Determining Pitch Movement from PITCHf/x Data

Alan M. Nathan, University of Illinois

October 18, 2012

Updated October 21, 2012

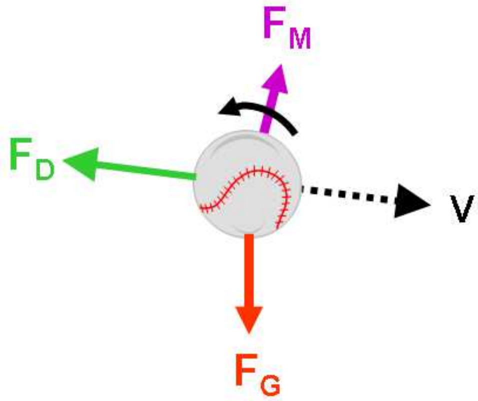
One of the more remarkable developments in baseball analysis in recent years is the availability of pitch-tracking data, courtesy Sportvison, MLBAM, and the PITCHf/x tracking system. These data allow us to record with unprecedented precision such quantities as the pitch speed and the location at home plate. But even more importantly, we have measures of quantities that we never had before. As a result, we now have new and novel ways to study the art of pitching.

A key ingredient of the new data is the movement, and that is the topic I address in this article. Here is a preview of what follows. I first will define exactly what is meant by the term "movement." I then will discuss how that movement is extracted from the PITCHf/x data and show that the current technique produces values for the movement that are systematically shifted from the true values. I will discuss a different technique to calculate movement that better conforms to our definition and give some examples of how using this technique might affect our analysis of pitches. An appendix contains all the technical details.

Before delving into this topic, I should point out that I first wrote about this in an article I posted on my web site in the very early days of the PITCHf/x era (December 2007): http://webusers.npl.uiuc.edu/~a-nathan/pob/Magnus.pdf. I also discussed it at the first PITCHf/x Summit in May 2008. The earlier article was written very much in an academic style, with lots of technical details included. In this one, I will strive for a far more conversational writing style so that it will hopefully be accessible to a wider readership.

So, what is meant by "movement"? Let me first give you a tentative definition. It is the amount by which the trajectory deviates from a straight line. That is about the simplest definition you could imagine and agrees with our everyday concept of what we mean by movement. I want to take this tentative definition as my starting point and move on from there. But first I have to bring up some physics…and I promise you I will keep it simple.

Newton's First Law says that in the absence of forces, objects in motion will move at constant velocity, meaning constant speed and in a straight line. Said a little differently, without forces acting on the ball, there is no movement, from which we can conclude that movement is the result of forces. It is then natural to ask what those forces are. The figure shows a side view of a ball moving horizontally from left to right, with a slight downward angle, and spinning in the counterclockwise direction (commonly called backspin). There are three forces acting on the ball. First, gravity (F_G in the figure) points directly downward and is the force responsible for the apple hitting Sir Isaac on the head. Next, air drag (F_D) acts opposite to the direction of motion. Finally, the so-called Magnus force on the spinning baseball (F_M) is perpendicular to the direction of motion, so it points upward and slightly to the right.



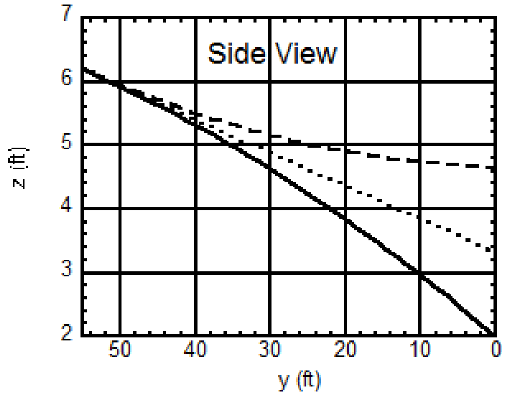
Newton's Second Law says that forces cause accelerations; that is, they cause the velocity—the speed and/or the direction—to change. So let's examine the effect of the three forces in detail. Since F_D always acts opposite to the direction of motion, it reduces the speed of the ball but does not change its direction. For F_M, the opposite is true. Since it always acts perpendicular to the direction of motion, it changes the direction of the ball but not its speed. Gravity is different in that it always acts in the downward direction, regardless of the direction the ball is moving. Thus, it changes both the speed and the direction of the ball. For a pitched baseball that is moving mainly in the horizontal direction, the change in speed is tiny, so the primary effect of gravity is movement in the downward direction.

