A mathematical excursion in the isochronic hills

Philip Scarf, Centre for OR and Applied Statistics, University of Salford, Salford M5 4WT, UK

Introduction

Looking out over the South Pennines in mid-winter, through the driving rain, we can see flat topped hills, drab and uninviting. The heather, dark at this time of year, gives them a foreboding aspect. When walking and faced with a choice to go over or around, their flat tops mean that the over-the-top route is usually quicker. Of course, we may be out for a leisurely day in the hills and may not be concerned with speed. Then again, the rain may be stinging our face and a warm fireside and a pint of ale may be beckoning. Muddy, coal-black paths may define our route. Perhaps the choice is not so simple. There may be a multitude of paths, relics of local industry and agriculture with greater human input than at present. The population density on the moorland fringe was significantly higher in 1900 than at the end of the millennium. As elsewhere, the human population has given way for sheep. These hill flocks play a role in maintaining the upland path network. They also assist the mountain runner and navigator—grazed upland makes good running. This runner, when competing in events such as the Original Mountain Marathon (2006), will be concerned with fastest routes and not with route aesthetics. Then, a good over-or-around decision will save time. Choosing the best line between checkpoints, one that minimizes climb and distance travelled and has good running, is part of the art of mountain navigation. The Original Mountain Marathon is an "adventure race" for teams of two. It takes place over two days in late October. Competitors in the elite class cover over 80km in two stages in mountainous terrain, navigating from point to point, and camping overnight at a remote location. All pairs have to be self-supporting, carrying all their food and gear for the two stages.

In the Lake District, the fells have more of a conical quality. Here it may be quicker to go around than over. The Howgill Fells, sandwiched between the Lakes and the Pennines, are geographically and topographically intermediate—they are more rounded in character. Well grazed and offering a multitude of routes over, around or in between, difficult route choices abound. With a pair of checkpoints carefully selected by an event planner, perhaps the execution times for all routes between them are approximately equal. That is, all routes may be isochronic (of equal duration). A mountain runner and navigator may then ask the question: can isochronic routes be characterized? We may imagine a hill such that all routes over it are equal it time terms. What does such a hill look like? We might call such a hill, if it existed, isochronic. A cone is a simple hill. Does there exist an isochronic cone? If a runner knew what an isochronic hill looked like when represented topographically (with contours), he or she could potentially make faster decisions regarding over or around when competing in mountain navigation races. The runner might be trained to spot a hill that was flatter than isochronic (go over) or steeper than isochronic (go around). The recreational walker may also want to find the quickest way to the pub! In this paper, we attempt to find descriptions of simple hills (cones, pyramids, and domes) that are isochronic.

First we have to consider a rule that relates climb to distance. (Note, in this article, climb will refer to the vertical component of distance. Distance will mean the horizontal component of distance.) Consider the problem of travelling on foot from one side of a hill to the other in the shortest time. The obvious solution is to run faster. Therefore, consider this problem for an athlete who runs at a fixed speed. Then the solution will be to go by the shortest route—over the top. But the effect of the climb is to slow the runner on the ascent. If the over-the-top route involves significant climb, it may not be quicker than going round. Therefore, to pose the problem more usefully, we will assume that the athlete travels horizontally, on level ground, at a constant speed, and ascends (travels vertically) at a rate that implies 1 unit of distance vertically is equivalent in time terms to α units of distance horizontally. This equivalence between distance and climb was proposed by Scarf (1998), and is based on Naismith's rule (Naismith, 1892) which states that "men in fair condition should allow for easy expeditions, namely, an hour for every three miles on the map with an additional hour for every 2000ft of ascent". Thus, 2000 feet of climb is equivalent to 3 miles (=15,840 feet) of distance and so Naismith's rule implies $\alpha = 7.92$. We call $\alpha = 7.92$ Naismith's number. If a route comprises of a horizontal distance component of x units and a vertical distance component of y units, then Scarf calls $x + \alpha y$ the equivalent distance of the route.

