Myth or fact? You can increase the force of your solenoid by removing turns. (Part 2 of 3)

Paul H. Schimpf
Eastern Washington University

From Energy to Force

Force is the derivative of energy with respect to the position of the armature. For this we are going to need an expression of the variation of inductance with position, *L(x)*. Let's begin by defining a coordinate system, as shown in **Figure 4**:

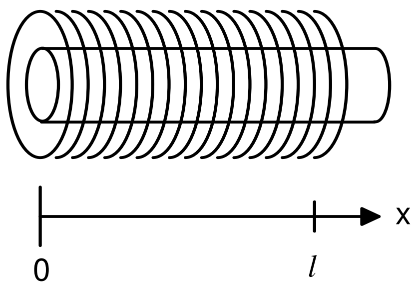

Figure 4: Armature coordinate system

At x=0 the armature is all the way inside the coil. At x=l the armature is at the entry edge of the coil. These are two positions for which the inductance is easy to define using the standard approximation for the inductance of a coil:

$$L(0) = L_0 = \frac{\mu_r \mu_0 N^2 A}{l} \tag{13}$$

$$L(l) = L_l = \frac{L_0}{\mu_r} = \frac{\mu_0 N^2 A}{l} \tag{14}$$

In actuality, you might want to define (14) as somewhat farther outside the coil in order for the armature to avoid the fringe of the B field, thereby improving the accuracy. For the sake of simplicity, however, we use the length of the coil in (14) and provide an alternative way to make that adjustment below. *L(0)* is much larger than *L(l)* because μ_r, the relative permeability of the iron armature, is much larger than μ_0, the permeability of free space. *L(x)* will vary monotonically between these two extremes, and the precise shape of this variation with position depends on the construction and shape of the solenoid.

A precise model of this variation requires development of a detailed geometric model (such as a finite element model) of the solenoid's magnetic field. For our purposes here, all we require is a reasonable approximation for a cylindrical solenoid, and for now we will assume an exponential decay:

$$L(x) = L_0 e^{-\frac{\alpha}{l}x} \tag{15}$$

As required, this has the value L_0 at $x=0$, and we want to choose the parameter α such that $L(l)=L_0/\mu_r$:

$$\begin{aligned} L_0 e^{-\frac{\alpha}{l}l} &= \frac{L_0}{\mu_r} \\ e^{-\alpha} &= \frac{1}{\mu_r} \\ -\alpha &= \ln\left(\frac{1}{\mu_r}\right) \\ \alpha &= \ln(\mu_r) \end{aligned} \tag{16}$$

A value for α somewhat less than this, but larger than unity, will have the same effect as defining (14) outside the end of the coil. We can now derive the force, as follows:

$$\begin{aligned} E_L &= \frac{V^2}{2R^2} L(x) \\ F &= \frac{dE_L}{dx} = \frac{d}{dx}\left[\frac{V^2}{2R^2} L(x)\right] = \frac{V^2}{2R^2}\frac{d}{dx} L(x) \\ \frac{d}{dx} L(x) &= L_0 \frac{d}{dx} e^{-\frac{\alpha}{l}x} = -L_0 \frac{\alpha}{l} e^{-\frac{\alpha}{l}x} \\ F(x) &= -\frac{V^2}{2R^2}\frac{\alpha}{l} L_0 e^{-\frac{\alpha}{l}x} \end{aligned} \tag{17}$$

At this point we make several substitutions in order to get the force in terms of coil design parameters. We start by computing R from the resistance per unit length of the wire, γ, and the total length of wire, which is computed from the number of turns times the average circumference of the turns:

$$R = 2\pi r_a N\gamma \tag{18}$$

It should be noted that r_a is not the same as the inside radius of the coil, which we shall call r_0. The difference between these is illustrated in **Figure 5**.

