Analysis of Baseball Trajectories

Alan M. Nathan*
Department of Physics, University of Illinois, Urbana, IL 61801
August 28, 2017

Fly ball trajectories games played during the 2015-2017 seasons at Tropicana Field are investigated. A randomly selected half of the 2016 data are used to establish a model for the drag and lift coefficients. That model is then used to calculate the trajectories for the remaining data, with fixed initial conditions determined by Statcast. The calculated distances are then compared to the actual distances.

I. INTRODUCTION: THE HOME RUN SURGE

There has been a marked increase in the rate of home run production in MLB starting approximately with the All-Star break in 2015. Three possible reasons have been proposed for the surge:

1. The Coefficient of Restitution (COR, or "bounciness") of the ball has increased, resulting in harder-hit ball (i.e., higher exit velocity). All other things equal, a higher exit velocity will result in longer fly ball distances and therefore more home runs.
2. The properties of the baseball affecting its aerodynamic properties have changed, resulting in better fly ball "carry" and therefore longer distances and more home runs.
3. The batters are consciously trying to hit more home runs by swinging the bat harder and altering the swing plane to elevate the batted ball. The former would result in higher exit velocities. The latter would result in more balls hit in the range of vertical launch angles that result in the longest distances, resulting in more home runs.

In this article, I will only consider the 2nd of these reasons. In particular, I will use actual fly ball trajectories from MLB games to determine whether or not the "carry" on a fly ball has changed in a way that would lead to more home runs over the period 2015-2017, as suggested by Lindbergh[1] and Arthur.[2] I will start with some physics background before discussing the actual analysis.

II. PHYSICS BACKGROUND

A. Drag and Lift

When a baseball travels through the air, it experiences various forces, shown in Fig. 1, and it is these forces that determine the trajectory of the baseball. The most familiar of these forces is the downward pull of gravity F_G. Less familiar are the aerodynamic forces, namely the drag force F_D and the Magnus force F_M. The drag force, or in everyday language "air resistance," is due to the fact that the ball has to push the air out of the way, thus slowing down the ball. The conventional way to express the magnitude of F_D is through the expression

$$\vec{F}_D = -\frac{1}{2}\rho A C_D v^2 \, \hat{v} \, , \tag{1}$$

where A is the cross sectional area of the ball and ρ is the density of the air. The direction of the drag is exactly opposite to the direction of the velocity (the -$\hat{v}$ direction), so that the force always retards the motion. The factor C_D is called the drag coefficient. If the baseball is spinning, it also experiences the Magnus force F_M, which is conventionally written as

$$\vec{F}_M = \frac{1}{2}\rho A C_L v^2 \, (\hat{\omega} \times \hat{v}), \tag{2}$$

where C_L is called the lift coefficient. The direction of the Magnus force is always perpendicular both to the velocity and the spin axis and is in the direction that the leading edge of the ball is turning.

One final note: In both Eqs. 1 and 2, the velocities are actually the velocity of the ball with respect to the air. It is identical to the velocity with respect to the ground if there is no wind.

B. Coordinate System

The origin of the coordinate system is at the point of home plate, $\hat{y}$ points towards the pitcher, $\hat{z}$ points vertically upward, and $\hat{x} = \hat{y} \times \hat{z}$ (i.e., the x axis points to the catcher's right). In that coordinate system, the instantaneous velocity $\hat{v}$ is most easily expressed in terms of a speed v, a vertical angle θ, and a horizontal angle ϕ, where $\phi = 0$ points towards 2B, $\phi = 45°$ points toward 1B, and $\phi = -45°$ points

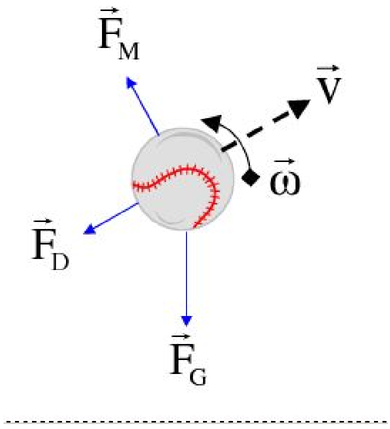

FIG. 1: Forces on a baseball in flight, including gravity (F_G), drag (F_D), and the Magnus force (F_M).

toward 3B. Therefore

$$\begin{aligned} v_x &= v\cos\theta\sin\phi \\ v_y &= v\cos\theta\cos\phi \\ v_z &= v\sin\theta \end{aligned} \tag{3}$$

Recalling that the y axis points from home plate toward 2B, v_y is positive for a batted ball and negative for a pitched ball. In the latter case, $\phi \approx 180°$.

