

An introduction to the art of
Solenoid Inductance Calculation
With emphasis on radio-frequency applications
By David W Knight¹

Version 0.20 (unfinished), 4th Feb. 2016.

Please check the author's website to ensure that you have the most recent versions of this document and its accompanying files: <http://g3ynh.info/zdocs/magnetics/> .

* * *

Overview

Information on the subject of solenoid inductance calculation is somewhat scattered in the literature and fraught with difficulties caused by differences of approach, inconsistencies of notation, errors, forgotten approximations, and the frequent need to translate between cgs and rationalised mks (SI) units. There are also issues of possible inaccuracy (and sometimes straightforward error) that result from applying of a body of information developed in the age of AC electrification to systems operating at radio frequencies.

In this article, the relevant information is collected, translated into SI units, reviewed and, where necessary, augmented. We start by setting-out the often neglected difficulties in solenoid parameter definition and showing how to make a rigorous separation between internal and external inductance. This allows us to view the traditional static magnetic model for what it is: an approximation for the principal component of the solenoid partial inductance. This model is, of course, suitable for correction to work at radio frequencies below the principal self-resonance; and we examine the various requirements in that respect.

For basic inductance calculation, three methods are compared. The first two are the Rosa-Nagaoka method of the American National Bureau of Standards (NBS) and the Kirchhoff summation method based on Maxwell's mutual inductance formulae. These give accurate results for coils with closely-spaced turns, but underestimate generally because they assume that current can only flow in the radial direction (i.e., exactly perpendicular to the coil axis). In other words, they lack helicity and so fail to include the inductance due to the axial component of current in the coil. The third method includes helicity. It was developed by Chester Snow of the NBS between 1926 and 1932, but has recently been revisited by Robert Weaver. By using a numerical integration method, Bob Weaver has succeeded in eliminating approximations that Snow (working without electronic computers) was forced to make for practical reasons. The result is a program that works for coils of any pitch; to the point that it calculates the straight-wire partial inductance when the pitch angle reaches 90°. The program is however computationally-intensive and thus unsuitable for general use. We therefore use its output as a source of data for the purpose of devising additional corrections for the summation and Rosa-Nagaoka methods.

¹ Ottery St Mary, Devon, England. <http://g3ynh.info/>

© D. W. Knight 2012, 2013, 2015, 2016.

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

Solenoid Inductance Calculation

By David W Knight

Table of Contents

Overview	1
Preface	3
Introduction	4
1. The current-sheet solenoid	8
2. Equivalent current-sheet length	10
3. Effective current-sheet diameter (LF)	11
4. Effective current-sheet diameter (HF)	13
5. Conductor length and pitch angle	16
5a. Minimum possible pitch	18
5b. Maximum number of turns for a given wire length	20
5c. Diameter, length and turns as derived parameters.....	22
6. Internal inductance	23
6a. LF-HF transition frequency	26
6b. Using the internal inductance factor	28
6c. Effective current sheet diameter linked to internal inductance	29
7. Magnetic field non-uniformity (Nagaoka's coefficient)	30
8. Approximate methods for calculating Nagaoka's coefficient	35
8a. Lundin's handbook formula	35
8b. Analytic asymptotic approximations for Nagaoka's coefficient	37
8c. Wheeler's long-coil (1925) formula	40
8d. Wheeler's 1982 unrestricted formulae	44
8e. Weaver's continuous formula	52
9. A note on the calculation of current-sheet inductance	54
10. Rosa's round-wire corrections and the summation method	55
10a. Geometric Mean Distances.....	55
10b. Loop self and mutual inductance formulae	57
10c. Solenoid inductance by the summation method	62
10d. Rosa's self-inductance correction	65
10e. Rosa's mutual inductance correction	76
11. Helicity	92
12. Combined static magnetic corrections	93
13. Apparent inductance and equivalent lumped inductance	95
14. Solenoid inductance calculation vs. measurement	96
99.	97

Preface

This document is a much expanded version of an earlier HTML article called 'Solenoids', which was first made publicly available in 2007. Its intention is to address a number of common misconceptions relating to the physics of inductive devices and to present accurate methods for calculating the inductance and other impedance-related parameters of a solenoid coil.

The original work was triggered by some encounters with misleading and incorrect information; and by the observation that, of a number of solenoid inductance calculators that were offered via the Internet at the time, there were none of any merit. Indeed, most programs were (and many still are) based on Wheeler's 1925 long-coil formula that, although widely assumed to be *the* formula for solenoid inductance, provides only an approximation to within a few % for coils of $\ell/D \geq 0.4$.

The original article was written as a supplement to information on the subject of impedance matching and measurement. Its subsequent revision and improvement involved going through a number of old books and papers, relating primarily to the early 20th Century work of Edward B Rosa and Frederick W Grover of the American National Bureau of Standards (NBS), and then devising or searching the literature for convenient ways of calculating the various infinite-series-form inductance functions and correction parameters. I have also translated everything into rationalised mks (i.e., SI) units, adopted what is (to my mind) an easy-to-remember notation (D for diameter, r for radius, ℓ for length, etc.), and made a few additional contributions in areas that I felt to be inadequately covered. This dry and dusty subject was not expected to arouse much interest; but somewhat surprisingly, it has attracted a steady stream of correspondence. It transpires that I was not the only one frustrated by the choice between methods that offer minimal physical insight and require the payment of software licence fees beyond the reach of private individuals², and the traditional semi-analytical methods that, although excellent, have needed to be updated for the age of the electronic computer.

There was interest moreover, not only in using the methods discussed, but developing them and checking their accuracy. In this respect I would particularly like to thank Bob Weaver for numerous helpful discussions and for writing and making available the various inductance-related functions and algorithms that are discussed and used in this and related documents. In many areas, Bob's efforts have surpassed mine, and although I try to summarise his findings here, I must also recommend the original material³. I would also like to thank Rodger Rosenbaum, who provided both Bob and me with a large part of the NBS archive on DVD before it became available online⁴, and who spent much time checking and analysing not only our work, but also that of Grover⁵ and others. Finally, I would like to thank Mark Kennedy⁶ for critical review and for supplying some hard-to-find reference materials.

DWK, July 2012, Sept. 2012.

Note

Important references cited on multiple occasions are given an alias at the first occurrence, as indicated in [square brackets].

2 I refer, of course to finite-element analysis techniques. Such methods are extremely valuable, and convenient for corroborating the findings of less brutal analytical approaches; but not exactly what you want to have running in the background of a general-purpose circuit simulation. FE modelling is, incidentally, now available at zero cost thanks to Dr David Meeker. See: <http://www.femm.info/>

3 <http://electronbunker.ca/CalcMethods.html>

4 <http://nistdigitalarchives.contentdm.oclc.org/>

5 **Grover's 'Inductance calculations'. Supplementary information & errata.** D W Knight and R Rosenbaum 2009. [**Grover errata**] Available from <http://g3ynh.info/zdocs/magnetics/>

6 <http://www.metallurgy.no/>

Introduction

When modelling and using inductive devices, it is important to be aware that the concept of lumped inductance is only strictly applicable at low frequencies. The construction of an inductor involves cramming a large amount of wire into a small volume, and at radio frequencies, this means that the wavelength is likely to be comparable to the length of the wire. In such circumstances, it cannot be said that any given point within the device is in instantaneous communication with every other part of the device; in which case, the lumped component theory cannot provide an accurate description. This does not necessarily preclude the use of simple approaches to circuit design; but it does mean that lumped element analysis should be applied with caution.

What particularly undermines the validity of the lumped approach is the propensity for inductors to exhibit dispersive behaviour. The term 'dispersion' comes from the field of optics, where a 'dispersion region' is a range of frequencies over which the refractive index of a medium changes; this being the the reason why a prism disperses white light into its component colours. The refractive index is the geometric mean of the relative permeability and permittivity; i.e.,

$$n = \sqrt{(\mu_r \epsilon_r)}$$

and so, in an electrical context, where a dispersion region is a frequency range over which permeability or permittivity is changing, the meaning is exactly the same. Coils with magnetic cores are inevitably dispersive, due to the complicated behaviour of ferromagnetic materials. What is less well recognised however, is that simple coils of wire are dispersive also.

The term 'refractive index' is not much used in electrical engineering; but many will be familiar with 'velocity factor', which is its reciprocal. This begs the question; "what has velocity got to do with inductance?" to which the answer is; "rather a lot". The traditional understanding of coils depends on the idea that they are effectively electromagnets, and that they have reactance because energy is stored in the surrounding magnetic field. This picture is mostly wrong, even though it suffices at low frequencies. If we may take the liberty of using the word 'light' to mean electromagnetic radiation of any frequency; what a coil really does is to modify the refractive index of space in its vicinity in such a way as to bend light and force it to follow the electrical conductor. All electrical circuits do that of course, but in inductors, the path is deliberately made long. Hence a coil is a waveguide or transmission line, which stores energy by trapping and detaining 'light' that would otherwise have made a much shorter journey.

The static magnetic conception of inductance works at low frequencies because the length of the wire used to make the coil is much shorter than the wavelength. This means that a wave entering the coil at one terminal will emerge from the other terminal with almost exactly the same phase. Thus an instantaneous view of the magnetic field surrounding the coil will be almost identical to the field produced by a direct current; in which case, the energy stored ($\frac{1}{2}LI^2$) will be the same as in the DC case and the inductance can be calculated accordingly. From an electrical point of view therefore, a coil operating at low frequencies looks like a lumped inductance in series with the DC resistance of the wire.

The first dispersion-related impedance variation (assuming that there are no ferromagnetic materials or lossy dielectrics to complicate matters) occurs at the onset of the skin effect; i.e., when the current ceases to be distributed uniformly throughout the wire cross-section and starts to concentrate at the surface. The frequency at which this change occurs depends on the diameter, resistivity and permeability of the wire, but it is usually somewhere between the audio and low short-wave radio regions. We can go part of the way towards understanding what happens by separating the total inductance into external and internal parts: where **external inductance** is that due to energy stored in the magnetic field that permeates the surrounding medium; and **internal inductance** is that associated with the field within the body of the wire itself. Inductance in electrical circuits is associated with current, and where there is no current there is no inductance.

Hence, as the current within the bulk of the conductor diminishes with increasing frequency, so too does the internal inductance. There is a little more to it than that however, because the redistribution of current is also affected by the magnetic fields produced by adjacent turns. This leads to a substantial second-order effect, known as the *proximity effect*, which gives rise to a reduction in the effective area enclosed by each turn of wire, and hence a reduction in the *external* inductance.

Thus the onset of the skin effect gives rise to a distinct transition from low-frequency to high-frequency behaviour; after which both the inductance and the resistance become frequency dependent. This does not necessarily preclude the use of the lumped component model however; because most of the decline in inductance occurs in the first two decades of frequency above the onset. Once out of the dispersion region, the inductance (now, strictly; the equivalent lumped inductance) settles down for a few octaves, and becomes reasonably (but never quite) constant.

In the high-frequency region, it is no longer possible to treat the coil as though its reactance is purely inductive; the reason being that a wave emerging from the coil is now significantly delayed, and therefore has a phase that differs from its phase on entry. One observable outcome is that the impedance at the coil terminals looks the same as that of an inductance (with series loss resistance) in parallel with a capacitance. This capacitance is known as the '*self-capacitance*' (or sometimes, misleadingly, as the 'distributed capacitance') of the coil. Presuming that the measured impedance has been corrected for strays, and that the coil is not extremely short in comparison to its diameter and is wound in a single layer (i.e., there are no overlapping turns), then the principal component of self capacitance is not of electrostatic origin. It is hypothetical (a way of expressing time delay), evoked in order to repair the lumped component model, and should be accorded no existence beyond that. It remains reasonably constant over several octaves however, it can be predicted, and it is therefore a perfectly valid parameter for the purpose of circuit analysis and simulation.

Unfortunately, the electrical literature abounds with articles that claim that the self-capacitance of a coil is due to the capacitance between adjacent turns. This hypothesis is easily refuted, because it makes the wholly incorrect prediction; that coils with closely-spaced turns will have much greater self-capacitance than those that do not. Unless the coil is short relative to its diameter (in which case the curvature of the field at the ends gives rise to an end-to-end capacitance), the static component of self capacitance is small in single-layer coils. The component due to inter-turn capacitance moreover is barely measurable. This is because a wave travelling along the wire does so with its electric vector nearly perpendicular to the coil axis, i.e., the electric field component parallel to the axis is negligible in comparison to the radial component. Nevertheless, the inter-turn capacitance idea appears to be so compelling, that there are at least two examples, in the peer-reviewed literature, where researchers have been motivated to fabricate or selectively report experimental evidence in order to support it.

The inclusion of self-capacitance into the lumped-component model gives rise to the prediction that a coil will still exhibit parallel resonance in the absence of an external circuit. This is indeed correct; except that, unless the coil is short relative to its diameter, the actual self-resonance frequency (SRF) is considerably greater than predicted. This failure of the lumped-component theory is mainly due to the onset of another dispersion-related effect; this time in which the apparent inductance declines (presuming that we adopt the view that the self-capacitance is constant) in such a manner that the SRF occurs at a frequency at which the wire in the coil is approximately one half-wavelength long (at least in the case of coil geometries suitable for making high-Q radio inductors).

This time, there is no reprieve for the lumped-element theory. The SRF occurs at the electrical half-wavelength point because that is the frequency at which a wave, trapped in the coil by reflection from the impedance discontinuities that occur at the terminals, arrives back at its starting point in phase with itself. We can partially explain the decline in inductance that leads to this coincidence by noting that a disconnected coil cannot have a uniform current distribution along its

conductor (there must be current nodes associated with the ends of the wire). This view is consistent with the fact that short coils have less of an apparent inductance decline than longer ones, the reason being arguably that the end-to-end capacitance helps to maintain the current. Explanations based purely on LC parallel resonance however can give no insight into what is obviously a dimensional resonance.

The self resonance occurs at the frequency at which the physical length of the winding wire is approximately one half-wavelength in coils for which the length and diameter are similar. This however is not true for short coils, because the SRF is reduced as a result of end-to-end capacitance and so tends to approach the value predicted by the lumped-equivalent self-capacitance. In very long coils moreover, the SRF occurs at a higher frequency than might be expected from the wire length, a peculiar phenomenon that can be referred to as the '*slow-wave effect*'.

The slow-wave effect can be understood by considering the overall field pattern as the superposition (vector addition) of two waves, one travelling along the coil axis and the other following the helix. In short coils, the coupling between the two modes is weak, and so the coil properties are dominated by helical propagation, with the result that the SRF is close to the principal resonance frequency of a half-wave wire antenna of the same conductor length. In longer coils however, excitation of helical propagation excites axial propagation, and since the two waves must remain in lock-step, the phase-velocity (i.e., the apparent velocity) of one must go up while that of the other goes down. What happens is that the helical phase-velocity increases to reach a limiting value somewhere in the vicinity of twice the velocity of light in very long coils, while the axial phase-velocity goes down; so that the ratio of velocities is the same as the ratio of the wire-length per turn to the inter turn distance (i.e., the winding pitch). Thus, while the helical wave has a phase-velocity greater than c , the axial wave, the slow wave, has a phase-velocity much less than c .

The ability of a long solenoid to produce an axial slow-wave is exploited in the travelling-wave tube (TWT) amplifier, where coupling between the slow-wave and an axial electron beam is used to produce gain in signals injected at one end of the helix and extracted at the other. The slow wave effect however is of little significance in the behaviour of typical inductance coils, because long coils of large pitch often have less inductance than a straight wire, and to get a good ratio of inductive reactance to loss resistance (i.e., high Q) it is necessary to keep the coil relatively short and use plenty of turns.

The lumped-element theory has another limitation of course, in that it assumes that the behaviour of a coil is undefined once the SRF has been reached. In reality, the coil will exhibit a sequence of alternating parallel and series resonances, which occur whenever the electrical length of the wire corresponds to a half-integer multiple of wavelengths. These overtone resonances incidentally, are not in exact harmonic sequence, because the phase velocity for helical propagation varies with frequency (the helix is a dispersive transmission line). From the lowest parallel resonance (which we must now refer to as the 'fundamental' SRF) to the first series resonance, the reactance is capacitive. It then switches back to being inductive until the next parallel resonance; and so on, almost ad infinitum, except that the length of a single turn will eventually become comparable to the wavelength and further complexities will arise. It follows, that coils have interesting properties at frequencies around and above the fundamental SRF, but lumped component theory is of no help in understanding or exploiting the resulting phenomena.

That coils are best regarded as transmission lines has long been known, but the behaviour of helical conductors is much more complicated than that of parallel wires, and the art of deducing their properties from first principles is hampered by the difficulty in solving Maxwell's equations for practical coils of arbitrary geometry. The problem is not completely intractable however; and can be usefully addressed by treating the coil as a surface waveguide constrained to conduct only in the helical direction. This model is known as the *Ollendorff sheath-helix*. An overview of this

subject is given by the Corum Brothers⁷, and additional information is given by Ramo et al.⁸ and elsewhere^{9 10 11}. The sheath-helix model points to a unification of the static magnetic and the transmission-line approaches, it accounts for the general variation of phase-velocity with frequency (overtone resonances not in exact harmonic sequence), and it also explains a useful but widely unrecognised phenomenon; which is that the resonant voltage magnification of a coil with minimal external capacitance is much greater than the lumped component theory predicts.

The downside of the sheath helix approach is that it involves simplifying assumptions and lacks certain important corrections. This limits its utility as an impedance calculation method. Also, it has to be said that traditional modelling methods, when properly applied, are very accurate at frequencies well below the SRF. Consequently, in the discussion to follow, we will adopt the view that a modified static-magnetic approach to coil modelling (albeit without the misconceptions) is adequate in the majority of situations, and that transmission-line concepts are best used to extend rather than replace what is well established.

-
- 7 **RF Coils, Helical Resonators and Voltage Magnification by Coherent Spatial Modes**, K L and J F Corum, Microwave Review, Sept 2001 p36-45. <http://www.ttr.com/TELSIKS2001-MASTER-1.pdf>
Class Notes: Tesla Coils and the Failure of Lumped-Element Circuit Theory, Kenneth and James Corum. <http://www.ttr.com/corum/>
Multiple Resonances in RF Coils and the Failure of Lumped Inductance Models. K L Corum, P V Pesavento, J F Corum. 6th International Tesla Symposium 2006. <http://www.nedyn.com/TeslaIntlSymp2006.pdf> .
- 8 **Fields and Waves in Communication Electronics**, Simon Ramo, John R. Whinnery, Theodore Van Duzer, 3rd edition. Publ. John Wiley & Sons Inc. 1994. ISBN 0-471-58551-3. [Ramo et al. 1994] Section 9.8: The idealised helix and other slow-wave structures.
- 9 **Theory of the Beam-Type Travelling-Wave Tube**. J R Pierce. Proc. IRE. Feb. 1947. p111-123. See Appendix B, p121-123, "Propagation of a wave along a helix", which gives Schelkunoff's derivation of propagation parameters for the Ollendorff sheath-helix.
- 10 **Coaxial Line with Helical Inner Conductor**. W Sichak. Proc. IRE. Aug. 1954. p1315-1319. Correction Feb. 1955, p148.
- 11 **The self-resonance and self-capacitance of solenoid coils**. David W Knight. g3ynh.info/zdocs/magnetics/

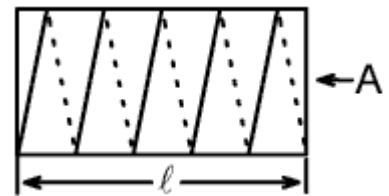
1. The current-sheet solenoid

In the design of high Q inductors for radio-frequency applications, the physical configuration most commonly adopted is the *single-layer solenoid*. The word 'solen' is an old-fashioned term meaning 'drainage channel', which eventually came to acquire the additional meaning 'drain-pipe'. The word 'cylinder' comes from the same root. Hence a solenoid is a pipe-like coil, usually wound with the aid of an actual pipe known as the coil-former. Winding the wire in a single layer produces an inductor with minimal parasitic capacitance, and hence gives the highest possible self-resonant frequency (SRF). Striving to obtain a high SRF and low losses is the key to producing coils that have radio-frequency properties bearing some useful resemblance to pure inductance.

