The PES Pareto Method: Uncovering the Strata of Position Error Signals in Disk Drives

Abstract— This paper and the two that accompany it will describe a method of breaking down the Position Error Signal (PES) of a magnetic disk drive to its contributing components. Once these components are identified, they can be ranked in terms of their overall effect on PES and thus the most critical ones can be worked on first. This method is based on three things: an understanding of how Bode’s Integral Theorem[1] ties into noise measurements, a measurement methodology that allows for the isolation of individual noise sources, and a system model that allows these sources to be recombined to form the drive’s Position Error Signal. We have found this method to be dramatically useful in identifying the key contributors to PES noise.

1. Introduction

This paper and the two that accompany it[2, 3] will describe a method of breaking down the Position Error Signal (PES) of a magnetic disk drive to its contributing components. Once these components are identified, they can be ranked in terms of their overall effect on PES and thus the most critical ones can be worked on first. In order to do a practical analysis of the contributors to PES, the fundamental question that must be answered is: What can be measured? While this may seem whimsical at first, it should be noted that in any real system, we will not have access to all the measurement points that we desire. Furthermore, although many different analysis tools might theoretically be available, they are useless to us if they cannot make use of the actual laboratory measurements available to us.

In order to guide our measurements and our modeling, it is useful to have a map of the system. The block diagram in Figure 1 will serve as the map for our tour of noises in the system. Starting at the left of this diagram, the reference position that the actuator arm must follow is the position of the magnetic track written on a disk, turning on a spindle. Only the position error – the difference between the reference track position and the readback head position – is sensed by the readback head, and this error signal is sent to the demodulator. The demodulator outputs a set of numbers at the sample rate, and these are combined electronically to form PES. This PES signal is then converted to a digital format via an analog to digital converter (ADC), filtered by the compensator and then sent back out to the power amplifier via a digital to analog converter (DAC). The power amp converts the desired voltage into a current to drive the voice coil actuator (with torque constant $K_t$). The actuator itself has rigid body behavior as well as resonances. Through this, the head position is set. The position error is then sensed by the head. Absolute head position is not generally known from what is read off the disk surface, but can be obtained in the laboratory by shining a laser spot from a Laser Doppler Vibrometer (LDV) off of the side of the head. While this nominally measures velocities, the result can be accurately integrated in time (for the frequencies we are concerned with) to obtain position.

There are several measurement points that can be accessed around the loop: $X_{out}$, $I_{sense}$, PES, and head velocity (and position) via the LDV. In general test signals can be injected into the loop only at $X_{in}$.

There are several likely noise input points on a disk drive. First of all, there are the noises associated with the moving disk and the readback process. These all enter the loop at the same point, but have different root causes. The noise due to the motion of the disk attached to a ball bearing spindle creates both Repeatable Run Out (RRO) (typically at orders of the spindle rotational frequency) and Non-Repeatable Run Out (NRRO). One of the interesting properties of servowritten disks is that one pass of the NRRO is usually locked into the servoposition information when it is written. Thus, this written in NRRO is repeated at every revolution of the disk. The other noise source that enters at this point is the noise from the readback process of position information, called Position Sensing Noise (PSN). This noise can be due to the magnetic domains on the disk, the behavior of the magnetic readback head, the interaction of these two, or the action of the demodulator. (We lump demodulator noise into PSN for our current analysis.) Downstream in the loop, there are potential noise sources at the ADC and DAC (due to quantization), noise at the power amp, and finally windage. Windage is caused by the air flow generated as the disk spins. This air flows over, under, around, and into the actuator arms and the readback head, disturbing the head position. Given all these potential noise sources, there is a fundamental need to identify which of these – if any – are the most significant contributors to PES. With this information, the effort to reduce the noise in PES can be concentrated on the critical few.

It is worth noting that we purposely ignore external shock and vibration in this analysis for two reasons. First of all, external shock and vibration is heavily influenced by the drive’s operating environment while the above noises are a function primarily of the drive. The second is that prior work in this area[4] gives us some confidence that we already have a reasonable engineering solution to many types of external shock and vibration. Thus, this work will focus on internal noises.

