A. Terry Bahill

The Science of Baseball

Batting, Bats, Bat-Ball Collisions, and the Flight of the Ball

Second Edition

Springer

The Science of Baseball

A. Terry Bahill

The Science of Baseball

Batting, Bats, Bat-Ball Collisions, and the Flight of the Ball

Second Edition

Springer

A. Terry Bahill
Systems and Industrial Engineering
University of Arizona
Tucson, AZ, USA

ISBN 978-3-030-03031-5 ISBN 978-3-030-03032-2 (eBook)
https://doi.org/10.1007/978-3-030-03032-2

Library of Congress Control Number: 2018958952

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Dedicated to my always smiling,
always laughing Karen

Foreword by Greg Rybarczyk

I first “met” Dr. Terry Bahill in 2005 while researching aerodynamic characteristics of batted baseballs as part of a personal project which would become the ESPN Home Run Tracker. I didn’t speak to him at the time (that would come later) but rather downloaded and read many of the papers which he had posted on his website. Dr. Bahill’s explanations and calculations were a great help to me at a time when my career in baseball analytics was just beginning, but as we’ve corresponded over the years, my admiration for his work, particularly his gift for communicating ideas, has only increased. His latest publication, *The Science of Baseball: Batting, Bats, Bat-Ball Collisions, and the Flight of the Ball* is a worthy contribution to his prodigious body of baseball research, compiled over four decades and presented with extraordinary clarity. It will serve as a valuable reference for scholarly fans, as well as baseball analysts who aspire to compete at the highest level.

Major League Baseball clubs are, as of 2018, in the midst of a revolution. The ranks of analysts employed by Major League Baseball clubs have swelled in recent years, as teams try to at least keep pace, and hopefully realize competitive advantages through the creative use of the data which is being generated and presented to teams at an unprecedented rate. Every MLB front office now employs people who scrutinize not only traditional statistics such as batting averages and home run totals, but also play-level summary metrics like pitch speed or batted-ball exit speed. The most analytically enthusiastic clubs study ball- and player-tracking data collected at rates as high as 100 data points per second, and disseminated by commercial vendors such as Baseball Info Solutions, Sport vision, Trackman, MLB Advanced Media and others. MLB’s demand for new forms of baseball analysis has inspired a large and rapidly growing pool of independent analysts who conduct research via publicly available sources, hoping to earn the opportunity to offer their services as consultants to or employees of Major League front offices. More people and companies are doing more baseball-related analytical work than ever before.

Throughout my dozen years of baseball-related work, both as an individual and in my current role as an analyst with the Boston Red Sox, I've found that the best research originated with people who possessed not only thorough baseball knowledge but also a solid understanding and a proper deference to the other governing principles of the situation under study. For contract and compensation issues, these principles are those of economics; for discretionary tactical moves such as stolen base or bunt attempts, or for pitch type selection, these principles are those of game theory; for issues related to the movement of the baseball, these principles are those of physics.

Unfortunately, too often these days we see analytical work that neglects, or even runs counter to, the underlying principles, because the analyst's mastery of the relevant principles is faulty or incomplete. For some, analysis of baseball data consists of arranging it in columns and performing statistical tests on it until something "pops." I was once offered a detailed analysis that rated elite closer Koji Uehara as the 16th best pitcher on the Red Sox roster, and further opined that his devastating splitter was among the weaker individual pitches on the entire team. After I stopped laughing, I asked a few questions and learned that these dubious results could be traced to a faulty premise about the value of pitch locations. It was, essentially, a lack of understanding of one of the most important elements of pitching analysis: how to judge the results of a pitch.

More knowledgeable analysts who are familiar with the applicable principles can better detect and avoid bad data, more efficiently set up and perform the most promising statistical tests, and can more reliably interpret the results. Dr. Bahill's expert dissection of the bat–ball collision (Chaps. 1–5) and the flight of pitched and batted baseballs through the air (Chap. 7) should be read by all who wish to enhance their expertise at analysis of ball-tracking data by first understanding why the baseball moves the way it does. Complete derivations have been provided for those who wish to delve deeply into the equations, but they need not present a persistent barrier to those readers who prefer to skim the line-by-line mathematics and skip ahead to the conclusions. A prime example is the sensitivity analysis presented in Chap. 7, which describes the change in batted-ball range which follows a given change in various inputs such as batted-ball speed, batted-ball spin or air density.