A convenient basis for the calculation of the properties of practical coils is the inductance of a theoretical solenoid constructed using infinitely thin conducting tape wound, in a single layer, with zero spacing (but no electrical connection) between turns. This model is mathematically straightforward (at least, relatively so), because the infinitesimal radial thickness permits precise definition of the diameter, and the infinitesimal inter-turn gap eliminates small-scale field non-uniformities. Such a coil is known as a *current-sheet inductor*.

A very long current-sheet inductor (operating at low frequencies) has the property that the magnetic field along its length is uniform, in which case its inductance is given by a very simple expression:

$L_s = \mu \frac{A}{\ell} N^2 \quad \text{[Henrys]}$	1.1
--	------------



Where the constant of proportionality μ (in Henrys/metre) is the magnetic permeability of the environment outside the conductor ($\mu = \mu_0 \mu_r$) and can be replaced with the permeability of free-space, μ_0 ("mu nought") in the absence of ferromagnetic material. A is the cross-sectional area of the cylinder, N is the number of turns, and ℓ is the cylinder length.

Recall that the inductance of a coil can be expressed as an inductance factor A_L , defined by the relationship:

$$L = A_L N^2$$

For the long current-sheet therefore:

$$A_L = \mu \frac{A}{\ell} \quad \text{[Henrys/turn}^2\text{]}$$

Since turns are dimensionless and can be omitted from the units, this is analogous to the expression for the capacitance of a capacitor:

$$C = \epsilon \frac{A}{h} \quad \text{[Farads]}$$

Note that permeability, like permittivity, is strictly complex; but for the sake of simplicity we can consider it to be real when not taking magnetic losses into account. Hence we should use the symbol μ (in **bold**) when including losses in the permeability factor, and the symbol μ when not. Notice also that the factor A/ℓ has units of $[\text{length}^2 / \text{length}] = [\text{length}]$, and since A_L is an inductance, it is this that dictates that the units of μ are Henrys/metre.

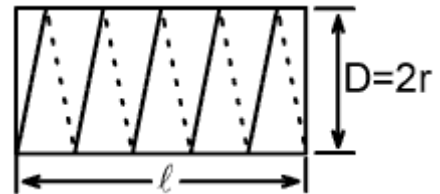
Equation (1.1) tells us that inductance is proportional to the cross-sectional area of a coil

(strictly, the area enclosed by the current loop). The optimal cross-sectional shape is that which gives the maximum amount of inductance using the minimum length of wire (maximum ratio of reactance / resistance), i.e., a former of circular cross-section is best. For a cylindrical coil, where $A = \pi r^2$, r being the coil radius, the long-current-sheet formula can be written:

$L_s = \frac{\mu \pi r^2 N^2}{\ell} \quad \text{[Henrys]}$	1.2
--	-----

We can also write this expression using the coil diameter D instead of the radius; noting that, since $D=2r$, the appropriate substitution is $r^2 = D^2 / 4$, i.e.:

$L_s = \frac{\mu \pi D^2 N^2}{4 \ell} \quad \text{[Henrys]}$	1.2a
--	------



Although the long current-sheet provides a starting-point for the calculation of inductance from physical dimensions, the equations given above require considerable modification if we are to obtain expressions accurate for practical coils. This entails the inclusion of various correction terms and factors, as will be explained in the discussion to follow. At least five distinct types of correction are required in principle; although the self-inductance corrections in particular are best split into sub-classes (wire-shape, axial current, curvature, and internal). The main corrections are listed below with the parameters that will be introduced in order to apply them. Some corrections can, of course, be neglected under appropriate circumstances; but the point is to understand what those circumstances are.

'Frequency independent'

k_L	field non-uniformity correction for short coils.
k_m	mutual inductance correction for round wire.
k_s	self-inductance correction for round wire.
	axial inductance for wide-spaced coils
	conductor curvature correction for thick-wire coils

Frequency dependent

L_i	internal inductance of the wire.
D or r	effective loop diameter (or radius).
C_L	self-capacitance (i.e., phase-delay modelled as a negative parallel reactance).

Note that the 'frequency independent' corrections are only so in the sense that the errors inherent in failing to include frequency dependence are reasonably small (or controlled by yet more corrections). Also bear in mind that inductance is only defined for complete current loops with their terminals coincident in space (i.e., in practice, close together). Since a solenoid has a finite separation between its terminals, its inductance is strictly a *partial inductance*. It is necessary to apply corrections for the connecting wires in order to obtain the total (measurable) inductance.

Notice also that the quantities listed above relate only to the problem of reactance calculation. The impedance of a coil must also include a resistive element to account for losses.

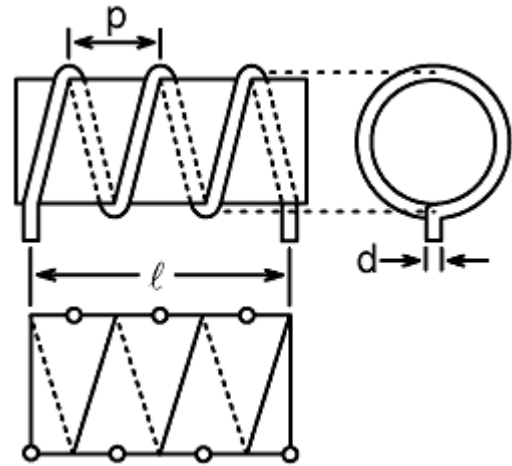
2. Equivalent current-sheet length

In the extensive literature on the subject of inductance calculation, one recurrent omission is that of an unambiguous definition for the coil length. The length required is that of the equivalent current-sheet from which the major part of the inductance will be calculated; but the problem is that a current-sheet inductor, being a hypothetical structure, can be defined without considering the method of connection. The correct definition is given by Grover¹², but requires interpretation.

The equivalent current-sheet length ℓ is obtained by considering each turn of the coil to lie at the centre of a corresponding turn of the current sheet. This means that if the length of the coil is measured on the side where the connecting wires are brought out (assuming a whole-number of turns) then the distance required is that from centre to centre of the emerging wires, i.e., it is the length of the coil measured from the outside of the winding less the diameter of the wire. This length is equal to $N \times p$, where N is the number of turns and p is the winding pitch-distance.

Rosa and Grover¹³ appear to give a different definition, but an ambiguity arises because the electrical termination is not considered. The instruction given is effectively; that the length can be obtained by measuring to the outside of the winding, then subtracting the wire diameter and adding the pitch. This length is stated to be equal to $N p$ as above, but it is only so if the measurement is made on the side of the coil opposite to the side where the connecting wires are brought out.

Note incidentally, that all of the expressions for solenoid inductance so far given (and to be given) contain a factor $1/\ell$. This factor goes to infinity as the length of the coil goes to zero, whereas the field non-uniformity correction (k_L , to be introduced shortly) goes to zero at this point. Hence the inductance of a zero length coil tends to $0/0$ and is undefined. This condition does not happen in practice, because the length of the equivalent current sheet can never be less than the diameter of the wire. The ambiguity arises because winding pitch (and hence solenoid length) is not strictly defined unless a coil has more than one turn. The inductance of a single turn coil is best obtained using a loop inductance formula (see [section 10b](#)).

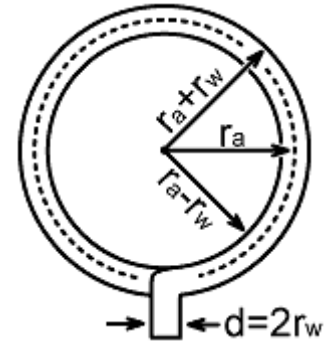


12 **Inductance Calculations: Working Formulas and Tables**. Frederick W Grover, 1946, 1973. [Grover 1946] Dover Phoenix Edition 2004. ISBN: 0 486 49577 9. p149.

13 **Formulas and Tables for the Calculation of Mutual and Self-Inductance**. E B Rosa, F W Grover. Bureau of Standards Scientific Paper No. 169 [BS Sci. 169]. 1916 with 1948 corrections. p119. [g3ynh.info/zdocs/magnetics/]

3. Effective current-sheet diameter (LF)

When a coil is wound with a thin flat conductor (broadside to the coil former), its radius ($r = D/2$) is well defined. When a coil is wound with round (i.e., cylindrical) wire, the equivalent current sheet radius will obviously be obtained by measuring from the solenoid axis to some point that lies within the body of the wire, but it is by no means obvious where that point should be. Referring to the diagram: If the radius of the wire (excluding any insulation) is r_w , and the average radius of the helix (measured from the solenoid axis to the wire axis) is r_a ; then there is a radial conduction zone that extends from $r = r_a - r_w$ to $r = r_a + r_w$. The effective current sheet radius must lie within that range.



It is traditional to assume that the effective radius is the same as the average radius r_a (at least at low frequencies), and that is the basis for most inductance calculations. It should be noted however, that the conduction path on the outside of the coil (at $r = r_a + r_w$) is longer than the path on the inside (at $r = r_a - r_w$). This means that the current-density in the wire will be biased towards the inside of the coil; and the equivalent current sheet radius will be consequently less than r_a . To that observation, we can also add, that the act of winding the wire around a cylindrical former causes the metal on the outside of the coil to become stretched relative to the metal on the inside. When metal wire is stretched (particularly in the case of soft copper), it does not so much shrink in diameter as increase in resistivity; i.e., the microcrystals within the material tend to rearrange and become less densely packed (until the yield point is reached). Hence the solenoid develops a radial resistivity gradient, the bulk resistivity being greatest at $r = r_a + r_w$ and something close to the native value at $r = r_a - r_w$. The effect, once again, is to bias the current distribution towards the inside, with consequent reduction in the effective radius.

This issue is investigated in a separate article¹⁴ in which the effective radius is assumed to lie at a distance from the coil axis chosen so that the total current outside that distance is equal to the total current inside it¹⁵. The low-frequency difference between the average radius and the effective radius, as calculated using that definition, is fairly large; being about +1% when $r_a = 8 r_w$, and only falling to about +0.1% (the point at which the difference might reasonably be neglected) when $r_a = 25 r_w$. Thus we can expect a systematic error in the generally adopted approach to inductance calculation when the coil is wound with relatively thick wire. The article gives methods for calculating the effective radius according to the adopted model, but there is no closed-form analytical solution for the round-wire problem, and so a numerical approach is used. A program routine accurate to within 0.01% is given as an Open Office Basic macro, which is used in the example inductance calculation spreadsheet ([Lcalcs.ods](#)) accompanying this article.

If a computationally straightforward approximation is required however, note that, for most coils, the strain of the wire is fairly small. In that case, the change in effective radius for a round-wire coil is not greatly different from that for a coil wound with wire of rectangular cross-section. An analytical solution exists for the rectangular wire case when the pitch of the winding is small relative to the circumference. This can be applied to the round wire case by defining $r_w = d/2$ as half the radial wire thickness. The formula is:

$r_0 = r_a \left[1 - \left(\frac{r_w}{r_a} \right)^2 \right]$	Equivalent current-sheet radius at low frequencies. Strained rectangular wire. $r_a / r_w > 4$, $2\pi r_a \gg p$	3.1
---	--	------------

¹⁴ **LF effective radius of a single-layer solenoid.** D W Knight. g3ynh.info/zdocs/magnetics/.

¹⁵ A better definition has been suggested by Mark Kennedy (private e-mail communications 31st Aug. and 5th Sept. 2012): The equivalent current sheet diameter defines as area that, if occupied homogeneously by the average solenoid magnetic flux density, would yield the true total flux.

This can also be stated in terms of coil and wire diameters:

$D_0 = D_a \left[1 - \left(\frac{d}{D_a} \right)^2 \right]$	Equivalent current-sheet diameter at low freq. Strained rectangular wire. $D_a / d > 4$, $\pi D_a \gg p$	3.1a
---	--	-------------

By inspecting the formula, it is apparent that when the average coil diameter D_a is much greater than the wire diameter d , then $D_0 \approx D_a$. High Q coils however, tend to be wound with relatively thick wire; in which case, inductance calculations that use D_a instead of D_0 will exhibit a systematic error. Such error, although usually small, is exacerbated by the fact that inductance is proportional to D^2 . Note however, the use of the rectangular wire formula will slightly underestimate D_0 .

4. Effective current-sheet diameter (HF)

In an isolated straight wire, at low frequencies, any current is distributed uniformly throughout the material. At high frequencies however, due to the inability of a good conductor to support an electric field within its bulk, the current is confined to a thin layer close to the surface. This is the well-known skin effect.

In coiled-wire inductors, the skin effect is perturbed (i.e., modified); not only by the conductivity gradient discussed in the previous section, but by an interaction with the external magnetic field known as the *proximity effect*. Presuming that the number of turns is reasonably large; the currents in adjacent turns are very nearly in phase, even at frequencies approaching the SRF. Under such conditions, there is a repulsion between adjacent current streams and a further interaction with the overall magnetic field. The result is that the current, in turns close to the middle of the coil at least, tends to crowd towards the coil axis^{16 17}. This means, of course, that there will be a further reduction in the effective current sheet radius at high frequencies.

It is important to be aware that dispersive phenomena have both real and imaginary parts. In the case of the proximity effect; the real part is that which causes the AC resistance of the wire to be greater than that predicted from the skin effect alone. The imaginary part is that which reduces the internal inductance of the wire (section 6) and reduces the effective current-sheet radius. It follows that the skin and proximity effects are not strictly separable. When the transition from low frequency to high frequency behaviour occurs (usually somewhere in the high audio to low radio frequency range); it is the proximity of other conductors, and the phases of the currents in them, that dictates how the current is distributed over the surface of the wire once it can no longer penetrate significantly into the body.

From its name, it should be obvious that the proximity effect is greatest in coils with closely-spaced turns. That part of it associated with a reduction in effective current sheet radius is also greatest when the wire diameter is significant in comparison to the coil radius. In high Q coils; which require the use of relatively thick wire to keep the AC resistance down and have plenty of turns to maximise the inductance obtained in a given volume; variation between the actual and the effective diameter can cause a difference of several percent between the low-frequency and the high-frequency inductance.

Due to the complexity of the underlying physics, the effective coil radius at high frequencies is difficult to predict from first principles. Fraga et al.¹⁸ (for example) approximate the situation by treating the coil as a modified current sheet with finite conductor thickness and resistivity. This approach has considerable merit, but is not completely realistic. It has also been suggested that, for modelling purposes, the wire can be considered to shrink towards the inner radius ($r_a - r_w$) as the frequency increases; but this is unconvincing. For those who are interested in this problem, it is important to understand that current still flows all over the surface of the wire when the proximity effect is present. Shrinkage of the wire implies a conduction layer that sinks below the surface, which is highly unrealistic. Thus it is a matter of current redistribution, rather than of parts of the wire ceasing to conduct. A more accurate determination of the effective radius might therefore involve finding an expression that defines the current density at any point in the wire cross section, and then setting the integral of the current density from the inner radius ($r_a - r_w$) to the effective radius to be the same as the integral from the effective radius to the outer radius ($r_a + r_w$).

Since the effective solenoid diameter at high-frequencies is difficult to determine, and since the difference between the average diameter (D_a) and the low frequency effective diameter (D_0) is not generally appreciated; inductance calculations are usually based on the average diameter. We can

16 Grover 1946. See Ch. 24.

17 H. F. Resistance and Self-Capacitance of Single-Layer Solenoids. R G Medhurst . Wireless Engineer, Feb. 1947 p35-43, Mar. 1947 p80-92. Corresp. June 1947 p185, Sept. 1947 p281. [Medhurst 1947]

18 Practical Model and Calculation of AC resistance of Long Solenoids. E Fraga, C Prados, and D-X Chen. IEEE Transactions on Magnetics, Vol 34, No. 1. Jan 1998.

do a little better than that however; there being no great difficulty in determining limits within which the actual inductance must lie. We start by noting that the current-sheet diameter D should be replaced by a mathematical function that depends on the winding-pitch to wire-diameter ratio (p/d), and on the wire diameter to solenoid diameter ratio (D_a/d), and varies between the low frequency value (D_0) and some high frequency limiting value, which we will call D_∞ . We cannot easily determine D_∞ ; but we can at least say that, for turns in the middle of the winding, it can never be smaller than the inner diameter (i.e., D_a-d). Furthermore, for the two turns at the ends of the coil, the current stream will be repelled from the next turn in, and so the effective diameter will remain close to D_0 . Hence we can define an absolute minimum effective diameter as the average of $N-2$ turns with a diameter of D_a-d and 2 turns with a diameter of D_0 , i.e.;

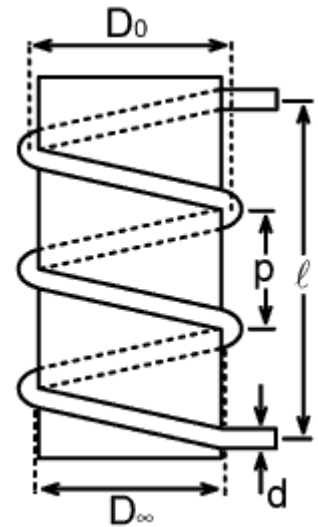
$$D_{\min} = \frac{[(N-2)(D_a-d) + 2D_0]}{N}$$

This expression will always underestimate D_∞ , and it will continue to do so if we use the approximation $D_0 = D_a$, i.e.;

$$D_{\min} = \frac{[(N-2)(D_a-d) + 2D_a]}{N}$$

which simplifies to:

$D_{\min} = D_a - d + \frac{2d}{N}$	4.1
-------------------------------------	-----



It is a straightforward matter (with the aid of a computer) to perform two inductance calculations; one with D set to D_0 , and one with D set to D_{\min} . From this we will obtain two inductances, L_0 and L_{\min} (say); the former being accurate at low frequencies and providing an upper uncertainty boundary for the high frequency inductance (L_∞), and the latter (presuming that the model is otherwise correct) giving the lower uncertainty boundary for L_∞ .

It is, of course, tempting to try to define a semi-empirical formula for D_∞ . For that, it is useful to know that D_∞ is fairly close to D_{\min} when the p/d ratio is close to 1, and almost the same as D_0 when $p/d > 10$. It follows that the accuracy of the inductance prediction will always be improved by taking the weighted average of D_0 and D_{\min} in such a way that D_{\min} dominates when $p/d \rightarrow 1$ and progressively loses its influence as p/d increases. Such a formula can be obtained by direct deduction, i.e.;

$D_\infty = \frac{D_0 + D_{\min} \frac{a}{[(p/d)-1]}}{1 + \frac{a}{[(p/d)-1]}}$	$a = 2$ $d \gg \delta_i$ (see section 6)	4.2
---	--	-----

Note that in practical inductors, $(p/d)-1$ is always > 0 because the angle between the pitch distance and the winding direction is $< 90^\circ$ (forcing p always to be greater than d) and because coils with closely spaced turns must be wound with insulated wire. Hence using the reciprocal of $(p/d)-1$ in the weighting coefficient will not cause divide-by-zero errors. The constant a is determined empirically, and setting it to 2 allows the HF inductance of typical radio coils to be

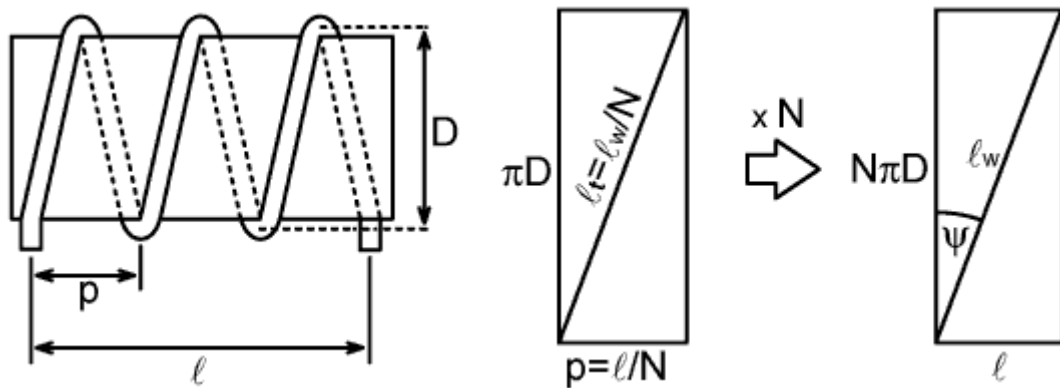
predicted to within a few parts per 1000 when the radius of the wire is greater than 3 times the skin depth. Note however that the value for a given above is for estimating L_∞ only. A frequency dependent, method for estimating the effective diameter D by taking the weighted average of D_0 and D_∞ is described later and requires a different value.

