Behind Gamma's Disguise

Thor Olson
Management Graphics Inc.
Minneapolis MN

A simple power law is often used to describe the light output of cathode ray tubes (CRTs). Its exponent gamma (γ) has been measured many times in many ways, with varying results. The failure of this simple formula to accurately characterize high-resolution CRTs used in film recording applications has led to a model that explains the observed light output from very low beam currents to high drive levels It comprises a cathode model and an anode current "trim" function. The relationship between anode and cathode currents establish the key parameters. The cathode current is represented by a blend of two power laws. The anode current is a fraction of the cathode current determined by the intersection of a gaussian beam with a circular aperture. The resulting model successfully predicts the behavior of high resolution CRTs even at low output levels. It also explains the many varying results when attempting to measure a single value for gamma.

Gamma is a highly misunderstood concept in today's world of digital imaging. This is in part due to the merging of previously disparate fields, each with their own definition and useage of the term [Poynton 93]. This paper will discuss in some detail the power law whose exponent has been graced with the name of the ancient Greek character, gamma:

$$i = kv^{\gamma} \tag{1}$$

In particular, the use of this formula to describe the light output of a CRT will be examined as a function of its control voltage (see the bibliography in [Berns 93]). The close scrutiny of the behavior of the CRT near its "cutoff" (black) region has not been dexcribed previously.

We are interested in CRT light output, but accurate and highly precise light measurements are notoriously difficult to make. Instead, we will treat the phosphor screen of the CRT as a direct "electron to photon" converter. This allows us to make current measurements, which, while still challenging, provides the necessary precision.

Our motivation in this study came about from the desire for a model that would be useful in calibrating and controlling CRTs used in high resolution film recording. Using a gamma value of 2.2 to characterize the light output was observed to be not merely a poor approximation, but a distinctly bad one. Having decided that CRT lore would be inadequate for our goals of characterizing and calibrating our devices, we set out to measure and model the behavior of the CRT in the regions where we operate. This paper summarizes the results.

Historical notes on gamma

The current-voltage relation in vacuum tubes was derived in 1911. It is called the Langmuir-Child law [Ryder 47] and shows that the current density reaching an anode from a thermionic cathode follows a 3/2 power law as follows:

$$j = kv^{3/2} \qquad (2)$$

This applies well to vacuum diodes, triodes and other electron guns. The current is limited to this relationship not because of any shortage of electrons emitted from the cathode surface, but rather the self repelling field generated by the electrons themselves. Attempting to increase the current (and thus electron density) will cause fields that drive the electrons back to the cathode or any other nearby electrode that can remove electrons from the vacuum. In the case of CRTs, the narrow beam that is formed diverges.

It would seem that this is the end of the story. Gamma is 1.5 as derived from fundamental physical constants. Measurements of real cathodes however, show that the current emitted from their surface follows a similar power law with an exponent of at least 3! [Moss68] Where is all the current going? And where did the ubiquitous assumption come from that gamma is 2.2?

There is room for deviation from the 3/2 Langmuir-Child law. There are different triodes compared to the idealized geometries assumed by the 3/2 law, especially those used in CRTs. But the most significant thing to note is that the 3/2 law applies to current *density*, not total current. The law can be applied to current measurements only if the cross-sectional area of the electron beam is held constant.

Taking the measurements

Eight assorted film recorder class CRTs were assigned to this project. They were set up in their usual operating configuration with their cathode grounded, anode at 18 to 20 KV, and control grid below cutoff. Currents in this state, if any, are leakage currents or ion drift currents and are measured in picoamps.

The voltage between the cathode and the control grid is controlled by a low ripple DC supply, and measured with a high impedance voltmeter. Currents are measured at the cathode and at the anode by floating ammeters. The anode current measurement in particular is sensitive to leakage currents and the meter must be carefully insulated against them (and the operator guarded against touching the meter!).

As the control grid voltage is increased from its initial (negative) value, the triode starts to conduct. Starting with small voltage steps, the currents are measured. As the grid voltage increases further, the currents increase from nanoamps to microamps to milliamps, covering six decades of operation. By the time the last data point is taken, the cathode has been driven long and hard. Several coulombs of electrons have transferred from cathode to anode representing a significant fraction of the useful cathode capacity. In this regard, these tubes donated their lives to science.

Making sense of the measurements

Here are some common ways to display current-voltage data. The straight plot of a typical data set (figure 1) shows clearly an exponential or power law relationship. Diodes follow exponential laws (e^v); CRT's follow power laws (v^γ). The distinction is seen by plotting log i versus voltage. An exponential relationship will plot as a straight line here; the power law will "bend over" as seen in figure 2. There is often confusion among technical workers over the use of the terms "exponential" and "power law" [Poynton92]. This is a good place to point out the distinction.

Figure 1. Linear plot of anode current vs. grid voltage.