The tools available to us are a set of laboratory instruments that can make both time and frequency domain measurements. In particular, Digital Storage Oscilloscopes (DSO) can record time domain data as can certain spectrum analyzers. The spectrum analyzers are most useful, though, for measuring lin-

---

1In this case, made by Polytec.
ear spectra, power spectra, power spectral densities (PSDs), and frequency response functions of systems. In particular, the spectrum analyzers that we use are the HP 3563A Control Systems Analyzer and the HP 3567A Multi-Channel Analyzer. The latter instrument has the advantage of allowing more than two signals from the system to be measured at once.

For analysis, we have the standard set of matrix based tools. In particular, we are using Matlab and Simulink. As has been the practice in our laboratory for several years, the measurements are made with a conscious thought of transferring them into Matlab/Simulink for analysis.

Given these tools, there are basically three types of measurements on which we could base our analysis. Their features are listed below:

**PSDs and Power Spectra:**

- These are easy to measure.
- It is straightforward to average PSDs.
- If the processes are independent, then the PSDs can be added. This allows us to decompose a PSD into its component parts.
- They contain no phase information \(\Rightarrow\) we cannot inverse FFT the PSD into a time signal to drive simulations (such as Simulink).

- We can do frequency response function filtering (in Matlab) using the following scheme:
  - Take the linear model and generate Bode plots at the same frequency points as PSD. Alternately (or in combination) measure the frequency response functions of the system. In either case, we get a frequency response function at the same frequencies as the PSDs. Call it \(H\) for now.
  - Multiply \(H\) times its complex conjugate to get \(|H|^2\).
  - Multiply \(|H|^2\) times the power spectrum or PSD to get effect of the loop on noise.
  - Note that \(|H|^2\) must be the appropriate units for filtering the PSDs.
  - The resulting output is another power spectrum or PSD.
  - We can use superposition to build up contributions from many sources.
  - We need to do some “loop unwrapping” to extract proper input noise levels for model.
  - This is limited to linear system models.

**Linear Spectra:**

- These are harder to measure: filtering and averaging of linear spectra is less straightforward on the HP 3567A.
- Linear spectra contain phase information:
  - We can inverse FFT averaged linear spectrum to get representative time domain input.
• We can drive Simulink with this.
• We can then use Simulink to generate PSDs from time domain data.
• This is limited to linear system models.

• Linear spectra cannot be added, therefore we cannot decompose noise component factors as with PSDs.

Time Domain Measurements:
• These are not limited to linear system models i.e., we can measure responses of nonlinear phenomena.
• When averaging, we can make use of an index signal which is generated once per revolution of the disk.
• Without synchronizing to the index, using time domain averaging drives all the signals 0.
• With synchronizing to the index, using averaging removes the Non-Repeatable Run Out (NRRO) of the disk, leaving only the Repeatable Run Out (RRO).
• This type of measurement does generate data for time domain simulations, but without using averaging it is hard to know if the data is representative of general system behavior.

Looking over all of these features, Power Spectra/PSDs appear to be the most promising measurements. The chief restriction of doing so is that we will have to limit ourselves to linear models of the disk drive. However, by doing so we are able to actually add and subtract PSDs. In order to do so, we formally should require some knowledge that the noise sources are independent. It turns out that there is no way to verify this for all sources, but it is very likely true. While any measured signal in the loop is correlated to several noise sources, each source arises from an independent physical phenomenon. Furthermore, without allowing for superposition of noise measurements, it would be next to impossible to analyze the noise of a measured system. Thus, it is a starting point we must choose. As far as the linearity assumption goes, it is well known that quantizers (ADCs and DACs) are nonlinear, however Widrow was able to model them using uniformly distributed white noise into a linear system. Furthermore while the actuator pivot friction is also nonlinear[7, 8, 9, 10], it has been shown[9] that the rotation of the spindle provides an operating point for the nonlinear system where the behavior can be modeled using linear components.

