Proof of an ellipse-drawing program due to Doug McIlroy

Jayadev Misra
Dept. of Computer Science
The University of Texas
Austin, Texas, 78712

March 8, 2023

Abstract

This note contains a proof of an ellipse-drawing program described in McIlroy [2].

Background The problem treated in this note is to draw a given ellipse on a discrete raster plane, as on a digital monitor that has pixels only at grid points, or printed pages that can apply ink at specific points. The ellipse is a continuous curve that only rarely passes through grid points. So, only an approximation to the ellipse can be drawn. McIlroy [2] proposes choosing those grid points that are near enough to certain points on the ellipse. He establishes a number of properties of the chosen grid points, and develops a sequence of efficient programs through refinement.

1 Mathematical properties of ellipse

Consider an ellipse with defining equation $x^2/a^2+y^2/b^2 = 1$, where both a and b are positive integers. The equation is symmetric in x and y, so, many properties regarding x-coordinate applies analogously to y-coordinate. We restrict ourselves to drawing the ellipse in the first quadrant; symmetric procedures apply to other quadrants.

In the first quadrant, the ellipse equation yields a function *ecl* from the points in the closed interval $[0, a]$ to the closed interval $[0, b]$. Specifically, $ecl(x) = b/a \times \sqrt{(a^2 - x^2)}$. Note that *ecl* is 1-1 and continuous. Also *ecl* is *antimonotonic*, so that for points p and q on the ellipse, $p.x < q.x \;\equiv\; p.y > q.y$, where $p.x$ and $q.y$ denote the x and y coordinates of p and q.

grid line A *horizontal grid line* is the set of points with a fixed integral value of their y-coordinates. Similarly, a *vertical grid line* is the set of points with a fixed integral value of their x-coordinates. We restrict the values of the x-coordinates to the closed interval $[0, a]$, and y-coordinates to $[0, b]$. Thus the region of interest in the first quadrant is restricted to a rectangle R which is shown with bold borders in Figure 1, the dashed lines are the horizontal and vertical grid lines. A horizontal grid line with $y = k$ is shown in the figure.

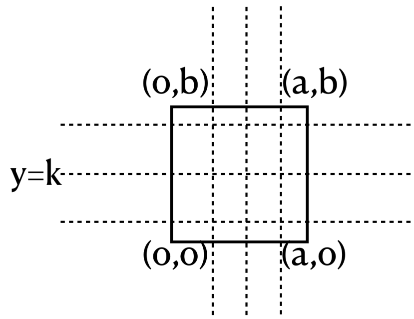

Figure 1: Rectangle R is within bold lines. Grid lines are dashed.

An *interior* point is a grid point within rectangle R that is not on any of the boundary grid lines.

ipoint An *intersection point*, or *ipoint*, is the point of intersection of the ellipse in the right quadrant with a grid line. The ipoint is vertical if it intersects a vertical grid line, horizontal otherwise. Points $(0, b)$ and $(a, 0)$ are both ipoints, and both horizontal and vertical.

Observation 1 There is a unique ipoint on every vertical and horizontal grid line within rectangle R.

Proof: The proof uses the intermediate value theorem. To see this for a horizontal grid line with $y = k$, where $0 \leq k \leq b$, observe that the $ecl(0) = b$ and $ecl(a) = 0$. So, ecl attains the intermediate value k for some x, $0 \leq x \leq a$. The proof for vertical grid lines is analogous.

The uniqueness of the ipoint follows from the 1-1 property of the ellipse function ecl. □

It follows from this observation that the only ipoint on the horizontal grid line $y = 0$ is $(a, 0)$, and the vertical grid line $x = 0$ is $(0, b)$.