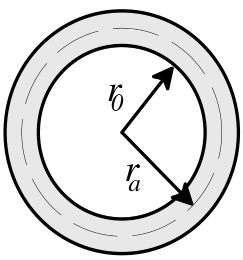

Figure 5: Coil radius vs. average turn radius

The parameter γ may be obtained from a table of wire gauges (Reference 4 and Reference 5), and is also computable from the resistivity of copper, ρ, and the cross-sectional area of the wire, a:

$$\gamma = \frac{\rho}{a} \tag{19}$$

We also substitute into (17) the inductance at the stop position, L_0, along with the cross-sectional area of the coil opening, A:

$$L_0 = \frac{\mu_r \mu_0 N^2 A}{l} \tag{20}$$

$$A = \pi r_0^2 \tag{21}$$

Substituting (21) into (20), and substituting the result along with (18) into the last line of (17), results in:

$$\boxed{\begin{aligned} F &= \frac{-V^2 \mu_r \mu_0}{8\pi\gamma^2 l^2}\left(\frac{r_0}{r_a}\right)^2 \alpha e^{-\frac{\alpha}{l}x} \\ &= \frac{-V^2 \mu_r \mu_0}{8\pi\gamma^2 l^2} W_f \alpha e^{-\frac{\alpha}{l}x} \end{aligned}} \tag{22}$$

In (22) we have defined the squared ratio between r_0 and r_a as a new parameter, the *winding factor*, W_f. The winding factor is always less than unity, and represents a reduction on what might otherwise be considered the nominal force:

$$\begin{aligned} F &= F_{nom} W_f \\ W_f &= \left(\frac{r_0}{r_a}\right)^2 \end{aligned} \tag{23}$$

Discussion

A comparison between the predictions of (22) and a commercially-available cylindrical solenoid are illustrated in **Figure 6**. The simulated coil used the following parameters: l=27 mm, r_0 = 2.3 mm, r_a = 4.5 mm, AWG = 30, N=572. Both the simulated and actual coil resistances were 5.3 Ω. As can be seen, the simulated curve matches the high power curve for the solenoid quite well. At lower powers, the force of the actual solenoid drops off at a higher exponential rate. This could be simulated by substituting a higher power of x in the inductance formula of (15).

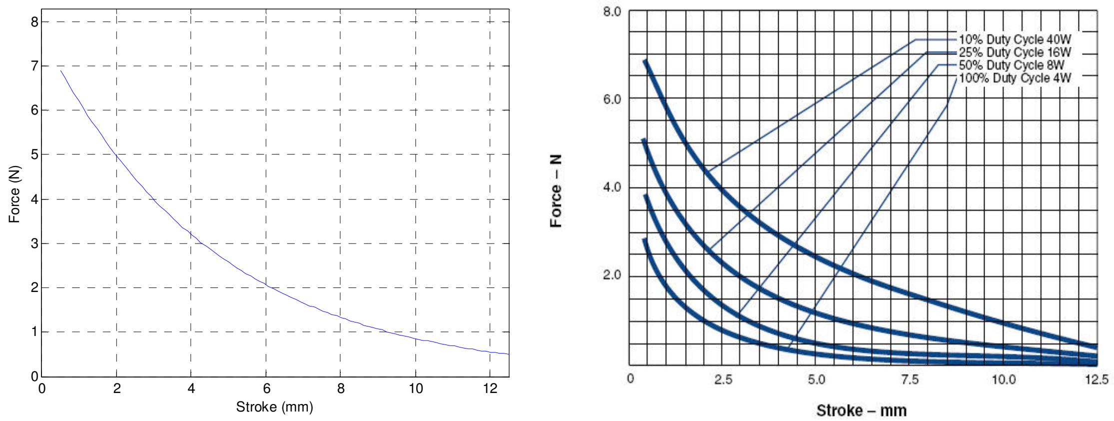

Figure 6: Simulated (left) vs. actual (right) force

It is worth noting at this point that assuming a linear variation in inductance from *L(0)* to *L(l)* would result in a force that is constant vs. stroke. While there are some actual solenoids that approximate that, the more usual case is the exponential variation shown in Figure 6. This is why we modeled the variation in inductance as an exponential decay, rather than linear, in (15). As stated earlier, the precise shape of *L(x)* depends on the geometric construction of the solenoid.

Notice that *N* does not appear in (22) at all. An N^2 term shows up in the denominator as a part of the squared resistance, but N^2 also shows up in the numerator as part of the inductance, and the two effects thus cancel out. This means that while removing turns will increase the current, it will also decrease the magnetomotive force of the *B* field by the same factor. Another way to recognize this canceling effect is to look at the estimation of the magnetic field in (1). In that equation, removing turns decreases *B*, but it also increases *I* due to a proportional reduction in *R*, and the two effects cancel each other.

