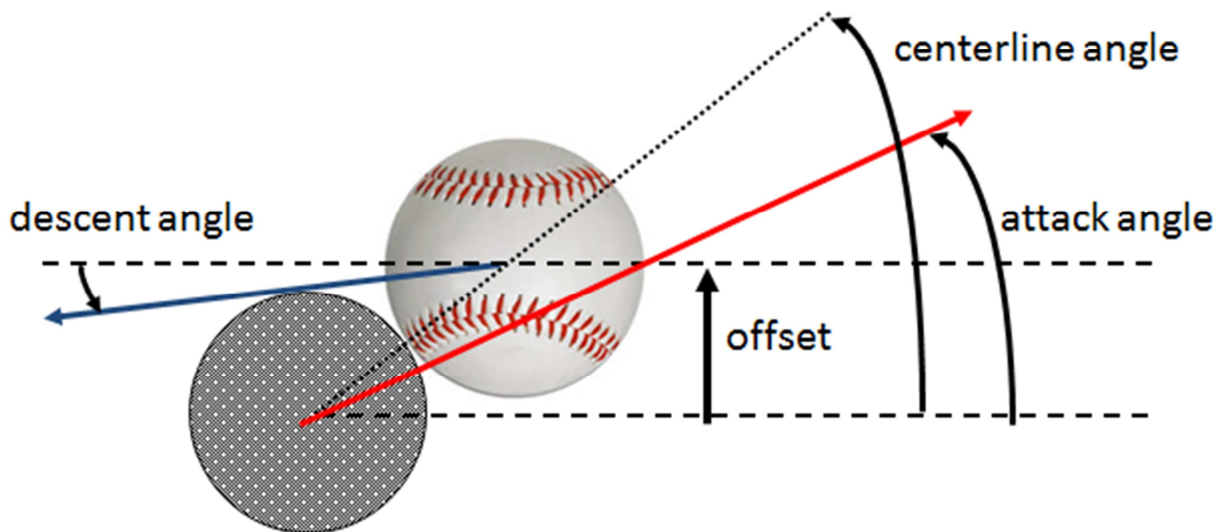


Optimizing the Swing, Part Deux:
Paying Homage to Teddy Ballgame
Alan M. Nathan
Department of Physics
University of Illinois at Urbana-Champaign

This is a follow-up to my article [Optimizing the Swing](#), which appeared in The Hardball Times on November 11. The initial motivation for the article was to determine whether a curveball can be hit farther than a fastball, assuming both are hit optimally. That led to a discussion of the parameters that define a batter's swing and how those parameters can be adjusted to lead to maximum distance. But there are other definitions of "optimum" besides maximum distance, such as maximum on-base probability. That topic will be the primary focus of this article.

Before going on, I want to remind you of the diagram in the earlier article showing the geometry of the ball-bat collision. Recall that the important parameters are the descent angle of the incoming pitch, the attack angle of the batter's swing (sometimes called the "swing plane"), and the offset. The latter is the vertical distance between the centers of the ball and bat and is a measure of the batter's aim. Also shown is the dotted centerline connecting the centers of the bat and ball, with an orientation (the centerline angle) determined uniquely by the offset. "Squared-up" collisions are those where the attack angle approximately coincides with the centerline angle.