Baseball analysts past, present and future are indebted to Dr. Bahill for the efforts he has made to make understanding of the complex underlying physics of baseball accessible to all at each person's chosen level of detail. His precise yet eminently accessible explanations of the physics of the bat–ball collision and the flight of the ball are more useful than ever in an era when MLBAM's Statcast system tells 30 and 100 times per second **what** has happened but leaves to the observer the task of figuring out **why** it happened (which is, of course, the key to predicting what will happen in the future, the ultimate objective of all analysts). If you wish not only to understand the game of baseball better but to contribute to the body of knowledge of the game of baseball, read this book carefully, and then read it again. For the moment, knowledge of baseball physics can still differentiate an analyst from his or her peers but in the field of baseball analytics, no competitive advantage persists for long.

Southborough, MA, USA

Greg Rybarczyk
Senior Analyst, Baseball R&D
Boston Red Sox
Creator of ESPN HR Tracker

P. O. BOX 481
ISLAMORADA, FLORIDA KEYS
33036

January 23, 1984

Prof. A. Terry Bahill
Electrical & Computer Engineering
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dear Mr. Bahill:

Received your letter and have also had a chance to read your research, and I fully agree with your findings.

I always said I couldn't see a ball hit the bat except on very, very rare occasions and that was a slow pitch that I swung on at shoulder height. I cam very close to seeing the ball hit the bat on those occasions.

As to participating in your other experiments; at this time, I can't tell you that I can comply with your request.

Regarding the current theories of some of the present batting coaches (with which I absolutely disagree) to watch the ball go into the catcher's mitt - by doing that, you don't give yourself a chance to swing and open up properly. Try it yourself - look down at the plate and try to make a full swing. I hope you don't throw your back out of joint!

In any event, good luck with your projects.

Sincerely,

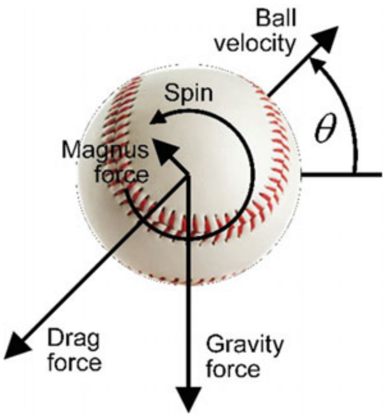
Ted Williams

Ted Williams

TW/shg

Preface

Collisions between baseballs, softballs and bats are complex and therefore their models are complex. The first purpose of this book is to show how complex these collisions can be, while still being modeled using only Newton's axioms and the conservation laws of physics. This book presents models for the speed and spin of balls and bats. These models and equations for bat–ball collisions are intended for use by high school and college physics students, engineering students, the baseball analytics community and most importantly nonprofessional students of the science of baseball. Unlike models in previous books and papers, these models use only simple Newtonian axioms and the conservation laws to explain simple bat–ball collision configurations. It is hoped that this book will help readers develop an understanding of the modeling of bat–ball collisions. The second purpose of this book is to help batters select or create baseball or softball bats that would be optimal for them. The third purpose is to show what affects air density and how air density affects the flight of the ball.



Chapter 1 lays the groundwork for analyzing bat–ball collisions and previews the theme that alternative models help you understand the system.

Chapter 2 introduces nine basic configurations of bat–ball collisions using words and figures.

Chapter 3 starts developing the equations for these configurations. It starts with the simple configurations having the ball collide with the center of mass of the bat. Then it moves on to configurations that are more complex using the same equations and development. The notation developed here will be used throughout the book.

Chapter 4 is the pinnacle of this book. It contains our most comprehensive model, which is a collision at the sweet spot of the bat with spin on the pitch. It shows which parameters are the most and least important. It also has advice for selecting and modifying each person's optimal bat. Such a bat does not have its barrel end cupped out. This chapter is unique in the science of baseball literature. It is also self-contained. You need not read previous chapters to understand it. In other words, a teacher could use this chapter in a physics or engineering course and the students would only have to buy this one chapter. The BaConLaws model presented in this chapter also describes the motion of the *bat* after the collision. Many models describe the motion of the ball after the collision but few (if any) describe the motion of the bat. When you see a batter hit a ball, do you see the recoil of the bat? Can you describe it? Well, these equations do.