5. Conductor length and pitch angle

The length of wire used in an inductor is required when determining its AC resistance, its internal inductance, and its SRF. This length is commonly referred to as the 'line-length', but it is advisable to abandon that term. The problem is that, at its SRF, a coil behaves as a $\lambda/4$ -wave transmission-line resonator, whereas the electrical length of the wire at that frequency is one half-wavelength. This incidentally, is not a paradox. A transmission line is a go-and-return circuit, and so any $\lambda/4$ line has $\lambda/2$ of conductor.

Consequently, if we refer to the line length, it is not clear whether we mean the length of the wire, or the length of the equivalent transmission-line (which is about one-half as great). Hence the terms **conductor-length** and **wire-length** are strongly recommended as alternatives.

Shown below-left is a coil of diameter D and length ℓ , with a winding pitch (axial turn spacing) of p .



The length of the coil is equal to the number of turns multiplied by the pitch, i.e.;

$$p = \ell / N$$

The length of wire in the coil (ℓ_w) is the length of a single turn (ℓ_t) multiplied by the number of turns, i.e.;

$$\ell_t = \ell_w / N$$

The middle diagram above represents a single turn unwrapped and laid flat. The length of the turn is the diagonal of a rectangle having the circumference of the coil (πD) as one dimension, and the pitch as the other. If this map is scaled-up by the number of turns (i.e., every dimension is multiplied by N), then the diagonal becomes the wire length, and the dimensions of the rectangle are $N\pi D$ and ℓ . Hence, using Pythagoras's theorem:

$$\ell_w = \sqrt{(\pi D N)^2 + \ell^2} = \sqrt{(2\pi r N)^2 + \ell^2}$$

5.1

we can also remove a factor $(2\pi r N)^2$ from the square root bracket to obtain:

$$\ell_w = 2\pi r N \sqrt{1 + \left(\frac{\ell}{2\pi r N}\right)^2}$$

but

$$\frac{\ell}{2\pi r N} = \tan \psi$$

where ψ (psi) is known as the 'pitch-angle'. Hence:

$$\ell_w = 2\pi r N \sqrt{1 + \tan^2 \psi}$$

Now making use of the relations: $\tan \psi = \frac{\sin \psi}{\cos \psi}$ and $\sin^2 \psi + \cos^2 \psi = 1$; we get:

$$\tan^2 \psi + 1 = \frac{1}{\cos^2 \psi}$$

Hence:

$\ell_w = \frac{2\pi r N}{\cos \psi} = \frac{\pi D N}{\cos \psi}$	5.2
---	-----

If the pitch-angle is small (i.e., if the turns are closely spaced), then $\cos \psi \rightarrow 1$ and the wire length can be approximated as:

$$\ell_w \approx 2\pi r N = \pi D N$$

Note incidentally, that the factor $1/\cos \psi = \sec \psi$ will crop up frequently in helix-related problems. Therefore it is useful to have it in a convenient form. Referring to the diagram above:

$$\cos \psi = \frac{2\pi r}{\sqrt{(2\pi r)^2 + p^2}} = \frac{1}{\sqrt{1 + \left(\frac{p}{2\pi r}\right)^2}}$$

Therefore:

$$\frac{1}{\cos \psi} = \sec \psi = \sqrt{1 + \left(\frac{p}{2\pi r}\right)^2}$$

The **effective conductor-length** of a coil will always be slightly less than the physical wire length, and it will vary with frequency. This is due to the difference between the average coil diameter and the equivalent current-sheet diameter, as discussed in **sections 3 - 4**. Hence, when using the conductor length to determine the RF properties of coils, it should be as calculated by using some sensible estimate of the effective solenoid diameter (see **section 6c**). A possible exception to that rule is when using the approximation $\ell_w = \pi D N$, in which case, the neglect of the $1/\cos \psi$ factor in equation **(5.2)** can be partly offset by using the average diameter D_a , i.e.;

$$\ell_w \approx 2\pi r_a N = \pi D_a N$$

5a. Minimum possible pitch

The minimum possible distance between two round wires lying side-by-side (assuming insulation of infinitesimal thickness) is the wire diameter d . The pitch distance of a coil however, is defined in a direction that lies parallel to the coil axis. The winding direction of a helical coil is not perpendicular to the pitch direction, it is tilted away from the perpendicular by an angle ψ (the pitch angle). This causes the minimum pitch distance (p_{\min}) to be slightly greater than the wire diameter. Since p_{\min} is a limiting value for solenoid optimisation problems, it is important to define it correctly.

The diagram below shows two turns from a coil, with zero spacing, which have been unwound and laid out flat. The distance between the axes of the two wires is d , and the circumference of the coil is πD_a , where D_a is the coil diameter taken from wire centre to wire centre. Since the two wires are as close as they can possibly be; as each turn is wound, the wire advances along the coil axis by a distance p_{\min} . The length of wire in each turn is defined as l_t , and we can immediately write a relationship between p_{\min} and l_t using Pythagoras's theorem:

$$p_{\min}^2 + (\pi D_a)^2 = l_t^2$$

We can also define l_t as the sum of the lengths marked on the diagram as l_a and l_b . Thus:

$$p_{\min}^2 + (\pi D_a)^2 = (l_a + l_b)^2$$

where, using Pythagoras again:

$$l_a^2 + d^2 = p_{\min}^2$$

i.e.;

$$l_a = \sqrt{p_{\min}^2 - d^2}$$

and

$$l_b^2 + d^2 = (\pi D_a)^2$$

i.e.:

$$l_b = \sqrt{(\pi D_a)^2 - d^2}$$

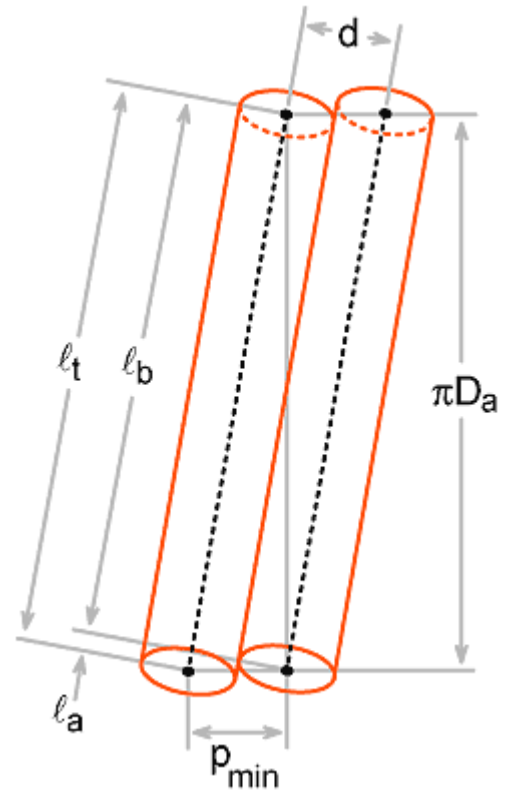
Hence, combining expressions we get:

$$p_{\min}^2 + (\pi D_a)^2 = \left[\sqrt{p_{\min}^2 - d^2} + \sqrt{(\pi D_a)^2 - d^2} \right]^2$$

Multiplying-out the right-hand side gives:

$$p_{\min}^2 + (\pi D_a)^2 = p_{\min}^2 - d^2 + (\pi D_a)^2 - d^2 + 2\sqrt{(p_{\min}^2 - d^2) [(\pi D_a)^2 - d^2]}$$

which after regrouping and squaring becomes:



$$(p_{\min}^2 - d^2) [(\pi D_a)^2 - d^2] = d^4$$

This can be multiplied out again to give:

$$p_{\min}^2 (\pi D_a)^2 - p_{\min}^2 d^2 - (\pi D_a)^2 d^2 + d^4 = d^4$$

i.e.:

$$p_{\min}^2 [(\pi D_a)^2 - d^2] = (\pi D_a)^2 d^2$$

Thus, rearranging and taking the square root:

$$p_{\min} = \frac{\pi D_a d}{\sqrt{(\pi D_a)^2 - d^2}} \quad \dots \dots \dots (5.3)$$

which simplifies to:

$p_{\min} = \frac{d}{\sqrt{1 - \left[\frac{d}{(\pi D_a)} \right]^2}}$	5.4
--	------------

Note that, as $D_a \rightarrow \infty$ or $d \rightarrow 0$, $p_{\min} \rightarrow d$; but for all finite coil and conductor dimensions, $p_{\min} > d$ always.

Minimum pitch angle

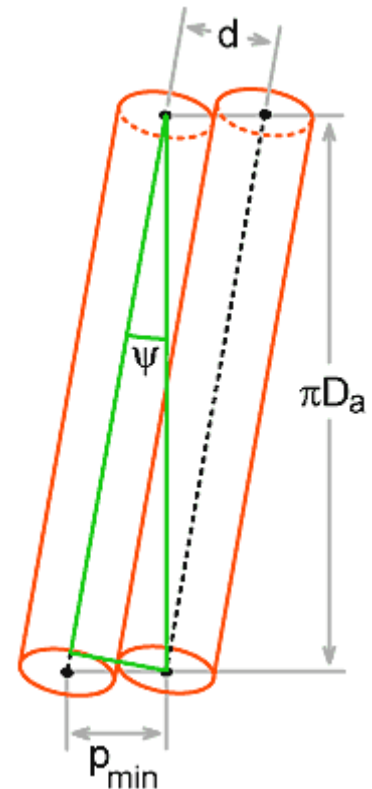
When $p = p_{\min}$, $\psi = \psi_{\min}$

Referring to the diagram on the right:

$$\sin \psi_{\min} = \frac{d}{\pi D_a}$$

Therefore:

$\psi_{\min} = \arcsin \left(\frac{d}{\pi D_a} \right)$	5.5
--	------------



5b. Maximum number of turns for a given wire length

Coils designed for radio frequency applications often have an upper limit on the allowed wire length (because that dictates the SRF). Consequently, another important limiting value for inductor modelling problems is the maximum number of turns that can be wound in a single layer using a given length of wire on a given diameter of coil former.

Recall that solenoid length is defined as: $\ell = N p$

Now observe that the maximum number of turns also corresponds to the case where the wire is most tightly packed (because each turn takes the shortest possible route around the former); in which case, the pitch will be at its minimum and the length of the solenoid will be at its minimum. Hence:

$$\ell_{\min} = N_{\max} p_{\min}$$

The relationship between coil diameter, wire length and solenoid length (5.1) can be written:

$$\ell_w^2 = (N \pi D_a)^2 + \ell^2$$

and at minimum pitch:

$$\ell_w^2 = (N_{\max} \pi D_a)^2 + \ell_{\min}^2$$

Hence, substituting for ℓ_{\min} ,

$$\ell_w^2 = (N_{\max} \pi D_a)^2 + (N_{\max} p_{\min})^2$$

i.e.:

$$N_{\max} = \frac{\ell_w}{\sqrt{(\pi D_a)^2 + p_{\min}^2}}$$

Substituting for p_{\min} using (5.3) then gives:

$$N_{\max} = \frac{\ell_w}{\sqrt{(\pi D_a)^2 + \frac{(\pi D_a)^2 d^2}{(\pi D_a)^2 - d^2}}} = \frac{\frac{\ell_w}{\pi D_a}}{\sqrt{1 + \frac{d^2}{(\pi D_a)^2 - d^2}}} = \frac{\frac{\ell_w}{\pi D_a}}{\sqrt{\frac{(\pi D_a)^2}{(\pi D_a)^2 - d^2}}}$$

i.e.;

$N_{\max} = \frac{\ell_w}{\pi D_a} \sqrt{1 - \frac{d^2}{(\pi D_a)^2}}$	5.6
--	------------

Also note from (5.4) that:

$$\frac{p_{\min}}{d} = \frac{1}{\sqrt{1 - \left(\frac{d}{\pi D_a}\right)^2}}$$

Hence

$N_{\max} = \frac{\ell_w}{\pi D_a} \frac{d}{p_{\min}}$	5.7
--	------------

5c. Diameter, length and turns as derived parameters

Equation (5.1) gave the coil conductor length as a derived parameter:

$$\ell_w = \sqrt{(\pi D N)^2 + \ell^2} = \sqrt{(2\pi r N)^2 + \ell^2}$$

From this we can see that the major parameters defining coil inductance are D (or r), ℓ and N. What is also apparent however, is that there are four major parameters defined in such a way that specifying any three of them fixes the fourth. It seems most natural, when modelling a coil, to measure the length, the diameter and the turns number, and then allow the conductor length to be determined as a consequence. This however, is not necessarily the best way to do it, especially when making coils for RF experiments and applications.

If an RF coil has a few metres of wire in it, then it is not difficult to measure the length of the wire to an accuracy of about ± 1 mm. This means that the wire length can often be determined to better than $\pm 0.1\%$. Likewise, the turns number N can be determined by counting; and by looking along the end of the coil and using a clear-plastic protractor, the angular difference between the start and the end can be determined to within a few degrees. This gives the non-integer value of any incomplete turn to about 1%, and if there are more than 10 turns, this brings the error down to $\pm 0.1\%$ or better. The difficult dimensions are the solenoid length and the diameter, both of which require engineer's callipers for reasonable accuracy; and even then may not be particularly well defined due to the tendency of the coil to deform when handled, or due to unevenness and non-circularity of the cylinder, etc.. In general, in situations in which all four principal parameters are obtainable, a coil should be specified in terms of the three parameters that can be measured to the greater accuracy.

So, from equation (5.1), in deriving D (usually the hardest to measure accurately), we have an average diameter:

$D_a = \frac{\sqrt{\ell_w^2 - \ell^2}}{\pi N}$	5.8
--	------------

The solenoid length can be obtained from:

$\ell = \sqrt{\ell_w^2 - (\pi D_a N)^2}$	5.9
--	------------

and, from (5.8), the number of turns is:

$N = \frac{\sqrt{\ell_w^2 - \ell^2}}{\pi D_a}$	5.10
--	-------------

6. Internal inductance

The 'external inductance' of a coil is the inductance due to the storage of energy in the magnetic field that permeates the surrounding medium. The 'internal inductance' is due to the magnetic energy stored within the body of the conductor itself. Internal inductance diminishes with frequency because it depends on the current distribution within the wire; i.e., its corresponding reactance is the imaginary counterpart of the skin effect.

The conducting strip in the theoretical current-sheet is infinitely thin and therefore has no internal inductance. Wire, on the other hand, does have internal inductance. The internal contribution to overall inductance is generally small, and is therefore usually neglected in approximate calculations; but it can amount to several % of the total under certain circumstances. The following points may help when considering its importance:

- Internal inductance is proportional to the conductor-length, and therefore to the number of turns N ; whereas external inductance, being enhanced by winding the wire into a helix, is proportional to N^2 . Hence, internal inductance is most likely to be significant in coils that have a low number of turns.
- External inductance is enhanced by the use of a magnetic core, whereas internal inductance is unaffected. Hence internal inductance is not usually significant if the coil has a high-permeability core.
- Internal inductance diminishes with frequency more rapidly in thick wire than it does in thin wire; i.e., thick wire coils have the skin effect dispersion at lower frequencies than thin wire coils. For coils made from wire of less than 1 mm diameter, internal inductance can still be significant at the low end of the HF radio region (see next section).

The general problem of calculating internal impedance is discussed in detail in a separate article¹⁹. Here we will summarise only those sections relevant to the calculation of solenoid inductance.

The internal inductance of a round wire at DC is given by:

$$L_{i(dc)} = \ell_w \frac{\mu_{(i)}}{8\pi} \quad [\text{Henrys}]$$

where ℓ_w is the length of the wire, and $\mu_{(i)}$ is the permeability of the wire material (i.e., the internal permeability). For non-ferromagnetic conductors, $\mu_{(i)}$ can be taken to be the same as μ_0 , i.e., $4\pi \times 10^{-7}$ H/m, which means that the low-frequency internal inductance of any non-magnetic round wire is 50 nH/m. Note that, for the construction of high Q inductors, only non-magnetic wire (preferably copper or silver) should be used. Due to the generally high resistivity, and the fact that skin depth is a function of permeability, skin-effect losses are extremely high in wires made from ferromagnetic materials.

The internal inductance of a wire at high frequencies is given by:

$$L_{i(hf)} = \ell_w \left(\frac{\mu_{(i)}}{2\pi} \right) \left(\frac{\delta_i}{d} \right) \quad [\text{Henrys}]$$

where d is the diameter of the wire, and δ_i is the skin depth given by:

¹⁹ **Practical continuous functions for the internal impedance of solid cylindrical conductors**. D W Knight, 2010. g3ynh.info/zdocs/comps/

$$\delta_i = \sqrt{\frac{\rho}{\pi f \mu_{(i)}}}$$

ρ being the resistivity of the wire. Hence, at high frequencies, internal inductance becomes proportional to the reciprocal of the square root of the frequency.

A suitable internal inductance formula for solenoid modelling is the PACAML approximation²⁰, which is accurate at all frequencies to within ± 151 ppm. Since the internal inductance contribution to the total inductance is typically $\approx 1\%$, the error in the PACAML approximation contributes less than 1 part in 10^5 to the overall error. The formula uses the skin depth δ_i (as given above) and calculates an internal inductance factor Θ (Theta), which lies between 1 and 0. This quantity is multiplied by the DC internal inductance to obtain the internal inductance at the specified frequency.

$q = \frac{d}{\delta_i \sqrt{2}} \quad ,$	
$\Theta_\infty = \frac{4}{q} \frac{1}{\sqrt{2}} \left[1 + \frac{0.01209}{(q+1)} - \frac{0.63523}{(q^2+1)} + \frac{0.16476}{(q^3+1)} \right] \quad , \quad \text{[polynomial]}$	
$\Theta_{da} = \Theta_\infty \left[1 - \exp\left(-\Theta_\infty^{-1.5819}\right) \right]^{\frac{1}{1.5819}} \quad , \quad \text{[asymptote correction]}$	
$z = 0.38691q \quad ,$	
$y = \frac{-0.198584}{\left[1 + 0.25741(z^{1.2652} - z^{-0.39709})^2 \right]^{2.62343}} \quad , \quad \text{[modified Lorentzian]}$	
$\Theta = \Theta_{da} (1 - y) \quad ,$	
$L_i = \ell_w \frac{\mu_{(i)}}{8\pi} \Theta \quad \text{[Henrys]}$	Li-PACAML (6.1)

Basic macro code for the calculation is shown in the following box. Θ is calculated by a function 'Flpml(q)' which is called by a front-end routine 'Lintern()'. Lintern requires the wire length and diameter, the frequency, the relative permeability of the wire (1 for Cu or Ag), and a quantity ' K_{iacs} '. K_{iacs} is the wire conductivity as a proportion of the International Annealed Copper Standard^{21 22}. For electrical copper wire at 20°C, $K_{iacs} = 1$. Note that at 20°C, the resistivity ρ of IACS standard copper is $17.241 \times 10^{-9} \Omega\text{m}$, whereas the resistivity of silver at this temperature is $15.87 \times 10^{-9} \Omega\text{m}$. Therefore, for silver wire (bearing in mind that conductivity is the reciprocal of resistivity):

$$K_{iacs} = 17.241 / 15.87 = 1.086$$

20 Basic routines for internal impedance calculation are given in the macro library of the spreadsheet accompanying the author's article. The acronym 'PACAML' stands for 'polynomial asymptotically correct approximation with modified-Lorentzian correction'.

21 **Copper wire tables.** Bureau of Standards Circular No. 31. 3rd edition. 1914. International Annealed Copper Standard. pages 8 - 13. Available from <http://g3ynh.info/zdocs/comps/>

22 **Coppers for electrical purposes.** V A Callcut. Proc. IEEE, Vol. 133, Pt. A, No. 4. June 1986.

i.e., at 20°C, silver is 1.086 times more conductive than IACS copper.