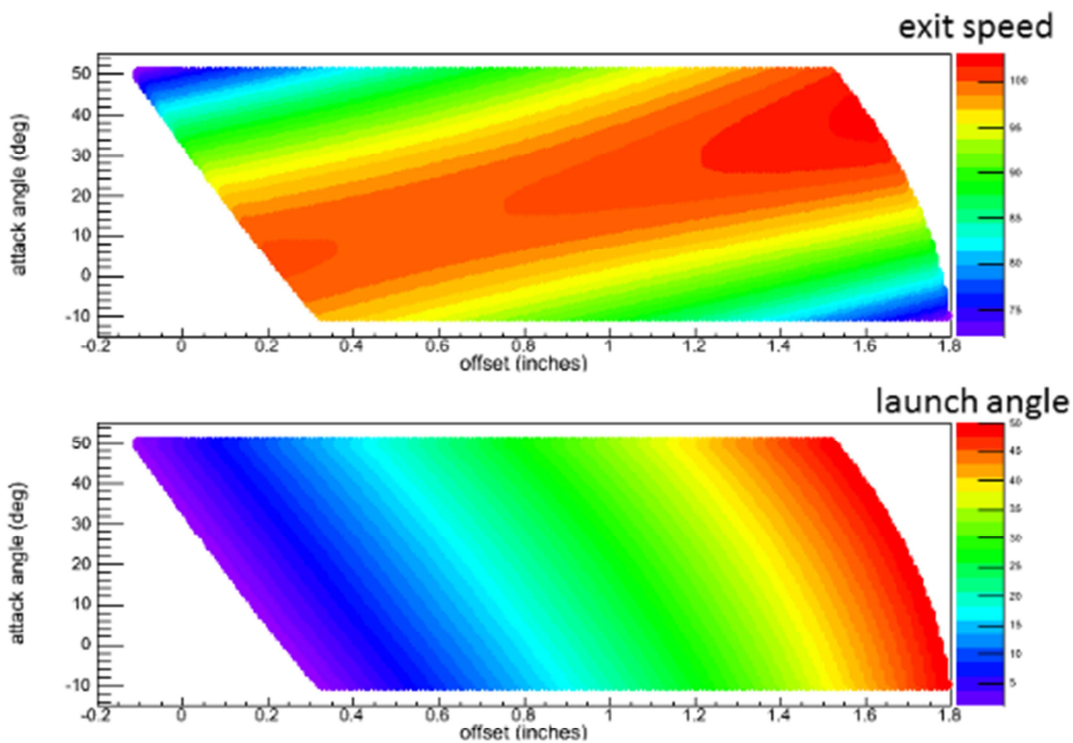


Aside from timing—i.e., getting the bat in the right place at the right time—the two parameters under the batter's control are the attack angle and the offset, both of which can be adjusted to

produce particular characteristics of the batted ball: exit speed, launch angle, and spin. In this article, I want to connect the batted ball characteristics with the probability of actual outcomes. The study is partly computational and partly statistical. The computational part utilizes a model for the ball-bat collision to calculate the exit speed and vertical launch angle for given pitch parameters (speed, spin, and descent angle) and for given swing parameters (attack angle and offset). It is the same model used in the first article. The statistical part uses MLB Statcast data to obtain the probability of any particular outcome for given values of exit speed and launch angle. For the purposes of this article, I will consider only a fastball, with a speed at contact of 86 mph, a backspin of 2000 rpm, and a descent angle of 6° . It would be very straightforward to apply the analysis to other pitch characteristics. As before, I assume a fixed bat speed of 70 mph.

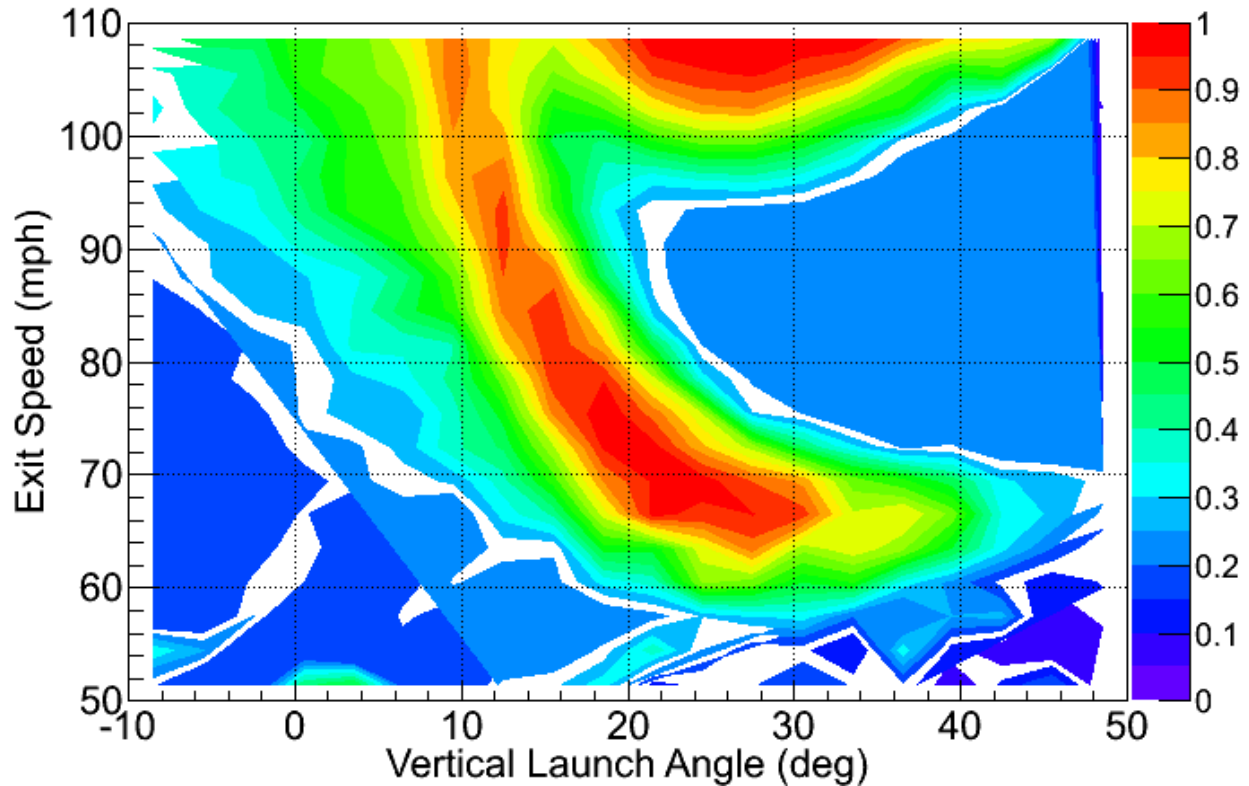
I want to emphasize several caveats in the analysis. First, I am ignoring other characteristics of the batted ball besides exit speed and launch angle, such as spin and spin axis. It's not that spin is not important. It's simply that there are not enough Statcast data available to determine with good statistical precision the dependence of outcome on spin for given values of exit speed and launch angle. The results I will present are therefore averaged over all possible spins. Hopefully a future analysis can include the spin variable. Second, as in the earlier article I will assume the bat is horizontal and that contact occurs at the front plane of home plate. The latter therefore assumes perfect timing—the bat and ball meet at a precisely defined point. Later in the article, I will explore the consequences of imperfect timing, where the batter is a little behind or a little ahead of the pitch. Third, I am assuming the contact occurs close to the sweet spot of the bat (although I forgot to state this assumption in the earlier article). Impacts off the sweet spot (or with a lower bat speed) will result in exit speed below those calculated in my model.