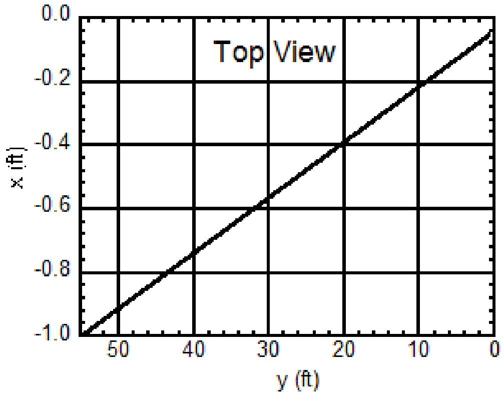
From the preceding discussion, we see that both the Magnus force (F_M) and gravity (F_G) result in the pitch deviating from a straight line. However, it has become conventional among PITCHf/x analysts to remove the effect of gravity so that the vertical movement is due entirely to the spin. The rationale for doing this is that the spin on the baseball, both the rate of spin and the spin axis, is something under control of the pitcher whereas gravity is not. Indeed, it is largely the combination of movement due to F_M and release speed that is used to classify pitch type for a given pitcher. Accordingly we modify our tentative definition to arrive at the following precise definition of movement:

- *Movement is the deviation of the trajectory from a straight line with the effect of gravity removed*.

Now, I realize that removing the effect of gravity seems a bit arbitrary, even though there might be good reasons to do it. Indeed, there may well be analysis contexts for which it is more useful not to remove the effect of gravity. However, for the present context, removing gravity helps me illustrate my main point more clearly. In any case, the effect of gravity is as easy to restore as it is to remove, so I will proceed with my gravity-removed analysis.

With these ideas in mind, take a look at the next figures which show the trajectory of a typical pitched baseball both from a side view (z vs. y) and a top view (x vs. y). The standard PITCHf/x coordinate system is used in which the origin x=y=z=0 is at the corner of home plate, the y axis points toward the pitcher, the z axis vertically up, and the x axis to the catcher's right. The pitch is typical of a four-seam fastball thrown directly overhand by a right-handed pitcher, with a release angle slightly downward and gloveside and with backspin but no sidespin. The figures show the trajectory from the release point (y=55 ft) to home plate (y=0 ft). First look at the side view, in which the solid curve is the actual trajectory and the dotted line is the trajectory expected if the ball moved in a straight line from release to home plate. The difference between the actual trajectory and the dotted line at the front edge of home plate is about -1.4 ft and is due to the combined effect of gravity (which produces a large negative number) and the backspin (which produces a smaller positive number).





The dashed curve is the trajectory the ball would have followed in the absence of gravity. The difference between the dashed and dotted curves is the quantity we are defining to be the movement, the deviation from a straight line with gravity removed, and is approximately +1.2 ft. The fact that the movement is positive is exactly what one expects for a pitch thrown with backspin, since the Magnus force is upward.