2 Approximating ellipse intersections points

Freeman approximation [1] chooses grid point p to be *lighted*, written as $lighted(p)$, if it is within 1/2 unit of an ipoint p' in both coordinates:

$$p.x - 1/2 \leq p'.x < p.x + 1/2, \text{ and } p.y - 1/2 \leq p'.y < p.y + 1/2 \qquad [L(p)]$$

Note that p' is on a grid line, so either $p.x = p'.x$ or $p.y = p'.y$. Call ipoint p' a *witness* to p. Observe that $(0, b)$ and $(a, 0)$ are lighted, and they are both their own winesses.

segment A horizontal grid line can be partitioned into unit length segments called *horizontal segment*. Vertical segments are similarly defined. Write pq for a segment whose end points are grid points p and q: for a horizontal segment $p.x = q.x - 1,\ p.y = q.y$, and for a vertical segment $p.x = q.x,\ p.y = q.y + 1$. Point r is on horizontal segment pq if $p.x \leq r.x < q.x,\ p.y = r.y$ and on vertical segment pq if $p.x = r.x,\ p.y > r.y \geq q.y$. That is, each segment is a half-open interval of points that is open at the right end for a horizontal segment and the top end for a vertical segment.

Henceforth, write $ipt(pq)$ to denote that there is an ipoint on segment pq.

Observation 2 An ipoint is a witness to one of the end points of its segment. That is, $ipt(pq) \ \Rightarrow\ (lighted(p) \vee lighted(q))$.

Proof: The ipoint on pq, assuming it is not exactly midway between p and q, is closer than 1/2 unit to either p or q. Then it is a witness to the closer end point. If it is exactly midway then, from $L(q)$, it is witness to q. □

Observation 3 Consider the unit square $pqrs$ whose upper left corner is an interior point p; see Figure 2.

1. If there is an ipoint on segment pq, there is an ipoint on the segment qr or sr. That is,

$$ipt(pq) \ \Rightarrow\ (ipt(sr) \vee ipt(qr)).$$

2. If there is a witness for q on segment pq, then $ipt(qr) \vee lighted(r)$.

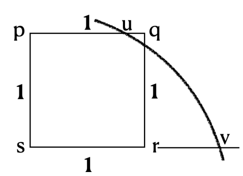

Figure 2: Unit grid $pqrs$ whose upper left corner p is an interior point.

Proof: I prove the result for ipoint u on segment pq, as shown in Figure 2.

1. From Observation 1

there is an ipoint v on the horizontal gridline that includes r
$\Rightarrow$ {antimonotonic property of the ellipse function ecl}
$u.x < v.x$
$\Rightarrow$ {rewrite: $u.x < v.x < r.x \vee r.x \leq v.x$}
v is on segment sr or $r.x \leq v.x$
$\Rightarrow$ {definition of ipt}

$ipt(sr) \vee r.x \leq v.x$

$\Rightarrow$ {Apply Intermediate value theorem when $r.x \leq v.x$
$u.x < q.x = r.x \leq v.x$ and $u.y = v.y + 1$ implies
there is y', $u.y > y' \geq v.y$ such that $ecl(r.x) = y'$. So, $ipt(qr)$.}

$ipt(sr) \vee ipt(qr)$

2. If u is a witness for q, from the above argument $ipt(qr) \vee ipt(sr)$. I show that $ipt(sr)$ implies $lighted(r)$. Any ipoint v on sr is closer to r than u is to q because

$$r.x - v.x < \{r.x = q.x, v.x > u.x\}\ q.x - u.x \leq 1/2.$$

So, v is a witness for r, and r is lighted. □

There is a dual result corresponding to Observation 3 for ipoint on a vertical grid line:

1. If there is an ipoint on segment ps, there is an ipoint on the segment qr or sr. That is,

$$ipt(ps) \;\Rightarrow\; (ipt(qr) \vee ipt(sr)).$$

2. If there is a witness for s on segment ps, then $ipt(qr) \vee lighted(r)$.□

Given grid points p and q, q is to the: (1) south of p if $q.x = p.x,\ q.y < p.y$, (2) east of p if $q.x > p.x,\ q.y = p.y$, (3) southeast of p if $q.x > p.x,\ q.y < p.y$. Other directions, such as north, west, northeast, southwest and northwest can be similarly defined.

Write $E(p)$, $S(p)$ and $SE(p)$, respectively, for the points that are immediately (i.e. are at unit distance) to the east, south, southeast of p.