What remains to be seen is how all of these noise sources can affect the Position Error Signal. The fundamental concept that ties them together comes from what is known as Bode’s Integral Theorem[1]. The next section will give a thumbnail sketch of Bode’s Integral Theorem and discuss what its implications are for measurements of control loops.

2. Bode’s Theorem on Sensitivity Functions

There is is an old theorem by Bode[1] which deals with what he calls regeneration. It turns out that this theorem has some very interesting applications to control systems. This has only recently come to light as a tool for evaluating control systems[11]. However it is the starting point for design methodologies such as QFT[12]. There is even a discrete time version of this theorem[13] that gives some insight on how this theorem is affected by sample rates.

2.1 Sensitivity Functions: The block diagram for the following discussion is in Figure 2. The closed loop transfer function from $u_1$ to $y_2$ is given by the standard form:

$$T = \frac{PC}{1 + PC}. \quad (1)$$

The sensitivity function is also known as the disturbance rejection function. Designated $S$, it is given by:

$$S = \frac{e_1}{u_1} = \frac{1}{1 + PC} = \frac{y_2}{d} = -\frac{e_1}{d}. \quad (2)$$

Note that

$$S + T = \frac{1}{1 + PC} + \frac{PC}{1 + PC} = 1, \quad (3)$$

hence $T$ is commonly called the complementary sensitivity function. Note that $S = H_{gd}$ (= the transfer function from $d$ to $y = y_2$).

The sensitivity function is important because it shows how disturbances, $d$, go through the system and show up at the output, $y$, or at the error signal $e$. For a unity feedback system

$$S \triangleq H_{gd} = -H_{e1d} = H_{e1}u_1. \quad (4)$$

Thus, the transfer function from disturbance, $d$, to the output, $y$ is the same as the transfer function from the input $u_1$, to the error (PES), $e_1$, and the transfer function from disturbance, $d$, to error (PES), $e_1$. In other words it is very good gauge of how noise will be filtered through the system.

2.2 Bode’s Integral Theorem: While the mathematics used to prove both versions of Bode’s theorem can be fairly...
complicated, the result is fairly simple and extremely powerful. We will leave the proofs to those papers[1, 11, 13]
and talk simply about the interpretation. Looking at Figure 3 it says simply that:

\[ \text{the area of disturbance amplification} = \text{the area of disturbance rejection} + \text{a non-negative constant}. \]  

(5)

Looking at Figure 4, we see that for discrete time the main difference is that the Nyquist frequency limits the space we have to work with. In both cases, if we want to attenuate disturbances at one frequency, we must amplify them at another. There is no way to get around this.

**Theorem 1 (Bode's Integral Theorem for Continuous Time, Open Loop Stable Systems)** For a stable, rational \( P \) and \( C \) with \( P(s)C(s) = O(s^2) \) (i.e. they fall off as \( \frac{1}{s^2} \))

\[ \int_{0}^{\infty} \log |S|d\omega = 0. \]

Consequences: “Sooner or later, you must pay for every good deed.” (Eli Wallach in the The Magnificent Seven)

Translation: If you make the system less sensitive to noise at some frequencies, you then make the system more sensitive at other frequencies.

Typical control designs attempt to spread the increased sensitivity (noise amplification) over the high frequencies where noise and/or disturbances may be less of an issue.

A wonderful treatment of this theorem and the importance of it was given as the Bode Lecture at the 1989 IEEE Conference on Decision and Control (Tampa, FL)[14]. The talk, by then Honeywell Researcher and MIT Professor, Gunter Stein, was entitled “Respect the Unstable.” Unfortunately, no papers accompanied Bode Lectures at that time, although there is a video distributed by the IEEE. Stein used this theorem to show how tightly control engineers are dancing when we deal with unstable systems. Stein described the net effect of control systems design as trying to get a certain amount of disturbance rejection at some frequency span while trying to thinly spread the amplification over a large frequency span. Stein referred to this as shoveling dirt. An attempt to recreate his drawing is in Figure 5. The guy shoveling dirt is moving around the disturbance amplification. He is doing classical control. He can move the dirt around, but the dirt does not go away. Even with our modern, sophisticated control tools, in Figure 6, the dirt is still there.