Naismith did not provide any empirical evidence to substantiate his rule. However, the record times for fell races provide support for Naismith's number (Scarf, 2007). Norman (2004) finds evidence for a smaller value of α (in road running and treadmill experiments). Others (e.g. Langmuir, 1984; Rees, 2004) have proposed refinements to the rule particularly for steeper ground. α may vary between runners. It is important to note that in the analysis of Scarf (2007), and in Naismith's original proposal, there is a presumption that the rule applies to routes that start and finish at the same elevation—what goes up must come down—and therefore empirical values of α should be based on the times for journeys or events that start and finish at the same elevation. Thus the effect of ascent is confounded with the effect of descent. While we can therefore only calculate the equivalent distance of a route that starts and finishes at the same elevation, we can compare competing routes between two points at different elevations because Naismith's rule implies that the difference in time only depends on the difference in climb and the difference in distance between the routes. Alternatively, one can make an additional assumption that descent has no effect on speed. Whatever the exact underlying nature of the relationship in time terms between climb and distance, we will assume that the equivalence rule holds for our idealized athlete.

For our idealized athlete, the shortest route in time terms from one side of a hill to the other will be the route with least equivalent distance. There will be an infinite number of routes and so this approach is only helpful if one can narrow down the choices to a small number. The two simplest routes are (i) over and (ii) around. There may be reasonable routes that lie between these.

During an event, an athlete may attempt to calculate the equivalent distance of a number of competing routes by measuring the distance of and counting the number of contour lines crossed by a route. This is often impractical except in the simplest cases. Instead athletes rely on experience based on the representation of the hill on a map. An experienced navigator will claim the ability to recognize when it is quicker to go over than around. However, the over or round choice can be misleading even on the simplest hills. Consider determining the fastest route over the leg A to B in figure 1. The hill here is a cone. Suppose this cone is constructed so that the direct route AB (over) has the same equivalent distance as the route on the semicircular contour AB (around). Then there is an intermediate route which is faster—we establish this later. The cone is not isochronic with respect to A and B even if the over and around routes are equivalent.

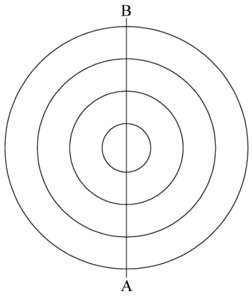

Figure 1. Equally spaced contours of a cone with base diameter AB.

Consider determining the fastest route between opposite base corners of a regular pyramid whose base diagonal is 1 unit (figure 2). Further, let the over and around routes be equivalent so that the height of the pyramid is $h = (\sqrt{2}-1)/\alpha$ and both routes have equivalent distance $\sqrt{2}$. Consider now a route ACB which goes over the shoulder of the pyramid at C. Let the perpendicular distance of the point C from the direct route AB be x. Then the equivalent distance of route ACB, $|ACB|_{eq}$, independent of α, is

$$2\sqrt{x^2+1/4}+2(\sqrt{2}-1)(\tfrac{1}{2}-x),$$

since the vertical height, y, of the point C above the base is given by $y = 2h(\frac{1}{2}-x)$ and $|ACB|_{eq} = 2|AC|+\alpha y$. $|ACB|_{eq}$ is a minimum When $x = 0.23$, and $\min|ACB|_{eq} = 1.325$. Therefore the pyramid is not isochronic with respect to A and B, and furthermore, while the calculation of the shortest route is straightforward mathematically, it is difficult to imagine using navigational experience to arrive at this choice of best route. The best route would also give a time saving of the order of 6% or 4 minutes in every hour. While few hills are shaped like pyramids, real valleys often appear like "half-pyramids" and so this result is of practical interest.

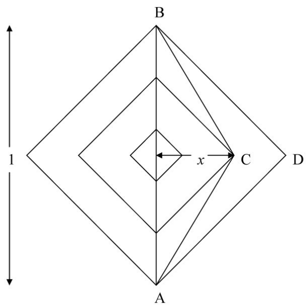

Figure 2. Equally spaced contours of a regular pyramid with base diagonal 1 unit.

The above initial analysis then brings us back to our central question: what is the shape of an isochronic hill? To answer this question, we first define what is meant by an isochronic hill. Roughly speaking, a hill is isochronic if all sensible routes on it have the same equivalent distance. To be more precise, we need to consider two points at equal elevation on the hill, A and B say, and the shortest route from A to B that climbs to height y above A, R_y. Then we say that the hill is isochronic with respect to A and B if the equivalent distance of R_y is the same for all $y (0 \le y \le y_{max})$. Note that the height of the summit above A (and the shape of the hill) defines y_{max}, and $R_{y_{max}}$ is the shortest route with maximum height gain. On a "regular hill", $R_{y_{max}}$ will be the route over the top. R_0 is the route around the base. By specifying A and B, we are suggesting that a hill can only be "relatively" isochronic. That is, the shape of an isochronic hill depends on where one starts and finishes on it.