Chapter 5 contains four alternative models for bat–ball collisions. Their purposes are different and are they based on different fundamental principles. The Effective Mass model was created by physicists independent of the author of this book. Therefore, comparisons to it are important for validating the model of Chap. 4. The second and third models are data-based, not theory-based. They use a different approach and they use a different *type* of data. The fourth model considers friction during the collision. It is shown that this type of collision cannot be modeled thoroughly using only the conservation laws. Our modeling technique could not handle the Collision with Friction model because our technique is only good for a point in time before the collision and a point after the collision: it cannot handle behavior during the collision. Chapter 4 fulfilled part of the first purpose of this book. It showed a complex configuration for which our technique did work. Chapter 5 completed the fulfillment of this purpose by showing a configuration for which our technique was too simple.

Nothing in Chaps. 1–5 is controversial. There are no unstated assumptions. Important equations have been derived with at least two techniques. In Chaps. 2–5, the equation numbers are the same. In other words, Eq. (2.3) is the same as Eq. (3.3) is the same as Eq. (4.3) and is the same Eq. (5.3). The equations in Chaps. 2–5 were derived using only Newton's axioms and the conservation laws of physics. The equations in Chap. 7 for the drag and Magnus forces are original and are based on more than Newton's' axioms.

Chapter 6 summarizes Chaps. 1–5. Chapters 1–6 deal with bat–ball collisions. They solve equations in closed form. There are no approximations. Chapter 7 deals with messy real systems. It uses experimental data and gives approximations.

Chapter 7 contains derivations for equations governing the flight of the ball. It shows what affects air density and how air density affects the flight of the ball. It shows that a home run ball might go 26 feet farther in Denver than in San Francisco. It also answers the question, “Which can be thrown farther a baseball or a tennis ball?” This chapter can be read independently from the rest of the book.

Chapter 8 discusses the accuracy of baseball simulations. When the television announcer says, for example, that home run went 431.1 feet. You, our reader, will know that he should have said, that the *true* range of that home run was 430 plus or minus 30 feet.

Chapter 9 presents the vertical sweetness gradient of the baseball bat. It shows that the sweet spot of the bat is one-fifth of an inch high.

Chapter 10 tackles the differences between right-handed batters and left-handed batters. It shows that neither is better than the other. Finally, it explains that cross-dominant batters *do* have an advantage on some pitches. Because for non-cross-dominant batters, the blind spot of their dominate eye can obscure the bat–ball collision.

Chapter 11 summarizes the insights and wisdom of the book. Chapter 12 presents our modeling philosophy.

We need people who can explain this book to baseball managers and general managers.

Teachers might challenge their students to try finding mistakes in this book. The author will give $25 to the first person/group to find a logical, algebraic, physics or engineering mistake in this book. Spelling, punctuation, grammar, fuzzy inconsistencies, typographical errors and broken links do not count. Send discoveries to Terry Bahill, 1622 W. Montenegro, Tucson AZ, USA 85704-1622.

Tucson, AZ, USA A. Terry Bahill

Acknowledgements

I am indebted to Al Nathan for preventing me from publishing a book with mistakes in it. Ferenc Szidarovszky ensured that the equations had no mistakes. I thank Bob Watts, Rod Cross, Bruce Gissing and Jim Close for helpful comments on the manuscript. This book is written in the first-person plural. Plural because my graduate students did all the work. Major contributions were made by Tom La Ritz, Bill Karnavas, Miguel Morna Freitas and J. Venkateswaran. Extra special thanks go to Dave Baldwin (17-year MLB pitcher with a 3.08 ERA) for inspiring my science of baseball papers.

Conflict of Interest

In the 1990s, Terry Bahill received research grants from Worth Sports Co. and Easton Sports Inc. for research on human–bat relationships. Since then he has received no money, renumeration, speaking fees, consulting contracts or research grants from any sources concerning the Science of Baseball. Bahill has no conflicts of interest regarding the content of this book.

