

The RMS Average.

By David W. Knight*

Version 1.01 (minor update), 4th April 2013.

(1st pdf version: v1.00, 11th Nov. 2009)

© D. W. Knight, 2009, 2013.

David Knight asserts the right to be recognised as the author of this work.

* Ottery St Mary, Devon, England.

The most recent version of this document can be obtained from the author's website:

<http://www.g3ynh.info/>

Introduction

Most people who work with electricity are aware that the RMS of a sinusoidal voltage or current waveform is $1/\sqrt{2}$ times the peak value. After that however, knowledge tends to become a little hazy. Design and measurement errors due to an incorrect understanding of the RMS can lead to expensive and sometimes dangerous failures in power supplies, amplifiers and transmission equipment. This article attempts to clarify the issue.

The heating effect of a sinusoidal voltage or current

Let us begin by considering a pure sinusoidal voltage. The RMS average of the wave is the equivalent DC voltage that, when connected across a resistance (R say), will produce the same heating effect. The heating effect is the average energy delivered per unit-of-time, i.e., the average power, which is given by:

$$P_{av} = V_{rms}^2 / R \quad \dots \quad (1)$$

The alternating voltage is described by the expression:

$$V = V_{pk} \sin\phi$$

Where V_{pk} is the peak or maximum instantaneous voltage magnitude, and $\phi=2\pi ft$ is the time-varying phase angle. If the voltage is applied to a resistor, the instantaneous power will be:

$$P = V^2 / R$$

i.e.:

$$P = V_{pk}^2 \sin^2\phi / R$$

but, using the standard trigonometric identities:

$$\sin^2\phi = 1 - \cos^2\phi$$

and

$$\cos(2\varphi) = 2\cos^2\varphi - 1$$

i.e.,

$$\cos^2\varphi = [1 + \cos(2\varphi)] / 2$$

we obtain the substitution:

$$\begin{aligned} \sin^2\varphi &= 1 - [1 + \cos(2\varphi)] / 2 \\ &= [1 - \cos(2\varphi)] / 2 \end{aligned}$$

Hence the instantaneous power function is:

$$P = V_{pk}^2 [1 - \cos(2\varphi)] / (2R)$$

This is a sinusoid at twice the frequency of the voltage wave, with an offset so that it just touches zero and is always positive. The average of a pure sinusoid is zero, and so the average power is obtained simply by deleting the cosine term:

$$P_{av} = V_{pk}^2 / (2R)$$

comparing this with (1) we have:

$$V_{rms}^2 / R = V_{pk}^2 / (2R)$$

i.e.

$$V_{rms} = V_{pk} / \sqrt{2}$$

Also, by a similar argument using

$$P_{av} = I_{rms}^2 R$$

it can easily be shown that

$$I_{rms} = I_{pk} / \sqrt{2}$$

The square-Root of the Mean of the Squares

Finding the average power delivered by a pure sinusoidal voltage or current is easy; but we also need a general method that works for any waveform. Without it we cannot even calculate the power ratings for the resistors in simple narrow-band amplifiers, because the combination of a direct and a sinusoidal current is no longer purely sinusoidal. It turns out that the required technique is analogous to the method used to calculate the standard deviation of a set of measurements.

First square the voltage (or current) waveform to obtain a time-varying function proportional to the instantaneous power. Now sample that power function at a large number of equally spaced

intervals over a whole cycle. Add the sample values together and divide by the total number of samples. This gives a quantity proportional to the mean (i.e., average) power. Now take the square root to get back to the original units of voltage (or current). The voltage (or current) obtained is the square-Root of the Mean of the Squares, commonly known as the RMS.

To find the RMS with perfect accuracy, of course, we need to take an infinite number of samples. The summation used to obtain the average thus becomes an integral. We need to take the average over the repetition interval of the waveform, so if we use ϕ as our relative measure of time, we need to integrate over the interval between 0 and 2π radians (i.e., over 360°). The interval between samples is the infinitesimal $d\phi$, and so the (infinite) total number of samples is defined as $2\pi/d\phi$. Hence, dividing the infinite sum by the infinite total number of samples, we have:

$$V_{\text{rms}}^2 = \frac{1}{2\pi} \int_0^{2\pi} V^2 d\phi \quad (2)$$

Of course, this doesn't look quite as easy as the rule we know for dealing with pure sine waves, but this will not prove to be a serious problem. As anyone who has ever used a radio receiver ought to have noticed, and J B J Fourier discovered long before spectrum analysers were invented, an arbitrary waveform can always be represented as a sum of sinusoids plus a constant. Squaring a series can be a messy business, but it turns out that the cross-terms generated by squaring a Fourier series cancel in the integration process. The physical reason is that a resistive load is a linear network. The various components of the signal do not mix, and so can be considered to exist independently.

