

Baserunner's Optimal Path

DAVIDE CAROZZA, STEWART JOHNSON AND FRANK MORGAN

When you hit that final long ball in the World Series of Baseball and know you need the home run, what is your optimal path around the bases? If you run straight for first, you either have to slow to a near stop or go sailing far beyond into the outfield. The standard recommended “banana” path follows the baseline maybe halfway and then veers a bit to the right to come at first base from a better angle to continue toward second. That cannot be ideal. It would have been better to start at an angle to the right to head directly to an outer point on the banana path.

So what is the optimal path? Using a very simple model, we obtain the path of Figure 1. You start out heading about 25° right of the base line and run with acceleration of constant maximum magnitude σ , as illustrated by the vectors decorating the path. You slow down a bit coming into first, hit a local maximum speed as you cross second, and start the final acceleration home a bit before crossing third base (see Fig. 2). The total time around the bases is about $52.7/\sqrt{\sigma}$, about 16.7 seconds for $\sigma = 10 \text{ ft/sec}^2$, about 25% faster than following the baseline for 22.2 seconds (coming to a full stop at first, second, and third base), and about 6% faster than following a circular path for 17.8 seconds. The record time according to Guinness [G] is 13.3 seconds, set by Evar Swanson in Columbus, Ohio, in 1932. His average speed around the bases was about 18.5 mph or 27 ft/sec.

Is it legal to run so far outside the base path? The relevant official rule of Baseball says:

7.08 Any runner is out when—(a) (1) He runs more than three feet away from his baseline to avoid being tagged unless his action is to avoid interference with a fielder fielding a batted ball. A runner's baseline is established when the tag attempt occurs and is a straight line from the runner to the base he is attempting to reach safely. The rule just says that *after a tag attempt* the runner cannot deviate more than three feet from a straight line *from that point*. The rule doesn't apply until the slugger is almost home, when our fastest path is nearly straight. So our path is legal.

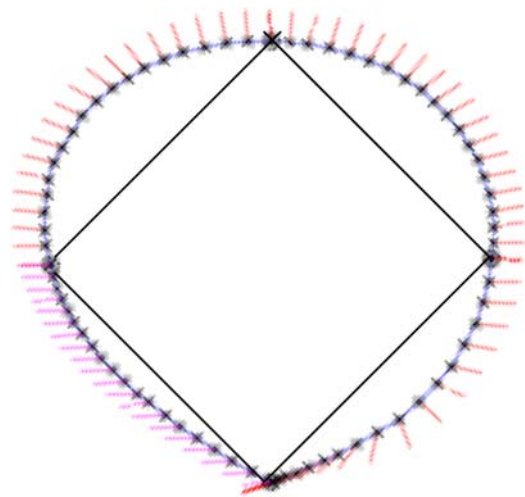


Figure 1. Second picture shows the fastest path around the bases given a bound σ on the magnitude of the acceleration vector, shown at each point. First picture from <http://www.bsideblog.com/images/2008/03/baseball-diamond.jpg>.

Our model simply assumes a bound σ on the magnitude of the baserunner's acceleration (which includes deceleration and curvature). The locus of the fastest path around the bases is independent of σ because you can scale velocity by λ , acceleration by λ^2 , and time by $1/\lambda$. So slow runners should follow the same route as fast ones. At first you might think that a very slow, awkward runner should just walk directly from base to base, except that he'd likely fall down trying to make the sharp turn at first.

To find the fastest path around the bases, we consider the simpler problem of finding the fastest path between two points, given the initial and final velocities, which has a unique solution. Intriguingly enough, for this problem, *total time is not continuous in the prescribed conditions*. Even on the line, consider starting at the origin with initial velocity 1 and going at maximal acceleration for a second, ending with velocity 2; now if, instead, the prescribed final velocity were increased a bit, you would have to start out by decelerating to velocity 0, go backward to well left of the origin, and then accelerate right to the terminus. (See Remark after Lemma 1. Fortunately time is lower-semicontinuous, which is what we need to prove the existence of fastest paths.)