There is, however, another factor that we must consider. Eq. (22) does not account for the fact that r_a is actually a function of *N*. For a fixed length coil, increasing *N* will eventually increase r_a. But since r_0 remains unchanged, the ratio between the two is reduced. Since this ratio is always less than unity, this means that adding turns will eventually decrease the force. In order to examine the extent of this effect, we must substitute into (22) a model for the dependence of r_a on *N*. **Figure 7** shows a cross-section of wires coming around the solenoid core.

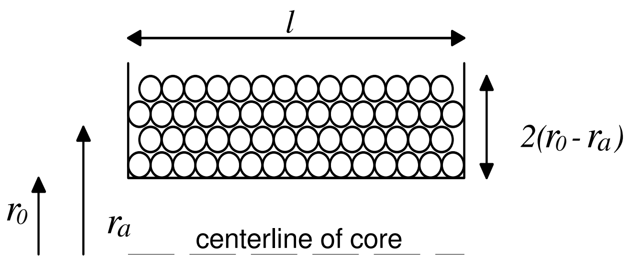

Figure 7: Turn packing

The space available for these *N* turns is $2l(r_0-r_a)$. The area taken up by the wires is a proportion of this defined by the packing density, λ. The theoretical maximum packing density for the lattice arrangement shown above is $\pi/\sqrt{12} \approx 0.907$ (Reference 6). If the wires are stacked in a grid instead of a lattice, the packing density is easily shown to be $\pi/4 \approx 0.785$. We can now express r_a as a function of N, as follows, where a is the cross-sectional area of a wire, available from an AWG table [References 4 and 5]:

$$\begin{aligned} \lambda A_{avail} &= A_{wires} \\ \lambda 2(r_a - r_0)l &= Na \\ r_a &= \frac{a}{2\lambda l}N + r_0 \qquad \lambda \le \frac{\pi}{\sqrt{12}} \\ &= \beta N + r_0 \end{aligned} \tag{24}$$

For brevity, we have collapsed λ, l, and a into a single parameter, β.

Substituting (24) into (22), and further simplifying by examining the force at x=0, we obtain:

$$\boxed{F_0 = \frac{-V^2\mu\alpha}{8\pi\gamma^2 l^2}\frac{r_0^2}{(\beta N + r_0)^2}} \tag{25}$$

This reveals a refinement on the winding factor in (23):

$$W_f = \frac{r_0^2}{(\beta N + r_0)^2} \tag{26}$$

When N is small, W_f will approach the squared ratio of r_0 to r_a in (23). As N increases, the ratio becomes smaller. Thus, N does play a role in the force, but it is a secondary effect having to do with the increase in r_a, not the increase in resistance. The rate at which an increase in N will decrease the force depends on the value of β, which depends on the packing density, the wire gauge, and the length of the coil. In order to get a feel for this variation, we now look at the effect of reducing N on the winding factor, the relative force (compared to nominal), and the relative power, for a coil with the following parameters: l=36 mm, r_0 = 7 mm, AWG = 26, a=0.129 mm^2, N_{nom}=1305 (Reference 7).

Figure 8 shows the variation in r_a along with the winding factor. From 100 to 1305 turns, the average turn radius increases linearly from 7.2 mm to 9.5 mm. Although the denominator is quadratic, in this case the shape of the winding factor in the region of interest is nearly linear.