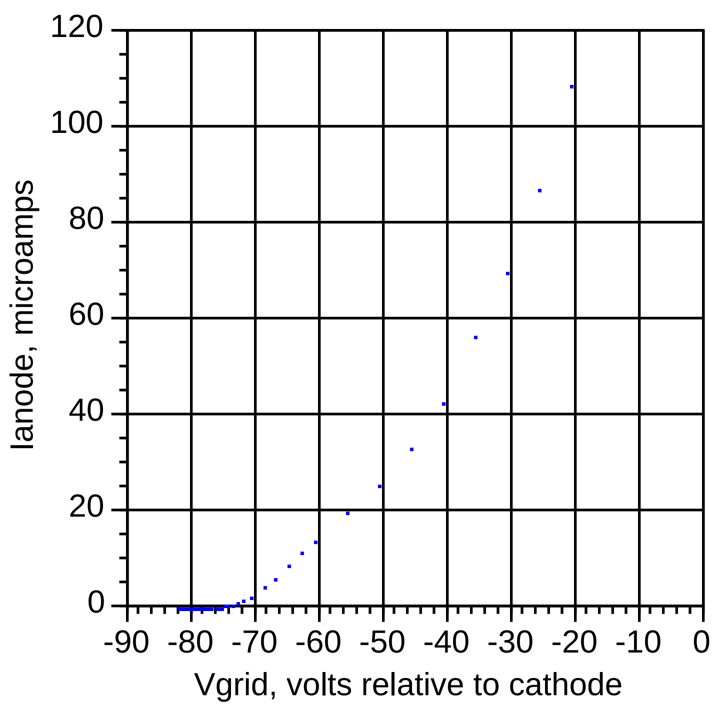
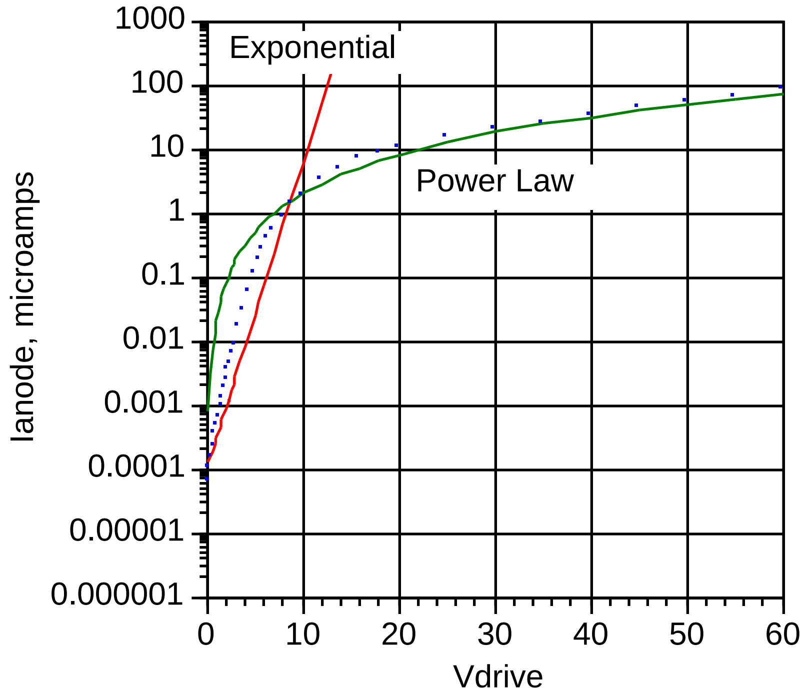


Figure 2.
A log plot of typical anode current data along with sample exponential and power law curves



The power law most often associated with CRT's predicts the output luminance as a function of voltage and has the form:

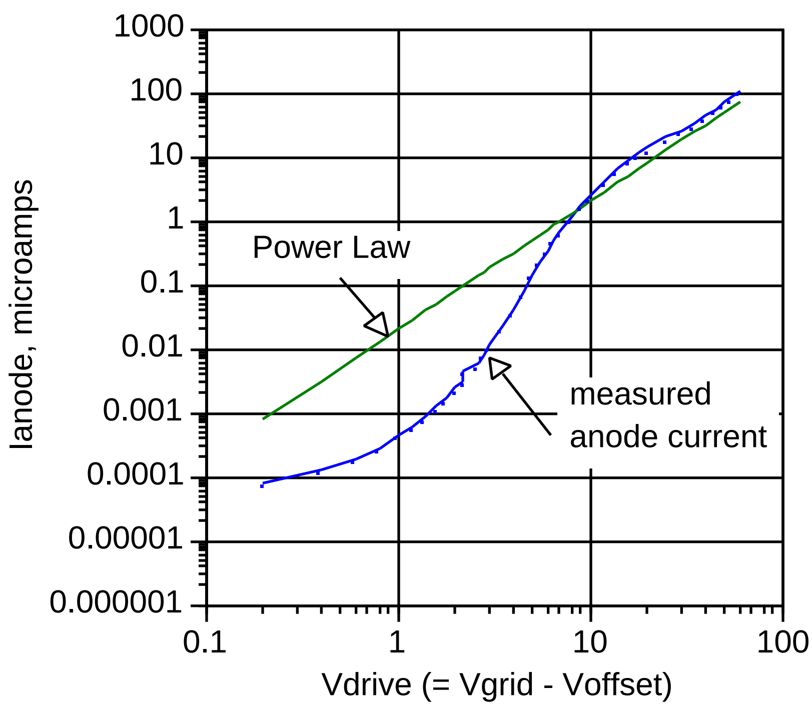
$$L = k(v - V_0)^{\gamma} \qquad (3)$$

where V_0 is the "cutoff voltage" (or offset), and k is a proportionality constant (the gain). The quantity $(v-V_0)$ will be referred to as the device drive level, V_{drive}, or V_d.

Measuring light is often a challenge, especially at low levels. The authors have found, for the operating regions we are interested in, that the light is directly proportional to anode current, i.e., that the phosphor screen is a linear electron-to-photon converter. Measuring anode current, while not trivial, is an excellent way to investigate the behavior of the CRT, thereby yielding quantitative results at very low input levels. The power law above applies immediately, replacing luminance with anode current and adjusting the proportionality constant appropriately.

The power law curve straightens out by plotting both axes as a log scale (figure 3). The slope of the line obtained is the value of the exponent, gamma, in the power law relation. It can be seen from this plot that the value of gamma is not constant over the entire regime.

Figure 3. Typical anode data plotted on log-log scales with sample power law relation.

