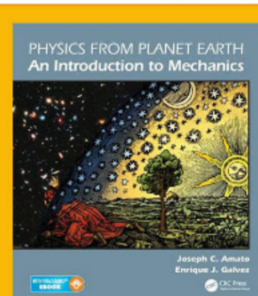
Using Statcast to lift the discussion of projectile motion

P. B. Siegel

Citation: American Journal of Physics **85**, 313 (2017); doi: 10.1119/1.4975302
View online: http://dx.doi.org/10.1119/1.4975302
View Table of Contents: http://aapt.scitation.org/toc/ajp/85/4
Published by the American Association of Physics Teachers



NOTES AND DISCUSSIONS

Using Statcast to lift the discussion of projectile motion

P. B. Siegel[a)]
Department of Physics and Astronomy, California State Polytechnic University Pomona, Pomona, California 91768

(Received 25 July 2016; accepted 17 January 2017)

Home run data from Major League Baseball's Statcast can be described by adding a lift force to the equations of projectile motion commonly used in undergraduate computational physics courses. We discuss how the Statcast data can be implemented in the classroom.

[http://dx.doi.org/10.1119/1.4975302]

I. INTRODUCTION

Projectile motion in sports is an interesting, realistic example that can bring enthusiasm to a physics classroom. In particular, the flight of a baseball has drawn the attention of many physics educators and scientists as evidenced by the numerous articles in this and other journals (see Refs. 1–8 to mention a few). Recently, Major League Baseball added a "Statcast" feature to their website that lists data on the flight of home runs, as well as other situations. We point out here how one can use this home run data in a computational physics course.

The data we examine are from Ref. 9, which are the longest 50 home runs from the 2015 postseason. The site lists the data in a convenient format that can be copied and pasted into a text file with four columns to be read and analyzed by a computer code. The data consist of the initial speed v_0, the launch angle θ_0, the maximum height $h_{\max}$, and the range R of the home run. The data span a fair spread of values with $395\,\text{ft} < R < 459\,\text{ft}$, $100.1\,\text{mph} < v_0 < 112.5\,\text{mph}$, $17.9^\circ < \theta_0 < 35.9^\circ$, and $51.6\,\text{ft} < h_{\max} < 136.0\,\text{ft}$. The v_0, θ_0, and $h_{\max}$ data are actually measured, and are given to the tenth's place in their respective units. The range R data are given to the nearest foot and represent the projected distance since the ball usually lands in the stands.[10]

II. MODELING THE TRAJECTORY

The home run data can be modeled using ideas from Refs. 1–8. The force that the air exerts on a flying baseball can be separated into a component opposite to the direction of the velocity $\vec{v}$, and one perpendicular to the velocity. The component opposite to $\vec{v}$ is referred to as drag. It is common to include this air frictional drag force in numerical exercises, and to take its magnitude proportional to the speed squared (v^2). The component of force perpendicular to $\vec{v}$ is due to the Bernouli effect and is referred to as the Magnus force; its direction is perpendicular to $\vec{v}$ and $\vec{\omega}$, where $\vec{\omega}$ is the angular velocity vector of the ball. For a classroom treatment, we use the most basic equations that best describe the motion. Since we are trying to determine two data points, R and $h_{\max}$, we should have at most two free parameters in our equations. For these considerations, we use the following ansatz:

$$a_x = -v_x g \frac{v}{v_T^2} - l_T g \frac{v_y}{v_T}, \tag{1}$$

$$a_y = -g - v_y g \frac{v}{v_T^2} + l_T g \frac{v_x}{v_T}, \tag{2}$$

where a_x and a_y are the acceleration components, and v_x and v_y the components of the velocity in the x and y directions. (The y direction is vertical and the x direction is horizontal, pointing in the direction the ball is traveling.) The acceleration g due to gravity enters only in the equation for a_y. The parameters v_T and l_T take into account air friction and the Magnus force as follows.