Basic routine for internal inductance of a wire

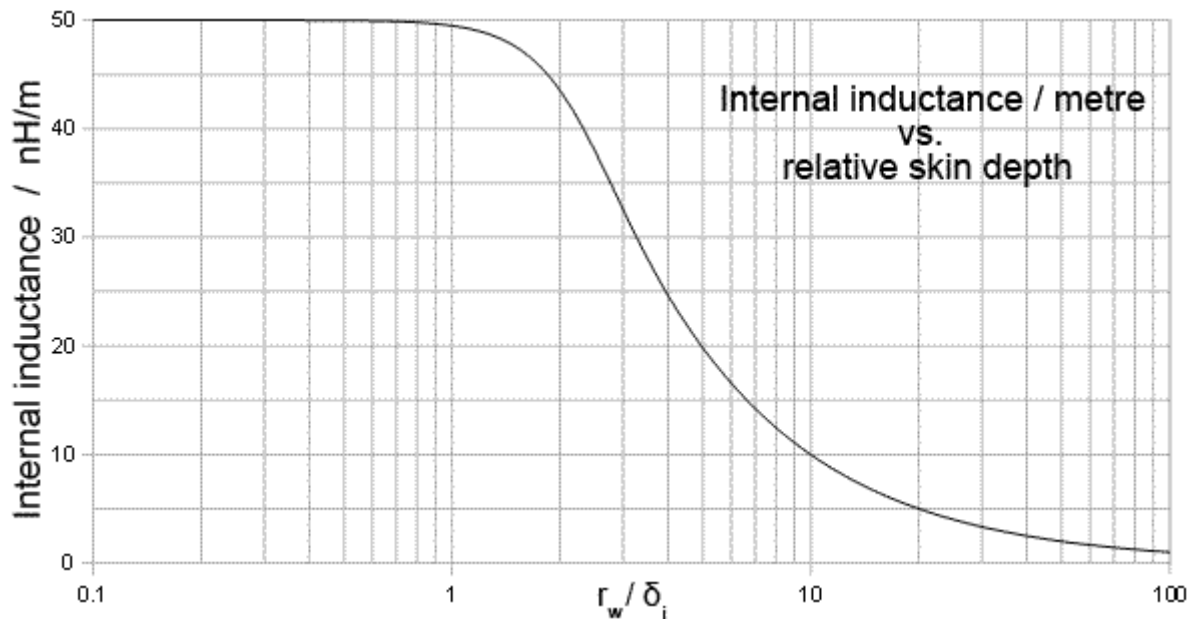
```
Function Lintern(ByVal lw as double, d as double, f as double, Kiacs as double, mur as double)
'Internal inductance of a round wire. Version 1.00, 9th Jan 2016.
'Calls function Flpml(q) for inductance factor Theta.
' lw = length of wire / m
' d = wire diameter / m
' f = frequency / Hz
' Kiacs is proportionate conductivity relative to IACS. For electrical Cu at 20 deg. C, Kiacs = 1
' mur = relative permeability of the wire = 1 for non magnetic conductors
Dim rho as double, delta as double, q as double, Theta as double
rho = 17.241E-9 / Kiacs
delta = sqr( rho / ( 4E-7*pi()*pi() * f * mur ) )
q = d / ( delta * sqr(2) )
Theta = Flpml(q)
Lintern = lw * 0.5E-7 * mur * Theta
end function
```

```
Function Flpml(ByVal q as double) as double
'Calculates internal inductance factor within 151 ppM.
'see: Practical continuous funces for int. Z of cylindrical conductors.
'D. W. Knight, version 1.01, 25th Jan 2016.
Dim y as double, z as double, zz as double
if q<0.0001 then
Flpml=1
else
Flpml=(4/q)*(1/sqr(2))*(1+0.01209/(q+1)-0.63523/(q*q+1)+0.16476/(q*q*q+1))
Flpml=Flpml*(1-exp(-1*Flpml^-1.5819))^0.63215121
z=0.38691*q
zz=Z^1.2652-Z^-0.39709
y=-0.198584/(1+0.25741*ZZ*ZZ)^2.62343
Flpml=Flpml*(1-y)
end if
end function
```

The total inductance of a current loop is the sum of the internal and external inductances. For coils however, there will be an additional term (analogous to mutual inductance) due to the external fields of adjacent turns passing through the wire; i.e., there will be a perturbation due to the proximity effect (as mentioned in the previous section). The proximity of other current-carrying conductors has little effect at low frequencies, but reduces the internal inductance at high frequencies. Fortunately, internal inductance makes a relatively small contribution to the overall inductance, and so the error in using an isolated wire model for internal inductance (i.e., ignoring the perturbation caused by the proximity effect) is usually small.

6a. LF-HF transition frequency

When deciding whether to use a low or a high-frequency inductance formula, it is necessary to be able to locate the intervening dispersion region. A simple rule for doing so can be obtained by examining the graph below, which shows the relationship between internal inductance and the ratio of wire radius to skin depth. The calculation is for an isolated wire (see the accompanying Open Document spreadsheet²³: [Li_transition.ods](#), sheet 1), but while the proximity effect will steepen the inductance decline, it will not greatly affect the frequency at which the change begins.



The graph confirms a rather obvious proposition, which is that the current distribution within the wire will be substantially uniform until the skin depth becomes less than the wire radius. Hence we can define a transition frequency (f_s) at which DC inductance formulae begin to break down. Skin depth is given by:

$$\delta_i = \sqrt{\frac{\rho}{\pi f \mu_{(i)}}}$$

and, from the graph above, it is apparent that we need to start making high-frequency corrections when $r_w = \delta_i$. Hence, to work out the wire diameter needed to achieve a particular f_s (noting that $d = 2 r_w$):

$$d = 2 \sqrt{\frac{\rho}{\pi f_s \mu_{(i)}}} \quad (6.2)$$

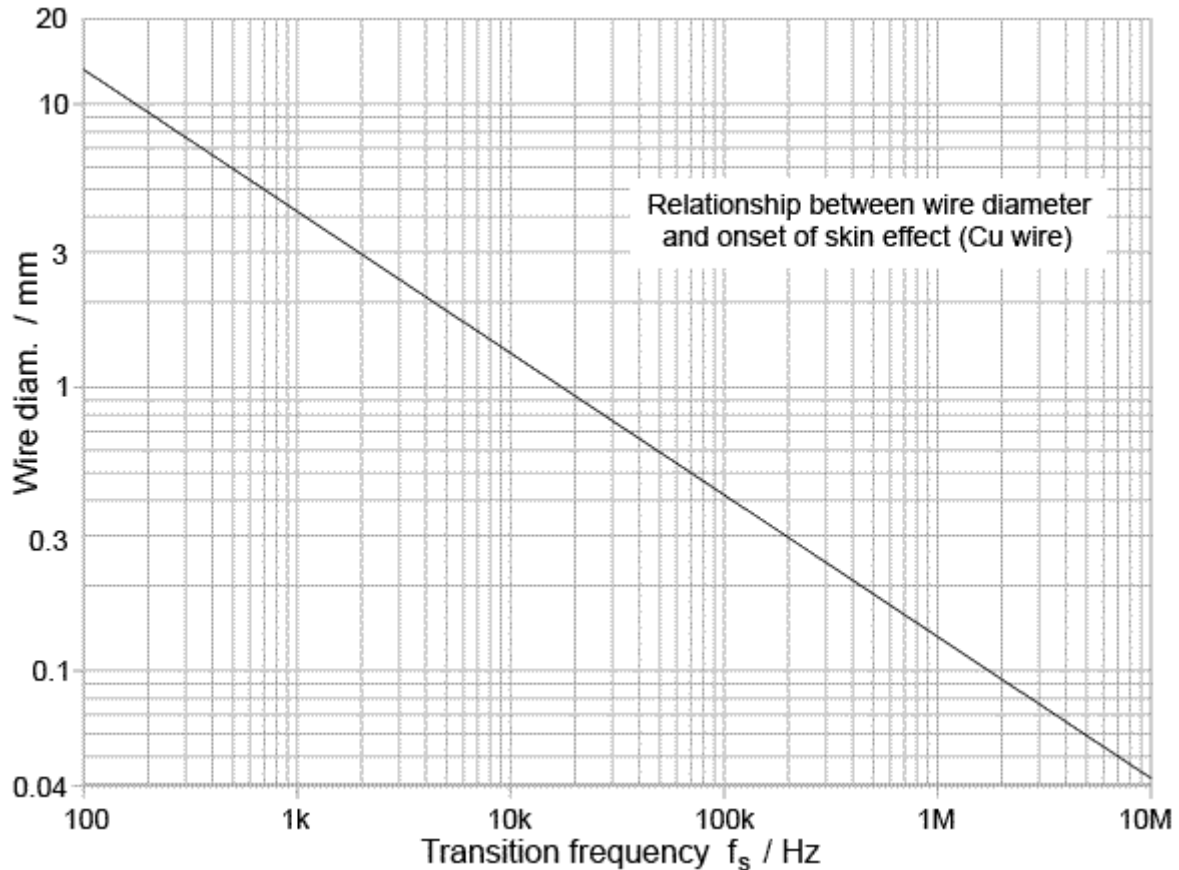
And to work out f_s for a particular wire diameter:

$$f_s = \frac{4 \rho}{\pi \mu_{(i)} d^2} \quad (6.3)$$

Where $\mu_{(i)} = \mu_0 = 4\pi \times 10^{-7}$ H/m for non-ferromagnetic wire (to within a few parts per 1000).

²³ Open Document spreadsheets can be opened and edited using **Open Office**, available from <http://www.openoffice.org/>. and **Libre Office**, available from: <http://www.libreoffice.org/>.

The relationship between the dispersion onset frequency f_s and wire diameter is shown below for IACS copper wire ($\rho = 17.241 \text{ n}\Omega\text{m}$ at 20°C , $\mu_{(i)} = \mu_0$). The calculation is given in the accompanying spreadsheet [Li_transition.ods](#), sheet 2.



With quick reference to the graph, and using equation (6.2) for accuracy, we find (for example) that a coil wound with 1 mm diameter copper wire will continue to exhibit DC behaviour up to 17.5 kHz, whereas using 0.1 mm wire will push the limit up to 1.75 MHz (although it will be necessary to include self-capacitance in the model to calculate the correct reactance in that case). While this type of information might be useful for the purpose of designing low-frequency reference inductors however, it is not so good for deciding the frequency above the dispersion region at which the inductance can once again be considered to be constant. A fair rule of thumb is to adopt the point at which $r_w/\delta_i=10$, which occurs two decades above f_s ; but if the objective is (say) to design an accurate high-frequency reference coil, it is better to minimise the proximity effect by using a large p/d ratio and include internal inductance in the model.

6b. Using the internal inductance factor

Although internal inductance is perturbed by the proximity effect, the frequency interval over which the major part of the dispersion occurs is not greatly affected. We can therefore use the internal inductance factor Θ to determine whereabouts we are in the dispersion region. Θ is easily obtained by, for example, making a direct call to the Basic routine given earlier:

$$\Theta = \text{Flpml}(q) \quad , \quad \text{where} \quad q = \frac{d}{\delta_i \sqrt{2}}$$

From the discussion at the beginning of this section, we have:

$$L_i = L_{i(\text{dc})} \Theta$$

where, with a rearrangement to separate a factor of $1/4$:

$$L_{i(\text{dc})} = \ell_w \frac{\mu_{(i)}}{2\pi} \frac{1}{4}$$

At low frequencies, where $r_w > \delta_i$, $\Theta = 1$; and we can find the high frequency limiting value for Θ from the expression (as was given earlier):

$$L_{i(\text{hf})} = \ell_w \left(\frac{\mu_{(i)}}{2\pi} \right) \left(\frac{\delta_i}{d} \right) \quad . \quad \left(\text{where } \delta_i = \sqrt{\frac{\rho}{\pi f \mu_{(i)}}} \right)$$

Thus, comparing the high and low frequency limits, we have:

$$\text{as } f \rightarrow \infty \quad , \quad \Theta \rightarrow 4 \frac{\delta_i}{d} \quad (6.4)$$

Notice here that δ_i is ever diminishing as frequency increases, so that the internal inductance becomes very small and effectively disappears.

Internal inductance of a solenoid

The wire length of a solenoid is given by (5.2):

$$\ell_w = \frac{2\pi r N}{\cos \psi}$$

Hence, the internal inductance of a solenoid expressed using the internal inductance factor is:

$$L_i = \frac{2\pi r N}{\cos \psi} \frac{\mu_{(i)}}{8\pi} \Theta$$

i.e.:

$L_i = \frac{\mu_{(i)} r N}{\cos \psi} \frac{\Theta}{4} \quad [\text{Henrys}]$	6.5
--	------------

6c. Effective current sheet diameter linked to internal inductance

The skin effect and proximity effect dispersions are interlinked and so occur on the same frequency interval. Therefore, at least to a reasonable first-order approximation, we can use the internal inductance factor Θ to weight the change in effective diameter from D_0 to D_∞ . This can be done as follows:

When $d/2 < \delta_i$, then $D \rightarrow D_0$,

and when $d/2 \gg \delta_i$, then $D \rightarrow D_\infty$.

Hence, to track the diameter change through the dispersion region:

$$D = D_0 \Theta + D_\infty (1 - \Theta)$$

i.e.;

$$D = \Theta (D_0 - D_\infty) + D_\infty \quad (6.6)$$

If D_a is the average coil diameter and $D_a \gg d$, then D_0 , as was discussed in [section 3](#), can be approximated as:

$$D_0 = D_a \left[1 - \left(\frac{d}{D_a} \right)^2 \right]$$

D_∞ is given by equation [\(4.2\)](#) as:

$$D_\infty = \frac{D_0 + D_{\min} \frac{a}{[(p/d)-1]}}{1 + \frac{a}{[(p/d)-1]}} \quad , \quad (a \approx 100, \text{ see text below}) \quad (6.7)$$

Where:

$$D_{\min} = \frac{[(N-2)(D_a - d) + 2D_0]}{N}$$

The empirical parameter a was given in [section 4](#) as 2, for HF only calculations. Now, since we are removing the requirement that $d \gg \delta_i$ when calculating the effective diameter, we need to bias D_∞ to be somewhat closer to D_{\min} . This can be done by increasing a to give a good average match to the most accurate HF inductance measurements we can obtain. It turns out, in practice, that for best results, D_∞ needs to be biased very strongly towards D_{\min} . A suitable value for a is around 100.

7. Magnetic field non-uniformity (Nagaoka's coefficient)

By far the greatest correction to the long-current-sheet formula is that which allows for the magnetic-field non-uniformity that appears when the length of the coil becomes comparable to its diameter (i.e., when the coil is short). This modification is analogous to the Maxwell fringing-field correction for a parallel-plate capacitor, but is a gross rather than a minor effect. It can be implemented by including a dimensionless factor (analogous to relative permeability), which we will here call k_L . Thus, for coils of arbitrary length/diameter ratio (ℓ/D):

$L_s = \frac{\mu \pi r^2 N^2 k_L}{\ell} \quad [\text{Henrys}]$	7.1
--	-----

where the inductance L_s retains its subscript as a reminder that it is still a current-sheet inductance and should only be regarded as an approximation to the inductance of a practical coil.

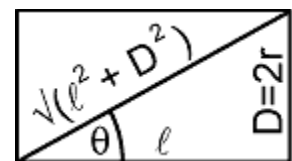
The subscript L in k_L can be taken to stand for 'Lorenz'; because it was Ludwig Lorenz, in 1879, who was the first to find an analytical expression for the inductance of current sheet solenoid of finite length²⁴. The factor k_L however (usually given elsewhere without a subscript) is most commonly known as *Nagaoka's coefficient*, because it was [Hantaro Nagaoka](#) who, in 1909, introduced it and developed a practical method for calculating it²⁵.

At this point note that; when reading early papers on electromagnetism, the cgs system of units is used. In that case inductance has units of length (cm \equiv nH). In the rationalised mks system²⁶, which is the basis of the SI, all inductance formulae are multiplied by $\mu_0/4\pi$ to put them into Henrys. (Also be aware of Nagaoka's use of turns per unit length, i.e., $n = N/\ell$). Here, to avoid confusion, we will use only SI units; and so, bearing in mind that it will look different in its original form, Lorenz's expression for the inductance of a current sheet (assuming that $\mu = \mu_0$) becomes:

$$L_s = \mu_0 N^2 \frac{8 r^3}{3 \ell^2} \left[\frac{2\kappa^2 - 1}{\kappa^3} E(\kappa) + \frac{1 - \kappa^2}{\kappa^3} K(\kappa) - 1 \right] \quad [\text{Henrys}] \quad (7.2)$$

where $K(\kappa)$ and $E(\kappa)$ are complete elliptic integrals of the first and second kind respectively²⁷, and the argument κ (kappa), known as the 'modulus' of the integral, is given by:

$$\kappa = \sin \theta = \frac{D}{\sqrt{D^2 + \ell^2}}$$



(for the definition of θ , see the diagram on the right).

Unfortunately, although the formula (7.2) looks reasonably straightforward, the complete elliptic integrals have to be calculated from infinite series that do not converge particularly rapidly. This meant that solenoid inductance calculation was impractical for the majority of engineers and

²⁴ **BS Sci.** **169**, pages 117 - 118.

²⁵ **The Inductance Coefficients of Solenoids.** H Nagaoka, J. Coll. Sci. Tokyo, Vol 27 (6), 1909. [**Nagaoka 1909**] [Available from [University of Tokyo Repository](#) and [g3ynh.info/zdocs/magnetics/](#)]

²⁶ **The mks or Giorgi system of units.** L H A Carr. Proc. IEE, Part I: General, 97(107), 1950. p235-240.

The process of 'rationalisation' is that of moving the factor of 4π out of the unit electric and magnetic fluxes and into the attached permittivity or permeability.

See also: **The position of 4π in electromagnetic units** (discussion between Oliver Lodge and Oliver Heaviside, 1892). Heaviside, Electrical papers, Vol II. p575 - 578. [available from [Internet Archive](#).]

²⁷ See, for example: **Tables of Integrals and Other Mathematical Data**, H B Dwight. 4th edition, Macmillan 1961 (10th printing 1969). LCCN 61-6419. [**Dwight 1961**] Articles 773.1 - 774.2 and tables 1040 - 1041.

scientists working in the first half of the 20th Century. Nagaoka's solution to that problem begins with an observation equivalent to saying that (7.2) can be put into the form of (7.1). Furthermore, note that, in the expression for κ , we can factor D from the denominator to get:

$$\kappa = \frac{1}{\sqrt{1+(\ell/D)^2}}$$

Alternatively, we can factor ℓ from the denominator and get:

$$\kappa = \frac{(D/\ell)}{\sqrt{(D/\ell)^2+1}}$$

Hence, by comparison of the two expressions for current-sheet inductance, we can observe directly that the value of Nagaoka's coefficient k_L in equation (7.1) depends only on a single dimensionless argument, which is usually chosen (by rearrangement of formulae) to be either ℓ/D (i.e., $\cot\theta$) or D/ℓ (i.e., $\tan\theta$). On that point, note that D/ℓ has the convenient property that it is zero for an infinitely long coil, and it can never become infinite because the finite thickness of the winding wire prevents a coil from ever having zero length. This means that the second choice helps in the avoidance of program errors. The first choice is however more intuitive, so we will tend to plot graphs showing ℓ/D as the abscissa (horizontal axis), but perform the calculations using D/ℓ .

Thus we find that k_L depends only on the shape of the coil, and not on its absolute physical size (within the limitations of the Lorenz model, which will be discussed later; and provided, of course, that the static magnetic field approach is valid for the system under consideration; i.e., the length of the conductor must be small in comparison to wavelength).

Putting the SI version of Lorenz's expression into the form of (7.1) we get²⁸:

$$L_s = \mu_0 N^2 \frac{\pi r^2}{\ell} \frac{8r}{3\pi \ell} \left[\frac{2\kappa^2-1}{\kappa^3} E(\kappa) + \frac{1-\kappa^2}{\kappa^3} K(\kappa) - 1 \right] \quad [\text{Henrys}]$$

Thus, noting that $2r/\ell = D/\ell$, we obtain a candidate for the evaluation of Nagaoka's coefficient:

$$k_L = \frac{4(D/\ell)}{3\pi} \left[\frac{2\kappa^2-1}{\kappa^3} E(\kappa) + \frac{1-\kappa^2}{\kappa^3} K(\kappa) - 1 \right] \quad . \quad \text{Lorenz form of Nagaoka's coefficient}$$

It is important to be aware however, that this expression has been variously transformed and non-trivially rearranged by investigators over the years, the preferred form in any given case being chosen to facilitate some particular attack on the problem of how to calculate it or approximate it. Nagaoka's preference was²⁹:

$$k_L = \frac{4}{3\pi} \frac{1}{\kappa'} \left[\frac{\kappa'^2}{\kappa^2} [K(\kappa) - E(\kappa)] + E(\kappa) - \kappa \right] \quad . \quad \text{Nagaoka's form}$$

where, as before:

²⁸ **Methods for the derivation and expansion of formulas for the mutual inductance of coaxial circles and the inductance of single-layer solenoids.** F W Grover, NBS J. Research. Vol 1, 1928, [BS RP16].
Page 503, Equation 72.