The RMS of a pure sine wave

Before complicating matters, let us first check that the general method really works for the case for which we already know the answer. A pure sinusoidal voltage is given by the expression:

$$V = V_{\text{pk}} \text{Sin}\phi$$

Substituting this into (2) gives:

$$V_{\text{rms}}^2 = \frac{1}{2\pi} \int_0^{2\pi} V_{\text{pk}}^2 \text{Sin}^2\phi d\phi$$

Earlier on we obtained the identity:

$$\text{Sin}^2\phi = [1 - \text{Cos}(2\phi)] / 2$$

and so we have:

$$V_{\text{rms}}^2 = \frac{V_{\text{pk}}^2}{4\pi} \int_0^{2\pi} [1 - \text{Cos}(2\phi)] d\phi$$

Now, as any good book of standard integrals will tell you:

$$\int \cos(ax) dx = (1/a) \sin(ax) + c$$

and so the solution is:

$$V_{\text{rms}}^2 = \frac{V_{\text{pk}}^2}{4\pi} \left[\phi - \left(\frac{1}{2}\right) \sin(2\phi) \right]_0^{2\pi}$$

i.e.:

$$V_{\text{rms}}^2 = \frac{V_{\text{pk}}^2}{4\pi} [(2\pi - 0) - (0 - 0)]$$

Hence:

$$V_{\text{rms}} = V_{\text{pk}} / \sqrt{2}$$

The RMS of an offset sine wave

Now let us consider the case when a sinusoidally varying voltage is added to a constant (i.e., DC) voltage. In this case the instantaneous voltage can be written:

$$V = V_{\text{dc}} + V_{\text{pk}} \sin\phi$$

where V_{pk} is defined as the peak value of the sinusoidal component on its own, not the peak value of the total. Substituting this into (2) gives:

$$V_{\text{rms}}^2 = \frac{1}{2\pi} \int_0^{2\pi} (V_{\text{dc}} + V_{\text{pk}} \sin\phi)^2 d\phi$$

Which can be multiplied out:

$$V_{\text{rms}}^2 = \frac{1}{2\pi} \int_0^{2\pi} (V_{\text{dc}}^2 + 2 V_{\text{dc}} V_{\text{pk}} \sin\phi + V_{\text{pk}}^2 \sin^2\phi) d\phi$$

Now we have produced a cross-term; but since the phrase "linear network" has already cropped-up, we must suspect that there will be no interaction between the components in the final reckoning. So, of course, we notice that the cross-term is a sinusoid, and all sinusoids average to zero over an integer number of cycles. Hence, provided that the integration is carried out over the repetition

interval of the waveform, any first-order (i.e., not squared) sine and cosine terms can simply be deleted. The mathematical reason is that, for integer n :

$$\sin(2n\pi) = 0 \quad , \quad \sin(0) = 0 \quad \text{and} \quad \cos(2n\pi) = \cos(0) = 1$$

Hence, deleting the first-order sine term, and substituting for $\sin^2\phi$ as before:

$$V_{\text{rms}}^2 = \frac{1}{2\pi} \int_0^{2\pi} \{ V_{\text{dc}}^2 + V_{\text{pk}}^2 (\frac{1}{2}) [1 - \cos(2\phi)] \} d\phi$$

Now the $\cos(2\phi)$ term dies out in the averaging process, and the solution of the integral is simply a matter of multiplying the remaining terms by a factor of 2π , which cancels with the $1/2\pi$ before the integral symbol. Hence:

$$V_{\text{rms}}^2 = V_{\text{dc}}^2 + V_{\text{pk}}^2/2$$

But now let us define a voltage V_{ac} that is the RMS value of the AC component on its own. Then, using the definition for the RMS of a pure sinusoid, we have:

$$V_{\text{ac}}^2 = V_{\text{pk}}^2/2$$

Hence:

$$V_{\text{rms}}^2 = V_{\text{dc}}^2 + V_{\text{ac}}^2$$

When an alternating voltage (or current) with a DC offset is applied to a resistance, the total power is given by calculating the power for the AC and DC components separately and adding the results .