The force of air resistance, or drag, is $|\vec{F}_d| = F_d = \frac{1}{2}C_D\rho A v^2$, where ρ is the density of air, A is the cross section of the baseball, and C_D is the drag coefficient. The drag force is in the $-\vec{v}$ direction and can be expressed in terms of the terminal speed v_T. At the speed v_T, the drag force equals the object's weight: $mg = (1/2)C_D\rho A v_T^2$. The resulting expression for the magnitude of the drag force is $F_d = mgv^2/v_T^2$; the force itself is therefore

$$\vec{F}_d = -\frac{mgv}{v_T^2}\left(v_x\hat{i} + v_y\hat{j}\right). \tag{3}$$

The magnitude of the Magnus force is $|\vec{F}_M| = F_M = (1/2)C_L\rho A v^2$, where C_L is the lift coefficient. The constant C_L depends directly on the spin factor $S \equiv r\omega/v$, where r is the radius of the ball. Consequently, to a good approximation F_M is proportional to ωv. The vector $\vec{F}_M$ can be decomposed into a component that is horizontal, $\vec{F}_H$, and one that lies in the vertical plane, $\vec{F}_l$. The component $\vec{F}_H$ can make the baseball curve left or right, while $\vec{F}_l$ can cause the ball to rise or sink. It is the $\vec{F}_l$ component that will affect $h_{\max}$ and R the most, and we only include this component in the equations of motion. We also assume that $\vec{\omega}$ is constant, or equal to its average value, throughout the flight. With these assumptions, the lifting component of the Magnus force will be proportional to v and can be included using only a single parameter, which we will call l_T. It is convenient to use the terminal speed in the parameterization by writing $|\vec{F}_l| = l_T mgv/v_T$. As such, the parameter l_T is the ratio of F_l to the ball's weight when the ball is traveling at

Table I. Results using typical values for the home runs we analyzed. In each case, $v_0 = 105$ mph and $\theta_0 = 28°$. To assist readers in checking their code, the Euler method was used with a time step of 0.01 s and the range R is the value of x when y becomes negative.

v_T (mph)	l_T	R (ft)	$h_{\max}$ (ft)
78	0.5	396	87
83	0.5	418	89
78	0.7	417	102

v_T. The vector $\vec{F}_l$ lies in the xy-plane and will be perpendicular to $\vec{v}$; it can therefore be written

$$\vec{F}_l = \frac{mgl_T}{v_T}\left(-v_y\hat{i} + v_x\hat{j}\right). \tag{4}$$

The acceleration a_{lift} caused by $\vec{F}_l$ is the last term in Eqs. (1) and (2). One can check that $|\vec{a}_{\text{drag}}| = gv^2/v_T^2$, $|\vec{a}_{\text{lift}}| = l_T g$ when $v = v_T$, and that $\vec{a}_{\text{lift}} \cdot \vec{v} = 0$ in these equations, as expected.

III. NUMERICAL METHODS AND RESULTS

Equations (1) and (2) can be solved numerically using finite difference methods, and without the a_{lift} terms are classic examples. We solved these equations using the Euler method with a time step of 0.01 s. The initial position of the ball was taken to be $x_0 = 0$ and $y_0 = 1$ m, assuming the batter hits the ball on average when it is around a meter above home plate. The initial velocity in the x-direction is $v_{x0} = v_0 \cos\theta_0$ and in the y-direction $v_{y0} = v_0 \sin\theta_0$. We determined the range R as the value of x when y becomes negative. For more accuracy, one can interpolate linearly between the positive and negative values of y to find where $y = 0$. The range could differ be as much as one foot, since the distance covered in 0.01 s by an object traveling at 100 mph is around 0.4 m. Typical values for the parameters of the home runs we analyzed are listed in Table I.

There are two parameters that determine the fate of the ball: v_T and l_T. One can match the range R using only v_T and setting $l_T = 0$. However, for every home run, the predicted $h_{\max}$ is well below the data. In fact, for 47 of the 50 home runs and all home runs where $\theta_0 < 31.5°$, the maximum height is larger than $v_{y0}^2/2g$, the value obtained neglecting the effect of the air. The trajectory of these 50 home runs cannot be accounted for without some lift. So, one has two parameters, v_T and l_T, to fit two data points, $h_{\max}$ and R. The students can determine v_T and l_T for each home run and examine if the values are consistent and reasonable. For some home runs, Statcast lists an estimated hang time. For these cases, the students can check their predictions, although the data are only approximate.