We could attempt to draw the isochrones for the point B (those lines joining points that have equal equivalent distances from B) superimposed onto a hill. In yachting, the study of isochrones is well developed (Philpott, 2005; Philpott and Leyland, 2006). If the hill is isochronic with respect to A and B, then that part of the isochrone near A that lies on the hill will be circular with centre at A. Then, all routes out of A are equal in time terms, since the optimum route is perpendicular to the isochrone. However, we cannot draw the isochrones, because to do so we need to know the relative speed of progress on all points of the hill. But these speeds are unknown because ascent is confounded with descent in Naismith's rule.

Descriptions of isochronic hills

To determine the shape of isochronic hills, it is necessary first to restrict the class of hill and then find the hill within the class that is isochronic with respect to A and B. A simple class consists of those hills with circular contours with a common origin (rotationally symmetric hills). This class includes the cone, the hemisphere, and the hill whose summit is a cusp. Let A and B lie at opposite ends of a diameter of the base contour of a hill in this class (figure 3).

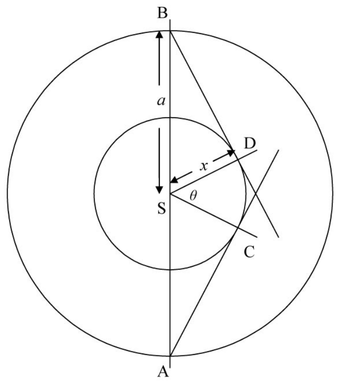

Figure 3. Plan view of a hill with circular contours with common centre at summit S.

The circular contours have a common origin at the summit, S. The topography outside the base contour need not concern us. On this hill, consider a contour with radius x and at height y above the base contour. Our intention is to determine how y is related to x when the hill is isochronic with respect to A and B. We will call the function $y = f(x)$ so determined the shape of the isochronic hill (in this class with respect to A and B). Let

$|AB| = 2a$. The shortest route from A to B that climbs to height y above A (with least horizontal distance component), R_y, is along the tangent to the contour radius x from A to the point C, then along the arc CD (maintaining a constant height y), and then along the tangent to B. The distance of this route $|R_y|$ is $2|AC|+x\theta$, where θ is the angle subtended by the arc CD at S. A little trig gives

$$|R_y| = 2\left\{\sqrt{a^2-x^2}+x\sin^{-1}(x/a)\right\}.$$

The equivalent distance of the route R_y, $|R_y|_{eq}$, is $|R_y|+\alpha y$. The hill is isochronic if $|R_y|_{eq}$ is the same as the horizontal distance on the base contour (πa) for all y. That is if

$$2\left\{\sqrt{a^2-x^2}+x\sin^{-1}(x/a)\right\}+\alpha y=\pi a.$$

Rearranging this gives the shape of the isochronic hill:

$$y=a\left\{\pi-2\sqrt{1-(x/a)^2}-2(x/a)\sin^{-1}(x/a)\right\}/\alpha, \qquad (1)$$
$$-a\le x\le a.$$

This function is drawn in figure 4a (with base radius $a=1$) with distance and climb on the same scale. It is remarkable how "flat" this hill appears in elevation view. Also drawn (figure 4b) is the quadratic hill ($y=(\pi-2)(1-x^2)/\alpha$, $-1\le x\le 1$.) with same summit height. We have had to reduce the vertical scale here to distinguish these functions.

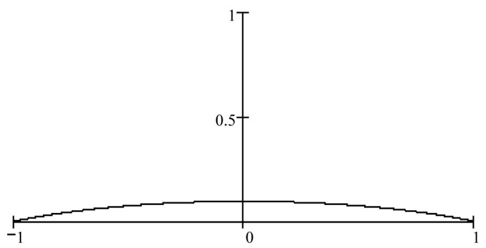

Figure 4a. Isochronic hill with circular contours and AB a diameter of base contour. Elevation view with equal vertical and horizontal scales.