For a critical path between bases, the acceleration vector a has constant magnitude σ and remarkably is given by $At + B$ normalized, for some constant vectors A, B . In velocity space, such paths are portions of catenaries (the famous least-energy shape of hanging cables as for suspension bridges), which in general can be absolute minima, local minima, or unstable critical points (see Remarks after Lemma 2).

It is easy to see that a fastest path for bounded $|a|$ also minimizes $\max |a|$ for given time, since if you could reduce $\max |a|$, then by increasing speed along an appropriate portion of the locus in space, you could reduce time. There are, however, more solutions to the second problem. In the example at the end of the Remarks after Lemma 2, all three paths minimize $\max |a|$ for given times $T_1 < T_2 < T_3$.

Given the fastest path between bases for prescribed velocities, we find the shortest path around the bases by minimizing over all choices of velocity at the bases, specifying velocity 0 at the start. We think that the solution is unique, but we know no proof.

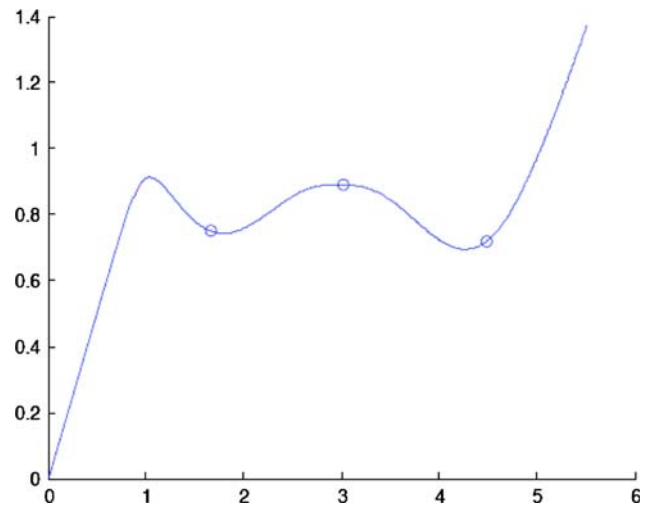


Figure 2. Speed as a function of time. For $\sigma = 10 \text{ ft/sec}^2$, each unit of time represents 3 seconds and each unit of velocity represents 30 ft/sec. The times for each segment are about 5.1, 4.1, 4.4, and 3.1 seconds, for a total of about 16.7 seconds.

Our model is, of course, an oversimplified one, since it assumes that maximum deceleration equals maximum acceleration and that maximum acceleration remains possible at high speeds; taking $\sigma = 10 \text{ ft/sec}^2$, it leads to a final speed coming into home of about 42 ft/sec, faster than the highest recorded human speed as of August 2009 of 40.5 ft/sec by Usain Bolt, even though his initial acceleration exceeded 18 ft/sec^2 [S].

Fastest Paths

Lemmas 1 and 2 provide existence and structure for the shortest path between two bases, given initial and final velocities. Proposition 1 considers the full baserunner problem with all four bases. We conclude by explaining our numerical solution of Figure 1.

LEMMA 1 *There exists a fastest path from one point to another in the plane, given initial velocity, final velocity,*



AUTHORS

DAVIDE CAROZZA graduated in 2009 from Williams College, where he presented his senior colloquium (a requirement for all senior math majors) on some preliminary analysis of optimal base-running. In his present job he teaches Algebra I.

Department of Mathematics and Statistics
Williams College
Williamstown, MA 01267
USA
e-mail: 09dgc@williams.edu



STEWART JOHNSON is a professor working in dynamics, optimal control, mathematical biology, game theory, differential equations, and statistics. He is also the Quantitative Skills Coordinator for Williams College, and as such he gives tutorials and courses for developing quantitative skills.

Department of Mathematics and Statistics
Williams College
Williamstown, MA 01267
USA
e-mail: Stewart.Johnson@williams.edu

and a bound $\sigma > 0$ on the acceleration. The minimum time is a lower-semicontinuous function of the initial and final positions and velocities.