Bahill had no preconceived notions of which models might be most appropriate for the Science of Baseball. He let the data, experiments and knowledge drive the development of this book.

Note on the Mathematics Used in This Book

Most of the mathematics in this book is simple high school algebra. If the reader prefers to just skip the equations, then my advice is to simply do so. Just read the text and you will still get a robust description of the dynamic physical interaction at work in hitting a baseball. The mathematics is there to illustrate the quantitative aspects of the phenomena described in this book.

Here is an example of a simple algebraic equation.

$$y = mx + b$$

This is the equation of a straight line. It has two variables and two constants. It says that the output variable, y, (plotted on the vertical axis in Fig. 1) is equal to the input variable, x, (plotted on the horizontal axis) multiplied by the slope, m, plus the intercept, b. A different way of writing this equation is

$$f(x) = mx + b$$

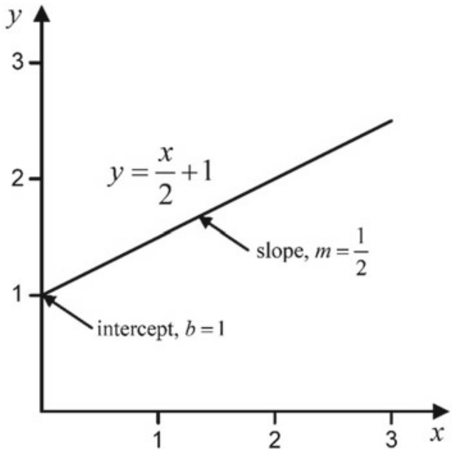


Fig. 1 Graph of a straight line

Here we have replaced y with $f(x)$, because we want to emphasize that the equation is a function of x. The naming is that the value of the dependent variable, y, depends on the value of the independent variable, x. Sometimes there can be two independent variables, like this:

$$x = w_y y + w_z z$$

This equation states that the variable x equals some weight w_y times y plus w_z times z. To emphasize the functionality, we could write it like this:

$$f(y, z) = w_y y + w_z z.$$

Variables and constants in equations are set in an *Italic font*.

OK, here is a big hairy equation. But it isn't scary.

$$v_{\text{ball-after}} = v_{\text{ball-before}} - \frac{(v_{\text{ball-before}} - v_{\text{bat-cm-before}})(1 + CoR_{1a})m_{\text{bat}}}{m_{\text{ball}} + m_{\text{bat}}}$$

It states that the velocity of the ball after its collision with the bat is equal to the velocity of the ball before its collision plus some other stuff. That other stuff includes the difference in velocities of the bat and ball before the collision times a bunch of constants. Now, that wasn't so scary, was it?

The next step up in mathematics is calculus and differential equations. However, whenever I used such equations, I 'hid them' from the reader—except for the partial differentials that were used in sensitivity functions.

Sometimes we want to know what change in the output would result from a small change in the input. For example, we want to know, if we change the input by a small amount, say Δx, what will be the change in the output, Δy. The Greek Delta, Δ, indicates change in a quantity expressed by a variable. Using simple rules that

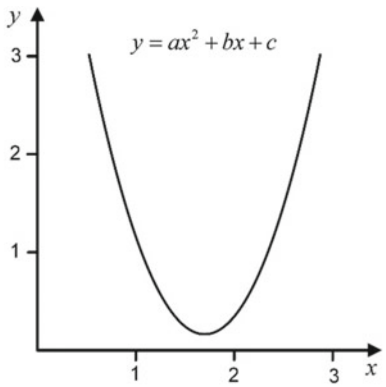


Fig. 2 Graph of a parabola

we look up in a table of derivatives on the Internet, we find that for this equation $y = mx + b$, $\frac{\Delta y}{\Delta x} = m$, the slope of the line. For small changes, we write $\frac{\Delta y}{\Delta x} = \frac{dy}{dx} = m$.

For a more complicated equation, let us consider the parabola

$$y = ax^2 + bx + c$$

shown in Fig. 2.

Once again, using rules that we look up in a table of derivatives, we find that

$$\frac{dy}{dx} = 2ax + b.$$

In some parts of this book, we use statistics. The following example will show the most common statistics that we use. Consider the following two sets of numbers: Set-1 = {3, 4, 5, 6 and 7} and Set-2 = {1, 3, 5, 7 and 9}.