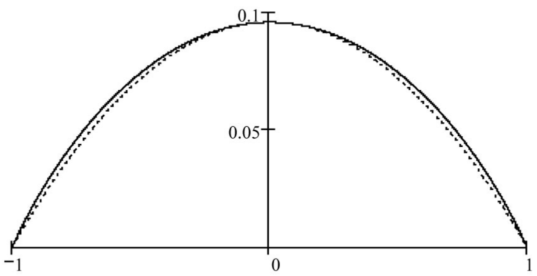

Figure 4b. Isochronic hill as above (solid line), quadratic hill with same base contour and summit height (dotted line). Both hills with circular contours and AB a diameter of base contour. Elevation view with unequal vertical and horizontal scales.

Taking a to be the unit of distance, and using the obvious property of rotational symmetry of the hill, implies that in three dimensions its functional form is given by

$$y=\left\{\pi-2\sqrt{1-x_1^2-x_2^2}-2\sqrt{x_1^2+x_2^2}\sin^{-1}\sqrt{x_1^2+x_2^2}\right\}/\alpha, \qquad (2)$$
$$x_1^2+x_2^2\le 1.$$

The surface, equation 2, is shown in figure 5 along with two map extracts of hills with approximately circular contours.

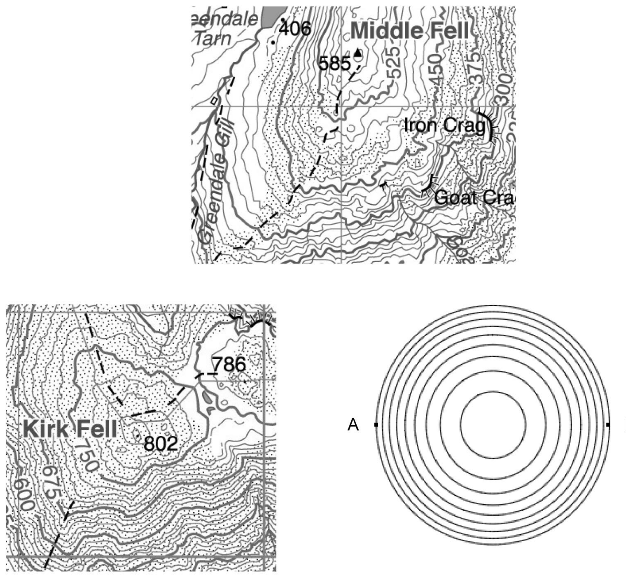

Figure 5. Top, Middle Fell, below left Kirk Fell in the Lake District. Below right, isochronic hill (with respect to A and B) with circular contours, with base diameter 1km and contour interval 15m, drawn to same scale and contour interval as map extracts (© Harvey Maps, 2008).

Suppose now that A is not on the base contour, and AB is an extension of the base diameter. Consider the hill drawn in figure 6, with the point B at height c above A. The shortest route that climbs to height y above the contour through A, R_y, travels along the tangent AC, then along the arc CD with radius x, and then along the tangent DB. Its equivalent distance is

$$|R_y|_{eq}=\sqrt{b^2-x^2}+\sqrt{a^2-x^2}+x\sin^{-1}(x/b)+x\sin^{-1}(x/a)+$$
$$\alpha(y-c)+g(c), \quad (-a\le x\le a)$$

where $g(c)$ is an unknown factor for the additional climb c (that does not have a matching descent). The route R_c travels along the tangent AE and then on the contour EB. Its equivalent distance is

$$|R_c|_{eq}=\sqrt{b^2-a^2}+a(\sin^{-1}(a/b)+\pi/2)+g(c).$$

Again setting $a=1$, we can now parameterize the isochronic hill with respect to A and B with two parameters b and c. The hill is isochronic if $|R_y|_{eq}=|R_c|_{eq}$ for all $c\le y\le y_{max}$. That is, if

$$y=c+\{\sqrt{b^2-1}-\sqrt{1-x^2}-\sqrt{b^2-x^2}+\pi/2+\sin^{-1}(1/b)$$
$$-x\sin^{-1}x-x\sin^{-1}(x/b)\}/\alpha, \quad (-1\le x\le 1).$$

If $c=0$ and $b=1$, we obtain equation (1) above. If $c=0$ and $b>1$, then we are approaching the hill on a flat plane from a distance $(b-1)$, and the further A is from the foot of the hill (at F), all else being equal, the more likely one is to go around. Or, as b increases, the isochronic hill will become flatter in comparison to equation (1). Note the topography between the contours through A and F need only be regular—"uphill only".