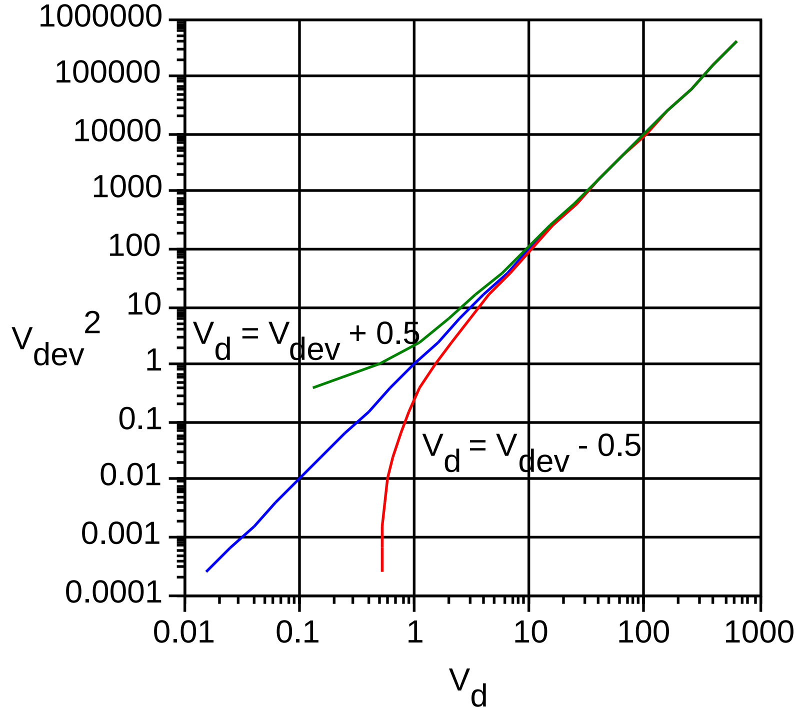


Where is cutoff?

It should be noted that the proper selection of the cutoff voltage V_0, influences the appearance of these plots considerably [Berns93]. Figure 4 illustrates idealized measurements from a perfect power law device when plotted assuming different baselines for the drive voltage.

Some of this is intrinsic to the use of logarithmic axes. The value zero will never be found with which to plot the zero drive condition. Instead, an ever closer approach to it is made, requiring ever higher precision in the data to hold to a perfect power law relation. Any real measurement of course has finite precision, and one of the two conditions in figure 4 must ultimately prevail.

Figure 4.
Effect of a 0.5 volt error in measuring the cutoff voltage for an idealized power law device.



So if we are condemned to always end up with one of the two non-ideal charts above, can we at least *infer* where the cutoff voltage would be in a perfect world? At a more practical level, there are some conditions where the drive voltage is deliberately offset from true cutoff. Can the resulting curve shape be accurately modelled?

Let's define the drive voltage V_d as an applied control voltage, starting at zero, that we think is relative to the cutoff voltage, but is really in error by some small bias amount V_b:

$$V_{dev} = V_d + V_b \qquad (4)$$

where V_{dev} is the voltage scale for which the CRT behaves as a power law device. Negative values of V_b imply that we can drive the CRT below its cutoff without V_d going negative itself; positive values of V_b indicate that we cannot reach cutoff.

When the bias voltage is negative, the response plummets on the log-log plot at that offset. The value can be read directly from the data and the cutoff voltage estimated. When the bias is positive, it is a little more difficult to guess where the cutoff is, because we can never get there on a log-log plot. A method to estimate the cutoff for this case is described in the Appendix.

Modelling cathode behavior

There is no fundamental relation similar to the Langmuir-Childs law that applies to total cathode current. There is considerable experimental data on cathode current however. Over a wide assortment of triodes and operating conditions, the cathode current follows a power law that has an exponent of about 3. At very high output levels the value seems to be 3.5 [Moss68]. This is not a discrepancy, since the Langmuir-Childs law specifies the current *density*, and what we can actually measure is the total current. The area of the electron beam must expand to make up the difference.

The data we collected for anode and cathode currents ranged from picoamps to milliamps and consistently showed two or three distinct operating regimes. A typical data set for one of our CRTs is shown in figure 5.

Consider just the cathode current curve. At low levels, gamma, (the slope of the curve) is approximately 6. It makes a transition to a value of 3 at the higher drive levels and shows signs of increasing to 3.5. If we replot this after estimating the natural device drive voltage V_{dev}, the curve is similar but now the slopes are 9.5 (!) and 3.5 (figure 6).

Figure 5.
Anode and cathode currents measured from CRT under study.