Theorem 1 Given $lighted(p)$, $p \neq (a, 0)$, at least one of $E(p)$, $SE(p)$ and $S(p)$ is lighted.

Proof: First, consider the case where p is an interior point. Assume that the witness u to $lighted(p)$ is on a horizontal grid line; the proof is similar if the witness is on a vertical grid line. See Figure 3 for the unit squares to the left and right of p. Here q, s and r are $E(p)$, $S(p)$ and $SE(p)$, respectively.

$lighted(p)$

$\Rightarrow$ {there is a witness for p on pq or $q'p$. From Observation 3,
witness on $pq \;\Rightarrow\; ipt(pq) \Rightarrow (ipt(qr) \vee ipt(sr))$}

$ipt(qr) \vee ipt(sr) \vee$ (there is a witness for p on segment $q'p$)

$\Rightarrow$ {apply Observation 3 to witness on $q'p$}

$ipt(qr) \vee ipt(sr) \vee ipt(ps) \vee lighted(s)$

$\Rightarrow$ {apply Observation 3 to the third term $ipt(ps)$}

$ipt(qr) \vee ipt(sr) \vee ipt(sr) \vee ipt(qr) \vee lighted(s)$

$\Rightarrow$ {apply Observation 1 and remove duplicate terms}

$lighted(q) \vee lighted(r) \vee lighted(s)$

Next, suppose p is not an interior point and $p \neq (a, 0)$. Then p is on a grid line, $x = a$ or $y = 0$. I prove that if p is on the vertical grid line $x = a$, $south(p)$ is lighted. The dual result for p on the horizontal grid line $y = 0$ is similarly proved.

p is on the grid line $x = a$ and u is a witness for p
$\Rightarrow$ $\{(a, 0)$ is an ipoint on $x = a$ and,
from Observation 1, $(a, 0)$ is the only ipoint on $x = a$.
So, witness u is not on $x = a\}$
u is on the horizontal segment $q'p$; see Figure 3
$\Rightarrow$ $\{q'$ is an interior point; apply Observation 3$\}$
$ipt(ps) \vee lighted(s)$
$\Rightarrow$ $\{$Only ipoint on $x = a$ is $(a, 0)$. So, $ipt(ps) \Rightarrow (s = (a, 0))\}$
$s = (a, 0) \vee lighted(s)$
$\Rightarrow$ $\{lighted(a, 0)$ and $s = south(p)\}$
$lighted(south(p))$ □

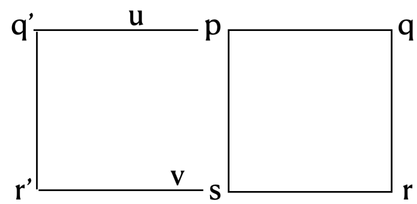

Figure 3: Lighted grid point p has a witness in its left or right segment.

3 An ellipse drawing program

The goal of the program is to compute the set of lighted points. It is easily seen that all the lighted points are within rectangle R, as shown in Figure 1. Therefore, it is sufficient to traverse R along horizontal or vertical grid lines, and identify the lighted points. Based on Theorem 1 Mcilroy proposes a far more efficient program.

Lemma 1 Grid points p and q, where q is to the southwest of p, can not both be lighted.

Proof: The constraint (SW), below, expresses that q is to the southwest of p: $q.x + 1/2 \leq p.x - 1/2, \;\; q.y + 1/2 \leq p.y - 1/2$ (SW)

Suppose both p and q are lighted. Let p' and q' be the corresponding witnesses.

$q'.x$
$<$ $\{$from $L(q)\}$
$q.x + 1/2$
$\leq$ $\{$from (SW)$\}$
$p.x - 1/2$
$\leq$ $\{$from $L(p)\}$
$p'.x$

Using the symmetry of x and y in $L(p)$ and $L(q)$, assert that $q'.y < p'.y$. Since both p' and q' are on the ellipse, $q'.x < p'.x$ implies $q'.y > p'.y$, contradiction. □

Lemma 1 implies that points p and q, where q is to the northeast of p, can not both be lighted because p is to the southwest of q.