If A and B lie on the base contour but AB is not a base contour diameter, then the isochronic hill will have a different shape. Suppose B is at a perpendicular distance $b(\le a)$ from the base contour diameter through A (figure 7). Then, the equivalent

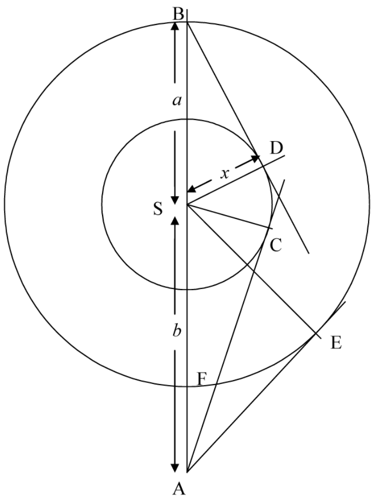

Figure 6. Hill with circular contours with common centre at summit S and base diameter AF extended to a point A, some height c below F.

distance of the shortest route that climbs to height y above the base contour through A and B is given by

$$|\mathrm{R}_y|_{eq} = 2\sqrt{a^2-x^2}+2x\sin^{-1}(x/a)-x\sin^{-1}(b/a)+\alpha y,$$
$$(-a \le x \le a).$$

The route on the base contour from A to B has length $\pi\alpha|2+\alpha\cos^{-1}(b|\alpha)$ and so the hill is isochronic if

$$y = \{\pi a/2 + a\cos^{-1}(b/a) - 2\sqrt{a^2-x^2} - 2x\sin^{-1}(x/a) + x\sin^{-1}(b/a)\}/\alpha. \tag{3}$$

This is only true for $a\sin(\frac{1}{2}\sin^{-1}(b/a)) \le x \le a$. This is because when $x = a\sin(\frac{1}{2}\sin^{-1}(b/a))$, the line AB is perpendicular to the contour with radius x, and therefore the line AB is the direct

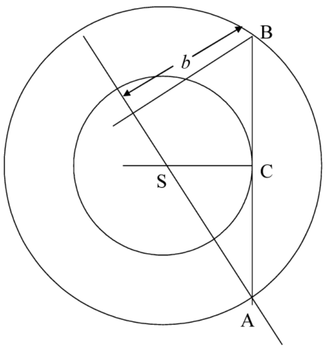

Figure 7. Hill with circular contours with common centre at summit S, with AB not a base diameter.

"over-the-top" route. A route that climbs higher than this direct route must also go further horizontally. Therefore, the hill with circular contours cannot be isochronic with respect to A and B for $x < a\sin(\frac{1}{2}\sin^{-1}(b/a))$. If it were then the height of the hill at S for example must be less than that at C. But then the contours would no longer be circular. This though suggests another extension—the description of an isochronic hill with contours that are circular arcs that are reflected in the line AB and that have common focus (figure 8). A contour of this hill is a (symmetric) lens.

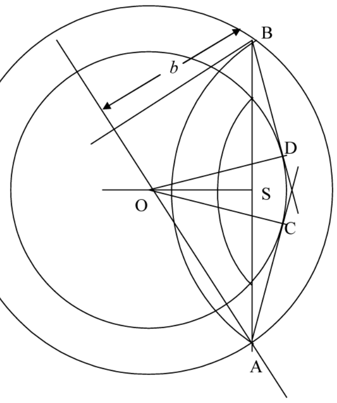

Figure 8. Symmetric lens contours of a regular hill with summit at S defined by the common chord AB of the circles centre O. Base contour diameter is 2α.

Again let the perpendicular distance to the point B from the base diameter be. Let the diameter through A have centre at O, and let the hill have summit at S. Then the horizontal distance OS is $s = a\sin(\frac{1}{2}\sin^{-1}(b/a))$. Replacing x in equation (3) by $x' = x - s$ we will obtain the shape of the isochronic hill with lens contours for $-(a-s) \le x' \le a-s$.

