \ldots, 1$. The true margin is dashed; solid line segments connect the output skirting points.

4 Example

The result of applying the adaptive skirting algorithm to predicate (5) is illustrated in Figure 4. The bisection tolerance τ was chosen so that the points lie on the margin to within graphical accuracy while the dogleg tolerance was tightened to exhibit the efficacy of the algorithm in penetrating into cusps and then recovering step length as the curve straightens again.

The bottom left item of Figure 4 (dogleg tolerance 8) contains 66 points and required 968 predicate evaluations. To obtain a similar resolution of the cusp without adaptivity required 8594 evaluations. Experience with the algorithm suggests that this ratio is typical, and that even greater savings would be obtained if the margin were to be followed a considerable distance beyond the cusp.

5   Discussion and conclusion

Some Orr–Sommerfeld marginal curves are easily traced by assuming that the marginal curve defines a function $R(\alpha) = 0$; for example, the cubic velocity profile [7]. This of course fails at turning points, but can be remedied by simple parameter switching [5, e.g.].

Allen [1] proposed applying pseudo-arclength continuation methods [2, p. 4], and found them useful in computing the marginal curves of the Ekman boundary layer. These methods require the Jacobian, and so are more complicated to derive and program than the present method; they also encounter difficulties in regions where f has discontinuous derivatives. This occurs, for example, even in the simple case of Poiseuille flow [10, Figure 3.8].

The special features distinguishing the simple skirting algorithm from the generic piecewise linear continuation method [2] follow from the restriction of the dimension of the domain to two, so that the simplices are triangles and their 'facets' line segments. The former renders the task of selecting the 'exit facet' trivial and the latter permits the use of Theorem 1. These features are common to PAT [8]; and the present method is distinguished from that by its efficient adaptivity and simpler bookkeeping (even the adaptive version is considerably shorter than PAT). PAT has the advantage of robustly recognizing closed curves, and is also suited to parallelization since each bisection is a separate job not influencing later execution.

Examples of the use of the algorithm to trace linear hydrodynamic stability margins can be seen in the accompanying paper [7].

References

[1] L. Allen. *Modelling Dolphin Hydrodynamics*. PhD thesis, University of Surrey, 2001.   C90

[2] E. L. Allgower and K. Georg. *Numerical Continuation Methods: An Introduction.* Springer, 1990. C80, C83, C84, C85, C90

[3] P. G. Drazin. *Introduction to Hydrodynamic Stability.* Cambridge University Press, 2002. C80

[4] J. W. Eaton. *GNU Octave Manual.* Network Theory, 2002. C87

[5] A. E. Gill and A. Davey. Instabilities of a buoyancy driven system. *J. Fluid Mech.*, 35:775–798, 1969. C90

[6] J. E. Hart. Stability of the flow in a differentially heated inclined box. *J. Fluid Mech.*, 47:547–576, 1971. C85

[7] G. D. McBain and S. W. Armfield. Natural convection in a vertical slot: accurate solution of the linear stability equations. In Jagoda Crawford and A. J. Roberts, editors, *Proc. of 11th Computational Techniques and Applications Conference CTAC-2003*, volume 45, pages C92–C105, March 2004. [Online] http://anziamj.austms.org.au/V45/CTAC2003/McB2/home.html [March 23, 2004]. C80, C81, C90

[8] D. Mezher and B. Philippe. PAT—a reliable path-following algorithm. *Numer. Algorithms*, 29:131–152, 2002. C80, C83, C84, C85, C90

[9] H.-O. Peitgen, H. Jürgens, and D. Saupe. *Fractals for the Classroom.* Springer, 1992. C85

[10] P. J. Schmid and D. S. Henningson. *Stability and Transition in Shear Flows.* Springer, 2001. C90

[11] M. Spivak. *Calculus on Manifolds.* Benjamin, 1965. C85