C. Spin Direction

To solve the equations of motion, the components of the spin in the x,y,z directions are needed. It is actually more intuitive to decompose the spin into backspin (ω_b), sidespin (ω_s) and gyrospin (ω_g) components, where $\hat{\omega_g}$ is along the initial velocity direction, $\hat{\omega_b}$ is perpendicular to the initial velocity direction and in the horizontal plane, and $\hat{\omega_s} = \hat{\omega_b} \times \hat{\omega_g}$. Therefore after a bit of algebra, we arrive at

$$\begin{aligned} \omega_x &= \omega_b\cos\phi - \omega_s\sin\theta\sin\phi + \omega_g\cos\theta\sin\phi \\ \omega_y &= -\omega_b\sin\phi - \omega_s\sin\theta\cos\phi + \omega_g\cos\theta\cos\phi \\ \omega_z &= \omega_s\cos\theta + \omega_g\cos\theta \end{aligned} \tag{4}$$

Note that the total spin $\omega = \sqrt{\omega_x^2 + \omega_y^2 + \omega_z^2} = \sqrt{\omega_b^2 + \omega_s^2 + \omega_g^2}$. The sign convention I am using is as follows:

- ω_b is positive for batted ball with backspin; opposite for a pitched ball
- ω_s is positive for a batted ball that breaks in the -x direction; opposite for a pitched ball
- ω_g is generally negative for a pitch thrown by a RHP, positive for LHP; it is generally 0 for a batted ball

In the trajectory analysis to follow, I will assume $\omega_g = 0$. Further, I will define the spin axis ϕ_s as follows:

$$\phi_s = \arctan(\omega_s/\omega_b) \tag{5}$$

As discussed in Sec. III, Trackman measures the total spin ω, from which the backspin and sidespin components can be derived from

$$\omega_b = \omega\cos\phi_s \qquad \omega_s = \omega\sin\phi_s \tag{6}$$

D. Equations of Motion

Given the forces, Eqs. 1-2, the equations of motion can be written as follows:

$$\begin{aligned} \frac{dv_x}{dt} &= -KC_D v v_x + KC_L v(\omega_y v_z - \omega_z v_y)/\omega \\ \frac{dv_y}{dt} &= -KC_D v v_y + KC_L v(\omega_z v_x - \omega_x v_z)/\omega \\ \frac{dv_z}{dt} &= -KC_D v v_z + KC_L v(\omega_x v_y - \omega_y v_x)/\omega - g\,, \end{aligned} \tag{7}$$

where g is the acceleration due to gravity (32.174 ft/s^2). The factor K is given by

$$K = \frac{1}{2}\frac{\rho A}{m}\,, \tag{8}$$

where m is the mass of the ball. MLB specifies that m is in the range $5-5\frac{1}{4}$ oz and that the circumference C is in the range $9-9\frac{1}{4}$ inches. Moreover, the density of air is nominally 1.225 kg/m^2 (or 0.0767 lb/ft^3). Therefore, K can be expressed as follows:

$$K = 5.509\times10^{-3}\,\text{ft}^{-1}\left[\frac{C}{9\frac{1}{8}\,\text{in}}\right]^2\left[\frac{m}{5\frac{1}{8}\,\text{oz}}\right]\left[\frac{\rho}{1.225\,\text{kg/m}^3}\right]. \tag{9}$$

In the equations of motion, I have implicitly assumed that both the spin rate ω and the spin axis ϕ_s are constant throughout the trajectory. Not a lot is known about either spin decay (which affects ω) or spin precession (which affects ϕ_s). Some years ago, I did some estimates of the spin decay based on measurement on golf balls,[3] concluding that the time constant is over 20 sec. Since then, unpublished measurements using Trackman put the time constant at closer to 30 sec. To my knowledge, there are no studies of spin precession on a baseball. For the present analysis, I simply ignored both spin decay and spin precession.

With Eq. 7, along with initial values for position (x_0, y_0, z_0) and velocity $(v_{x,0}, v_{y,o}, v_{z,0})$ and the spin components (Eq. 4), the full trajectory $(x(t), y(t), z(t))$ can be found using the Runge-Kutta technique. As remarked earlier, the velocities in the equations of motion are respect to the air.

III. THE TRACKMAN SYSTEM

The MLB Statcast system is composed of two elements:

- The Trackman phased-array Doppler radar, which is used to track both the pitched and batted baseball.
- The ChyronHego stereo cameras, which are used primarily to track the players on the field

For present purposes, only Trackman is important. My web site[4] contains a number of links to articles about how Trackman works,[5] including how well it works.[6] For our purposes, the following is a summary of the important points:

- With some notable exceptions,[6] Trackman measures the full trajectory of batted balls, from impact to landing (i.e., x, y, z as a function of t). The data are smoothed and reported in steps of 0.01 sec. In many cases, the trajectory is only partially measured and proprietary techniques are used to extrapolate to ground level.
- Trackman measures the initial velocity vector (speed, vertical launch angle, and horizontal direction)
- Trackman measures the initial spin of the batted ball
- Trackman does not measure but infers from the resulting trajectory the initial spin axis of the batted ball.