Now, if the plant or compensator are not stable — i.e., if \( P \) and/or \( C \) have finite number of unstable poles — then the result generalizes to

\[ \int_{0}^{\infty} \log |S|d\omega = 2\pi \sum_{k=1}^{K} \text{Re}(p_k) \]

(a positive number) where \( K \) is the number of unstable poles of \( C \) and \( P \) are those poles. Thus, any unstable poles in the system only make life worse in that more of the noise would have to be amplified.

Note that the integration is done on a linear scale even though these drawings may imply a logarithmic frequency scaling.

**2.3 Bode’s Integral Theorem for Discrete Time**

The paper on Bode’s Integral Theorem for discrete time systems[13] uses a slightly different notation than that used above.

**Theorem 2 (Bode's Integral Theorem for Discrete-Time Systems):** For all closed-loop stable, discrete-time feedback systems, the sensitivity function has to satisfy the following integral constraint:

\[ \frac{1}{\pi} \int_{0}^{\pi} \ln |S(e^{j\phi})|d\phi = \sum_{i=1}^{m} \ln |\beta_i| \]

where \( \beta_i \) are the open-loop unstable poles of the system, \( m \) is the total number of unstable poles, and \( \phi = \omega h \) where \( h \) is the sample period and \( \omega \) is the frequency in radians/sec.

There are some implications of this theorem dealing with discrete time systems. Basically, they say the following:

a) With \( h \) as the sample period, the ideal upper limit of the frequency spectrum is \( \frac{\pi}{h} \), the Nyquist frequency. Mohtadi assumes for this discrete time theorem that there are no frequencies in the loop above the Nyquist frequency[13]. This would imply that \( PC = 0 \) for \( \omega > \frac{\pi}{h} = \omega_N \) and \( S = \frac{1}{1+PC} = 1 \), which is in general false for a physical system. However, typical digital
control systems do assume that $PC$ is small at or above the Nyquist frequency. Thus, even though the exact assumptions of the theorem will not hold for most physical systems, it is reasonable to assume that some insight can be gained from this theorem.

Note: This interpretation does leave open the door for multirate control. With the actuator signal going out at a higher rate than the input sensor, it might be possible to do something (good or bad) at frequencies above the Nyquist rate, $\omega_N$, of the sensor.

b) Since we can only manipulate frequencies up to $\omega_N = \frac{\pi}{T}$, and since $|S| \approx 1$ above that frequency, the theorem says that if for some frequency $|S| < 1$ then at some other frequency $|S| > 1$. Unlike the continuous time result, though, there is not infinite bandwidth to spread this over. Thus, the $|S| > 1$ all happens below the Nyquist frequency (and therefore in a finite frequency range).

c) Loop Transfer Recovery, as shown in a famous paper by Doyle and Stein[15], cannot be done. LTR attempts to asymptotically match the LQR result that says that LQR provides $|S| < 1$ for all frequencies, in part because LQR uses full state feedback. LTR tries to the same for frequencies up to some point. The $|S| > 1$ part is dumped over the infinite frequency band above that point. The Nyquist limit eliminates the possibility of doing this.

2.4 What does it mean: Looking at Figure 7, the discrete time version simply states that (analogous to the continuous time theorem):

$$\text{the area of disturbance amplification} = \text{the area of disturbance rejection} + \text{a non-negative constant}, \quad (6)$$

and this all this must happen before the Nyquist frequency. The reason why this becomes important is that by working to reject noise at one frequency, we will dump noise amplification at another frequency, but now that the Nyquist frequency establishes a limit, we may end up putting noise amplification at frequencies that we care about.