The fact that a DC component combines with an AC component to increase the RMS is one reason why half-wave rectification should be avoided when building DC power-supplies. A half-wave rectified supply not only needs more smoothing than a full-wave supply, it also adds a DC component to the current drawn from the mains transformer. This increases the dissipation in the loss resistance of the transformer winding, and increases the core loss by introducing a static component into the magnetic field.

Some years ago, there was a period when certain power-hungry items of consumer electronic equipment (large colour-TV receivers in particular) were a serious fire-hazard. This problem was associated with conventional transformer power-supplies and half-wave rectification, but not in an obvious way. It is often said that the more pernicious the error, the more slavishly it is copied, and this is a case in point. The power-supplies in question used bi-phase (i.e., two-diode, centre-tapped winding) full-wave rectification, which is fine; but someone had come up with the idea of fitting a fuse in series with each diode to protect against diode short-circuit failure. The problem is that fuses are stressed by switch-on surges; and if one of the fuses fails, half-wave rectification occurs, and the equipment carries on working while the mains-transformer overheats. The correct way to design a bi-phase power-supply is to arrange things so that the primary-side fuse blows if a diode is connected across the transformer secondary.

The RMS of an arbitrary waveform

An arbitrary waveform can always be written as a Fourier series. This has the form:

$$V = V_{dc} + \sum_{n=1}^{\infty} [V_{c,n} \cos(n\phi) + V_{s,n} \sin(n\phi)]$$

Here we use the series to represent a voltage, but it can represent a current just as well. If we want to find the RMS of such a series we have to square it. This operation will generate a shower of cross-terms, but they must somehow vanish in the subsequent integration. Firstly, we have already seen that the cross-terms produced by multiplying the DC component with a sinusoid are of first-order, and so they average to nothing. The DC component does not interact with any of the AC components. Secondly, consider the following well-known trigonometric identities:

$$\sin(a) \cos(b) = (1/2)[\sin(a+b) + \sin(a-b)]$$

$$\cos(a) \sin(b) = (1/2)[\sin(a+b) - \sin(a-b)]$$

$$\cos(a) \cos(b) = (1/2)[\cos(a+b) + \cos(a-b)]$$

$$\sin(a) \sin(b) = (-1/2)[\cos(a+b) - \cos(a-b)]$$

Whenever two components of different frequency are multiplied together, they will produce a pair of first-order terms. Therefore *all* of the cross-terms will average to zero in the integration. None of the components interact. Hence the square of the RMS voltage comes out as:

$$V_{rms}^2 = V_{dc}^2 + \frac{1}{2\pi} \int_0^{2\pi} \left[\sum_{n=1}^{\infty} [V_{c,n}^2 \cos^2(n\phi) + V_{s,n}^2 \sin^2(n\phi)] \right] d\phi$$

We obtained the substitutions needed to solve this integral earlier. They are:

$$\cos^2\phi = [1 + \cos(2\phi)] / 2 \quad \text{and} \quad \sin^2\phi = [1 - \cos(2\phi)] / 2$$

When we make these substitutions, the first-order cosine terms will, of course, average to zero. Hence:

$$V_{rms}^2 = V_{dc}^2 + \frac{1}{2\pi} \int_0^{2\pi} \left[\sum_{n=1}^{\infty} [(1/2)V_{c,n}^2 + (1/2)V_{s,n}^2] \right] d\phi$$

Each remaining term in the summation, upon evaluating the integral, will come out multiplied by 2π , which will cancel with the $1/2\pi$ before the integral symbol. Thus:

$$V_{rms}^2 = V_{dc}^2 + \sum_{n=1}^{\infty} [(1/2)V_{c,n}^2 + (1/2)V_{s,n}^2]$$

The voltage coefficients after the summation symbol are, of course, peak values. If we replace each one with its RMS, then we lose the factors of $\frac{1}{2}$. Hence:

$$V_{\text{rms}}^2 = V_{\text{dc}}^2 + \sum_{n=1}^{\infty} [V_{c(\text{rms}),n}^2 + V_{s(\text{rms}),n}^2]$$