REMARK The minimum time is not continuous in the prescribed conditions. For example, for $\sigma = 1$, the fastest path from $(0, 0)$ to $(v_0 + 1/2, 0)$ with initial velocity $v_0 > 0$ and final velocity $v_0 + 1 + \varepsilon$ is for $\varepsilon = 0$ simply forward motion for 1 second at unit acceleration, but for small $\varepsilon > 0$ one must decelerate for v_0 seconds to velocity 0 at $(.5 v_0^2, 0)$, move backward, accelerating and decelerating for another $\sqrt{2} v_0$ seconds to come to rest just left of $(0, 0)$, and then move forward for a bit more than a second at unit acceleration, for total time a bit more than $1 + v_0(1 + \sqrt{2})$ seconds, a huge discontinuity if v_0 is large. See Figure 3 for the case $v_0 = 1$. In summary, increasing the final velocity of a linear path with maximum acceleration involves backing up and a discontinuous increase in total time.

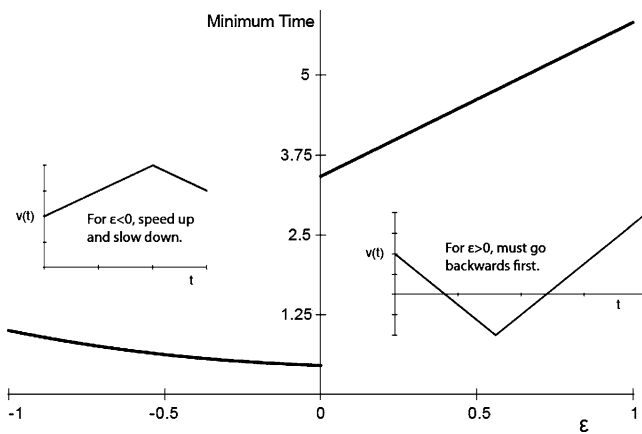


Figure 3. As the prescribed final velocity increases past that obtained by constant maximum acceleration, the fastest path has to back up, with a discontinuous increase in total time.

PROOF OF LEMMA 1 First we note that there exists some path satisfying the conditions. If the given velocities are 0,



FRANK MORGAN is a specialist in minimal surfaces. One of his books is “The Math Chat Book”, based on the call-in TV show he used to run and his column at MathChat.org. He is founder of Williams’s “SMALL” Undergraduate Research Project. As Vice-President of the American Mathematical Society, he launched the blog <http://mathgradblog.williams.edu/> by and for mathematics graduate students.

Department of Mathematics and Statistics
Williams College
Williamstown, MA 01267
USA
e-mail: Frank.Morgan@williams.edu

this is obvious. Otherwise just follow the given initial direction with maximum negative acceleration until obtaining velocity 0, and similarly backwards from the terminal point, to reduce to the obvious case. This path bounds the minimum time and hence the positions and velocities. Except for the trivial case when the initial and final position and velocities coincide, there is also a lower bound on the total time.

To prove simultaneously existence and lower-semicontinuity in the prescribed conditions, consider a sequence of paths with conditions converging to the prescriptions and times T_i converging to the infimum T . We may assume that the velocities are bounded functions from $[0, T_i]$ into \mathbb{R}^2 with Lipschitz constant at most σ and that $T_i \leq 2T$. Rescale time to change the domain to $[0, T]$. Now each velocity has Lipschitz constant at most $\sigma T_i/T \leq 2\sigma$ and the conditions still converge to the prescriptions. By the compactness of uniformly bounded Lipschitz functions, we may assume that the velocities and hence the paths converge; the limit has time T as desired.

LEMMA 2 A fastest $C^{1,1}$ path from one point to another in the plane, given initial velocity, final velocity, and a bound $\sigma > 0$ on the magnitude of the acceleration a ,

$$a = \sigma \frac{At + B}{|At + B|}$$

for some constant vectors A, B , is unique.