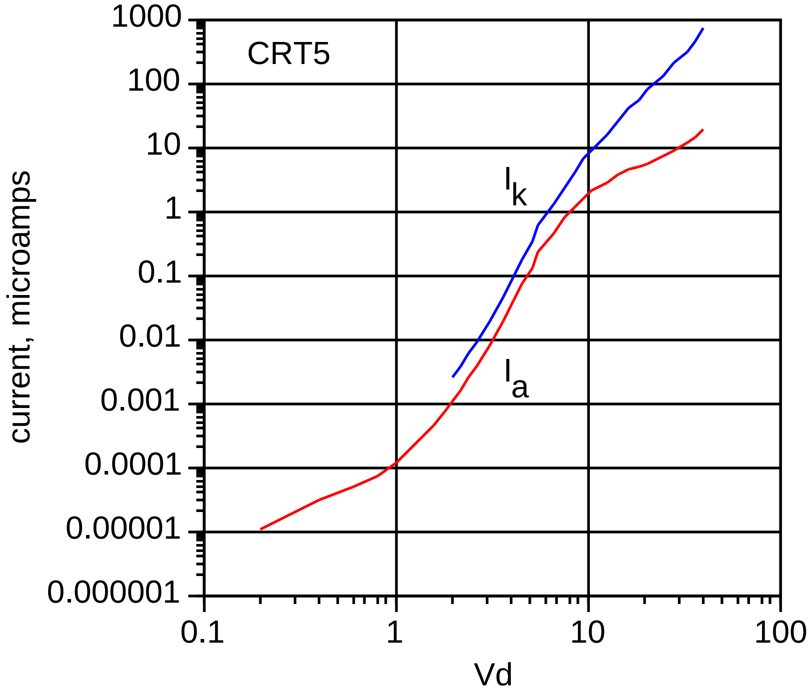
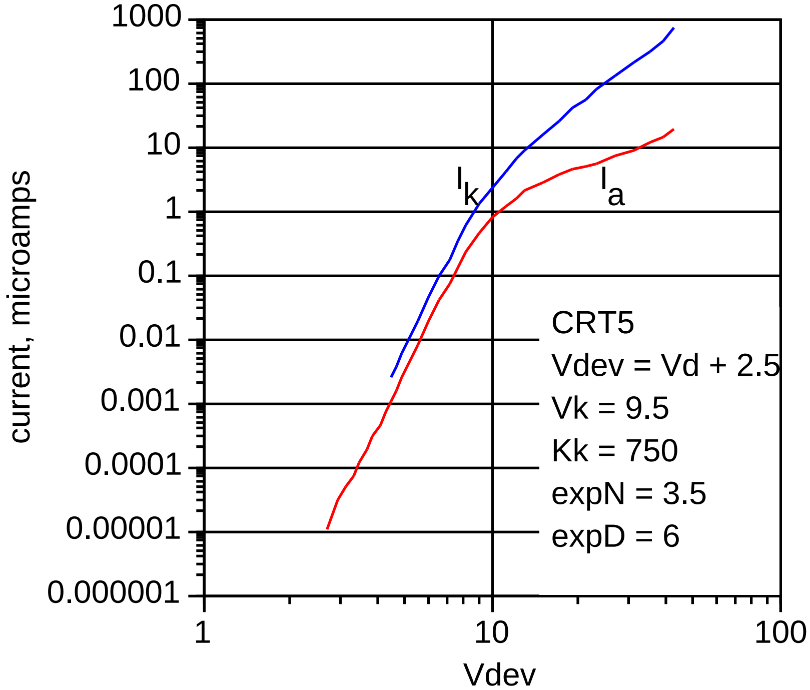


Figure 6.
Current data when plotted against "device voltage" where the drive voltage has been corrected by 2.5 volts.



The hint of two separate regions of slope 3.0 and 3.5 for the cathode current is gone, replaced by a single region of slope 3.5. By the proper choice of the voltage scale, a longstanding empirical observation is explained. There is no gamma 3.0

region for the cathode current, all of the upper operating area has a gamma of 3.5.

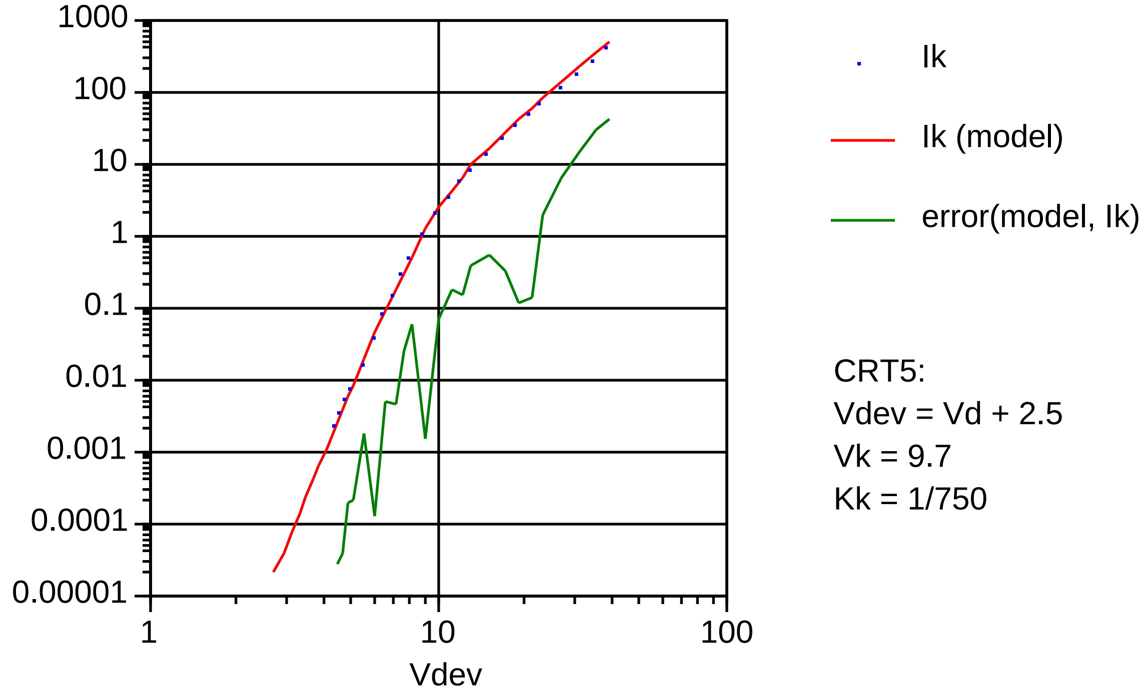
It is easy to see why this uncertainty existed before. The device does not follow the same power law all the way down to zero. Before reaching the cutoff voltage, the exponent changes and generates a sudden drop in ouput. This drop looks deceptively like the case shown in the previous section where the drive voltage had a negative bias on it. This would lead you to estimate the cutoff voltage significantly *higher* than it really is. Instead, the cutoff voltage (that point that causes the device to follow the power law accurately) is really *lower*, and a separate low-level operating region is exposed.

The voltage where the slope transition occurs may differ between tubes and will be a parameter in the model. An equation that describes this behavior is:

$$i_k = K_k \frac{v_{dev}^{3.5}}{1 + \left(\frac{V_k}{v_{dev}}\right)^6} \qquad (5)$$

where V_k is the "cathode transition voltage", and K_k is a proportionality constant for the cathode. This models the data quite well as shown in figure 7 which also shows the error of the fit to the measurements.