The primary results of the study are shown in a series of heat maps. The first plots show the results of the computational part of the study, in which exit speed (top) and launch angle (bottom) are given as a function of the two swing parameters, attack angle and offset. Lines of constant exit speed on the top plot are diagonals running from lower left to upper right. There is approximately a one-to-one correspondence between batted ball speed and the difference between the centerline and attack angles. The highest exit speeds (the red band) correspond to swings where the attack angle is approximately equal to the centerline angle. On the other hand, constant launch angles on the bottom plot are approximately diagonal lines extending from upper left to lower right. Actually the slope of these lines is such that the launch angle depends much more strongly on offset than it does on attack angle.



The second plot shows the results of the statistical study using Statcast data, in which on-base average is shown as function of exit speed and launch angle. The two regions of high on-base average are primarily red. One is the narrow strip extending from 110 mph/ 10° to ~ 66 mph/ 30° . The high exit speed part of this strip are hard-hit line drives with small launch angles, leading to a high on-base average. For the batter, these types of hits are highly desirable. On the other extreme, we have softly hit bloopers that manage to clear the infield but fall in front of the outfielders. While the outcome is still very favorable to the batter, it is not so clear to me that anyone actually tries to hit a ball like that. The second important region occurs at high exit speed and launch angles in the 20° - 35° range. These are largely extra-base hits, primarily home runs.