Next we can ask what is the effect if AB is the minor axis of such a hill (figure 9). A little more trigonometry gives the shape of the isochronic hill as

$$y = \{2a\cos^{-1}(s/a) - l(x)\}/\alpha,$$

where

$$l(x) = \begin{cases} 2\sqrt{a^2-(s^2+x^2)} + 2\sqrt{s^2+x^2}\{\sin^{-1}(x/\sqrt{s^2+x^2}) \\ -\sin^{-1}(\sqrt{a^2-(s^2+x^2)}/a)\}, \\ \qquad\qquad s(a-s) \le x^2 \le (a+s)(a-s), \\ 2\sqrt{x^2+(a-s)^2}, \qquad\qquad x^2 < s(a-s). \end{cases}$$

Such a hill is shown in figure 9, along with Yewbarrow in the Lake District, a hill with contours that are remarkably like a symmetric lens.

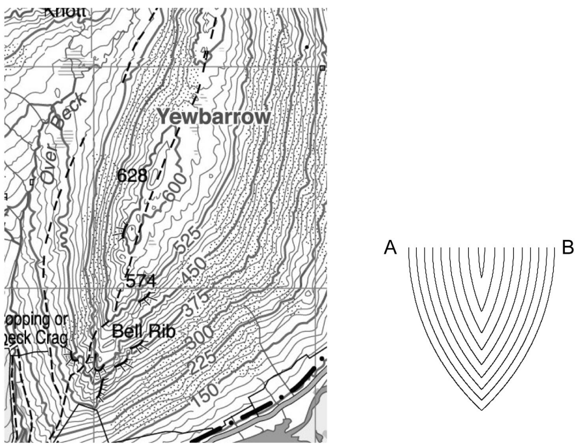

Figure 9. Left, Yewbarrow in the Lake District (© Harvey Maps, 2008). Right, isochronic hill (with respect to A and B) with symmetric lens contours with a=b, drawn to same scale and contour interval as map extract.

Haycock (figure 10) suggests consideration of a hill with contours that are equilateral triangles with a common centre. Let B lie at a vertex of the base and let A be the midpoint of the opposite side of the base, a distance a from B (figure 10). Consider a contour at height y above the base and let C be a vertex of this contour. Let the perpendicular distance from C to the line AB be x. Then the hill is isochronic with respect to A and B if

$$y = \{a\sqrt{3} - \sqrt{x^2 + (\tfrac{a}{3} - \tfrac{x}{\sqrt{3}})^2} - \sqrt{x^2 + (\tfrac{2a}{3} + \tfrac{x}{\sqrt{3}})^2}\} \,/\, a,$$

$$-\frac{a}{\sqrt{3}} \le x \le \frac{a}{\sqrt{3}}$$

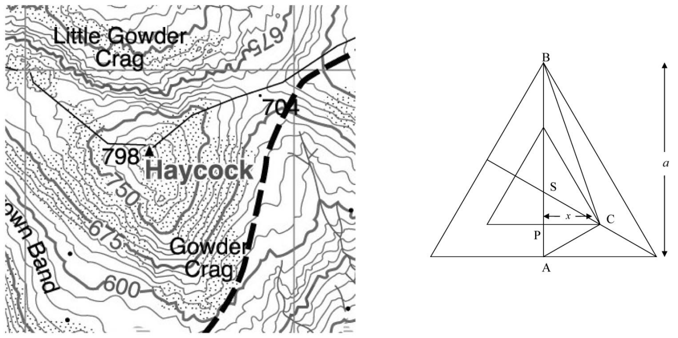

Figure 10. Left, Haycock in the Lake District (© Harvey Maps, 2008). Right, hill with contours that are equilateral triangles with common centre at the summit S.

This is obtained by noting that the route with no climb (y=0, with C a base vertex) has length $a\sqrt{3}$. The other term in equation (4) is the length of route ACB. The height of the isochronic hill is $a(\sqrt{3}-1)/\alpha$. Note here that this is the only hill we have considered in which the ascent and the descent have different gradients. This will have a small effect in practice since normally it will be marginally faster to go AB than BA. This is because a steep descent will be slower than a shallow descent. This is an example of a case where Naismith's rule is violated—hence the correction for steep ground (Langmuir, 1984).