Lighted points in sub-rectangles within R Define $R(p)$ to be the set of grid points within R that are to the south and east of p, i.e.

$$q \in R(p) \;\equiv\; p.x \le q.x \le a,\ 0 \le q.y \le p.y$$

Note that $R((0,b)) = R$ and $R(p)$ is empty if $p \notin R$. Write $R.lits$ and $R(p).lits$ for the set of lighted points in R and $R(p)$, respectively. The goal of the program is to compute the set $R.lits$.

Observation 4 Given $lighted(p)$,

1. $lighted(E(p)) \;\Rightarrow\; R(p).lits = \{p\} \cup R(E(p)).lits$
2. $lighted(S(p)) \;\Rightarrow\; R(p).lits = \{p\} \cup R(S(p)).lits$
3. $(\neg lighted(E(p)) \wedge \neg lighted(S(p))) \;\Rightarrow\; R(p).lits = \{p\} \cup R(SE(p)).lits$

Proof: I prove only part(1); the other proofs are similar. The set of grid points in $R(p)$ is the union of the set of points in $R(E(p))$ and in the vertical grid line from p to horizotal grid line $y = 0$. So,

$$
\begin{array}{ll}
 & R(p).lits \\
= & \{lighted(p)\} \\
 & \{p\} \cup R(E(p)).lits \cup \{q \mid lighted(q), q.x = p.x, q.y < p.y\} \\
= & \{\text{Point } q \text{ with } q.x = p.x, q.y < p.y \text{ is to the southwest of } E(p). \\
 & \ \text{Given } lighted(E(p)), \text{ from Lemma 1, the last set is empty.}\} \\
 & \{p\} \cup R(E(p)).lits
\end{array}
$$

□

An abstract ellipse-drawing program The skeleton of a program is given in Figure 4 that computes T, the set of lighted points. It has the invariant:

$$I :: \; lighted(p) \;\wedge\; T \cup R(p).lits = R.lits$$

Proof of correctness Observation 4 provides justification for the annotation in Figure 4. I show the proof of termination next.

The size of $R(p)$ decreases in each iteration unless the loop exit condition $p = (a, 0)$ holds. Since $R(p)$ is a finite set, it can not decrease forever, so, eventually $p = (a, 0)$.

Here is another proof of termination. Consider the pair $(a - p.x, p.y)$. In each iteration either or both components of the pair decrease; therefore, the pair decreases lexicographically. The minimum value of the pair in lexicographic ordering is $(0, 0)$. As long as the pair is different from $(0, 0)$, i.e. $p \neq (a, 0)$, the pair decreases. So, eventually $p = (a, 0)$.

Ellipse-drawing, Version 0

$\{a > 0 \wedge b > 0\}$
$T := \phi$ **;** $p := (0, b)$
$\{lighted(p),\ R(p) = R,\ T \cup R(p).lits = R.lits\}$
$\{I\}$
while $p \neq (a, 0)$ **do**
$\{p \neq (a, 0),\ lighted(p),\ T \cup \{p\} \cup R(p).lits = R.lits\}$
$T := T \cup \{p\}$**;**
$\{p \neq (a, 0),\ lighted(p),\ p \in T,\ T \cup R(p).lits = R.lits\}$
if
$[\!]\ lighted(E(p)) \longrightarrow p := E(p) \quad \{I\}$
$[\!]\ lighted(S(p)) \longrightarrow p := S(p) \quad \{I\}$
$[\!]\ \neg lighted(S(p)),\ \neg lighted(E(p)) \longrightarrow p := SE(p) \quad \{I\}$
endif
$\{I\}$
enddo ;
$\{p = (a, 0),\ T \cup R(p).lits = R.lits\}$
$T := T \cup \{p\}$
$\{T = R.lits\}$

Figure 4: Abstract version of ellipse drawing program, annotated

References

[1] Herbert Freeman. Computer processing of line-drawing images. *ACM Computing Surveys (CSUR)*, 6(1):57–97, 1974.

[2] M Douglas McIlroy. Getting raster ellipses right. *ACM Transactions on Graphics (TOG)*, 11(3):259–275, 1992.