REMARKS By a translation in time, we may assume that $B \cdot A = 0$ and that A is a unit vector. The path is real-analytic in time unless $B = 0$ and $t = 0$, when a flips direction. In addition, the path in space can have a singularity where the velocity vanishes, as in Figure 4B.

Up to rotation and translation in the plane and scaling in time and space, we may assume that

$$a = \frac{(1, t)}{\sqrt{1 + t^2}},$$

$$v = (\operatorname{arcsinh} t, \sqrt{1 + t^2}) + v_0,$$

$$x = (t \operatorname{arcsinh} t - \sqrt{1 + t^2}, .5 t \sqrt{1 + t^2} + .5 \operatorname{arcsinh} t) + v_0 t,$$

pictured for $v_0 = 0$, $v_0 = -(0,1)$, $v_0 = -(0,10)$, and $v_0 = -(1,0)$ in Figure 4; or in the degenerate case

$$a = (\operatorname{sign} t = \pm 1, 0),$$

$$v = (-1 \pm t, 0) + v_0,$$

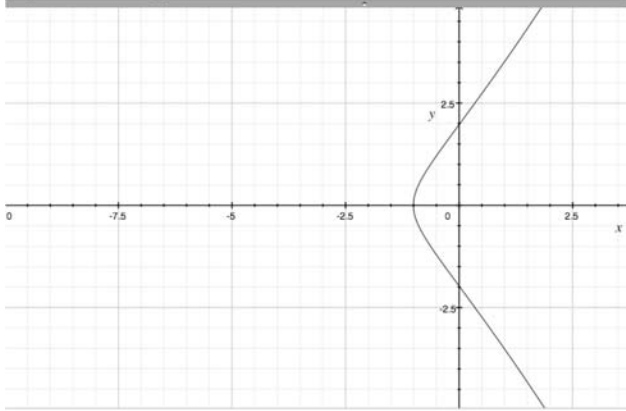
$$x = (-t \pm .5 t^2, 0) + v_0 t,$$

pictured for $v_0 = -(0,1)$ in Figure 5.

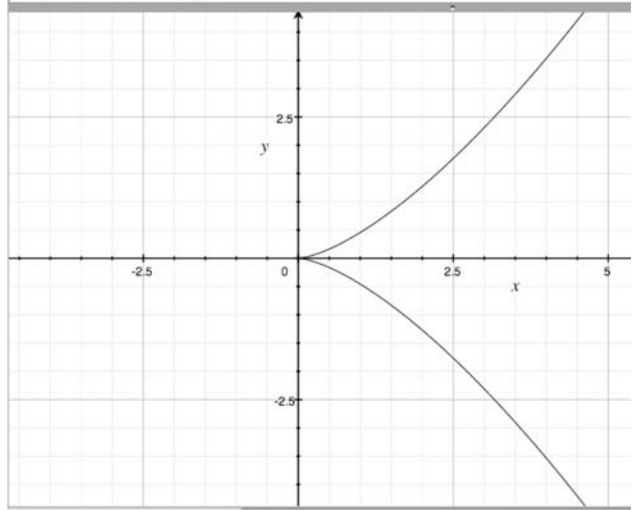
Some such critical paths are not minimizing. Indeed, the translation in velocity space of a minimizer need not be minimizing. For example, for $\sigma = 1$, the following path P is minimizing, but its translation P' by $v_0 = (1, 0)$ is not. The path P starts at $(0, 0)$, accelerates left for 1 second to $(-1/2, 0)$, decelerates for 1 second to $(-1, 0)$, and then accelerates

A

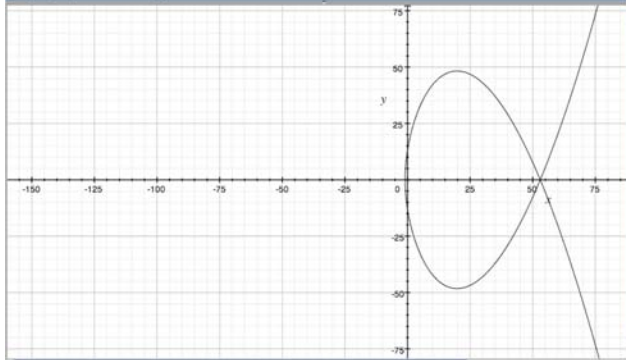
$$\begin{cases} x \\ y \end{cases} \begin{cases} (t \operatorname{arcsinh} t) - (1+t^2)^5 \\ 5t(1+t^2)^5 + 5 \operatorname{arcsinh} t \end{cases}, t=-100 \dots 100$$