In general, if a voltage (or current) waveform is made up of an arbitrary combination of components of different frequency, each having an RMS value V_n (or I_n) when considered on its own, the true RMS of the combination is given by:

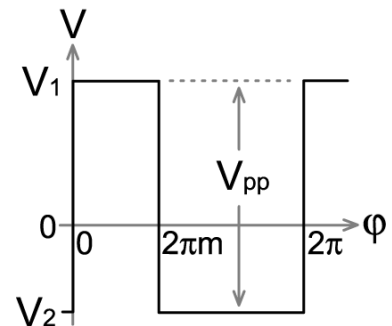
$$V_{\text{rms}} = \sqrt{\left[V_{\text{dc}}^2 + \sum_{n=1}^{\infty} V_n^2 \right]}$$

When a voltage is connected across (or a current is passed through) a resistance, the power in the resistance is given by the sum of the powers due to the individual frequency components.

The RMS of a floating square wave

We don't always have to know the Fourier coefficients in order to find the RMS. Notwithstanding the fact that we can use an electronic integrator circuit to find the mean of the squares (that's how a true-RMS meter works), there are some waveforms that can easily be integrated piecewise. The most obvious example is the square wave, which is also an important special case.

Represented on the right is a voltage square-wave of arbitrary mark-space ratio that is AC-coupled to a resistive load. The mark interval as a proportion of the whole cycle is defined as m ; so that over the course of one cycle, the wave is positive for $2\pi m$ radians, and negative for $2\pi(1-m)$ radians. The peak positive and negative voltages are V_1 and V_2 respectively, and since V_2 is by definition negative, the peak-to-peak voltage is defined as:



$$V_{\text{pp}} = V_1 - V_2$$

Because the voltage contains no DC component, its average is zero. Hence we can find the values of V_1 and V_2 relative to V_{pp} by taking the integral of the voltage between 0 and 2π and setting its value to zero, i.e.:

$$\int_0^{2\pi m} V_1 d\phi + \int_{2\pi m}^{2\pi} V_2 d\phi = 0$$

Thus:

$$[2\pi m V_1 - 0] + [2\pi V_2 - 2\pi m V_2] = 0$$

i.e.:

$$m (V_1 - V_2) = -V_2 \quad \dots \quad (3)$$

but

$$V_{pp} = V_1 - V_2$$

therefore

$$V_2 = -m V_{pp} \quad \dots \dots \dots (4)$$

also

$$V_2 = V_1 - V_{pp}$$

which can be substituted into (4) to give:

$$V_1 = (1-m) V_{pp} \quad \dots \dots \dots (5)$$

The RMS value of the waveform can be obtained by piecewise integration of the square of the instantaneous voltage:

$$V_{rms}^2 = \frac{1}{2\pi} \int_0^{2\pi m} V_1^2 d\phi + \frac{1}{2\pi} \int_{2\pi m}^{2\pi} V_2^2 d\phi$$

Which gives:

$$V_{rms}^2 = [m V_1^2 - 0] + [V_2^2 - m V_2^2]$$

i.e.:

$$V_{rms}^2 = m (V_1^2 - V_2^2) + V_2^2 = m (V_1 + V_2) (V_1 - V_2) + V_2^2$$

Using (3) as a substitution gives:

$$V_{rms}^2 = -V_2 (V_1 + V_2) + V_2^2$$

i.e.:

$$V_{rms}^2 = -V_2 V_1$$

This is one simple form for the RMS of a floating square-wave that does not require prior knowledge of m (note that V_{rms}^2 is positive because one of the voltages on the right-hand side has to be defined as negative). We obtain a more informative definition however by using equations (4) and (5) as substitutions:

$$V_{rms}^2 = m V_{pp} (1-m) V_{pp}$$

i.e.:

$$V_{\text{rms}}^2 = m(1-m) V_{\text{pp}}^2$$

The factor $m(1-m)$ has the property that it is zero when $m = 0$ or 1 (infinitely narrow pulses contain no energy) and is at its peak when $m = 1/2$. This tells us that, for a given peak-to-peak voltage, assuming that the waveform is floating (i.e., contains no DC component), the RMS value of a square-wave is at a maximum when the mark-space ratio is unity, i.e., when $m/(1-m)=1$.