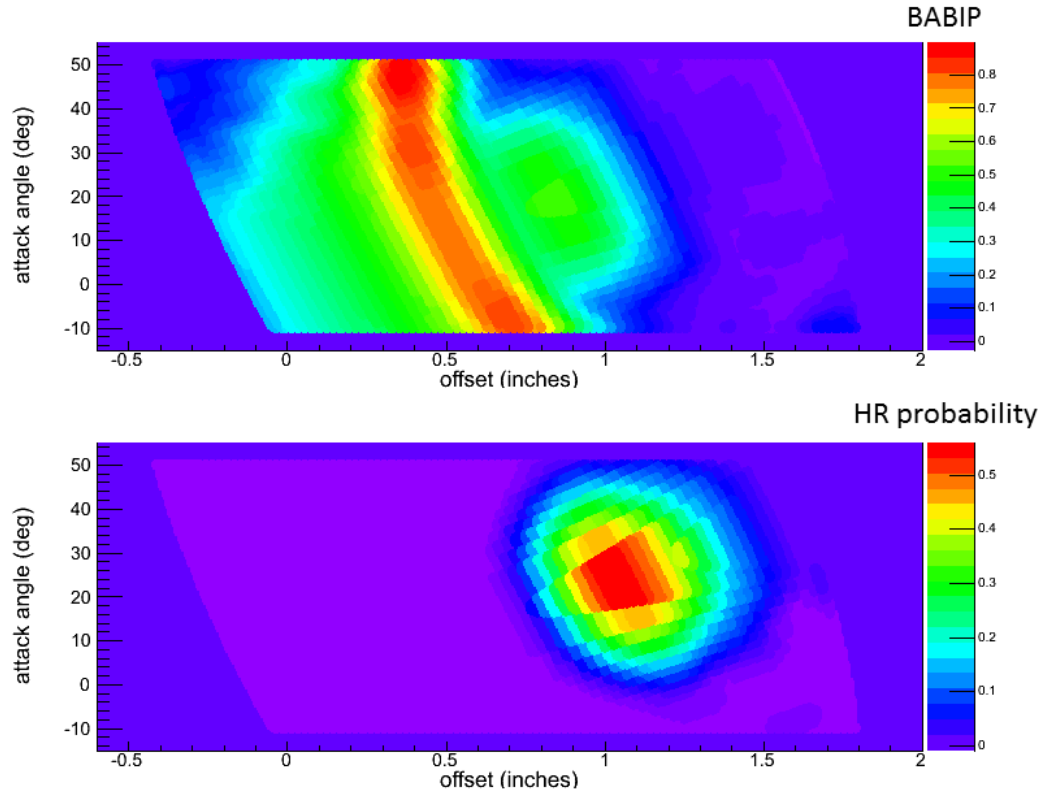
Batting Average for Batted Balls



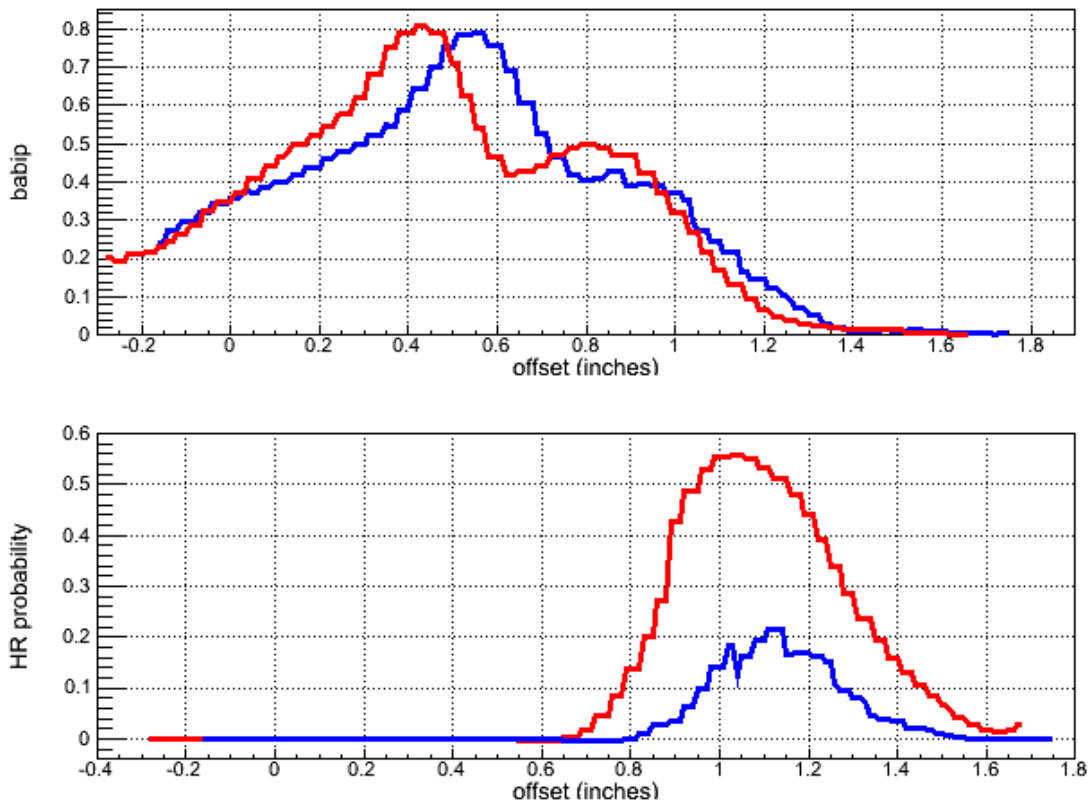
As an aside, it is interesting to look at balls hit with launch angles in the 20° - 30° range as a function of exit speed. At low speed we have bloop singles. At high speed we have home runs. In between, we have the “donut hole”, where the ball is hit at moderate exit speed but at too high launch angle to carry very far, leading to the proverbial “can of corn” for the outfielder.