IV. TRAJECTORY ANALYSIS

The goal of the analysis is to investigate the aerodynamic properties of the baseball from analysis of fly ball trajectories from the 2015-2017 seasons. As is evident from Eqs. 7-8, the drag and lift both depend on the quantities $\rho A/m$ times either C_D or C_L. Of these quantities, all but the air density ρ are properties of the baseball. If the goal is to determine whether the aerodynamic properties of the baseball have changed, it is important either to know the air density (and wind velocity) for every fly ball or to keep it constant. Since it is not very easy to know the former, I chose the latter by only investigating fly balls hit in Tropicana Field. Since the Trop is a covered stadium, the atmospheric effects can reasonably be expected to be constant, with no wind. I estimate the air density at the Trop to be 1.194 kg/m^3. In future investigations, one might also consider venues with retractable roofs, choosing data when the roof is known to be closed.

The properties of the ball that matter for drag (and similarly for lift) are $C_D A/m$. In the analysis that follows, I assumed nominal values for A and m, so that any variation in the drag properties are absorbed into C_D (and for lift, C_L). This process is necessary since we don't actually know the exact area and mass of each ball.

The trajectories and related Trackman information for all balls hit with launch angles in the range $10^\circ - 50^\circ$ and exit velocities greater than 60 mph in Tropicana Field during the 2015-2017 seasons were obtained. Not wanting to deal with only incomplete trajectories, only those chosen for study that were tracked to 90% of their extrapolated distance were marked for further study. Five separate data sets were used in the study:

- **y15E**: 2015 pre-ASG
- **y15L**: 2015 post-ASG
- **y16b**: half of the 2016 data, randomly selected as training data
- **y16a**: the remaining half of the 2016 data
- **y17E**: 2017 pre-ASG

The technique involved two distinct steps:

1. Use the **y16b** data to develop a model for the lift and drag coefficients, with parameters of the model adjusted to best fit those trajectories. I call this step the "Training Step".
2. With the model fixed from the previous step, calculate the trajectories for the remaining data sets using the Trackman initial conditions (velocity vector and spin rate), but with spin axis ϕ_s adjusted to best fit the data. Then compare the calculated fly ball distances with the actual distances measured by Trackman. I call this step the "Analysis Step".

In all steps involving fitting, the Levenberg-Marquardt nonlinear least-squares algorithm was used, where the fitting function was the numerical solution of the equations of motion using the RK4 technique.

A. Training Step

The first part of the training step was to establish the kinematic variables that C_D and C_L depend on. On theoretical grounds, and guided by similar studies on golf balls, one might expect C_D to depend on the instantaneous speed v (or, better, Reynold's number) and possibly on spin rate ω. Similarly, C_L is expected to depend on the spin factor $S = R\omega/v$, where R is the radius of the ball. To establish the functional dependences of C_D and C_L, the **y16b** data set were initially fitted over the first 0.5 seconds of the trajectory, with C_D, C_L, and ϕ_s as the fitted parameters. The results lead to functional forms as follows:

$$\begin{aligned} C_D &= C_{D,0} + C_{D,1}\left[\frac{\omega}{1000\,\mathrm{rpm}}\right] \\ C_L &= \frac{C_{L,2}S}{C_{L,0} + C_{L,1}S} \end{aligned} \tag{10}$$

Interestingly and perhaps surprisingly, after accounting for the dependence of C_D on spin, there was no further dependence on v, at least for speeds in the range 60-110 mph.

Next, the full trajectories of the **y16b** data set were fitted, with the five parameters of Eq. 10 as the fitting variables. Since we are primarily interested in fly balls that would lead to home runs, only trajectories with exit speeds at least 90 mph and launch angles in the range 20°-35° were included. Each trajectory was fitted individually and the optimum values of the five parameters were found. Then these optimum values were averaged over all trajectories to obtain the following values, which are taken to be the best description of the **y16b** data:

$$\begin{aligned} C_{D,0} &= 0.297 \;\; C_{D,1} = 0.0292 \\ C_{L,0} &= 0.583 \;\; C_{L,1} = 2.333 \;\; C_{L,2} = 1.120 \end{aligned} \tag{11}$$