Values for v_T and l_T that "best fit" each home run can be carried out by varying the parameters in expected ranges to minimize a χ^2 function. Using nested loops, we varied v_T from 70–100 mph in increments of 1 mph, and varied l_T from 0.2–1.0 in increments of 0.01. Our χ^2 function was $\chi^2 = (R_{\text{calc}} - R_{\text{data}})^2 + (h_{\max,\text{calc}} - h_{\max,\text{data}})^2$. The grid sizes of ± 1 mph for v_T and ± 0.01 for l_T were small enough to match, within one foot, the range and maximum height data. One could search a smaller grid for better predictions, however, the model is too crude to justify more accuracy.

The values we obtained for the longest 50 home runs for the 2015 postseason were as follows. The terminal speeds v_T were all in the interval 75 mph $< v_T <$ 88 mph, with an average of 81.5 mph and a standard deviation of $\sigma = 3.3$ mph. The terminal speed will depend upon air density and vary from park to park, though we note that none of the home runs in the 2015 post season were hit at high elevation. The relation $mg = (1/2)C_D\rho A v_T^2$ can be used to obtain C_D from the terminal speed. A value of $v_T = 81.5$ mph yields a value for C_D of 0.41. The values for l_T were all in the interval $0.42 < l_T < 0.71$, with an average of 0.56 and a standard deviation of $\sigma = 0.08$. Since $C_L = 2l_T mg/\rho A v_T^2 = l_T C_D$, the data have an average value for C_L of approximately 0.23. This value for C_L results in a spin factor[1,4] of around 0.25 or a rotation rate of $\approx$2400 rpm at v_T. Thus, both C_D and C_L are "in the ball park" of accepted values.[11]

The results can lead to interesting classroom discussions. How accurate are the equations? How much could ω change during the flight? What considerations were not included? How much might wind affect the range? What variation could v_T have at sea level? What is v_T for home runs hit in Denver? What rotation rates would give values of l_T between 0.4 and 0.7, and how reasonable are these rates? The students can speculate about how far the home run ball would travel without the Magnus force, and so on. Answers to some of these questions can be found in the references. Every year there will be new home runs for the next class to analyze, and perhaps Statcast will include accurate estimates of hang time. We hope the Statcast home run data are a hit with the students, and lift their interest in the Bernoulli effect in baseball as well as other sports.

[a]Electronic mail: pbsiegel@cpp.edu

[1]Alan M. Nathan, "The effect of spin on the flight of a baseball," Am. J. Phys. **76**, 119–124 (2008); also see <http://baseball.physics.illinois.edu/>.

[2]A. F. Rex, "The effect of spin on the flight of batted baseballs," Am. J. Phys. **53**, 1073–1075 (1985).

[3]R. G. Watts and S. Baroni, "Baseball-bat collisions and the resulting trajectories of spinning balls," Am. J. Phys. **57**, 40–45 (1989).

[4]G. S. Sawicki, M. Hubbard, and W. Stronge, "How to hit home runs: Optimum baseball bat swing parameters for maximum range trajectories," Am. J. Phys. **71**, 1152–1162 (2003).

[5]L. W. Alaways and M. Hubbard, "Experimental determination of baseball spin and lift," J. Sports Sci. **19**, 349–358 (2001).

[6]Rod Cross, "Measuring the effects of lift and drag on projectile motion," Phys. Teach. **50**, 80–82 (2012); also see <http://www.physics.usyd.edu.au/~cross/baseball.html>.

[7]David Kagan, "What is the best launch angle to hit a home run?," Phys. Teach. **48**, 250 (2010); also see <http://phys-webapps.csuchico.edu/baseball/>.

[8]Michael K. McBeath, Alan M. Nathan, A. Terry Bahill, and David G. Baldwin, "Paradoxical pop-ups: Why are they difficult to catch?," Am. J. Phys. **76**, 723–729 (2008).

[9]The URL for the home run data can be found at <http://m.mlb.com/statcast/leaderboard#hr-distance>. For this article, we used the 2015 post-season data.

[10]To determine the projected HR distance, the MBL Statcast website "[c]alculates the distance of projected landing point at ground level on over-the-fence home runs."

[11]The coefficient C_D for a baseball does depend on the speed and rotation rate of the ball. From Ref. 4, most measurements of C_D lie between 0.35 and 0.5. References 1 and 4 show plots relating a value for C_L of 0.23 to a spin factor of approximately 0.25. The average rotation rate for major league pitchers is around 2240 rpm.