²⁹ **Nagaoka 1909.** Equation 17. page 20.

$$\kappa = \sin\theta = \frac{D}{\sqrt{D^2 + \ell^2}} = \frac{(D/\ell)}{\sqrt{(D/\ell)^2 + 1}}$$

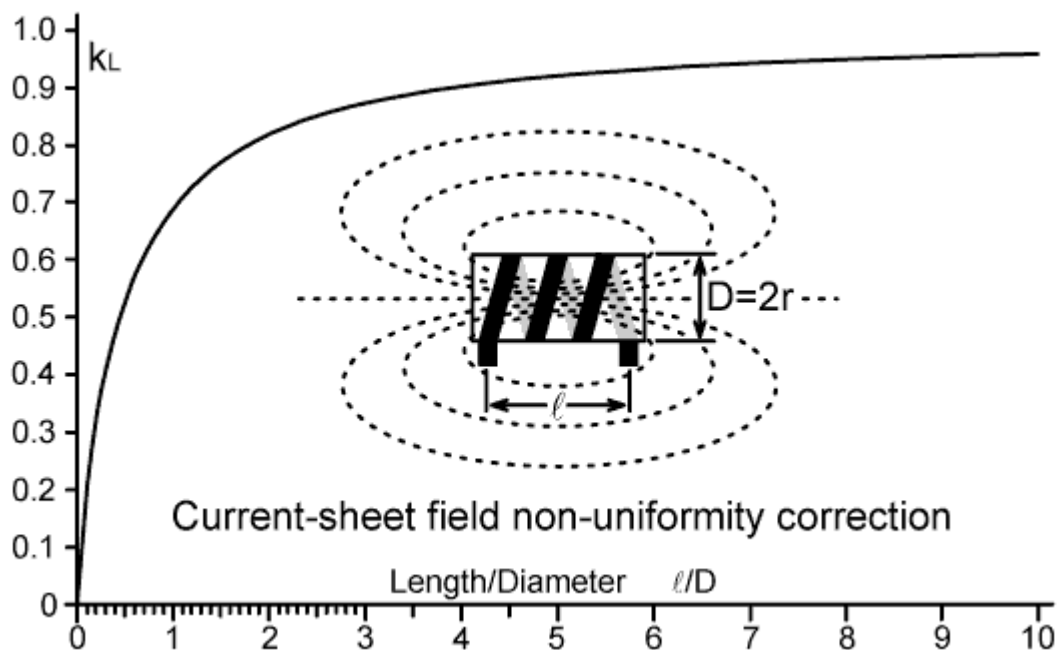
and

$$\kappa' = \cos\theta = \frac{\ell}{\sqrt{D^2 + \ell^2}} = \sqrt{1 - \kappa^2}$$

At first glance, Nagaoka's rearrangement seems to give rise to a proliferation of complete elliptic integrals, but bear in mind that the point was to evaluate the expression using series expansions. In that case, there are only two series; one for $E(\kappa)$ and one for $K(\kappa) - E(\kappa)$. Somewhat glossing over the details, we will merely note here that the combination series (i.e., the series obtained by term-by-term subtraction) provides faster convergence and suppresses roundoff error in regions of the argument range where the two integrals are similar in value.

Nagaoka tabulated his coefficient to 6 decimal places in his 1909 paper (for its symbol, he uses a Gothic form of the letter z). His calculations have also been checked and reproduced by Rosa and Grover (who use the symbol K)^{30 31}. One Australian manufacturer even produced an engineer's slide rule with a scale for Nagaoka's coefficient³².

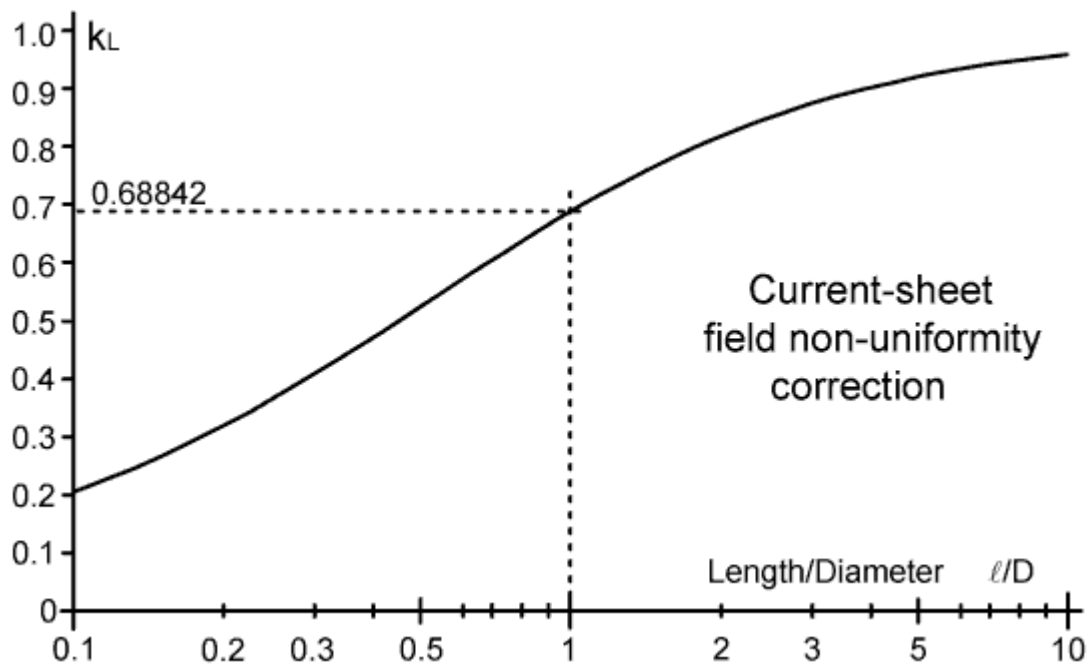
Graphs of k_L vs. ℓ/D are shown below, first on a linear scale, and then on a logarithmic scale for $0.1 \leq \ell/D \leq 10$.



30 Grover 1946. See Tables 36 and 37. p144 - 147. Some minor corrections are also given in Grover errata.

31 BS Sci. 169, page 224.

32 REED Riddle solved, David G Rance. J. Oughtred Soc. Vol 17(1), April 2008.



Notice that k_L varies between 0 and 1 as ℓ/D varies between 0 and ∞ . If we put both $\ell = 0$ and $k_L = 0$ into equation (7.1), we get $L_s \rightarrow 0/0$. Thus the long-coil formula modified by Nagaoka's coefficient correctly asserts that the inductance of a zero-length solenoid is undefined, whereas the unmodified formula (1.2) tells us that the inductance of a zero-length coil is infinitely large. Nagaoka's coefficient therefore serves to impose a necessary limiting condition.

If a one-off current-sheet inductance calculation is to be performed, the use of Nagaoka tables is not a bad idea. For the general business of electromagnetic modelling however, we require efficient algorithms for the situations in which a programming environment is available, and reasonably compact and accurate formulae otherwise. Since this is a long-standing problem, the methods available are many and varied; but the choice is often confused (and the use of computer programs of unknown provenance is fraught) by failure to distinguish between exact methods and approximations. Here we will review some of the options, and evaluate and improve certain well known formulae.

For the exact calculation of current sheet inductance (where 'exact' means; 'according to the Lorenz model' and 'within the precision of computer floating-point arithmetic'), the most obvious approach is to use Lorenz's formula or some rearrangement thereof. To that end Bob Weaver³³ has devised program routines based on Dwight's formulae for calculation of the the complete elliptic integrals³⁴ $K(\kappa)$ and $E(\kappa)$. These routines, which are also useful for loop inductance calculation and other magnetics problems, are given as Open Office Basic macro functions and can be copied from the accompanying spreadsheet file³⁵. An example Basic function for calculating Nagaoka's coefficient using the Lorenz form, which calls the separate $K(\kappa)$ and $E(\kappa)$ functions, is shown below.

33 Numerical methods for inductance calculation, <http://electronbunker.ca/CalcMethods.html>.

34 Practical considerations in the calculations of Kelvin functions and complete elliptic integrals, Robert S Weaver, 2009. <http://electronbunker.ca/DLpublic/KelvinEllipticCalcs.pdf>. [also available from g3ynh.info/zdocs/magnetics/]

35 Inductance examples. Bob Weaver 2009. g3ynh.info/zdocs/magnetics/appendix/InductanceExamples.ods
To view and edit macros, use the Open Office top menu and navigate to:
Tools/Macros/Organise Macros/OpenOffice.org Basic

```

Function Lorenz(ByVal x as Double) As Double
' Nagaoka's coeff using Lorenz form. x is solenoid Diam / length.
' Calls complete elliptic integral functions EllipticE and EllipticK.
If x = 0 then
  Lorenz = 1
else
Dim k as double, kk as double, kkk as double, c1 as double, c2 as double
  k = x/sqr(1+x*x)
  kk = x*x/(1+x*x)
  kkk = k*kk
  c1 = (2*kk-1)/kkk
  c2 = (1-kk)/kkk
  Lorenz=4*x*(c1*EllipticE(k) + c2*EllipticK(k) -1)/(3*pi)
end if
End Function

```

As mentioned earlier however, direct and separate use of $K(\kappa)$ and $E(\kappa)$ is not the most computationally efficient approach. The importance of elliptic integrals in the wider scientific context has also led to considerable research into the properties of the numerous possible series expansions. Bob Weaver³⁶ goes on to draw our attention to the arithmetico-geometric-mean (AGM) method^{37 38 39 40} for computing the complete elliptic integrals and linear combinations thereof. Using that approach, he gives an algorithm that calculates the AGM and the linear combination $[K(\kappa) - E(\kappa)] / K(\kappa)$ in a single program loop that requires no more than 9 iterations for $\ell/D > 0.001$. The AGM is a simple multiple of $1/K(\kappa)$, and so $E(\kappa)$ and the combination $K(\kappa) - E(\kappa)$ are then extracted by trivial arithmetic and Nagaoka's coefficient is calculated using Nagaoka's preferred form. Nagaoka's form, as mentioned earlier, gives less roundoff error than the approach using separately-evaluated complete elliptic integrals.

Bob reports that his AGM calculations (in double precision arithmetic) differ from Nagaoka's table by a maximum of 9 in the 6th decimal place at $\ell/D = 0.1$ (this is the shortest coil form for which Nagaoka gives data). Comparison of the AGM method with Lundin's short coil formula (discussed in the next section) however shows agreement to at least 6 decimal places in that region. Since Lundin's formula is asymptotically-correct for short coils, it appears that the discrepancy is due either to an error in the table or to the limited precision of the 1909 calculation. In fact, in the limits where Lundin's formulae tend towards the analytical values, deviation from the AGM result is less than 1 part in 10^8 .

In view of the forgoing, and since the code is available as an Open Office Basic macro, we will here use Bob Weaver's AGM algorithm as the standard method for calculating Nagaoka's coefficient and as a benchmark against which to evaluate the accuracy of other formulae.

36 <http://electronbunker.ca/CalcMethods1a.html>

37 **On some new formulae for the numerical calculation of the mutual induction of coaxial circles.** Louis V King. Proc. Roy. Soc.. Series A, Vol .100 (702), 1921, p60-66.

38 **BS RP16**, 1928. p496.

39 **Inductance formula for a single-layer coil**, H Craig Miller, Proc. IEEE, Vol 75 (2), 1987, Letters, p256-257.

40 **Mutual inductance calculations by Maxwell's Method.** Antonio Carlos M de Querioz, 2003, 2005. <http://www.coe.ufrj.br/~acmq/tesla/maxwell.pdf>

8. Approximate methods for calculating Nagaoka's coefficient

Given that it is nowadays possible to calculate Nagaoka's coefficient extremely accurately with minimal effort, it might seem that there is no longer any need for approximate formulae.

Approximations however can still be useful in a variety of ways. Approximate expressions that are precise (i.e. functions that produce smooth curves) can, for example, be used to check the performance of algorithms that are accurate but noisy (i.e., subject to roundoff error or abrupt truncation). Asymptotically-correct approximations can also be used to create simple analytical expressions applicable in special cases (i.e., for very short or very long coils). Some people also want to perform calculations using only built-in spreadsheet functions or hand calculators, and so it is useful to have formulae that are relatively simple.

8a. Lundin's handbook formula

An extremely accurate approximation formula for Nagaoka's coefficient is due to Richard Lundin⁴¹. This is known as 'Lundin's Handbook Formula', and is in the form of two expressions, one for $\ell/D \leq 1$ (short coils), and one for $\ell/D > 1$ (long coils). Both expressions can be used to calculate k_L for $\ell/D = 1$, but the short coil expression gives a value that agrees with the AGM result to 6 decimal places at this point ($k_L = 0.688423$). Lundin's formula agrees with the Lorenz model to better than 3 parts-per-million ($\pm 0.0003\%$), this being generally superior to the accuracy with which a coil can be made or measured, and less than the error due to dimensional variation with temperature. The two expressions are given in the boxes below:

Lundin's formula for short coils ($D \geq \ell$). Max. error: < 2 ppM (0.0002%).

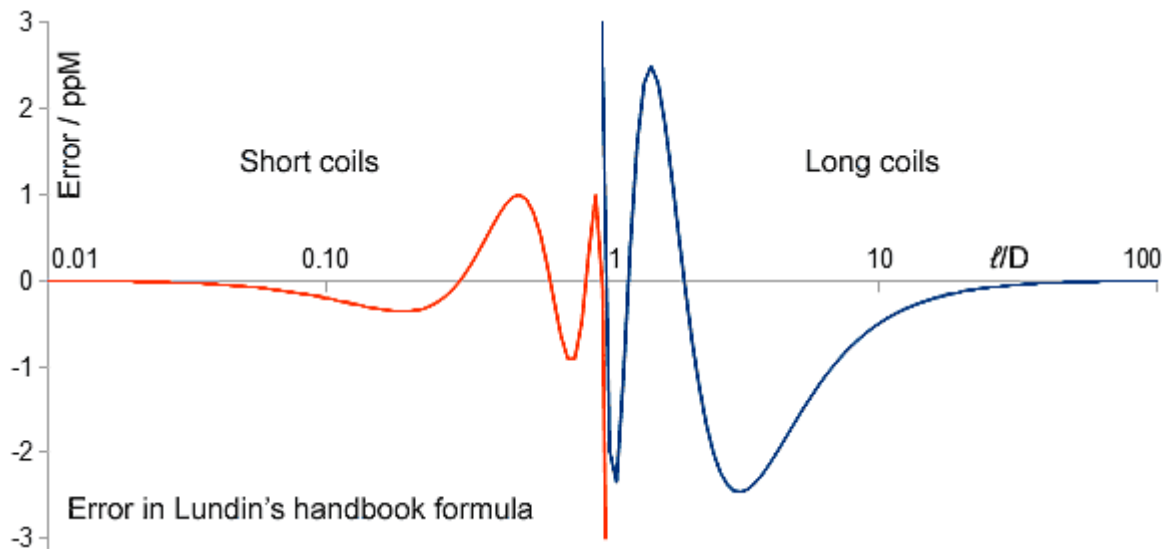
$$k_{LS} = \frac{2}{\pi} (\ell/D) \left[\frac{[\ln(4D/\ell) - \frac{1}{2}] [1 + 0.383901(\ell/D)^2 + 0.017108(\ell/D)^4]}{[1 + 0.258952(\ell/D)^2]} \right. \\ \left. + 0.093842(\ell/D)^2 + 0.002029(\ell/D)^4 - 0.000801(\ell/D)^6 \right]$$

Lundin's formula for long coils ($\ell > D$). Max. error: < 3 ppM (0.0003%).

$$k_{LL} = \frac{[1 + 0.383901(D/\ell)^2 + 0.017108(D/\ell)^4]}{[1 + 0.258952(D/\ell)^2]} - \frac{4(D/\ell)}{3\pi}$$

41 **A Handbook Formula for the Inductance of a Single-Layer Circular Coil**, Richard Lundin, Proc. IEEE, Vol. 73 (9), p 1428-1429, Sept. 1985.

A graph comparing Lundin's formula against the AGM calculation is given below (the calculation can be inspected by downloading the accompanying spreadsheet: [L_formulae.ods](#). See sheet 1). A Basic macro function for calculating Nagaoka's coefficient using Lundin's formula can be copied using the spreadsheet macro editor and is shown below the graph. Note that the calling argument used by the function is D/ℓ .



```
Function Lundin(byVal x as double) as double
' Calculates Nagaoka's coeff. using Lundin's handbook formula. x = solenoid diam. / length
Dim num as double, den as double, kk as double, xx as double, xxxx as double
xx = x*x
xxxx = xx*xx
if x = 0 then
  Lundin=1
elseif x<1 then
  num = 1 + 0.383901*xx +0.017108*xxxx
  den = 1 + 0.258952*xx
  Lundin = num/den -4*x/(3*pi)
else
  num = (log(4*x)-0.5)*(1 +0.383901/xx + 0.017108/xxxx)
  den = 1 + 0.258952/xx
  kk = 0.093842/xx +0.002029/xxxx -0.000801/(xx*xxxx)
  Lundin = 2*( num/den + kk )/(pi*x)
end if
End function
```

8b. Analytic asymptotic approximations for Nagaoka's coefficient

Shown below and on the following page are some analytical restricted-range approximations for k_L . These are truncated versions of infinite series representations involving no empirical constants. They are also singly asymptotically-correct; i.e., they converge with Nagaoka's coefficient in the limit of a very long or a very short coil, but are not exact for coils of intermediate length. They can provide a useful check on the coding or transcription of other formulae, since agreement within the stated limits between two different expressions is a very good test of correctness. They can also be truncated and otherwise modified to produce compact formulae for special applications. Note that, in each case, we show not the original formula, but an approximation for k_L taken from it.

Shown with the formulae are short Basic routines that can be copied from the macro library in the the spreadsheet [L_formulae.ods](#). These have been verified against the elliptic integral calculation and so are typographically more authoritative than the given formulae.

Rayleigh-Niven formula⁴² for short coils of zero radial conductor thickness.

Coincident with Nagaoka's coefficient as $\ell/D \rightarrow 0$, +0.08% when $\ell/D = 0.5$,
+0.28% when $\ell/D = 0.7$

$$k_{RN} = \frac{2}{\pi} (\ell/D) \left[\ln(4D/\ell) - \frac{1}{2} + \frac{1}{8} (\ell/D)^2 \left[\ln(4D/\ell) + \frac{1}{4} \right] \right]$$

Function Rayleigh(byVal x as double) as double

' Calculates Nagaoka's coeff. using Rayleigh-Niven short coil formula. x = Diam. / length

Dim lg as double, s as double

lg = log(4*x)

s = lg -0.5 + (lg +0.25)/(8*x*x)

Rayleigh = (2/pi)*s/x

end function

⁴² On the determination of the Ohm in absolute measure. **Rayleigh Scientific Papers, Vol II 1881-1887**, Cambridge UP 1900, p15 formula (12), for coils of zero radial thickness. Also **BS Sci. 169**, p116

Coffin's Formula⁴³ for short current-sheet coils.

Extended version of the Rayleigh-Niven formula.

Coincident with Nagaoka's coefficient as $\ell/D \rightarrow 0$, -0.21% when $\ell/D=1$

$$k_{CF} = \frac{2}{\pi} (\ell/D) \left[\ln(4D/\ell) - \frac{1}{2} + \frac{1}{8} (\ell/D)^2 \left[\ln(4D/\ell) + \frac{1}{4} \right] - \frac{1}{64} (\ell/D)^4 \left[\ln(4D/\ell) - \frac{2}{3} \right] \right. \\ \left. + \frac{5}{1024} (\ell/D)^6 \left[\ln(4D/\ell) - \frac{109}{120} \right] - \frac{35}{16384} (\ell/D)^8 \left[\ln(4D/\ell) - \frac{431}{420} \right] \right]$$

Function Coffin(byVal x as double) as double

' Calculates Nagaoka's coeff. using Coffin's short-coil formula. x = Diam. / length

Dim lg as double, xx as double, xxxx as double, s as double

lg = log(4*x)

xx=x*x

xxxx=xx*xx

s = lg -0.5 + (lg+0.25)/(8*xx) - (lg-2/3)/(64*xxxx) + (lg-109/120)*5/(1024*xx*xxxx) - (lg-431/420)*35/(16384*xxxx*xxxx)

Coffin = (2/pi)*s/x

end function

Webster-Havelock formula⁴⁴ for long current-sheet coils.