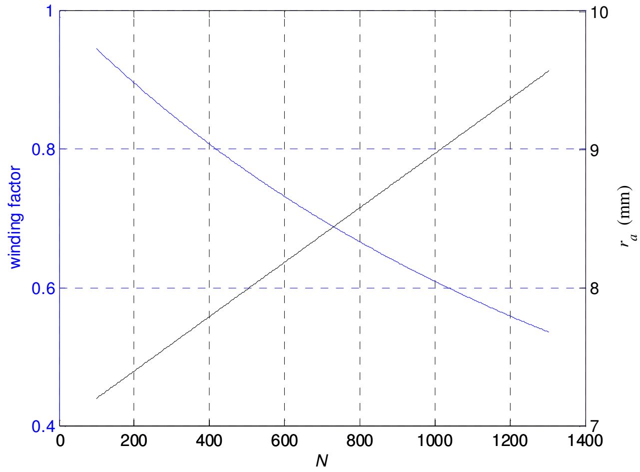

Figure 8: Winding factor and average turn radius vs. N

Figure 9 shows the variation in relative force and relative steady-state power as a consequence of varying *N*. The relative force has the same shape as the winding factor above, except that it is made relative to the nominal force at $N = 1305$. The relative power is inversely proportional to $R = 2\pi r_a N\gamma$, where r_a is a function of *N*. As an example, these curves predict that reducing the number of turns from 1305 to 400 will increase the force to 150% of nominal, but at the cost of an increase in power to 400% of nominal. A decrease of turns from 1305 to 700 will increase the force to 130% of nominal, at an increase in power to 210% of nominal.

Figure 10 plots the ratio of relative force to relative power as an efficiency measure. At 400 turns, the efficiency is 38% what it is at 1305 turns. This shows that *N* should be treated as a secondary factor in designing a coil for a desired force. Referring to (22) and (25), other coil parameters that are important in the force calculation are the resistance per unit length of the wire, γ, and the total length of the coil, *l*.

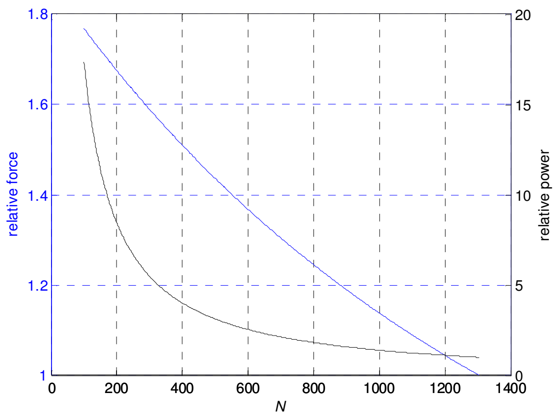

Figure 9: Relative force and power vs. N

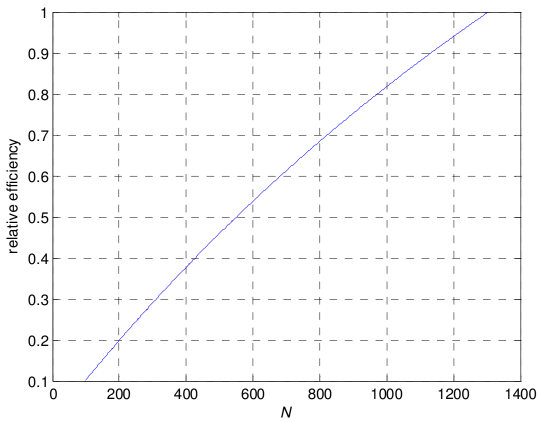

Figure 10: Relative efficiency vs. N

As shown in **Figure 11** and **Figure 12**, it is much more efficient to increase force by decreasing γ (decreasing wire gauge). These figures show the same quantities as Figures 9 and 10, for a 400 turn coil with γ (and a) ranging from the values for AWG16 through AWG25. Note especially that unlike decreasing N, decreasing γ not only increases the force, it also increases efficiency.