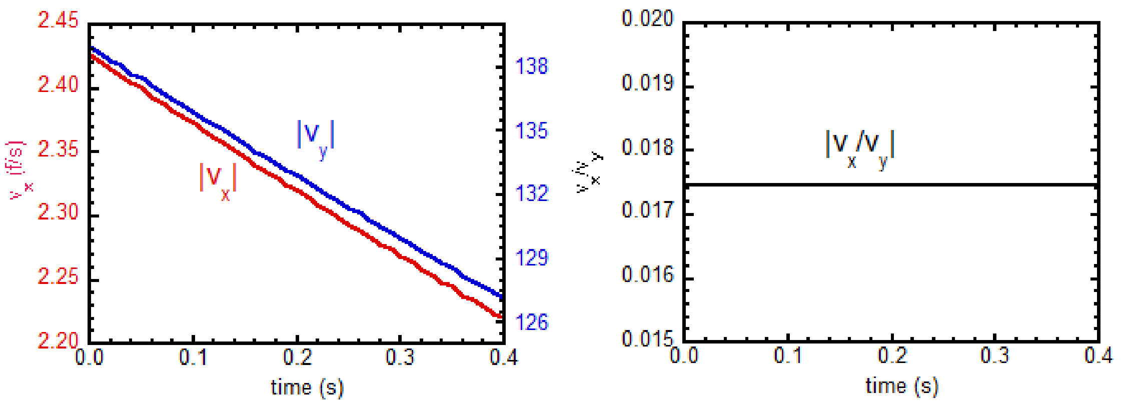
Next look at the top view showing the x-y plot. This trajectory is the ultimate in simplicity in that the ball follows a perfectly straight line from release to home plate. Therefore there is no movement in the horizontal direction. But that is exactly what is expected for a pitch with no sidespin. No sidespin means no sideways Magnus force which means no sideways movement.

In summary, the pitch shown in the figure has a vertical movement of +1.2 ft (14.4 inches) and a horizontal movement of 0 inches.

So, now that we have a precise definition of movement, how do we determine it from the PITCHf/x data? First a small digression. All the pitch information that is publicly available is determined from the so-called nine-parameter fit to the actual trajectory. That is, each trajectory is parametrized by nine numbers: an initial location, an initial velocity, and a constant acceleration for each of the three coordinates x,y,z. Using these nine numbers, the full trajectory can be calculated and all the interesting quantities (e.g., release velocity, home plate crossing, etc.) can be determined with the aid of some standard physics formulas. These formulas are also used to calculate the movement. Let's take a closer look at how that is done.

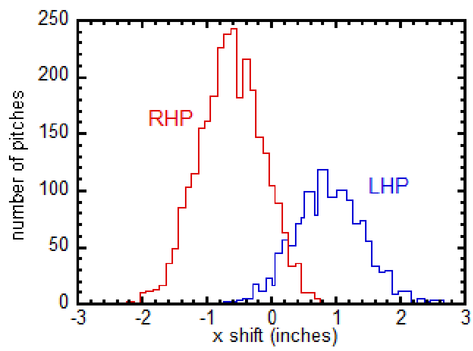
As we have discussed, forces cause accelerations which can result in movement. So it might seem natural to use the x and z accelerations (with gravity removed) to calculate the x and z movements. In fact, that is exactly what is done and I will hereafter refer to this as the “standard procedure.” Now comes the key point in this whole discussion. *The standard procedure is wrong!* That’s a pretty strong statement, so I want to go into some detail as to why it is wrong and how one can do better, using the trajectory we have already discussed as an example.

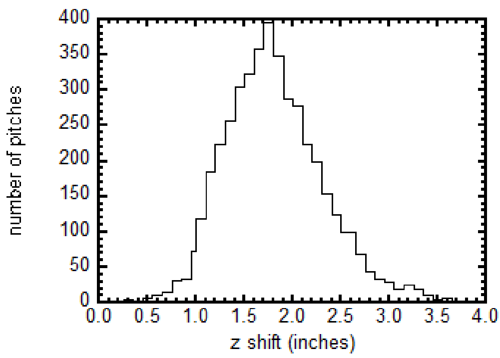
First look at the top view (the x-y plot) of the figure above. Remember that this trajectory was calculated assuming no sidespin. Therefore no movement is expected in the x direction. The fact that the x-y plot is a straight line shows that there is indeed no movement. Now look at the left figure below, where I have plotted the magnitude of the x and y velocities as a function of time. The fact that both are decreasing shows that there is an acceleration in both the x and y directions. Now look at the right figure, where I have plotted the ratio v_x/v_y plotted as a function of time. That ratio does not change. The constancy of v_x/v_y means that although the velocity in both the x and y directions is decreasing, the *direction* of the ball in the horizontal plane is not changing.