Figure 7. Model for the total cathode current.



The focus aperture trim function

Not all of the cathode current makes it to the CRT screen anode. Most of it is intercepted by the downstream focus grid or other trimming aperture. The focus grid in a CRT is a circular metal part with a small hole in its center to let the electron beam through. The beam is diverging as it encounters this electrode, but the electric field near the aperture causes the electrons to bend back. At the correct potential, the beam will converge to a small spot at the screen surface of the anode.

In high resolution CRTs, only the "core" of the diverging beam passes the focus electrode. In other CRT configurations, the beam is also trimmed by apertures, including the anode electrode in an Einzel type lens. The outside electrons are collected by the electrode and return to the cathode via any and all intervening circuitry. The *fraction* of cathode current that continues on to the screen is shown in figure 8.

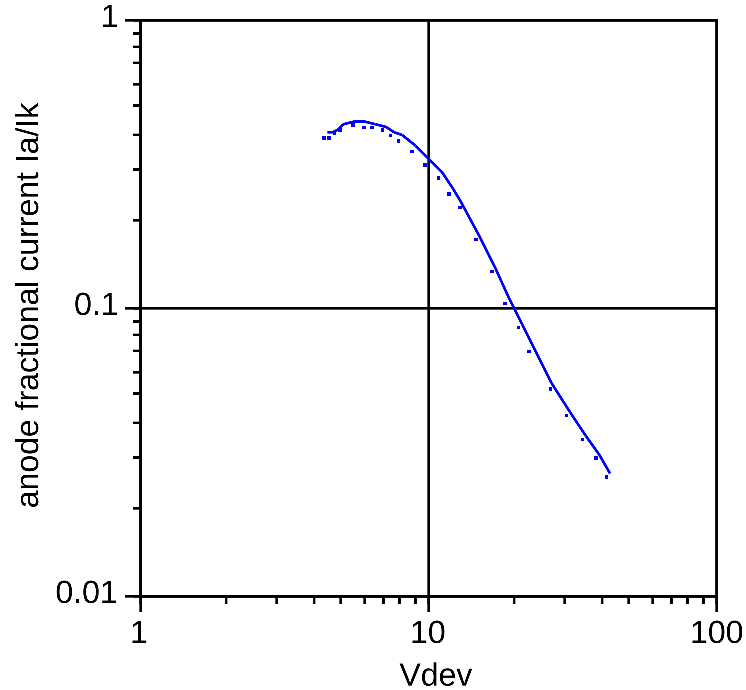


Figure 8. Fraction of cathode current that gets to the CRT screen (anode). From measured data I_a/I_k.

The authors believe that this shows the result of an electron beam which at low levels is initially "trimmed" a constant percentage by the focus aperture. At higher levels, the cross sectional area of the beam increases. The fixed diameter aperture passes a smaller fraction of the widening beam.

The electron density across the beam is often approximated as a gaussian with a characteristic gaussian radius R_g:

$$\rho(r) = Ae^{-\left(\frac{r}{R_g}\right)^2} \qquad (6)$$

The fractional amount of current f_a, which passes through an aperture of radius R_a will be called the aperture *trimming function*, and for a gaussian shaped beam is:

$$f_a = 1 - e^{-\left(\frac{R_a}{R_g}\right)^2} \qquad (7)$$

We can estimate the ratio R_a/R_g when the electron beam is at its initial size, from the ratio of anode current to cathode current. In the data set from CRT5 that we have been examining, the anode fraction starts out around 0.4, indicating that the focus aperture is 0.7 the gaussian size of the beam. (The focus aperture for this tube is 0.1 inch in diameter, making the electron beam about 0.15 in. where it intersects this grid).

At some drive level, the electron density becomes high enough so that the beam starts to diverge due to its own electric field. The beam radius R_g is replaced with the variable r_g. We find empirically that if the beam grows according to:

$$r_g = R_g \frac{1 + \left(\frac{v_{dev}}{V_t}\right)^3}{1 + \left(\frac{v_{dev}}{V_t}\right)^2} \qquad (8)$$

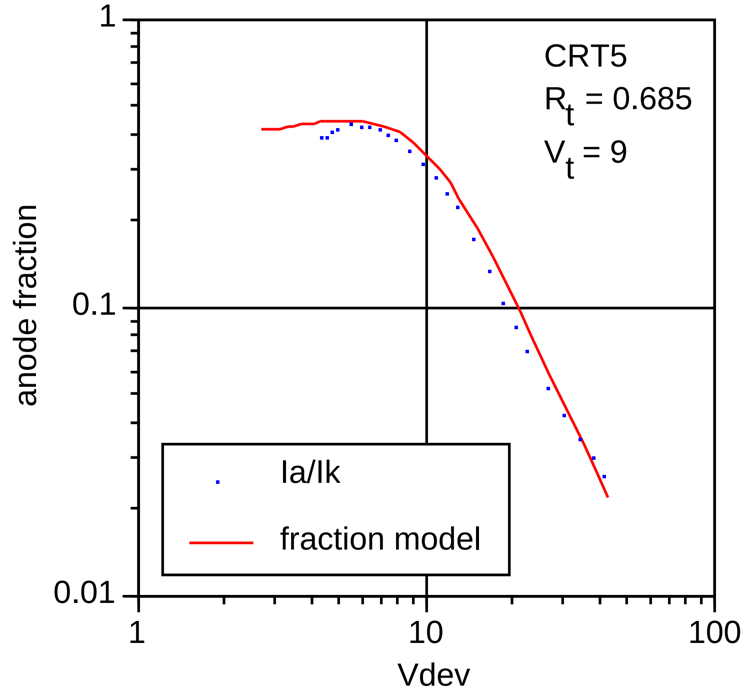
then the anode current falloff would be explained. Rather than maintain explicit knowledge about the focus aperture radius, it is more useful to create an effective *trim radius ratio*, R_t, which is then used in the trimming function as follows:

$$f_a = 1 - e^{-(r_t)^2} \qquad (9a)$$

$$r_t = \frac{R_a}{r_g} = \frac{R_a}{R_g} \frac{1 + \left(\frac{v_{dev}}{V_t}\right)^2}{1 + \left(\frac{v_{dev}}{V_t}\right)^3} = R_t \frac{1 + \left(\frac{v_{dev}}{V_t}\right)^2}{1 + \left(\frac{v_{dev}}{V_t}\right)^3} \qquad (9b)$$