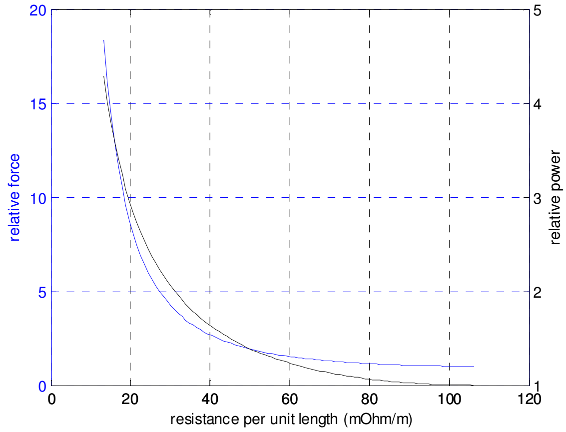

Figure 11: Relative force and power vs. γ

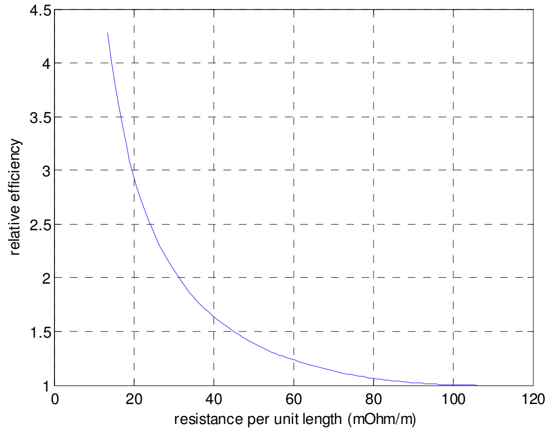

Figure 12: Relative efficiency vs. γ

Varying l offers another strategy for increasing force. **Figure 13** and **Figure 14** show these measures again for a 400 turn coil using AWG24 wire, with l varied from 25 to 50 mm. As with γ, decreasing l increases both force and efficiency. In this case, however, the relative force curve offers more steady, and modest, increases.

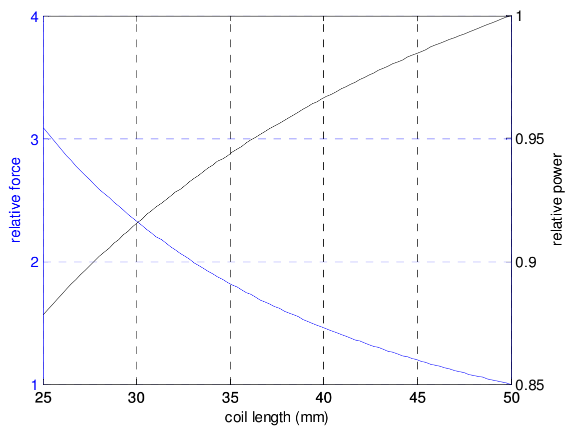

Figure 13: Relative force and power vs. l

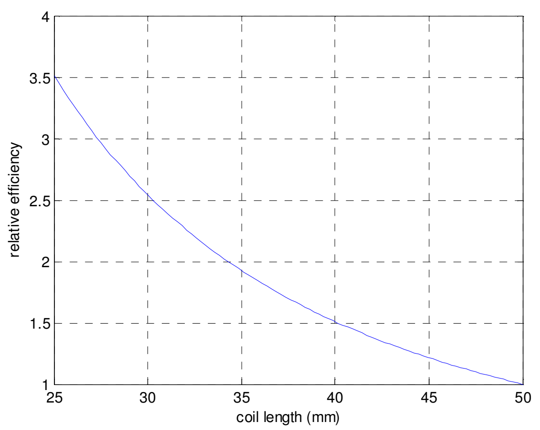

Figure 14: Relative efficiency vs. l

Figure 15 shows the measured force vs. turns from an actual solenoid using the same parameters as those simulated in Figures 8, 9, and 10. The force was sampled 10 times at each of several decrements in N, and a least-squares approximation was used to find the best-fitting inverse quadratic in N, which is superimposed on the data. The results agree quite well with the simulation in Figure 9. For example, decreasing the number of turns from 1503 to 700 increased the force by 126%. As in the simulation, this required twice the power. While this was a commercial solenoid, the coils were clearly hand-wrapped as they did not come off in uniform layers, which is the likely explanation for why the data does not span the best-fit curve at all samples of N.