This behavior is exactly what one expects if the acceleration is due entirely to drag, which results in a change of speed but no change of direction. There is an acceleration but no movement. The standard procedure would use that acceleration, which is in the –x direction, to infer a movement in the –x direction. So, we have a ball that has no sidespin and travels in a perfectly straight line in the horizontal plane, yet the standard procedure tells us there is a movement in the –x direction. The standard procedure surely has gotten it wrong. But the good news is that this example tells us exactly what we have to do to get it right. We need to remove the contribution of drag from the acceleration, just as we have already removed the effect of gravity, then use that to calculate the movement. My technique for doing that—call it the drag-corrected technique—is described in the Appendix.

Having established that the standard procedure is wrong, let’s now try to quantify how large a mistake is made if the drag is not removed when calculating the movement. I compare the current and drag-corrected techniques for each of the 4324 pitches thrown throughout MLB on September 15, 2012, with movements calculated between y=50 ft and the front edge of home plate.





The difference between the standard and drag-corrected movement is presented in the figures for both the x and z movements. For the z movement, there is a systematic shift toward positive values with a mean of about 1.7". The reason is very clear. Since the average z velocity is always negative (the ball always drops between release and home plate), the z component of drag is always positive, leading to a positive contribution to the z acceleration. For the x movement, there is a systematic shift of +1" for LHP and -1" for RHP. Once again, the reason is clear. LHP and RHP generally release the ball with a negative and positive horizontal velocity, respectively, leading to the results found. The relationship between the sign of the shift and the sign of the average velocity components is discussed in the appendix.

These systematic shifts have some interesting consequence for baseball analysis. For example, typically 4-seam and 2-seam fastballs have both upward and arm-side movement, both of which are overestimated (in magnitude) by the current technique, leading to a systematic overestimate of the rotation rate of the ball. On the other hand, a curveball typically has both downward and glove-side movement, both of which are underestimated by the current technique, leading to a systematic underestimate of the rotation rate of the ball. Another interesting example comes from splitters/forkballs, a topic that Mike Fast wrote about in his December 2010 Baseball Prospectus article (http://www.baseballprospectus.com/article.php?articleid=12558). Generally these pitches are thrown with very little spin, leading to very small movement, especially in the vertical direction. But the pitches are thrown very differently, as Mike discusses in the article: A splitter has backspin and a small amount of upward movement, while a forkball has topspin and a small amount of downward movement. With a systematic upward shift of a few inches using the current technique, it would be easy to incorrectly identify a forkball as a splitter. It is perhaps worth pointing out that Mike utilized drag-corrected movements in his analysis, even though he does not indicate so in his article. Finally consider a cut fastball, which typically has a small amount of glove-side movement, the magnitude of which is systematically underestimated by the current technique. Using the drag-corrected technique might well change our perspective about the effectiveness of this pitch.

So, let me summarize the contents of this article. First I give a careful definition of what is meant by movement. It is a common-sense definition: the deviation of the trajectory from a straight line, with the effect of gravity removed. Next, I show that the standard procedure calculates the movement in a manner that does not conform to that definition because the effect of drag on the accelerations has not been removed. In the appendix I show a technique for calculating the movement that corrects this shortcoming. I show the typical size of the mistake that is made by using the standard rather than the drag-corrected method. Finally, I discussed a few practical implications for baseball analysis.

I thank Dan Brooks and Harry Pavlidis for encouraging me to write this article. I thank both Dan, Mike Fast, and Tom Tango for a critical reading of an earlier draft. I also thank Mike for reminding me of his earlier work on splitters/forkballs. Finally I want to acknowledge Ike Hall, who contributed to my earlier article on this subject.

APPENDIX: Technical Details

To determine the movement of a pitched baseball, it is necessary to isolate the part of the acceleration that is due to the Magnus force. To do that requires removing the contributions due to gravity (this is easy) and due to drag (this is harder). I now outline how to do that.