Figure 9 shows the result of the trimming function using this beam growth model. It is satisfying to see that the anode fraction falls off as the square of the voltage. When coupled with the cathode gamma of 3.5 in this region, the result is that the anode gamma converges to 1.5, the number predicted by the Langmuir-Childs law.

Figure 9. Model for the amount of "trimming" on the electron beam.



The exact behavior of the beam growth and its trimming is very likely to be much more complicated. In particular, the gaussian profile is an idealization of a beam which may not be centered in the aperture, and may have various astigmatism and abberation components.

One feature deserves some discussion. At very high drive levels, the anode fraction does not fall off as much as this model predicts. This is seen on most of the CRTs we measured. We speculate that this could be due either to the beam profile being not perfectly gaussian near its center, or that the growth of the beam slows down at these levels.

Whatever the cause, the result of the anode current increasing at a power law rate greater than 1.5 means that the *minimum area* over which that current can be distributed must be increasing. This is because the Langmuir-Childs law applies to current *density*, the current per unit cross sectional area. If we are trying to maintain the smallest possible CRT spot size, we will eventually be restricted by this law of nature. In the high drive level regime, expect to see significant spot growth because of this.

Putting it all together

The goal here is to obtain a simple, yet accurate model for the light output behavior of high resolution CRTs, one that can be used confidently even in the low intensity region. We have discussed the components of such a model:

• The phosphor screen is a linearly proportional electron to photon converter. Luminous power output is directly proportional to electrical power supplied by the electron beam.

• Estimation of the true cutoff which establishes the input scale for which the CRT behaves as a power law device. The parameter of interest is V_b, the offset voltage which converts the applied drive voltage v_d, to the natural device voltage, v_{dev}:

$$v_{dev} = v_d + V_b \qquad (10)$$

- Total cathode current model. This has two parameters: a proportionality constant K_k, and a cathode transition voltage, V_k:

$$i_k = K_k \frac{v_{dev}^{3.5}}{1 + \left(\frac{V_k}{v_{dev}}\right)^6} \qquad (11)$$

- The fraction of cathode current reaching the anode, f_a, which requires two parameters: the trim radius ratio, R_t, and a beam growth transition voltage V_t:

$$f_a = 1 - e^{-r_t^2}$$

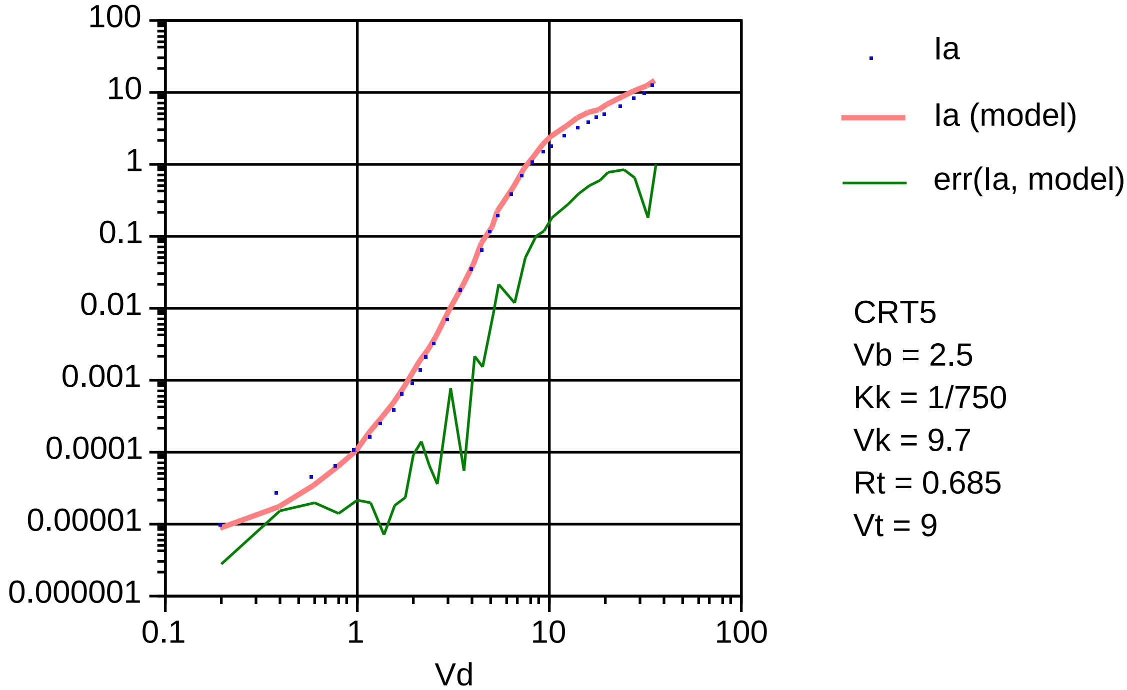
$$r_t = R_t \frac{1 + \left(\frac{v_{dev}}{V_t}\right)^2}{1 + \left(\frac{v_{dev}}{V_t}\right)^3} \qquad (12)$$

The total expression for the anode current becomes:

$$i_a = f_a i_k = K_k \frac{\left(1 - e^{-r_t^2}\right) v_{dev}^{3.5}}{1 + \left(\frac{V_k}{v_{dev}}\right)^6} \qquad (13)$$

Figure 10 compares this result to the data obtained from CRT5. It is a good approximation to the anode current even over the 6 decade span in our measurements.