	Set 1	Set 2
		1
	3	3
	4	
	5	5
	6	
	7	7
		9
Mean	5	5
Standard deviation	1.6	3.2

Both sets of numbers have the same mean or average. Set-1 is clustered, whereas Set-2 is spread out: it has more variation. A statistic that we use to model variation is the standard deviation. We put the numbers for Set-1 and Set-2 into a calculator and it produces the standard deviations in the above table. The standard deviation (or variance) is a measure of the spread or variation in the data. It is also used to prove that two distributions are statistically different.

We will now show an example of the most complicated mathematics that appears in this book. We will start with the equation for the ball velocity after the collision, v_{1a}, Eq. (4.8).

$$v_{1a} = v_{1b} - \frac{(v_{1b} - v_{2b} - \omega_{2b} d)(1 + CoR_{2b}) m_2 I_2}{m_1 I_2 + m_2 I_2 + m_1 m_2 d^2}$$

The subscript **b** is for *before* the bat–ball collision and **a** is for *after* the collision. The subscript 1 is for the ball and 2 is for the bat. To do a sensitivity analysis, we need the partial derivatives of v_{1a} with respect to the variables v_{1b} and v_{2b}.

To simplify let $K = \left(m_1 I_2 + m_2 I_2 + m_1 m_2 d^2\right)$
and $C = (1 + CoR_{2b}) m_2 I_2$
Therefore,

$$v_{1a} = v_{1b} - \frac{(v_{1b} - v_{2b} - \omega_{2b} d)C}{K}$$

The following partial derivatives of the function v_{1a} with respect to the variables v_{1b} and v_{2b} are easy to derive using a table of differentials. Substituting numerical values gives

$$\frac{\partial v_{1a}}{\partial v_{1b}} = 1 - \frac{C}{K} = 1 - 1.2 = -0.2$$
$$\frac{\partial v_{1a}}{\partial v_{2b}} = \frac{C}{K} = 1.2$$

This means that the velocity of the ball after the collision is influenced more by bat velocity before the collision than it is by ball velocity. That is the most complicated mathematics that we do in this book.

Just remember, the mathematics is there merely to prove that the text is correct. If you don't care about the proofs, then skip the equations. I wrote this book so that you can skip the equations and still understand the phenomena at hand using only the narrative description.

Contents

About the Author

A. Terry Bahill is an Emeritus Professor of Systems Engineering and of Biomedical Engineering at the University of Arizona in Tucson. He served as a Lieutenant in the United States Navy. He received his Ph.D. in electrical engineering and computer science from the University of California, Berkeley. He is the author of eight engineering books and over two hundred and fifty papers, over one hundred of them in peer-reviewed scientific journals. Bahill has worked with dozens of high-tech companies presenting seminars on Systems Engineering, working on system development teams and helping them to describe their Systems Engineering processes. He holds a U.S. patent for the Bat Chooser™, a system that computes the Ideal Bat Weight™ for individual baseball and softball batters. He was elected to the Omega Alpha Association, the systems engineering honor society. He received the Sandia National Laboratories Gold President's Quality Award. He is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE), of Raytheon Missile Systems, of the International Council on Systems Engineering (INCOSE) and of the American Association for the Advancement of Science (AAAS). He is the Founding Chair Emeritus of the INCOSE Fellows Committee. His picture is in the Baseball Hall of Fame's exhibition "Baseball as America." You can view this picture at http://sysengr.engr.arizona.edu/.

Acronyms

BA	Batting Average
BaConLaws	Baseball Conservation Laws
CoAM	Conservation of Angular Momentum
CoE	Conservation of Energy
CoM	Conservation of Momentum
CoR	Coefficient of Restitution
KE	Kinetic Energy
LHP	Left-Handed Pitcher
MLB	Major League Baseball
NCAA	National Collegiate Athletic Association
OPS	On-base Plus Slugging
RHP	Right-Handed Pitcher
SaD	Angle between Spin axis and Direction of motion
SaD Sid	Spin axis $\times$ Direction = Spin-induced deflection
VaSa	Angle between Vertical axis and Spin axis

List of Figures