Now let’s put it all together by looking at the third set of heat maps, which combines the previous two. As a function of attack angle and offset, the top plot shows BABIP (on-base average with home runs excluded) and the bottom home run probability. For BABIP, the red/orange strip results in an on-base average greater than about 0.75 and corresponds to launch angles in the range 9° - 17° . The optimum swing parameters for high BABIP are 10° and 0.5 inches for attack angle and offset, respectively, corresponding to an exit speed of 100 mph and launch angle of 12° . On the other hand, home runs require a larger attack angle and larger offset, the optimum being 24° and 1.05 inches, respectively, corresponding to an exit speed of 101 mph and a launch angle of 29° - 30° .

As an amusing aside, the bloop singles (~ 70 mph exit speed and 20° - 30° launch angle) are not shown on this plot. If you look at the first set of heat maps, you can easily see that these batted ball parameters are not accessible. Given that I have assumed an impact at the sweet spot with a fixed bat speed of 70 mph, I conclude those assumptions are not valid for these hits.



Armed with all this information, let's talk about hitting strategies. In particular, I want to compare a "level swing" with a "home run swing". The level swing is one where the attack angle coincides with the descent angle of the ball, which in our case is 6° . That angle is very close to the optimum angle for high BABIP. The home run swing is one where the attack angle maximizes the probability of hitting a home run, which we have just found to be 24° . In the next figure, I plot BABIP (top) and home run probability (bottom), as a function of offset for each of these two attack angles, blue corresponding to the level swing and red corresponding to the home run swing. Note that BABIP is essentially the same for the two different attack angles, albeit shifted in offset by about 0.1 inches. However, the home run probability is nearly a factor of three lower for the level swing than for the home run swing. Armed with this information, this raises the following question: Why not always go for the home run swing. If your aim is right, around 1 inch, you'll get your home run with a reasonably high probability. If your aim is a little off and the offset is a bit smaller, you'll get a line drive single. Either way, you can't lose.



That brings me to one of the key points I want to make in this article: Timing! Thus far I have assumed that the contact always takes place at the front edge of home plate, meaning that the bat and ball both arrive at that location simultaneously. Let's suppose that the batter swings at some attack angle, with the goal of having some desired offset at the front plane of home plate. If the batter is going for a line drive single, that offset might be about 0.5 inches; if he is going for a home run, it will be around 1 inch. Now suppose that the batter is a little late with the swing (say, a bit over 3 milliseconds late), so that the contact occurs 4 inches behind the front plane of home plate. If the attack angle is level, the actual offset will be identical to the desired offset. Such a swing is therefore very forgiving for small errors in timing.

However things are very different if the batter has a home run swing. Over those 4 inches, the ball will be about 0.4 inches lower and the bat will be about 1.8 inches lower. The net effect is that the offset will be larger than the desired offset by 1.4 inches. Likewise if the swing were too early by the same 4 inches, the offset would be smaller than desired by 1.4 inches. Think about that for a minute: 1.4 inches is about half the diameter of a baseball. It is huge! For example, if you are aiming for 1 inch and are late by 3 milliseconds, your offset will be 2.4 inches, which gives you a on-base probability of close to zero. Indeed, if the offset is larger as 2.7 inches (the sum of the ball and bat radii), the bat will miss the ball altogether.

So what's a poor batter to do. It depends on what the batter is trying to do. If getting on base is the primary goal, then swinging level would appear to be the best strategy. You'll still hit a lot of home runs if the bat speed is high and the aim is good. Indeed that is exactly the approach advocated by Ted Williams, as written in his classic book "The Science of Hitting." Someone who hit 521 career home runs with a 0.482 lifetime on-base percentage must surely know whereof he speaks.

But to get the highest home run probability requires swinging with a larger attack angle, keeping in mind that you will fail often if the swing is mis-timed. A possible compromise is to pick some intermediate attack angle. For example, an attack angle of 12° reduces the home run probability by about 35% but also reduces the 4-inch timing error to be about 0.4 inches. Moreover, such an attack angle is close to "level" for a typical curveball, which has a larger descent angle than a fastball.

There may well be other possible strategies, but my goal in this article is not to analyze these in detail. Rather, it is to set up a framework for doing such an analysis. While the development of the framework is still a work in progress, hopefully some useful initial steps have been taken.