![Figure 8: Sensitivity function at nominal sample rate, $\omega_{N_1}$.](image)

2.5 Effect of sample rate: If the control system is merely a process of shoveling the “disturbance amplification dirt” around, then what does the Nyquist rate signify? It can be thought of as a retaining wall which prevents the “dirt” from going out beyond the Nyquist frequency. Thus, the freedom to spread the dirt around is limited by the Nyquist “retaining wall.”

Graphically it is quite easy to see what the theorem implies with sample rate. Looking at Figure 8, say we have a certain amount of rejection, $|S| < 1$, for a compensator sampled at $2\omega_{N_1}$. This implies a certain area of $|S| > 1$. This has to be done before the Nyquist rate retaining wall.

Now, we double the sample rate. With the extra “space,” we can either do some extra filtering but keep the bandwidth of the closed-loop system constant (Figure 9) or demand more rejection at low frequency and higher bandwidth (Figure 10). Note that the effect of using the extra bandwidth to filter is essentially to spread the amplification, $|S| > 1$, over a broader frequency band. This shrinks the height of any amplification hills (Figure 9). By pushing the closed-loop bandwidth (Figure 10), better performance at low frequency may result in much worse performance at high frequency.

There are two reasons why Bode’s Integral Theorem is important in a discussion of a disk drive’s Position Error Signal. First of all, it gives us a very good gauge on what we can and cannot do with disturbance rejection and noise in a control system. Amazingly it comes from such an old and simple result that is generally not well known. This result tells us that whenever we improve with the noise rejection at one frequency we pay for it at another. If we are smart and put the noise amplification at places where there is only a small amount of noise, then we do well. If not, we may inadvertently be boosting much of the noise that we are trying to eliminate.

The second reason will become apparent in the next section. It turns out that when we measure PES from a closed loop system, we should actually open the loop and look at PES. The exact same effects that are the point of the above theorem affect our measurement of PES. We will see that when we measure PES that is flat in closed-loop, opening the loop (mathematically in Matlab or on a spectrum analyzer) will give us a PES spectrum that looks considerably different from the ones we are accustomed to.

3. Measurements of PES and Loop Unwrapping

Typically in a disk drive the Position Error Signal (PES) is only measured in closed-loop. This is in general due to the difficulties of obtaining a linear measurement of head position across multiple tracks while the loop is not closed. What seems less common is “opening the loop” as is often done with closed-loop transfer function measurements. While a PSD of closed-
loop PES might be a reasonable measure of loop performance, it is not the useful quantity for determining what the noise inputs to the system are. In order to obtain this quantity, we want to open the loop, either physically or mathematically.

Referring back to the block diagram in Figure 2, the closed loop transfer function from $u_1$ to $y_2$ is given in Equation 1 and the sensitivity function is which turns out to be the transfer function from $u_1$ to $e_1 (= \text{PES})$ is given in Equation 2. Both of these transfer functions represent many responses. We typically think of the sensitivity function as the error response, $e_1$, from either the reference, $u_1$, or a disturbance, $d$.

Now, to unwrap the closed-loop transfer function, $T$, there is a convenient button on some spectrum analyzers - such as the HP 3562A/3563A - which generates the open loop response, $PC$, via the waveform math operation $\frac{T}{PC} (= \text{PC})$ conversion. However, there is no such key to generate $S$. It can still be done simply, thought. The procedure is:

1) Measure closed-loop response from $X_{in}$ to $X_{out}$. Call this $T$. (Note that if the phase of the measurement does not start at 0 at low frequency, then there is an extra factor of $-1$ in the closed loop measurement which must be removed by negating the trace.)

2) Subtract $T$ from 1, i.e., $S = 1 - T = \frac{T}{1+PC}$. Here $PC$ does not start at 0 at low frequency, then there is an extra factor of $-1$ in the closed loop measurement which must be removed by negating the trace.

3) It is also useful to calculate $\frac{T}{PC}$. This is $PC$, a quantity that is very useful to have.