Coincident with Nagaoka's coefficient as $\ell/D \rightarrow \infty$, +0.06% when $\ell/D=1$

$$k_{WH} = 1 - \left(\frac{4}{3\pi} \right) (D/\ell) + \frac{1}{8} (D/\ell)^2 - \frac{1}{64} (D/\ell)^4 + \frac{5}{1024} (D/\ell)^6 - \frac{35}{16384} (D/\ell)^8 + \frac{147}{131072} (D/\ell)^{10}$$

Function Webster(byVal x as double) as double

' Calculates Nagaoka's coeff. using Webster-Havelock long-coil formula. x = Diam. / length

Dim xx as double, x4 as double, x6 as double, x8 as double, x10 as double

xx = x*x

x4 = xx*xx

x6 = xx*x4

x8 = x4*x4

x10 = x8*xx

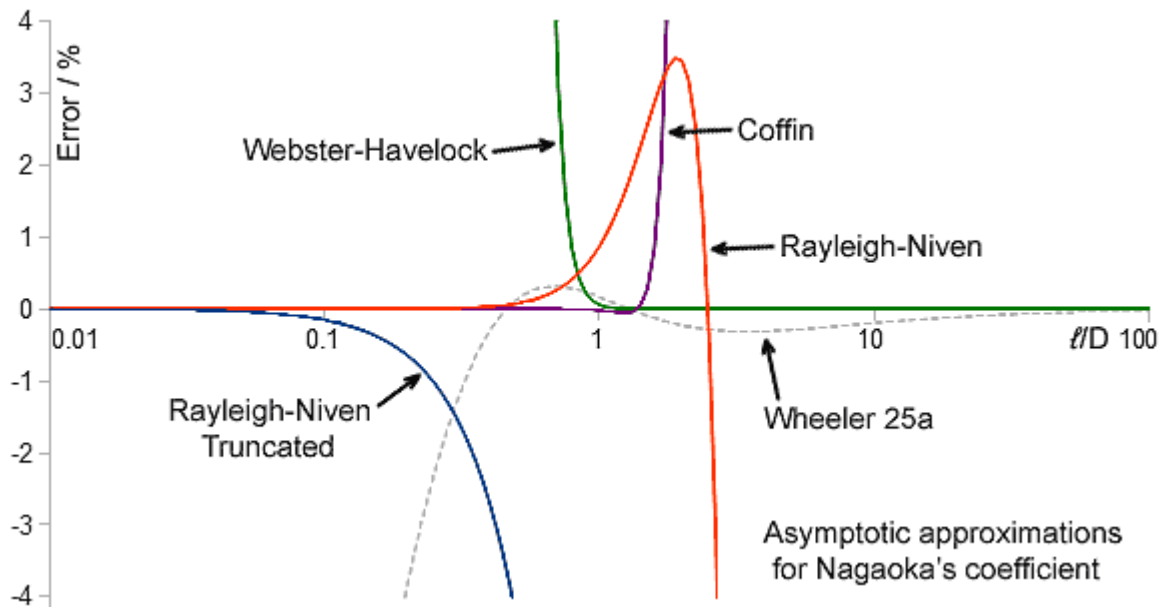
Webster = 1 -4*x/(3*pi) +xx/8 -x4/64 +5*x6/1024 -35*x8/16384 +147*x10/131072

end function

43 **BS Sci. 169**, p117. A truncated version of Coffin's formula is also given in **Grover 1948**, page 143, formula 119, but there is a typographical error in that case: see **Grover errata**.

44 **BS Sci. 169**, p121

The ways in which various restricted-range approximations deviate from the exact calculation is shown in the graph below (see the spreadsheet: [L_formulae.ods](#), sheet 2).



The curve marked 'Rayleigh-Niven truncated' is obtained by using only the first term of the Rayleigh-Niven-derived formula, i.e.;

$$k_{\text{RNT}} = \frac{2}{\pi}(\ell/D) \left[\ln(4D/\ell) - \frac{1}{2} \right]$$

(we refer to it as a single term because it is the first element of a power series in ℓ/D). This simple short-coil formula is accurate to within 0.16% up to $\ell/D = 0.1$.

The curve marked 'Wheeler 25a' is also due to a remarkably simple formula, which is discussed in the next section.

8c. Wheeler's long-coil (1925) formula

What is probably the best-known formula for single-layer solenoid inductance was published in 1928 by Harold A Wheeler⁴⁵. This is widely known as 'Wheeler's formula', but since there are numerous candidates for that title we will refer to it here as 'Wheeler's long coil formula', or 'Wheeler's 1925 formula' (that being stated to be the year of its derivation), or as W25 for short. The formula is given in its original form as:

$$L = \frac{a^2 N^2}{9a + 10b} \quad [\text{microHenries}]$$

Where a and b are respectively the radius and length of the coil in inches. The accuracy is claimed to be within $\pm 1\%$ for $b > 0.8a$.

The above is, of course, a current-sheet formula (even though it is not identified as such in the original paper). It is therefore interesting to rearrange it with a view to putting it into the form of equation (7.1) and extracting an expression for Nagaoka's coefficient. We start by factoring b from the denominator and multiplying by 10^{-6} to convert it to Henrys. This gives:

$$L_s = \frac{10^{-6} a^2 N^2}{b \left(10 + 9 \frac{a}{b} \right)} \quad [\text{Henrys}]$$

Now note that the quantity a/b (i.e., radius / length) is dimensionless. We can therefore immediately replace that part using the symbols preferred here (although we will use $D/2$ instead of r). Thus:

$$L_s = \frac{10^{-6} a^2 N^2}{b(10 + 4.5 D/\ell)}$$

and factoring 10 from the denominator gives:

$$L_s = 10^{-7} N^2 \frac{a^2}{b} \frac{1}{(1 + 0.45 D/\ell)} \quad [\text{Henrys}]$$

Now recall that Nagaoka's coefficient $\rightarrow 1$ when the coil becomes very long and thin, i.e., when $D/\ell \rightarrow 0$. Hence, according to this asymptotic behaviour, we can extract an approximation for Nagaoka's coefficient as:

$$k_{w25} = \frac{1}{(1 + 0.45 D/\ell)}$$

Reinserting this into equation (7.1) we get:

$$L_s = \mu N^2 \frac{\pi r^2}{\ell} \frac{1}{(1 + 0.45 D/\ell)} \quad [\text{Henrys}]$$

with an accuracy of $\pm 0.33\%$ for $\ell \geq 0.4D$.

45 **Simple inductance formulas for radio coils.** Harold A Wheeler, Proc. IRE, 1928, Vol 16 (10) p1398-1400.

This is fairly close to the optimal, but a small adjustment of the empirical coefficient from 0.45 to 0.4502 reduces the maximum error for the range $\ell/D \geq 0.4$ from 0.33 to 0.32%. Thus the best metric formula is:

$L_s = \frac{\mu \pi r^2 N^2}{\ell(1+0.4502 D/\ell)}$ [Henrys]	$\pm 0.32 \% \text{ for}$ $\ell \geq 0.4D$	W25a (optimised)
--	---	----------------------------

with

$$k_{W25a} = \frac{1}{(1+0.4502 D/\ell)}$$

The complete current-sheet expression (**W25a**), with the original coefficient value of 0.45 is a trivial rearrangement of a formula that appears in the 1965 first edition of Ramo et al⁴⁶. In that book, the formula is simply attributed to Wheeler's 1928 paper; but it has been pointed out by Rodger Rosenbaum⁴⁷, that the attribution is misleading because Wheeler's 1925 formula is not an asymptotic approximation. The sleight of hand involved in using the asymptotic behaviour of Nagaoka's coefficient, rather than the length conversion factor, results in a small discrepancy between the inch and metric forms. If we express the error as a proportion p we have:

$$L = \frac{\mu_0 N^2 \pi r^2 k_{W25}}{\ell} = \frac{10^{-7} p N^2 a^2 k_{W25}}{b}$$

and substituting $\mu_0 = 4\pi \times 10^{-7}$ gives:

$$\frac{4\pi^2 r^2}{\ell} = \frac{p a^2}{b}$$

The US inch to metric conversion factor⁴⁸ in use in 1928 was 1" = 25.400051 mm. Since most readers will be interested in the discrepancy obtained using the modern conversion factor however, we will use 1" = 25.4 mm . Thus:

$$r = 25.4 \times 10^{-3} a \quad \text{and} \quad \ell = 25.4 \times 10^{-3} b$$

Using these substitutions gives:

$$\frac{4\pi^2 (25.4 \times 10^{-3} a)^2}{25.4 \times 10^{-3} b} = \frac{p a^2}{b}$$

i.e.:

$$p = 4\pi^2 \times 25.4 \times 10^{-3} = 1.002\,751\,807$$

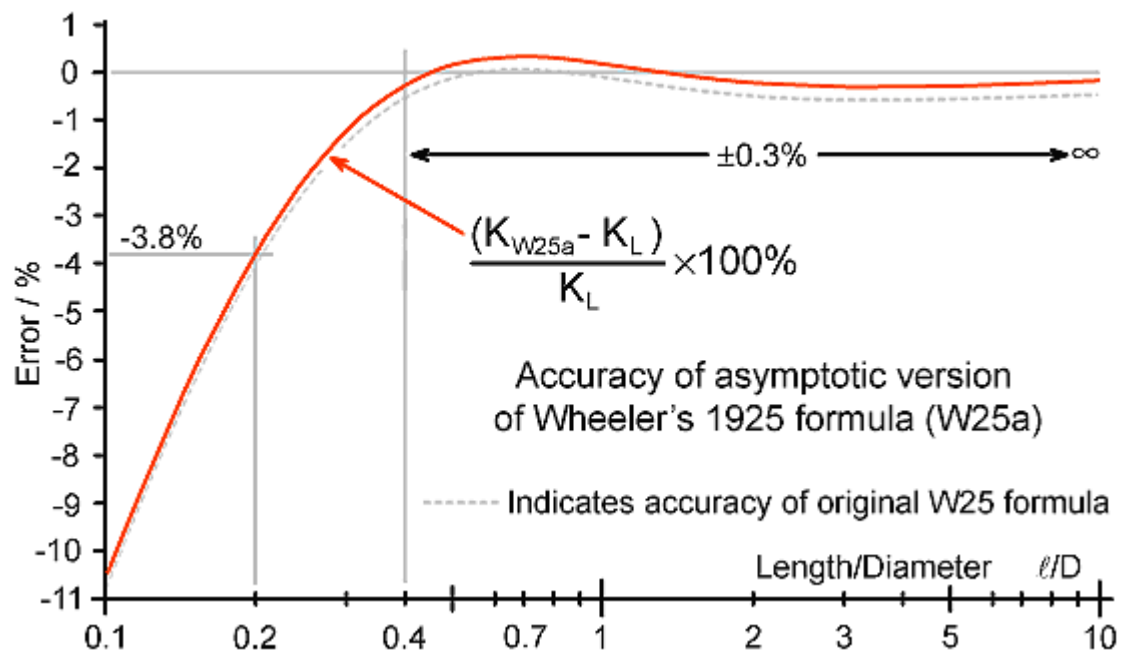
46 **Fields and Waves in Communication Electronics**, Simon Ramo, John R. Whinnery, Theodore Van Duzer, Publ. John Wiley & Sons Inc. 1965. Library of congress cat. card no. 65-19477. page 313. The same formula also appears in **Ramo et al 1994** (the 3rd edition) on page 195.

47 **Subtle error**. Rodger Rosenbaum, Private e-mail communications 27th & 28th March 2009.

48 **Physical and chemical constants**, Originally compiled by G W C Kaye and T H Laby, 12th edition, Longmans, 1959. Pages 2-3.

Thus Wheeler's 1925 formula comes out at $1 / 1.002751807 = 0.997255744$ of the long coil asymptotic value for the current sheet formula (i.e., 0.274% low). Due to the shape of the error curve however; this choice, makes it slightly better than the metric asymptotic version in the region around $\ell/D = 0.7$, but inferior for longer coils (see graph below). Such subtleties, of course, do not affect the use of either formula in the preferred application as a simple approximation for use with hand calculators.

The reworking of Wheeler's 1925 formula as detailed above provides a simple asymptotically-correct semi-analytical current-sheet formula giving good accuracy with only a single empirical parameter. It should be noted however, in view of Rodger's criticism, that it is not strictly Wheeler's formula because it does not return the same values as Wheeler's formula. Hence, since Ramo et al. do not identify its actual origin (presumably, one of them derived it), it is recommended that it should be referred to as the 'asymptotic version of Wheeler's long-coil formula', or 'Wheeler 25a'. Its accuracy is shown below.



The calculation used to produce the error curves can be examined in the spreadsheet [L_formulae.ods](#) (sheet 2). When examining the spreadsheet note that the result for Wheeler 25 is simulated by using the scale factor p rather than dimensions in inches.

The ℓ/D range from 0.4 to ∞ covers a wide range of practical situations. Hence Wheeler 25 and minor variants provide an excellent starting point for estimating the inductance of a coil from its dimensions, particularly when an approximation within a few % is all that is required. Note however, that the accuracy deteriorates rapidly for $\ell/D < 0.4$, the error (incurred using the original 1925 formula) being -1.4% when $\ell = 0.3 D$, -4% when $\ell = 0.2 D$, and -10.8% when $\ell = 0.1 D$. This should not be an issue for hand calculation because we can use the Rayleigh-Niven current-sheet formula for short coils, and this is accurate to within 0.3% for ℓ/D up to 0.7. There is a warning here however for people who use 'online inductance calculator' programs of unknown provenance. As was pointed out by Bob Weaver⁴⁹ after a survey carried out in 2010; nearly every

⁴⁹ Numerical methods for inductance calculation part 3. <http://electronbunker.ca/CalcMethods3b.html>.

Also private e-mail communications, 3rd and 4th May 2010.

program on offer used Wheeler's 1925 formula, despite the fact that any person finding out about the program would, by definition, be doing so using a computer. Worse still, many of those programs failed to check for input data corresponding to $l/D < 0.4$.

8d. Wheeler's 1982 unrestricted formulae

After an interval of 54 years, Harold Wheeler came back to the subject of solenoid inductance calculation in a short article⁵⁰ (this time in SI units) in which he reminds readers of the restricted shape-range of existing simple formulae and shows how to combine various long and short-coil expressions in order to remove such limitations.

Note that, in the discussion to follow, the mathematical notation differs from that used in Wheeler's 1982 paper, and the use of ℓ/D instead of ℓ/r has caused some of Wheeler's coefficients to change in value. The transformations involved however, are too trivial to warrant further discussion.

For his prototype long-coil formula Wheeler gives an expression equivalent to:

$$k_{w82.3} = \frac{1}{1 + \left(\frac{4}{3\pi}\right)(D/\ell)} \quad (\text{W82-3})$$

Notice here that the Maclaurin series for $1/(1-x)$ and $x < 1$ is:

$$\frac{1}{(1-x)} = 1 + x + x^2 + x^3 + \dots$$

Thus (W82-3) is related to the first term of the Webster-Havelock formula, i.e.;

$$k_{w82.3} = \frac{1}{1 + \left(\frac{4}{3\pi}\right)(D/\ell)} \approx 1 - \left(\frac{4}{3\pi}\right)(D/\ell)$$

More intriguingly however, it is in the same form as the expression for k_{w25} given earlier (i.e., the approximation for Nagaoka's coefficient extracted from Wheeler 1925). It therefore hints at the underlying deduction that led to Wheeler's most famous formula. Note however, that $4/3\pi = 0.4244$, and the use of this value in place of the original empirical coefficient reduces the accuracy of the approximation. The point however, is not to use this expression directly, but to obtain unrestricted range by combining it with others.

For his prototype short-coil expressions, he gives three examples; which can be written as follows:

$$k_{\text{RNT}} = \frac{2}{\pi}(\ell/D) \left[\ln(4D/\ell) - \frac{1}{2} \right] \quad (\text{W82-4.1})$$

which we have seen before as originating from the first term of the Rayleigh-Niven current sheet formula;

$$k_{w82.4.2} = \frac{2}{\pi}(\ell/D) \left[\ln\left(1 + \frac{\pi}{2}(D/\ell)\right) + \ln\left(\frac{8}{\pi}\right) - \frac{1}{2} \right] \quad (\text{W82-4.2})$$

and

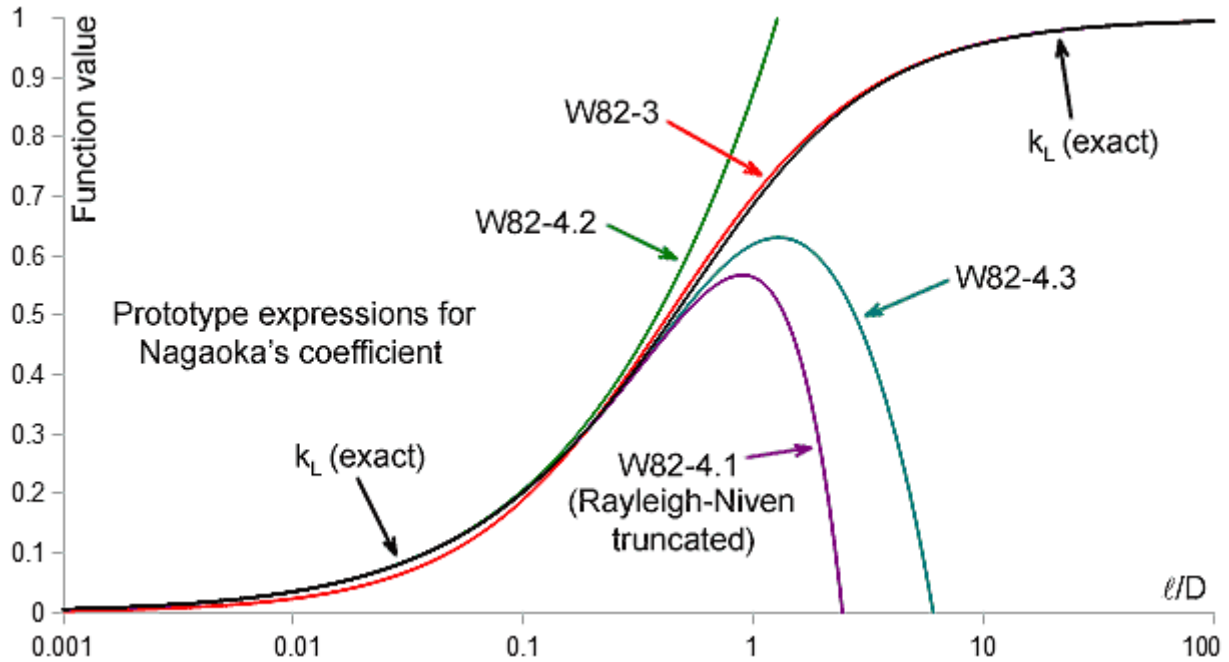
$$k_{w82.4.3} = \frac{2}{\pi}(\ell/D) \left[\operatorname{asinh}\left(\frac{\pi}{2}(D/\ell)\right) + \ln\left(\frac{4}{\pi}\right) - \frac{1}{2} \right] \quad (\text{W82-4.3})$$

⁵⁰ **Inductance Formulas for Circular and Square Coils.** Harold A Wheeler, Proc. IEEE (Letters), Vol 70, No 12, Dec 1982, p1449-1450.

Where asinh is the inverse hyperbolic sine operator, defined by the expression:

$$\operatorname{asinh}(x) = \ln\left(x + \sqrt{x^2 + 1}\right)$$

These curves are shown plotted below (on a logarithmic scale), with the exact curve for Nagaoka's coefficient shown for comparison (see spreadsheet [L_formulae.ods](#), sheet 4).