Figure 10.
The full anode current model compared to measurements on CRT5



Gamma revealed

We have defined gamma as the (local) slope of the current-voltage curve on a log-log scale. The curve in figure 10 shows that it seems never to be constant. We can plot the apparent gamma as computed from:

$$g = \frac{\Delta \log(i_a)}{\Delta \log(v_d)}$$

The comparison between the gamma derived from the measured data and the model are shown in figure 11.

The inflection at the low end of the scale is now understood to be an artifact of using a drive voltage which is offset from the natural device voltage where the power law holds. The value of the offset can be inferred using the technique in the Appendix. If we recompute gamma based on the device voltage scale, figure 12 results.

Figure 11. Apparent gamma as obtained from the model and from measurements, based on drive voltage, V_d

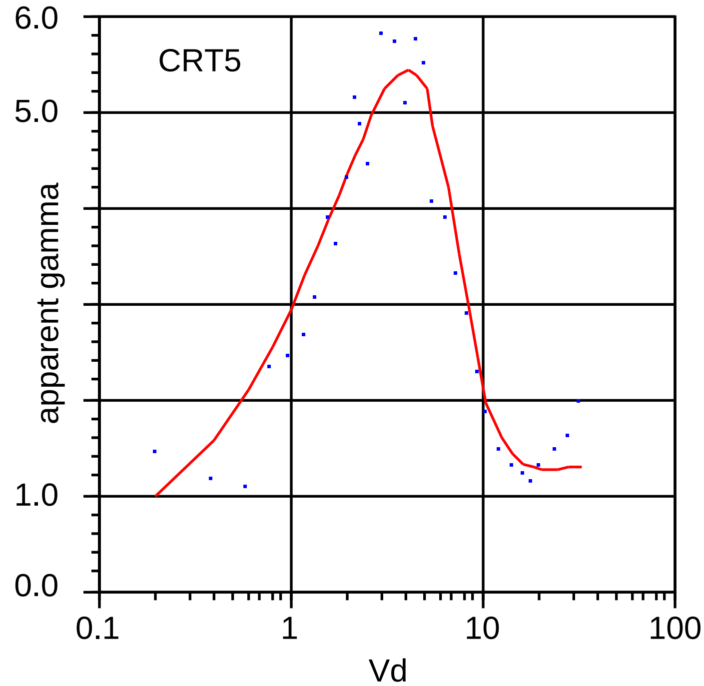
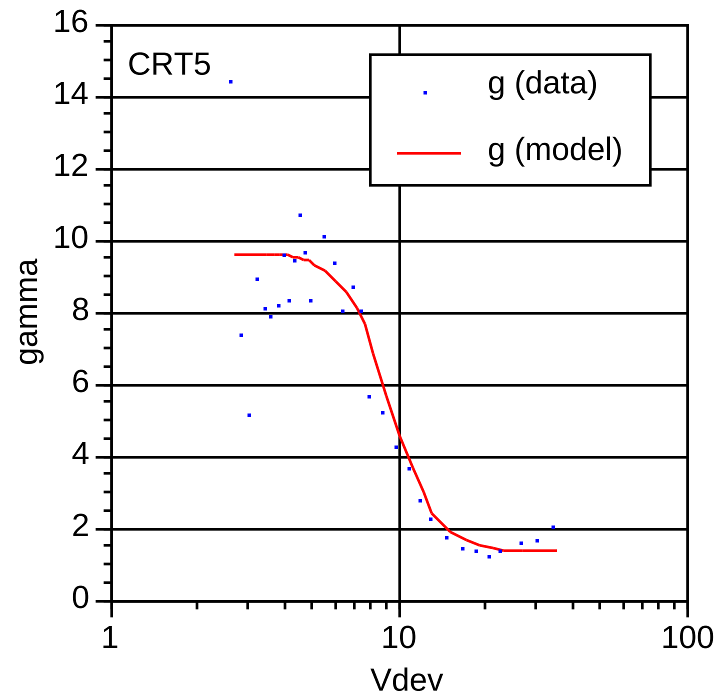


Figure 12. Apparent gamma, obtained from the model and from measurements, based on device voltage, V_{dev}



The modelled gamma matches the measured values nicely with the exception of the extreme high end. It is here that the measured gamma increases beyond the theoretical asymptote of 1.5. We believe this is accompanied by dramatic spot growth and is beyond the region where we operate our high resolution CRTs.

This also explains some of the gamma lore that has evolved. One of the reasons gamma has not been measured consistently is the difficulty of setting the correct drive voltage baseline. Further, the behavior of the tube makes a large transition of gamma between 9.5 and 1.5 right in the middle of most operating regions. No wonder a single consistent number is never obtained! This may even explain the origin of the standard gamma of 2.2 used in video systems.

Conclusion

A model has been presented that can describe the anode current delivered to a CRT screen as a function of its grid drive voltage. Insofar as the screen behaves as a linear electron to photon converter, the model also prescribes the light levels that will be observed.

The value of gamma, the slope of the response curve in log-log scales, is seen to vary from a surprisingly high value of 9.5 to the value 1.5 predicted by the Langmuir-Child law. This indicates that the CRT output will rise very rapidly as it comes out of cutoff. As it enters the normal operating region, the rate of increase moderates. The output is still steadily increasing, but the growth is limited to the 1.5 power of the input.

The fact that the electron beam current is influenced in a complex way explains why gamma is often mismeasured: there is no single value for it. In particular, the precipitous curve shape at low output levels looks deceptively like the cutoff voltage of the tube.