When a square wave is symmetric, $V_1 = -V_2$, and so the peak voltage is half the peak-to-peak voltage, i.e.:

$$V_{\text{pp}} = 2 V_{\text{pk}}$$

Hence, for symmetric square waves only, i.e., when $m(1-m)=1/4$:

$$V_{\text{rms}}^2 = (2V_{\text{pk}})^2 / 4$$

i.e.:

$$V_{\text{rms}} = V_{\text{pk}}$$

The RMS of a floating symmetric square wave is $\sqrt{2}$ times greater than that of a sine wave.

The spectrum of a symmetric square-wave contains only the odd harmonics of the fundamental repetition frequency. This is by no means the richest possible spectrum in terms of harmonics present, but it gives rise to the AC waveform that conveys the maximum amount of energy for a given peak magnitude, i.e., it achieves the limiting condition $V_{\text{rms}}/V_{\text{pk}} \rightarrow 1$. The RMS can be increased still further however, by adding a DC component.

The RMS of a zero-based square-wave

Given the ubiquity of logic circuits, the most commonly encountered type of square-wave is the zero-based or clamped variety. The effect of clamping one extreme of the voltage excursion to zero is to introduce a DC component equal to the extent to which the wave would have crossed over the zero line had it not been clamped. Thus the DC component is $-V_2$ from the previous section, which is given by equation (4) as:

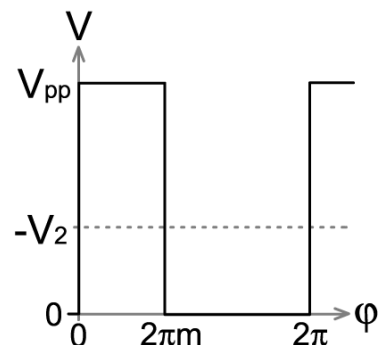
$$-V_2 = m V_{\text{pp}}$$

Hence the RMS of a zero-based square wave is given by:

$$V_{\text{rms}}^2 = m(1-m) V_{\text{pp}}^2 + m^2 V_{\text{pp}}^2 = V_{\text{pp}}^2 [m - m^2 + m^2]$$

i.e.:

$$V_{\text{rms}}^2 = m V_{\text{pp}}^2$$



$$V_{\text{rms}} = V_{\text{pp}} \sqrt{m}$$

We can also obtain exactly the same result by integration over the part of the cycle corresponding to the 'on' or 'mark' period:

$$V_{\text{rms}}^2 = \frac{1}{2\pi} \int_0^{2\pi m} V_{\text{pp}}^2 d\phi = m V_{\text{pp}}^2$$

Notice that when the mark and space intervals are equal ($m=1/2$), the zero-based waveform has twice the RMS of the floating waveform. The maximum RMS for the zero-based waveform however occurs when $m=1$; but this just means that maximum power is delivered to a resistive load when the generator is producing a constant output of V_{pp} , i.e., when the generator is stuck in the 'on' position.

The RMS of a noise spectrum

In the previous derivations, it has been convenient to assume that there is a harmonic relationship between all of the AC components of a waveform. This ensures that, if we integrate over the fundamental repetition interval, all of the components to be averaged will be sampled for an integer number of cycles. This is the analytical (i.e., mathematical) RMS; but it is also possible to determine a statistical or experimental RMS, which tends towards the true RMS when the measurement is repeated several times and the results are averaged.

For the statistical RMS, there is no need to synchronise the sampling interval to the period of the waveform. It is sufficient to sample the square of the voltage or current for a time that is long enough to capture a few cycles of the lowest alternating component of significant magnitude. If the sampling interval is randomly related to a sinusoidal component, there will be a maximum error of $\pm 1/2$ of a cycle. Hence sometimes the result will come out a little too high, sometimes it will come out a little too low, but the average of many measurements will tend towards the true value. Furthermore, the error incurred by the lack of synchronisation will become smaller as the number of cycles captured in the sampling window increases; which means that accurate measurements can be made by sampling over a relatively long period.

It follows, that by a process of electronic integration, we can determine the RMS of a voltage or current consisting of a set of harmonically-unrelated frequency components; i.e., it is possible to define and measure the RMS of pure noise.