**B**

$$\begin{cases} x \\ y \end{cases} \begin{cases} (t \operatorname{arcsinh} t) - (1+t^2)^5 + 1 \\ 5t(1+t^2)^5 + 5 \operatorname{arcsinh} t - t \end{cases}, t=-100 \dots 100$$

**C**

$$\begin{cases} x \\ y \end{cases} \begin{cases} (t \operatorname{arcsinh} t) - (1+t^2)^5 \\ 5t(1+t^2)^5 + 5 \operatorname{arcsinh} t - 10t \end{cases}, t=-100 \dots 100$$

**D**

$$\begin{cases} x \\ y \end{cases} \begin{cases} (t \operatorname{arcsinh} t) - (1+t^2)^5 - t \\ 5t(1+t^2)^5 + 5 \operatorname{arcsinh} t \end{cases}, t=-100 \dots 100$$

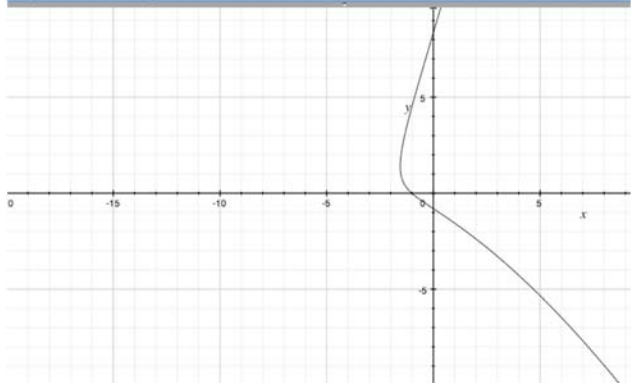


Figure 4. (A, B, C, D) Some critical paths with acceleration $At + B$ normalized.

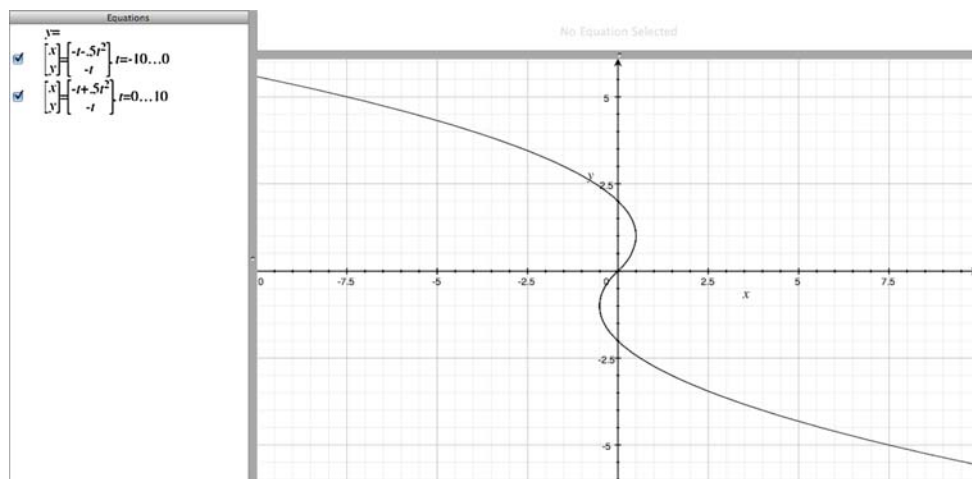


Figure 5. A symmetric critical path with acceleration $\pm (0,1)$, which is the special case $A = (0,1)$, $B = 0$.