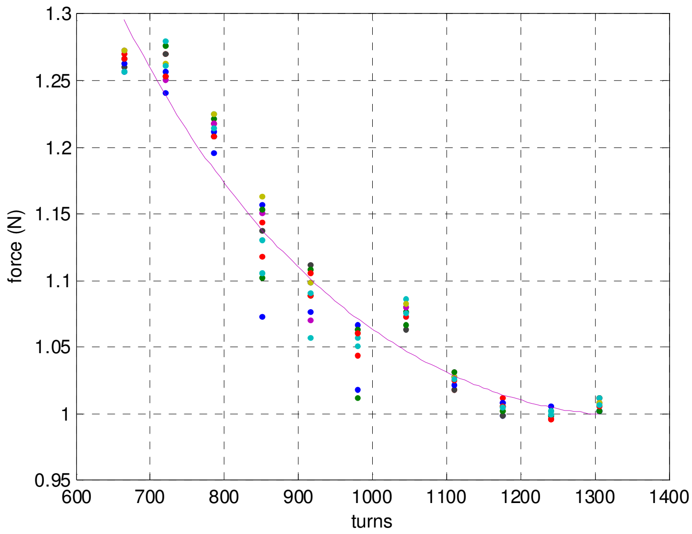

Figure 15: Force vs. turns for an actual solenoid

If turns should not be a primary design factor for force, then how should one determine the nominal number of turns to use for a given application? Note that (22), in theory, allows the design of a coil for a given force with very few turns. In actuality, however, the limited current-carrying capacity of the wire will prevent the use of too few turns. For a DC coil, the steady-state current is can be obtained from our estimate of total resistance in (18).

We require that to be less than some limit based on the current-carrying capacity of the wire, which can be obtained from a table of wire gauges. A fractional safety factor is probably desired here, and some coil designers simply establish a rule-of-thumb for maximum current based on the cross-sectional area of the wire. Here we'll use the latter approach and call that parameter η (Amps per square meter of cross-section):

$$I_{DC} = \frac{V}{R} \leq \eta a^2 \tag{27}$$

Substituting (18) for R and solving (19) for a^2 results in:

$$\begin{aligned} \frac{V}{2\pi r_a \gamma N} &\leq \eta \frac{\rho}{\gamma} \\ N r_a &\geq \frac{V}{2\pi\rho\eta} \end{aligned} \tag{28}$$

This gives a nice lower limit for the product of the number of turns and the average coil radius. Notice also that γ has been cancelled out, although the resistivity of copper remains in the equation.

It has been suggested to the author that η = 3.5 A/mm^2 = 3.5 x 10^6 A/m^2 is a reasonable value, but the author knows of at least one arcade machine design that goes substantially over that value (and gets rather hot), although it remains below the recommended wire ampacity with insulation.

To achieve a desired force, a reasonable design sequence would be to first use (22), with the assumption of some reasonable winding factor, say 0.7, in order to determine an appropriate γ (wire resistance per unit length) and coil length, l. These two parameters can be traded off against each other such that a lower resistance wire could be used by lengthening the coil, or a shorter coil could be used by increasing the wire resistance. The number of turns, N, and coil radius, r_a (or r_0) could then be determined from (28).

These two parameters can also be traded off against each other such the number of turns can be reduced with a corresponding increase in the radius, and vice-versa. Of course, the magnetic field and inductance approximations in (1) and (13) are based on long coils, so one would want to be wary of making a coil too short (l), in comparison to width (r_0 and r_a). These baseline parameter values could then be substituted into (25) which will include an estimate of the actual winding factor, and the parameters could then be tuned incrementally while ensuring that (28) remains satisfied.

About the author

Paul H. Schimpf holds the B.S., M.S., and Ph.D. degrees in Electrical Engineering from the University of Washington, with respective degree specialties in digital electronics, embedded signal processing, and modeling electromagnetic fields in complex domains such as the human body.

He previously worked in the electronics industry for 13 years at Honeywell Marine Systems in Mukilteo Washington (which later became a division of Alliant Techsystems, Hughes Aerospace, and Raytheon), where he attained the level of Senior Principal Development Engineer, prior to entering a career in academics. He was previously a tenured Associate Professor of Electrical Engineering and Computer Science at Washington State University, and is presently Professor and Chair of Computer Science at Eastern Washington University.

One of his hobbies is the restoration and repair of arcade games, of which he presently owns three: a Chicago Coin Twin Rifle (electro-mechanical), a Stern Meteor (early solid state), and a Williams Fish Tales (late solid state). Paul recently designed a printed circuit board to interface an Arduino Mega to 85 different Bally/Stern pinball machines so that his embedded systems students can program pinball game logic using a modern processor. The board includes the capability to play sound effects from WAV files stored on an SD card.