The value of a model is in its application. A fixed exponent model with gamma set to 2.2, or some other number is immensely appealing because of its simplicity. We have found however that it is too simple to be valuable in our application and advise others to determine just the right amount of detail needed. There is an abundant range of complexity behind gamma's simple disguise!

Acknowledgements

A generous thank you to Stan Kriz and Greg Schaffer at Resolution Displays Inc. for taking on the painstaking measurements of the electron beam currents within the cathode ray tubes. It was through the accuracy and low noise precision of the data that this model was derived.

References

[Benson] K. Blair Benson, editor in chief. *Television Engineering Handbook*, McGraw-Hill Book Co, New York, 1986.

[Berns 93] Roy S. Berns, Ricardo J. Motta, Mark E Gorzynski. "CRT Colorimetry. Part I: Theory and Practice" Color Research and Applications, v18, no5, October 1993.

[Poynton 92] Charles Poynton. Private correspondence regarding technical use and misuse of the terms "power law" and "exponential". 1992.

[Poynton 93] Charles Poynton. "'Gamma' and Its Disguises: The Nonlinear Mappings of Intensity in Perception, CRTs, Film and Video" SMPTE Journal, v102, no12, December 1993.

[Moss 68] Hilary Moss. *Narrow Angle Electron Guns and Cathode Ray Tubes*, Academic, New York, 1968.

[Ryder] John D. Ryder. *Electronic Engineering Principles, with revisions*. Prentice-Hall Inc., New York, 1947.

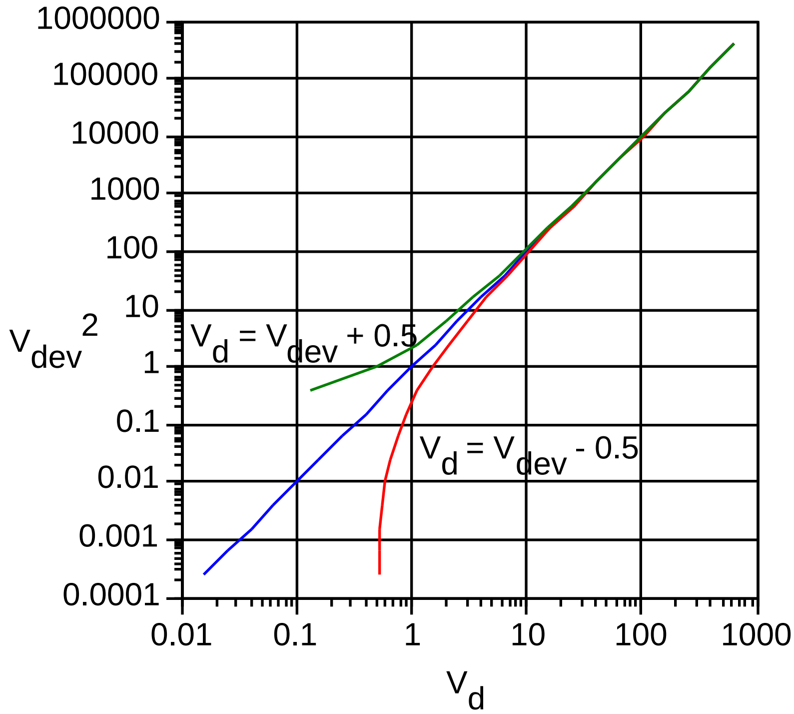
[Terman] Frederick Emmons Terman. *Radio Engineers Handbook*, First Edition. McGraw-Hill Book Company, New York, 1943.

Appendix

Estimating the cutoff voltage

It was noted in an earlier section that sometimes the stated drive voltage to a CRT is not really relative to the natural device cutoff voltage; there may be some additional offset. The effect of this offset on a true power law device was shown in figure 4; it is repeated as figure A1.

Figure A1. The effect of a 0.5 volt error in measuring the cutoff voltage for an idealized power law device.



When the offset is negative, the response plummets on the log-log plot and the value can be read directly from the data. When the error is positive, it is harder to tell where the cutoff is, because we can never get there on a log-log plot. Here is a method to estimate the cutoff for this case.

First, remember that what we have plotted is the power law device response against the log of the applied drive voltage. Whether this is done on common log scales or on natural logs, the apparent value of gamma, g, is the slope of the curve:

$$g(x) = \frac{d}{dx}\gamma \ln\left(V_{dev}\right) = \frac{V_d}{V_d + V_b}\gamma \qquad \text{(A1)}$$

where x is $\ln(V_d)$. This is comforting because it shows that the apparent value of gamma, g, approaches the real value γ, as the drive voltage increases.

If we were to plot the apparent value of gamma as a function of log output current, y, instead of input drive we would have the following:

$$y = \gamma \ln(V_d + V_b) + \ln(k)$$

$$g(y) = \gamma(1 - V_b\, e^{-(y-\ln(k))/\gamma}) \qquad \text{(A2)}$$

This expression is valid for both positive and negative values of V_b. Note that for large output values, g(y) approaches γ, as expected, but this is particularly useful for when V_b is positive. In this case, the apparent gamma crosses zero for:

$$V_b = e^{(y(0)-\ln(k))/\gamma} \qquad \text{(A3)}$$

which can be solved for V_b, and the true cutoff obtained, provided that the true gamma, γ, and the gain constant k is known. On common log scales use:

$$V_b = 10^{(y(0)-\log(k))/\gamma} \qquad \text{(A4)}$$

An example is shown in figure A2 which indicates the apparent gamma crossing zero at a log output value of -0.6. This corresponds to the ideal device curve shown above (figure A1) whose true gamma was 2, but was plotted with a bias offset of 0.5V ($0.5 = 10^{-0.6/2}$).

Figure A2. Apparent gamma plotted as a function of output. A positive cutoff voltage error can be found from the zero crossing.