to the right for 2 seconds, ending up at (1, 0) with velocity (2, 0). Its translation P' starts at (0, 0) with velocity (1, 0), decelerates for 1 second to (1/2, 0) and velocity (0, 0), and then accelerates for 3 seconds ending at (5, 0) with velocity (3, 0), for a total time of 4 seconds. A minimizer P'' accelerates for $\sqrt{10} - 1$ seconds and then decelerates for $\sqrt{10} - 3$ seconds for a total time of $2\sqrt{10} - 4 \approx 2.32$ seconds. In summary, the translation P' of a backtracking minimizer P may decelerate unnecessarily and fail to be minimizing.

Note that up to translation, rotation, and scaling, the path in velocity space is the famous catenary $v = \cosh u$ or in the degenerate case a line. It is well known that such paths minimize energy $\int v dt = \Delta y$ for given length $\sigma \Delta t$.

There are relative minima that are not absolute minima. Consider given velocities $(-1, 1)$, $(1, 1)$ and change in position $(0, \Delta y)$ vertical. Possible paths in velocity space are catenaries (or horizontal lines), a 1-parameter family. A horizontal line yields minimum time, but a catenary $v = a \cosh(u/a)$ yields local minimum Δy . Rotating such a catenary about the u -axis generates the famous minimal catenoid surface, with area 2π times the potential energy $\int v dt = \Delta y$ of the catenary. It is well known that for two relatively close congruent vertical circles about the same horizontal axis there are two catenoids, a slightly bowed area minimum and a deeply bowed unstable one [TF, chap. I, §3]. Bowing upward from the catenary generator of the stable catenoid, Δy increases; time decreases to the horizontal line, then increases. Downward, time increases; Δy decreases to the generator of the stable catenoid, then increases to the generator of the unstable catenoid, then decreases, eventually going very negative. So Δy values between the generators of the two catenoids are obtained three times, with times $T_1 < T_2 < T_3$. All have the same σ . The first is the global minimum. The third is a local minimum, since by the energy-minimizing property of the catenary, decreasing time requires increasing Δy .

All three paths minimize $\max |a|$ for given time, because if you could reduce $\max |a|$ for given time, you could rescale to reduce time and Δy in the same proportion instead, but for reduced time, the minimum Δy is the catenary in velocity space, for which Δy (the potential energy of the catenary) is reduced less than proportionately, because the average value of velocity increases.

PROOF OF LEMMA 2 For variable position $x(t)$ in $C^{1,1}$ and variable acceleration $a(t)$ in L^∞ , we want to minimize

$$\int_0^T dt$$

subject to the constraints $\ddot{x} = a$ (a.e.) and $|a| \leq \sigma$. Since T is smooth in a , for some Lagrange multiplier $\lambda(t)$, a minimizer is a critical point for $\int H dt$ where

$$H = 1 + \lambda \cdot (a - \ddot{x}).$$

The Euler conditions of vanishing first variation (see, e.g., [M, 29.2]) say first that weakly

$$0 = \frac{\partial H}{\partial x} - \frac{d}{dt} \frac{\partial H}{\partial \dot{x}} + \frac{d^2}{dt^2} \frac{\partial H}{\partial \ddot{x}} = 0 - 0 - \ddot{\lambda},$$

so that $\lambda(t) = At + B$ for constant vectors A, B , and second that $0 \leq \partial H / \partial a$. Since a is constrained to lie in the disc of radius σ , this second condition just says that

$$a = -\sigma \lambda / |\lambda| = \sigma(At + B) / |At + B|.$$

Suppose that there were two fastest paths $x_1(t), x_2(t)$. Then their average $x_3(t)$ would also be a fastest path. Since the acceleration $a_3(t)$ must, like a_1 and a_2 , have constant length σ , $a_1 = a_2$ and $x_1 = x_2$.