I start by writing the acceleration vector as the sum of the three contributing parts: $\vec{a} = \vec{a}_D + \vec{a}_M + \vec{g}$, where D and M refer to drag and Magnus and g is the acceleration due to gravity. The first thing to realize is that the drag is always opposite to the velocity whereas the Magnus force is perpendicular to the velocity. The tricky part is that neither the magnitude nor the direction of the velocity is constant during the trajectory, yet we have assumed the acceleration is constant by doing the 9P fit. In fact, the 9P fit determines the *average* acceleration. We will assume that the drag is the contribution to the average acceleration, with g removed, which is directly opposite to the *average* velocity vector $\langle\vec{v}\rangle$. Mathematically, the drag contribution to the average acceleration is the projection of the acceleration vector, with g removed, in the $-\langle\hat{v}\rangle$ direction. That is, $\vec{a}_D = -\left|(\vec{a} - \vec{g})\cdot\langle\hat{v}\rangle\right|\langle\hat{v}\rangle$. Therefore to remove the contribution of drag, we simply subtract this expression (along with g) from the acceleration. Putting this all together, we arrive at the following expression for the x and z components of the Magnus acceleration:

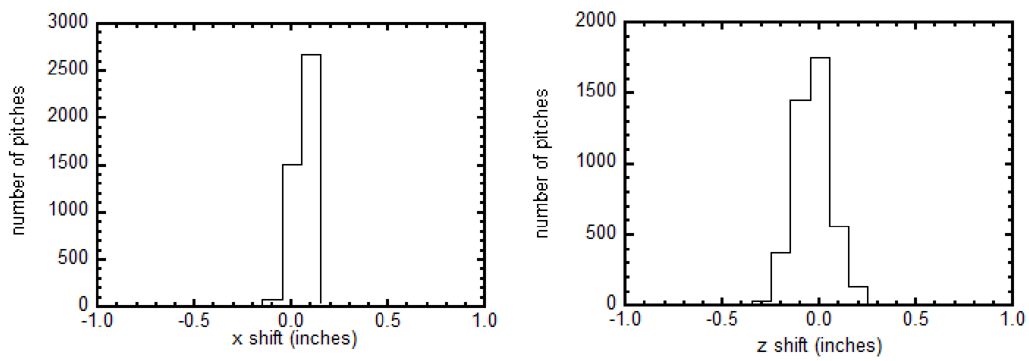
$$a_{Mx} = a_x + \left|(\vec{a} - \vec{g})\cdot\langle\hat{v}\rangle\right|\frac{\langle\vec{v}\rangle_x}{\left|\langle\vec{v}\rangle\right|}$$

$$a_{Mz} = a_z + g + \left|(\vec{a} - \vec{g})\cdot\langle\hat{v}\rangle\right|\frac{\langle\vec{v}\rangle_z}{\left|\langle\vec{v}\rangle\right|}$$

For either the x or z directions, the movement is $a_M t^2/2$, where t is the flight time. In the standard calculation of movement, the last term on the RHS of these expressions is omitted. It is precisely that term that is needed to remove the contribution of drag.

The structure of these equations makes it clear that the sign of the omitted terms is the same as the sign of the average x or z velocities. Or, said differently, the sign of the error made by omitting these terms is the opposite of the sign of the average x or z velocities. Given that the x and z components of the average velocity are simply equal to $(x_f - x_0)/t$ and $(z_f - z_0)/t$, the sign of the correction depends entirely on the difference between the final and initial locations of the ball in the x and z directions. By the way, if it is desired to include the effect of gravity in the vertical movement, simply omit the first g in the equation for z.

By solving exactly the equations of motion for the trajectory of the baseball, one can calculate the movement (as I have defined it) exactly. That calculation can be compared with the technique just described. I have made that comparison for the 4324 pitches thrown on September 15, 2012, and the results (drag-corrected minus exact) are shown in the figure below. While the drag-corrected technique is still not perfect, the errors have been reduced considerably compared to the current technique and are almost surely at a level suitable for baseball analysis.



I have put together a spreadsheet template for calculating the drag-corrected movement. To use the template, simply enter the 9P fit parameters (x_0, …,a_z) into columns A-I. The drag-corrected movement in inches (dx and dz) are calculated in columns X and Y. For comparison, the current technique for movement is calculated in columns Z and AA. Both are calculated from a starting point y_{start} to the front edge of home plate. The standard method uses y_{start}=40, but you are free to specify whatever value you like for y_{start} in cell E3, in units of feet. The template can be download at http://webusers.npl.illinois.edu/~a-nathan/pob/MovementTemplate.xls.