In plan view, hills and valleys are indistinguishable and therefore all the analysis above applies equally to valleys. The isochronic surface in figure 9 is a case in point. Also, the analysis extends to depressions, although in the Lake District, these tend to be of the meteorological kind. Other hill shapes could be investigated, for example, hills with elliptical contours. The lengths of elliptic arcs are more difficult to handle and so we have not considered these.

Conclusions

In mountain navigation events, maps are typically not seen in advance and so route choices have to be made "on the run". For the most part, competitors have to rely on experience. Even simply shaped hills can be misleading in terms of best route. In this article, we draw isochronic hills in a number of simple circumstances. The topographical representation of an isochronic hill will depend on the map scale and the contour interval, but they can be drawn for typical scales encountered (1:40000, 15m; 1:25000, 10m; etc.). Such isochronic hills could then be used for armchair training—mountain navigators could train themselves, in the warmth of their own sitting rooms, to recognize such hills.

However, a hill can only be isochronic with respect to some chosen points A and B. Therefore, spotting isochronic hills and valleys will be difficult in a race. Arm chair training will also be problematic due to the fact that most hills have non-standard shapes and a leg (defined by its start and end points) could be anywhere. However, we have been gained some insight. Isochronic hills are rounded. Isochronic hills are surprisingly flat. (As an aside, this suggests that the analysis here is robust to departures from Naismith's rule on steeper ground.) If the flat route and the straight route on a conical hill are isochronic then there exists a faster route over the shoulder of the hill. This result is more practically useful for half-cone and half-pyramid like valleys. When taking a shoulder route on a hill with smooth contours, on ascent one should choose to climb to a particular elevation, meeting the contour through the chosen elevation at a tangent.

The value of Naismith's number may differ from that assumed here. If it is smaller, then isochronic hills get higher—their shape remains the same. Naismith's number may vary between individuals. If the functional form of Naismith's rule changes then the shape of an isochronic hill will change. In principal we could seek that equivalence rule between climb and distance that makes a conical hill isochronic. We have only looked at very simple, regular hills. We have also ignored the existence of paths. Paths and their location have an important bearing on the route choice problem, and their effect in flat terrain has been studied (Kay, 2006). In terrain with varying topography their effect is more difficult to model. On more complex hills and real topography, route choice can be effected by using dynamic programming (Hayes and Norman, 1984).

Looking out again over the South Pennines, with their summits like flat-caps, perhaps straight over is not so quick after all. From the shape of the isochronic hills that we study and draw, it appears that climb will slow our progress to a surprisingly large degree. If we are determined to go over the top and admire the view, don't buy the round of ale just yet—we may be a little longer than you expect.❑

Acknowledgement

I am grateful to my colleague Rose Baker for discussions about the ideas in this article.

REFERENCES

Harvey (2008) Maps for walking, hiking, rambling cycling. www.harveymaps.co.uk

Hayes, M. and Norman, J.M. (1984) Dynamic programming and orienteering: route choice and the siting of controls. *Journal of the Operational Research Society* 35, 791-796.

Kay A. (2006) Orienteering route choice: running at the speed of light? *Mathematics Today* 42, 138-142

Langmuir, E. (1984) *Mountaincraft and Leadership.* The Scottish Sports Council/MLTB. Cordee, Leicester, UK.

Naismith, W.W. (1892) Untitled. *Scottish Mountaineering Club Journal* 2, 135.

Norman, J.M. (2004) Running uphill: energy needs and Naismith's Rule. *Journal of the Operational Research Society* 55, 308-311.

Philpott A.B. (2005) Stochastic programming in yacht racing. In "Applications of Stochastic Programming", Wallace S.W. and Ziemba W.T. (eds), SIAM.

Philpott A.B. and Leyland, G. (2006) Rowing to Barbados. ORMS Today 33(2), 42-49.

OMM (2006). Original Mountain Marathon event. www.theomm.com

Rees W.G. (2004) Least cost paths in mountainous terrain. *Computers and Geosciences* 30, 203-209.

Scarf P.A. (1998) An empirical basis for Naismith's rule, *Mathematics Today* 34, 149-151.

Scarf P.A. (2007) Route choice in mountain navigation, Naismith's rule and the equivalence of distance and climb, *Journal of Sports Sciences* 25, 719 – 726.

(email: p.a.scarf@salford.ac.uk)