B. Analysis Step

Using the parametrization of the drag and lift coefficients Eq. 10, the fitted parameters Eq. 12, and the initial velocity vector and spin from Statcast, each trajectory was calculated and the distance D_c calculated and compared to the actual distance D_a. For each data set, the mean difference $\Delta D = D_a - D_c$ was found along with its standard error, and those results are shown in Fig. 2. By construction, $\Delta D \approx 0$ for the training set, y16b. It is also close to zero for the **y16a** and **y17E** sets, but distinctly less than 0 for the **y15E** and, to a lesser extent, the **y15L** sets. A negative mean difference indicates that the calculated distance is larger than the actual distance, suggesting that the actual drag coefficient in 2015 is larger than in subsequent years; or equivalently, the drag on a baseball got smaller starting sometime in 2015 and continuing in subsequent years. This reduced drag is partially responsible for the increase in home runs since 2015. However, given that 2017 looks pretty much like 2016, the increase in home runs in 2017 relative to 2016, is likely due to something else. Table I presents the same results in tabular form, showing that the upward shift in distances starting in 2016 is statistically significant. A comparison of histograms of ΔD between y15E and all data from 2016 and 2017 is shown in Fig. 3.

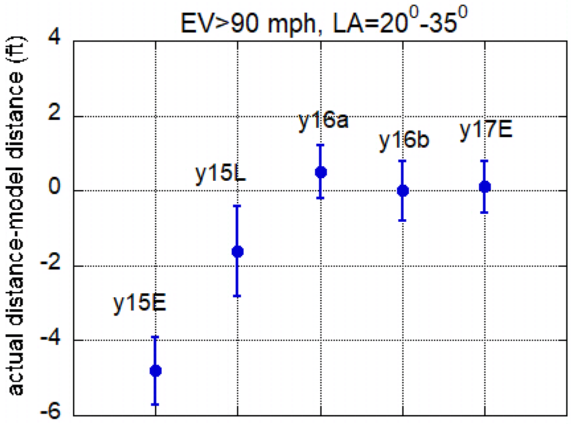

FIG. 2: Mean differences between the actual and calculated fly ball distance for each data set.

data set	N	$\Delta D \pm err$
y15E	133	-4.8±0.8
y15L	100	-1.6±1.2
y16a	169	0.5±0.7
y16b	160	0.0±0.8
y17E	199	0.1±0.7
≥2016	528	0.2±0.4

TABLE I: Mean difference between actual and calculated distance along with the standard error, where N is the number of fly balls in each data set. The last line includes all data from 2016 and 2017.

V. SUMMARY OF ASSUMPTIONS

It is useful to summarize here the assumptions about lift and drag used in this analysis.

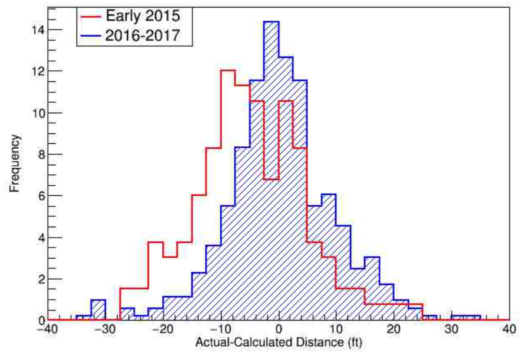

FIG. 3: Histogram of actual minus calculated distances, comparing **y15E** (red) and all data from 2016-2017 (shaded). These data show a clearcut shift to larger distances of the latter relative to the former by approximately 5 ft.

- The drag coefficient is independent of velocity but has a linear dependence on the spin rate
- The spin vector, both the magnitude and direction, is constant throughout the trajectory
- The lift coefficient depends on the spin factor $S = R\omega/v$, which is not constant since v is not constant
- The spin axis is initial perpendicular to the velocity vector

VI. THE TRAJECTORY CALCULATOR

I have created an Excel spreadsheet[7] for calculating baseball trajectories based on the model and assumptions for drag and lift described here.

* Electronic address: a-nathan@illinois.edu

[1] https://www.theringer.com/2017/6/14/16044264/2017-mlb-home-run-spike-juiced-ball-testing-reveal-155cd21108bc.
[2] https://theringer.com/2017-mlb-home-run-rate-is-the-ball-juiced-report-results-6e1dd0233203.
[3] http://baseball.physics.illinois.edu/spindown.pdf.
[4] http://baseball.physics.illinois.edu/trackman.html.
[5] http://baseball.physics.illinois.edu/trackman/NathanTrackmanIntro.ppt.
[6] http://www.baseballprospectus.com/article.php?articleid=23202.
[7] http://baseball.physics.illinois.edu/TrajectoryCalculator-new-3D.xlsx.