Wheeler's 1982 formulae (5) and (6)

It is by no means obvious how to combine the various prototype expressions, but as Wheeler points out; the first terms of the short-coil approximations (**W82-4.2**) and (**W82-4.3**) tend towards the correct value for long coils; i.e., using the substitution $z_k = \frac{2}{\pi}(\ell/D)$ to reduce the clutter:

$$\text{as } \ell/D \rightarrow \infty, \quad z_k \ln\left(1 + \frac{1}{z_k}\right) \rightarrow 1 \quad \text{and} \quad z_k \operatorname{asinh}\left(\frac{1}{z_k}\right) \rightarrow 1$$

(columns are given in the spreadsheet to demonstrate this point). Thus it is possible to combine these two terms with weighting coefficients chosen so that the correct offset is obtained for short coils; i.e.:

$$k_L \approx z_k \left[k_1 \ln\left(1 + \frac{1}{z_k}\right) + k_2 \operatorname{asinh}\left(\frac{1}{z_k}\right) \right] \quad (\mathbf{W82-5})$$

Wheeler gives $k_1 = 0.48$ and $k_2 = 0.52$ for an overall accuracy of $\pm 1.7\%$, and no restriction on solenoid shape. The empirical coefficients however can be improved slightly (at the expense of increasing the number of decimal places). The formula was therefore subjected to machine optimisation by comparison with the exact (AGM) calculation using a Nelder-Mead algorithm

written by Bob Weaver⁵¹ (see spreadsheet [L_formulae.ods](#), sheet 5). The following formula, accurate to within 1.6% results:

$k_L = z_k \left[0.53022 \ln \left(1 + \frac{1}{z_k} \right) + 0.47877 \operatorname{asinh} \left(\frac{1}{z_k} \right) \right] \quad \pm 1.6 \%$ <p>where $z_k = \frac{2}{\pi}(\ell/D)$</p>	<b style="color: red;">W82-5.1 (optimised)
--	---

Notice however, that due to the asymptotic behaviour of the individual terms, the empirical coefficients should strictly add-up to 1 if the formula is to be exact in the long-coil limit. Wheeler's coefficients are obviously intended to preserve that behaviour (indeed, he refers to the formula as a 'weighted average'); whereas the machine optimisation gives $k_1 + k_2 = 1.00899$. This means that the machine optimisation has sacrificed strict asymptotic behaviour for the sake of minimum runout (i.e., minimisation of the maximum error). This is, of course, a perfectly reasonable choice for a simple formula with no analytical pretensions; because most practical coils fall in the middle of the range.

If convergence with the exact expression in the long-coil limit is required; constraining the formula to have $k_2 = 1 - k_1$ gives minimum runout when $k_1 = 0.482$ and $k_2 = 0.518$ (i.e., very close to Wheeler's original values) the maximum error then being $\pm 1.63\%$. The asymptotic version (having effectively a single empirical parameter) is thus:

$k_L = z_k \left[k_1 \ln \left(1 + \frac{1}{z_k} \right) + (1 - k_1) \operatorname{asinh} \left(\frac{1}{z_k} \right) \right] \quad \pm 1.63 \%$ <p>where $z_k = \frac{2}{\pi}(\ell/D)$ and $k_1 = 0.482$</p>	<b style="color: red;">W82-5a (asymptotic, optimised)
--	--

Wheeler's next unrestricted formula (equation 6 in his paper) is based on the long-coil form (**W82-3**):

$k_L = z_k \left[\frac{2.78}{1.1 + 2\ell/D} + \ln(1 + 0.195D/\ell) \right] \quad \pm 0.91\%$	<b style="color: red;">W82-6
---	----------------------------------

With three empirical coefficients, this gives an accuracy of $\pm 0.91\%$ as written, but the maximum % error is not symmetric above and below zero and so, once again, further optimisation is possible. Also, since optimisation will increase the number of decimal places required in the parameters, we might as well start by factoring 1.1 from the denominator of the first term. Thus:

$$k_L \approx z_k \left[\frac{k_3}{1 + k_4 \ell/D} + \ln(1 + k_5 D/\ell) \right]$$

where, for initial fitting values:

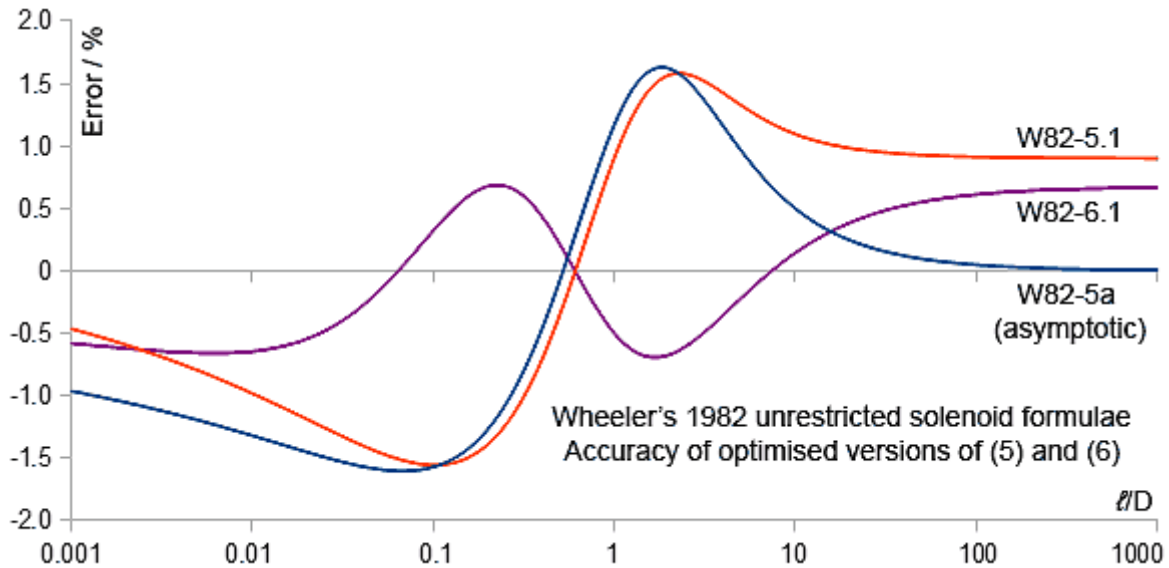
⁵¹ See: **Optimisation of multi-parameter empirical fitting functions:** <http://g3ynh.info/zdocs/math/>
+ Nelder-Mead demonstration (spreadsheet). Also: <http://electronbunker.ca/CalcMethods3b.html>.

$k_3 = 2.78 / 1.1 = 2.5273$, $k_4 = 2 / 1.1 = 1.8182$ and $k_5 = 0.195$ as before.

After machine optimisation (see spreadsheet: [L_formulae.ods](#), sheet 6), the result is:

$k_L = z_k \left[\frac{2.666}{1 + 1.877 \ell/D} + \ln(1 + 0.161 D/\ell) \right] \pm 0.7\%$	W82-6.1 (optimised)
where $z_k = \frac{2}{\pi}(\ell/D)$	

The error curves for the three optimised formulae (**W82-5a**), (**W82-5.1**) and (**W82-6.1**) are shown below:



What is particularly noticeable about these curves is that the error functions for formulae derived from (**W82-5**) and (**W82-6**) are broadly complimentary. This means that a new formula with greatly reduced error can be obtained by taking a weighted average. If we also re-float all of the empirical parameters during machine optimisation, a trading process that will reduce the error still further will occur. These observations lead to the following candidate expression:

$$k_L \approx z_k \left[\frac{k_1}{1 + k_2 z_k} + k_3 \ln \left(1 + \frac{k_4}{z_k} \right) + k_5 \ln \left(1 + \frac{1}{z_k} \right) + k_6 \operatorname{ashinh} \left(\frac{1}{z_k} \right) \right]$$

Notice however that the parameters k_3 and k_5 are likely to be strongly correlated in some regions of the parameter space, because the terms that they control can become the same. This implies that it may be possible to set one of them to zero with minimal penalty, which means that the formula might drop a term.

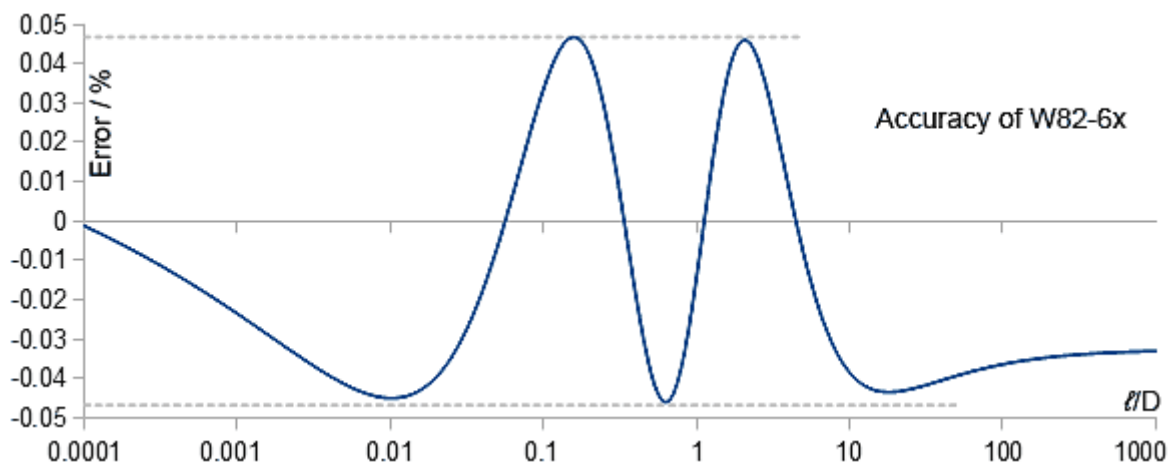
When major parameter correlations exist, the non-linear fitting process can become ill-conditioned. In such cases it is often necessary to apply constraints on parameter variation until the system begins to approach a candidate solution. For the problem here, an initial constraint was applied equivalent to forcing the asymptotic variant of the part corresponding to (**W82-5**); i.e.:

$$k_5 = k_7 k_8 \quad \text{and} \quad k_6 = k_7 (1 - k_8)$$

In this way, k_7 becomes a weighting coefficient for the whole of (W82-5a) and the correlation between k_5 and k_3 is eliminated. Thus the fitting proceeded without mishap and a good preliminary solution was found. Starting values for k_5 and k_6 were then calculated from the expressions above, and these parameters were allowed to vary. After that, as occurs when there is a correlation, the error function was found to have a very flat bottom, with a process of trading between k_3 and k_5 occurring without much improvement. k_5 was also noted to pass through zero at a point close to the error minimum. Hence it was set to zero and excluded from the fit. Finally, it was noticed that the formula obtained was being slow to reach the asymptote for small ℓ/D . This problem was overcome by including a single extra point in the fit at $\ell/D = 10^{-6}$, with a fitting weight of 10^6 (all of the other points having a weight of 1). This brought the overall error up a little, but guarantees the maximum stated runout (for details, see the spreadsheet [L_formulae.ods](#), sheet 8). The formula obtained, accurate to within 470 ppM, is shown below. Since it is effectively an extended version of (W82-6) it is designated (W82-6x).

$k_L = z_k \left[\frac{1.67405}{1 + 2.5382 z_k} + 0.8053 \ln \left(1 + \frac{0.1796}{z_k} \right) + 0.1955 \operatorname{ashinh} \left(\frac{1}{z_k} \right) \right] \pm 0.05\%$	W82-6x
where $z_k = \frac{2}{\pi}(\ell/D)$	

The error curve is shown below, and a Basic macro function for evaluating the formula is given after that.




```

Function W82eq6x(byVal x as double) as double
'Calculates Nagaoka's coeff. using optimised extended version of Wheeler 1982 eqn (6).
' Accuracy is +/- 470ppM. x = diam / length.
if x = 0 then
w82eq6x = 1
else
Dim u as double, v as double, w as double, zk as double
zk = 2/(pi*x)
u = 1.67405 / (1 + 2.5382*zk)
v = 0.8053*log(1 + 0.1796/zk )
w = 0.1955*log( 1/zk + sqr( 1/(zk*zk) + 1) )
w82eq6x = zk*(u + v + w)
end if
end function

```

Wheeler's 1982 formula (7)

The most accurate expression for solenoid inductance in Wheeler's 1982 paper is given as his equation (7). This is based on the short-coil form (**W82-4.2**), with the constant term replaced by the reciprocal of a polynomial. Expressed using ℓ/D as the argument, the formula is:

$k_L \approx z_k \left[\ln \left(1 + \frac{1}{z_k} \right) + \frac{1}{k_0 + k_1 (\ell/D) + k_2 (\ell/D)^2} \right]$ <p>where: $z_k = \frac{2}{\pi} (\ell/D)$, $k_0 = \frac{1}{\ln \left(\frac{8}{\pi} \right) - \frac{1}{2}} = 2.3004$, for k_1 see text ,</p> <p>and $k_2 = \frac{24}{3\pi^2 - 16} = 1.7636$</p>	W82-7
--	--------------

Note that with k_0 as given, the second term $\rightarrow \ln(8/\pi) - 1/2$ in the short-coil limit and the formula becomes convergent with (**W82-4.2**). k_0 therefore preserves the short coil asymptotic behaviour. Also, as noted earlier, $z_k \ln(1 + 1/z_k) \rightarrow 1$ in the long coil limit, and of course, the second term $\rightarrow 0$ when ℓ/D becomes large. Thus the formula is doubly asymptotic.

k_1 is the only empirical coefficient (in the original version at least), and Wheeler gave it the value 3.2, which gives a maximum overall error of 0.09%. Wheeler however, evidently had a liking for empirical constants with few decimal places; because this choice gives an error function that is always positive. A simple adjustment to $k_1 = 3.244$ distributes the maximum error equally above and below zero and gives an accuracy of ± 469 ppM (0.05%).

We should also note, of course, that although the value of k_0 is constrained if the asymptotic behaviour of the function is to be preserved, there is nothing particularly sacrosanct about k_2 . The enumeration of natural constants is also laborious if current-sheet inductance is to be evaluated using a hand calculator, and so k_2 might as well be replaced by an empirical constant. With that in mind, the formula was subjected to machine optimisation (using Bob Weaver's Nelder-Mead engine

as before) and the maximum error promptly fell to 265ppM ($\pm < 0.03\%$). The coefficients obtained are shown in the table below (**W82-7.2**).

It is also arguable that, since most coils fall in the middle of the ℓ/D range, there is not much point in preserving strict asymptotic behaviour. That view was put forward by Rodger Rosenbaum, who gave hand-tweaked coefficient values that reduce the error to 231ppM. A subsequent machine optimisation then reduced the error to 163ppM (**W82-7.3**), with coefficients as shown in the table.

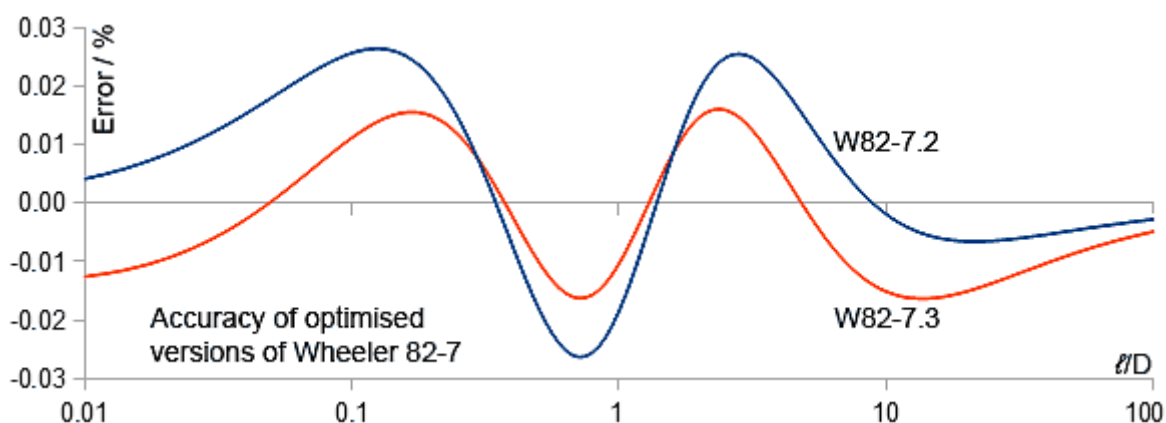
k_0	k_1	k_2	error / ppM		Source
$1/\{\ln(8/\pi)^{-1/2}\}$	3.2	$24/(3\pi^2 - 16)$	+879, -0	(W82-7)	Wheeler 1982
$1/\{\ln(8/\pi)^{-1/2}\}$	3.244	$24/(3\pi^2 - 16)$	± 469	(W82-7.1)	DWK 2007
$1/\{\ln(8/\pi)^{-1/2}\}$	3.2219	1.7793	± 265	(W82-7.2)	This work
2.303	3.213	1.784	± 231	(W82-7R)*	Rosenbaum ⁵² 2009
2.3056	3.2009	1.7904	± 163	(W82-7.3)*	This work

* Deviation from the analytical value of k_0 causes slow convergence with the short-coil asymptote.

The setup for performing the optimisations discussed can be examined in the spreadsheet **L-functions.ods** sheet 9.

Note, incidentally, that there is little to be gained by extending the polynomial to have terms of $(\ell/D)^3$ etc.. The reason is that most of the error occurs in the middle of the argument range, and so high-power terms end up with small coefficients and have minimal overall effect.

Thus the best optimisations of Wheeler's 1982 formula (7) (adherent to the original form) are (**W82-7.2**) if doubly-asymptotic behaviour is required, and (**W82-7.3**) otherwise. The error curves are shown in the graph below, and a Basic function that will calculate either according to the setting of a flag is listed on the following page.



```

Function W827opt(byVal x as double, a as integer) as double
' calculates Nagaoka's coeff. using optimised Wheeler 1982 eqn (7).
' x = Diam. / length
' a = 0 for asymptotic (+/-265ppM) version, otherwise gives min. runout (+/-163ppM).
if x = 0 then
  w827opt = 1
else
  Dim k0 as double, k1 as double, k2 as double, zk as double, p as double
  if a = 0 then
    k0 = 1/(log(8/pi)-0.5)
    k1 = 3.2219
    k2 = 1.7793
  else
    k0 = 2.3056
    k1 = 3.2009
    k2 = 1.7904
  end if
  zk = 2/(pi*x)
  p = k0 + k1/x + k2/(x*x)
  w827opt = zk*(log(1 + 1/zk) + 1/p)
end if
end function

```

The setup used for performing the optimisations is given in the spreadsheet [L-functions.ods](#) sheet 9, and the routine above, which was used to produce the graphs, is in the included macro library.