PROPOSITION 1. *Given $\sigma > 0$ and points x_1, x_2, \dots, x_n in \mathbb{R}^2 and optionally velocities v_1, v_n , there is a fastest path from x_1 to x_n passing in order through x_2, \dots, x_{n-1} with initial velocity v_0 , final velocity v_n , and acceleration bounded by σ . The acceleration is continuous of magnitude σ , with at most one possible exception from x_k to x_{k+1} : it may flip direction between x_k and x_{k+1} or it may change discontinuously at x_k or x_{k+1} ; the former can occur only if a is otherwise constant along the segment (as it is on the last segment), the latter only if a is constant along both incident segments. (If no velocities are prescribed, we must assume that the points do not lie in order along a line, the one case in which arbitrarily small time is possible.)*

PROOF. Since the set of all possible velocities at the points x_i is compact, existence follows from Lemma 1. Lemma 2 implies the asserted regularity except at the points x_2, \dots, x_{n-1} . Free velocity at x_i adds a boundary term

$$\lambda \cdot \delta v|_{x_i^-}^{x_i^+}$$

to the first variation, so that the Lagrange multiplier λ is continuous at x_i . Therefore the acceleration $a = \lambda / |\lambda|$ is continuous at x_i unless $\lambda(x_i) = 0$, in which case a is constant on both incident segments. At x_n , $\lambda = 0$, so on the last segment $\lambda = B(t - t_n)$ (see Remarks after Lemma 2), and a is constant on the last segment, except possibly for a flip.

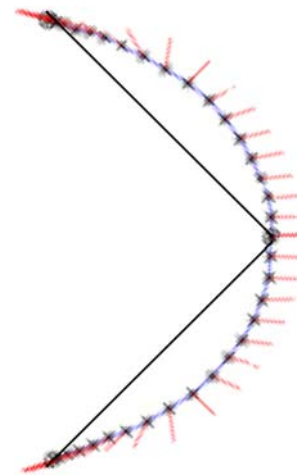


Figure 6. The fastest path to second base.

The fastest path (see Fig. 1). Computing the fastest path proceeds in two steps. First, for prescribed velocities at two sequential bases, we use Lemma 2, a finite difference boundary-value method [F, §14.2], and multidimensional Newton's method [F, §7.1], to find a solution with velocities that match the prescriptions. This problem can be highly nonlinear, and requires close guesses for Newton's method to converge, which we achieved by deforming an easily computed symmetric path. Second we minimize total time over varied choices of prescribed velocities at the bases as in the proof of Proposition 1, which we achieve with a gradient descent method [F, §7.2]. Since there is no general uniqueness result for relative minima, we cannot be sure that our solution reflects a global minimum. Our MATLAB code is given as an Appendix to the web version of this article, available at the blog entry at blogs.williams.edu/Morgan.

Figure 6 shows the fastest path to second base for a double, taking 10.4 seconds for $\sigma = 10 \text{ ft/sec}^2$, as compared to 12 seconds along the baseline, coming to a full stop at first and second base. The runner slows down a bit before rounding first base.

ACKNOWLEDGMENTS

This work stemmed from a Williams College undergraduate colloquium talk by Carozza advised by Morgan. Johnson,

who was in the audience, discovered the remarkably simple critical condition and computed the fastest path of Figure 1. Morgan acknowledges NSF support.

REFERENCES

- [F] Laurene V. Fausett, *Applied Numerical Analysis using MATLAB*, 2nd ed, Prentice-Hall, 2008.
- [G] Guinness World Records, http://www.baseball-almanac.com/recbooks/rb_guin.shtml.
- [M] Frank Morgan, *Real Analysis and Applications*, Amer. Math. Soc., 2008.
- [N] Johannes C. C. Nitsche, *Lectures on Minimal Surfaces*, Cambridge Univ. Press, New York, 1989.
- [S] <http://speedendurance.com/2009/08/19/usain-bolt-10-meter-splits-fastest-top-speed-2008-vs-2009/>
- [TF] Dao Trong Thi and A.T. Fomenko, *Minimal Surfaces, Stratified Multivarifolds, and the Plateau Problem*, American Mathematical Society, 1991.