Note that while the PSD of PES is typically measured as a closed loop quantity, we are now in a position to extract the input to the loop that would yield that PES PSD. Since the transfer function from $u_1$ to $e_1$ is given by $S$, a noise PSD input at $u_1$ would be filtered by $||S||^2$ by the time it showed up at $e_1$. Thus, if we start with a measurement of noise at $e_1$, we can filter backwards by $\frac{1}{PC} = ||1+PC||^2$ to get the input at $u_1$ that could have generated it. The procedure above shows how to measure the exact filters needed to “open” the loop.

3.1 Practical Considerations: It is useful to understand that closed loop frequency response functions are generated by injecting a signal at $X_{in}$ and reading the response at $X_{out}$. A swept sine (also known as sine-dwell) scheme is used to get the cleanest possible measurement[16]. Due to traditional design methods, the frequency response functions are often measured using a logarithmic frequency spacing although a linear spacing is also possible. On the other hand the measurement PSDs is almost always done with a linear frequency axis. Furthermore, there is no need to inject anything into the loop to get the nominal track following levels of PES PSD.

Note that in general the frequency spacing between swept sine and linear resolution mode does not match. However, by doing the swept sine in linear resolution and by carefully choosing matching frequency bands for both the swept sine and the PSD measurements, we are able to use the measured frequency response functions to filter the measured PSDs. The extra work to match frequency bands with the swept sine mode is justified by the improved signal to noise ration of swept sine frequency response function measurements over broadband measurements.

This allows us to measure the frequency response function, with the cleaner mode (swept sine) and still use it on the linear resolution PES data.
In order to do the appropriate filtering, certain frequency response functions need to be either generated from a model or measured in the laboratory. The following building blocks can either be obtained from laboratory measurements or models and used to construct any of the necessary filters. Since we are filtering PSDs, it is useful to remember that the operation will actually involve the magnitude squared of the filter response.

Note that obtaining open loop frequency response function measurements on a closed loop system can be done in one of two ways. The first, a 3-wire measurement involves injecting a signal at one point in the loop and reading the output at two other points in the loop. This allows for the direct measurement of open loop quantities from a closed loop system, but does a poor job if eliminating noise from the measured response. On the other hand a closed loop measurement is made by allowing the first of the loop measurement points to be the injected signal. In order to obtain open loop quantities from such a measurement, the loop must be mathematically opened. This does a much better job of decorrelating noise in the loop from the desired frequency response function. However, because of the loop unwrapping, this method of obtaining open loop dynamics can be susceptible to quantization errors at frequencies where the closed loop response, \( T \approx 1 \). In this case the loop opening operation \( \frac{1}{1-T} \) can be dominated by the quantization error in the denominator. The point of the above discussion is to point out why neither method is perfect and often both types of measurements are done on the same system.

\[
\begin{align*}
\text{Given that we have the elements to construct the appropriate filters, there is a common theme for each noise source:} \\
\bullet \quad P(s) &= \frac{1}{K_t \left( \frac{LDV_{sense}}{X_{in}} \right)} \quad \text{3-wire} = \frac{1}{K_t A(s) \left( \frac{LDV}{X_{in}} \right)} \quad \text{opened} \\
\bullet \quad A(s) &\text{ from model} \\
\bullet \quad C(s) &= \left( \frac{X_{out}}{PES} \right) \quad \text{meas} \\
\bullet \quad D(s) &= \left( \frac{PES}{LDV \text{ Position}_{sense}} \right) \quad \text{meas} \\
\end{align*}
\]

Given that we have the elements to construct the appropriate filters, there is a common theme for each noise source:

- Isolate a measurement of a noise source as described in [2] (“common mode reject”).
- Filter backwards from the measurement point to the noise input to obtain the noise source input PSD.
- Filter forwards from the noise source input to PES to obtain the effect of this noise on the PES PSD.
- Compare the PSDs at PES and add to cumulative PES PSD.
- Integrate across frequencies to obtain power spectra and total variances for each noise source.

As simple as this methodology might seem, it can yield surprisingly profound results in the area. Two papers by the authors illustrate its use[2, 3]. The net result is to identify which noise sources in a disk drive are truly limiting the servo performance.
References