8e. Weaver's continuous formula

An error of < 270ppM from a compact continuous doubly-asymptotic expression, as obtainable using the original form of Wheeler 1982-7, is remarkable enough; but in early 2010 Bob Weaver decided to see if there could be any improvement⁵³. As we have noted above, there is little advantage in adding high-power terms to the inverted polynomial, and so he considered using terms of fractional power. In fact, if terms of $(\ell/D)^{1/2}$, $(\ell/D)^{3/2}$ and $(\ell/D)^{5/2}$ are added, then the error can be reduced to about 91 ppM ; but after some experimentation he was able to reduce the error still further by adopting the following form:

$k_L = z_k \left[\ln \left(1 + \frac{1}{z_k} \right) + \frac{1}{k_0 + k_1(\ell/D) + k_2(\ell/D)^2 + \frac{w_1}{(w_2 + D/\ell)^v}} \right]$	W82-7W
<p>where: $z_k = \frac{2}{\pi}(\ell/D)$, $k_0 = \frac{1}{\ln\left(\frac{8}{\pi}\right) - \frac{1}{2}} = 2.30038$,</p>	
<p>$k_2 = \frac{24}{3\pi^2 - 16} = 1.76356$, and for other coefficients, see below.</p>	

Notice that the coefficient w_2 is shown as a modulus. This is because it must be constrained to be positive during machine optimisation, otherwise the variable power term can become complex and cause the process to crash. k_0 is set to Wheeler's analytical value in order to preserve the formula's asymptotic behaviour. Also it transpires that, if all of the parameters (apart from k_0) are floated, then the value for k_2 comes out so close to the analytical value that it makes no significant difference if it is set to the analytical value. Thus the formula ends up with four empirical parameters, and the following values are found to reduce the maximum error to ± 20.5 ppM :

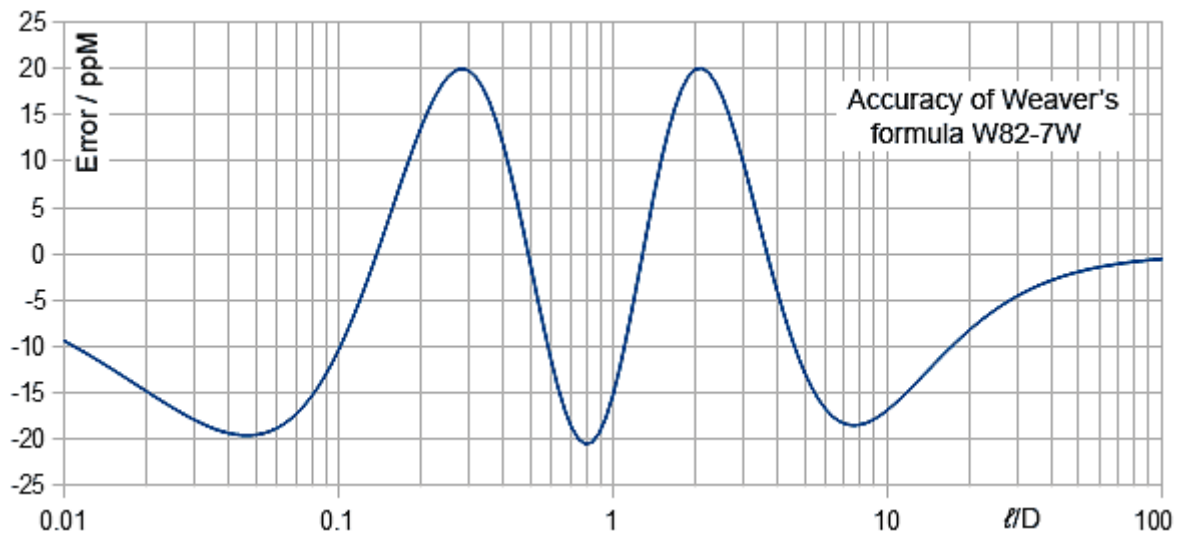
$k_1 = 3.437$, $w_1 = -0.47$, $w_2 = 0.755$, $v = 1.44$.	W82-7W fitting parameters
--	---------------------------

This formula is a serious rival for Lundin's formula because, although it is not quite as accurate, it has the advantage of being continuous, and ± 21 ppM is still, for all practical purposes, exact. The accuracy is maintained, incidentally, if the 5 decimal-place values for k_0 and k_2 are used in lieu of evaluating them from natural constants.

(Continued on next page →)

⁵³ **Site update - at last**, Bob weaver. Private e-mail correspondence, 3rd may 2010. Also discussed at <http://electronbunker.ca/CalcMethods3b.html> .

The error curve is shown below, and a verified Basic macro is given below that.



```
Function W82W(byVal x as double) as double
' calculates Nagaoka's coeff. using Wheeler 1982 eqn (7) as modified by
' Bob Weaver. Max error is +/- 21ppM. x = Diam. / length
if x = 0 then
  W82W = 1
else
  Dim zk as double, k0 as double, k2 as double, p as double, w as double
  zk = 2/(pi*x)
  k0 = 1/(log(8/pi)-0.5)
  k2 = 24/(3*pi*pi -16)
  w = -0.47/(0.755 + x)^1.44
  p = k0 + 3.437/x + k2/(x*x) + w
  W82W = zk*(log(1 + 1/zk) + 1/p)
end if
end function
```

For the sake of completeness, it is worth noting that the maximum runout error in Bob's formula can be reduced very slightly by increasing the number of decimal places in the empirical coefficients, and reduced a little more by floating k_0 (which sacrifices the asymptotic behaviour). The error using the coefficient set presented above however is extremely small in comparison to other sources of inaccuracy in inductance modelling, and so we will not pursue the matter here.

Necessary accuracy

The circumstances that strictly require accuracy better than about $\pm 0.1\%$ in the calculation of the Lorenz current-sheet inductance are likely to be extremely rare; but it is so easily obtained that we might as well avail ourselves of it as a matter of course. The first 0.1% empirical formula was given by Wheeler in 1982, and Lundin followed with a 3 ppM formula in 1985. Consequently, there is no conceivable excuse for using Wheeler's 1925 formula in so-called 'online inductance calculators' (or any other computer program for that matter).

9. A note on the calculation of current-sheet inductance

Note that many of the formulae given above for k_L are in the form:

$$k_L = \frac{2}{\pi}(\ell/D) \text{ [expression]}$$

This can also be written:

$$k_L = z_k \text{ [expression]}$$

If we call the expression in square brackets k'_L , then we have:

$$k_L = \frac{2}{\pi}(\ell/D) k'_L = z_k k'_L$$

Now turning our attention to the general form of the current sheet solenoid equation (7.1) we have:

$$\mu \frac{\pi r^2}{\ell} N^2 k_L \quad , \quad \text{[permeability } \times \text{ area / length]}$$

which, using $D=2r$, can be rearranged:

$$L_s = \mu r N^2 \left(\frac{\pi}{2}\right) (D/\ell) k_L = \frac{\mu r N^2 k_L}{z_k}$$

Substituting for k_L we have:

$$L_s \frac{\mu r N^2 z_k k'_L}{z_k}$$

i.e.,

$L_s = \mu r N^2 k'_L$	[permeability \times length]	9.1
where $k'_L = \frac{\pi}{2}(D/\ell) k_L$		

The point is that the factor $\frac{2}{\pi}(\ell/D)$ does not always have to be evaluated explicitly when calculating current-sheet inductance. It appears when the inductance is expressed using Nagaoka's coefficient because Nagaoka took the view that solenoid inductance is related to area divided by length (as is capacitance). In other words, Nagaoka in 1909, as we do now, preferred formulae that can be easily parsed to ensure that they make sense physically. Those who worked or began their careers in the 19th Century, Maxwell included, took the more fundamental view that inductance, in general, has units of length.

One last thing to note on this issue is that, for many of the functions given earlier, $z_k = \frac{2}{\pi}(\ell/D)$ would actually have been a more natural choice of argument than ℓ/D .

10. Rosa's round-wire corrections and the summation method

Apart from the inclusion of internal inductance; all of the formulae given so far have been aimed at determining the inductance of a coil from the inductance of the equivalent Lorenz current-sheet. Practical coils however, no matter how good the choice of effective diameter, do not conform exactly to the Lorenz model. This means that additional corrections are required; and in particular it is necessary to allow for the use of physically realistic wire. The so-called 'round wire corrections' were given for the low-frequency case by E B Rosa of the American National Bureau of Standards in 1906⁵⁴ and result in an extended expression for solenoid inductance that takes the form:

$L = L_s - \mu r N (k_s + k_m)$ [Henrys]	10.1
--	-------------

where L_s is the current-sheet inductance as given by equation (7.1); k_s is a dimensionless correction coefficient for the difference between the self-inductance of a round-wire loop and that of a single-turn current sheet; and k_m is a dimensionless correction coefficient for the difference in the total mutual inductance of a set of round-wire loops as compared to that of a set of current-sheet loops. In Grover's 1946 monograph⁵⁵, k_s and k_m are called G and H ; and in original NBS papers and elsewhere they are called A and B . Since all of the letters used in those early publications have other preferred meanings in a modern electromagnetic context, the notation has been altered here.

The Rosa coefficients (actually approximations, as we will see) are tabulated in the publications mentioned above and a simple approximate function for k_s , including DC internal inductance, is given. In a modern computer-based modelling context however, the use of approximations is to be avoided wherever possible; and we require functions, not tables. Internal inductance also needs to be treated separately so that it can be allowed to vary with frequency. We will therefore re-investigate the Rosa corrections in detail with a view to updating them. To that end; we can begin by noting that they are derived from the difference between the current-sheet inductance and the the round-wire solenoid inductance as obtained by summation of self and mutual inductances obtained using Maxwell's method of Geometric Mean Distances. An introduction to the Maxwell approach is therefore appropriate at this point.

10a. Geometric Mean Distances

The problem of calculating the inductance of an electrical circuit composed of an arbitrary system of conductors is the same as that of calculating all of the self and mutual inductances in the system and (bearing in mind that mutual inductance can be positive or negative) adding them all together. The problem of determining all of the individual inductance contributions moreover, was shown by Maxwell to be reducible in each instance to the problem of calculating the mutual inductance of a pair of equivalent infinitesimal filaments (i.e., conductors of finite length but zero cross-sectional area) spaced at the Geometric Mean Distance (GMD) for the corresponding conductor or conductor pair.

The geometric mean of two numbers is simply the square-root of their product. Similarly; the geometric mean of m numbers is the m^{th} root of their product. Thus, if we wanted to find the approximate GMD between two conductors lying side-by-side, we could define a representative set of points within the cross-section of each conductor, measure the distance between each point in one conductor and every point in the other to obtain a total of (say) m distances, then take the m^{th}

54 E B Rosa, BBS Vol 2 (2), 1906, p161-187 [BS Sci. 31]. See also; BS Sci. 169, p122.

55 Grover 1946, p149

root of their product. To obtain the exact GMD, of course, we would need to have an infinite number of points in each conductor; but that problem can be surmounted by taking logarithms to turn the product into a summation, so that the summation becomes a set of nested integrals as the number of points in each conductor goes to infinity.

The GMD of a conductor from itself is similarly given by taking the distance from every point within the conductor cross-section to every other point enclosed by the same area. That distance is never zero for a finite cross-section (even if it is only one-dimensional, like a current sheet), and indeed, there would be no such thing as the self-inductance of an isolated linear conductor if it was.

The subject of GMD determination is discussed by Maxwell in his Treatise on E&M⁵⁶, and was further expanded by Andrew Gray⁵⁷ in his magnum opus 'Absolute Measurements in E&M'. Brief discussion and numerous standard solutions can also be found in BS Science Paper 169⁵⁸ and in Grover's monograph⁵⁹. For an excellent modern introduction to the GMD concept, with detailed derivations of most of the important results, see 'Inductance' by Clayton Paul⁶⁰.

Note that in early articles, GMD is usually given the symbol R . In modern electrical documents however, the use of the capital letter R for anything not directly-related to resistance is likely to cause confusion. Clayton Paul gets around that problem by using D , but in this article we have already allocated D for diameter. Hence we allocate lower-case g and commend it as the initial letter of the concept to which it relates.

The determination of all of the self and mutual inductances in a system using the GMD approach, and the subsequent addition of the various contributions to obtain the total inductance is known as the summation method⁶¹. Given standard solutions for the self and mutual GMDs of commonly-encountered conductor cross-sections, and standard formulae for the mutual inductances of parallel filaments; the problem of inductance calculation for structures composed of parallel conductors (such as coils and cables) becomes conceptually straightforward. When the number of parallel filaments is large however, as it often is in coils; the summation method is computationally inefficient. That inefficiency is a drawback even in the age of electronic computers; but in the 19th and early 20th Centuries, it rendered the method largely impractical.

Rosa's insight into the computation problem was to note that, given solutions for the GMDs of current-sheet segments, the inductance of a current-sheet solenoid can be obtained by both the summation and the Lorenz methods. The Lorenz method is highly efficient by comparison; but it provides only a first approximation for the inductance of a practical coil and, being based on a model that is radically unlike any actual coil, cannot be directly adapted for accurate calculation. Rosa therefore decided to analyse the differences between the round-wire solenoid calculation and the current-sheet calculation when using the summation method for both. The point was to obtain corrections that could be backwardly applied to the Lorenz formula. The result was a pair of corrections; the reason being that there are distinct analytical differences between the mutual inductance and self-inductance parts of the problem. Rosa thereby obtained a method of round-wire solenoid inductance calculation that retains the efficiency of the Lorenz method, but is practically equivalent to the summation method.

56 **A Treatise on Electricity and Magnetism, Vol 2**, James Clerk Maxwell, 3rd edition 1892.

[Maxwell E&M, Vol 2] OUP reprint 2002. ISBN 0198503741.

On the GMDs of two figures in a plane. Articles [691] and [692], pages 324-328.

57 **Absolute measurements in electricity and magnetism**, Andrew Gray, 2nd edition, rewritten & extended (in single volume), Macmillan 1921. (<http://archive.org/details/absolutemeasrem00grayuoft>) [**Gray 1921**] Chapter XIII, Calculation of Inductances. page 475 - .

58 **BS Sci. 169**, Section 9. Formulas for geometrical and arithmetical mean distances. p166 - 170.

59 **Grover 1946**. Ch 2. The GMD method, p. 14-16. Ch.3. GMDs. p. 17-25.

60 **Inductance, Loop and Partial**, Clayton R Paul. Wiley 2010. ISBN 978-0-470-46188-4. Section 6.4. Concept of Geometric Mean Distance, p 266-272; 6.4.1. GMD between a shape and itself and the self partial inductance, p273 - 285; 6.4.2. GMD and mutual partial inductance between two shapes, p 285 - 291.

61 **BS Sci. 169**, The summation formula for L , p123. (apparently attributable to Kirchhoff).

10b. Loop self and mutual inductance formulae

The use of the summation method for solenoid inductance calculation is a matter of repeated application of loop inductance formulae. Such formulae, of course, are also useful in their own right.

The self inductance of a circular loop is the same as the mutual inductance between the loop and itself. That can be found by calculating the mutual inductance of a pair of filaments separated by the geometric mean distance (GMD) of the conductor from itself. For the construction of coils using ordinary round magnet wire, the quantities of interest are⁶²:

For a solid conductor of circular cross section: $g = r_w \exp(-1/4) = 0.7788 r_w$

For a thin-walled cylinder: $g = r_w \exp(0) = r_w$

The reason we are interested in these two specific cases, is that the first corresponds to a wire with uniform current throughout its cross section, which means that its self-inductance will include low-frequency internal inductance. The second corresponds to a wire with current flowing only at the surface, which means that its self-inductance will be without internal inductance. Thus we can control the inductance calculation process, using the exponent argument of $-1/4$ to compare results with existing methods and tables, but changing to 0 so that the properly-calculated internal inductance can be added separately.

Maxwell gives an expression for the mutual inductance of two coaxial circles^{63 64} in complete elliptic integrals. Converted to rationalised mks this becomes:

$$M = -\mu \sqrt{r_1 r_2} \left[\left(\kappa - \frac{2}{\kappa} \right) K(\kappa) + \left(\frac{2}{\kappa} \right) E(\kappa) \right] \quad [\text{ Henrys}] \quad (\mathbf{M701.1})$$

where r_1 and r_2 are the radii of the circles and s is the axial separation. The elliptic integral modulus κ (kappa) is given by:

$$\kappa = 2 \sqrt{\frac{r_1 r_2}{(r_1 + r_2)^2 + s^2}}$$

A slightly less opaque version is obtained by rearranging terms to get rid of the forward minus sign and factoring-out $2/\kappa$:

$$M = \mu \sqrt{(r_1 + r_2)^2 + s^2} \left[\left(1 - \frac{\kappa^2}{2} \right) K(\kappa) - E(\kappa) \right] \quad [\text{ Henrys}] \quad (\mathbf{M701.1a})$$

This can be used to obtain a well-known approximation for wire-loop inductance, a derivation for which is given by Ramo et al⁶⁵. Firstly; note that we can calculate the external self-inductance by placing two notional filaments at a separation of r_w , that being the radius of the wire. We can do that straightforwardly either by placing two circles of the same radius at an axial separation of r_w ,

62 See for example **Maxwell E&M Vol 2**. Article [692] (9) page 328.

63 **Maxwell E&M, Vol 2**, Article [701], To find M by elliptic integrals, page 339. Note that the elliptic integral of the first kind, nowadays called K, is written as F in Maxwell's notation.

64 **BS Sci. 169**, p6. Rosa also states that the formula is absolute, giving M for coaxial circles at any distance.

65 **Ramo et al. 1994**, 'Self inductance of a circular loop through mutual inductance concepts'. Example 4.7b, p192 - 193. The same derivation is also given in the 1st edition: Wiley, 1965, LCCN 165-19477. section 5.25, 'Self-inductance by selected mutual inductance' pages 309 - 311.

or we can set the axial distance to zero and place one circle inside the other. The two choices give slightly different results, the concentric filament pair giving slightly less inductance than the parallel pair when we use the average radius and the inner radius, and slightly more when we use the average radius and the outer radius. The parallel filament choice is the most realistic; but the difference is supposedly not great when $r \gg r_w$, and the concentric choice gives a considerable simplification.

If we put: $s = 0$, $r_1 = r$ and $r_2 = r - r_w$, we get:

$$L_{1x} = \mu (2r - r_w) \left[\left(1 - \frac{\kappa^2}{2} \right) K(\kappa) - E(\kappa) \right] \quad [\text{Henrys}] \quad (10.2)$$

with

$$\kappa = \frac{2\sqrt{r(r-r_w)}}{2r-r_w}$$

We can, of course, now calculate the loop external inductance using complete-elliptic-integral tables or functions, but as pointed out by Ramo et al:

when $r \gg r_w$, $\kappa \rightarrow 1$ and $E(\kappa) \rightarrow 1$

Thus:

$$L_{1x} \approx \mu r [K(\kappa) - 2]$$

Also, as $\kappa \rightarrow 1$;

$$K(\kappa) \approx \ln \left(\frac{4}{\sqrt{1-\kappa^2}} \right)$$

Substituting for κ^2 and rearranging, we get:

$$K(\kappa) \approx \ln \left(\frac{4(2r-r_w)}{r_w} \right)$$

which, since $r \gg r_w$, can be further approximated as:

$$K(\kappa) \approx \ln \left(\frac{8r}{r_w} \right)$$

Hence:

$L_{1x} \approx \mu r \left[\ln \left(\frac{8r}{r_w} \right) - 2 \right]$	[Henrys]	External inductance of a round-wire loop (approximate)	10.3
---	----------	--	-------------

The expression above is widely known and used. It was given by Maxwell, not as a loop-inductance formula, but as a first approximation for the mutual inductance of parallel coplanar circular filaments in close proximity⁶⁶. In BS Science paper 169, the same expression is obtained by rearrangement and truncation of a Maxwell series formula for the mutual inductance of coaxial circles⁶⁷ (and there are other routes⁶⁸). Hartshorn also gives this formula without derivation⁶⁹. It is alleged to be accurate for $r > 5r_w$. Note incidentally, that it corresponds to a truncated version of a formula for short round-wire coils given by Rayleigh and Niven⁷⁰. The full Rayleigh-Niven formula, excluding internal inductance, is:

$$L_x \approx \mu r N^2 \left[\ln \left(\frac{8r}{r_w} \right) - 2 + \frac{1}{8} \left(\frac{r_w}{r} \right)^2 \left(\ln \left(\frac{8r}{r_w} \right) + \frac{1}{3} \right) \right] \quad \mathbf{10.3a}$$

For the accurate calculation of loop mutual or self inductance however, we need to evaluate the complete elliptic integrals properly. The form given above (**M701.1a**) will work perfectly well, but as mentioned earlier (in **section 7**), best computational efficiency is obtained by using the K-E combination elliptic integral and calculating it by the AGM method. A formula suitable for that approach was also given by Maxwell⁷¹:

$$M = 2\mu\sqrt{r_1 r_2} \frac{1}{\sqrt{\kappa_1}} [K(\kappa_1) - E(\kappa_1)] \quad [\text{Henry}] \quad \mathbf{(M701.2)}$$

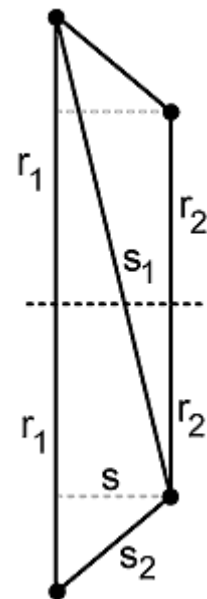
where

$$\kappa_1 = \frac{s_1 - s_2}{s_1 + s_2}$$

s_1 and s_2 being the furthest and nearest distances between the filaments, as defined in the cross-sectional diagram on the right. Hence, using Pythagoras:

$$s_1 = \sqrt{(r_1 + r_2)^2 + s^2} \quad \text{and} \quad s_2 = \sqrt{(r_1 - r_2)^2 + s^2}$$

A Basic macro function that performs the calculation has been written by Bob Weaver⁷²; although for the work described here, a version of Bob's program routine that has been trivially modified to accept input in mm and return M in nH is used.



As mentioned above, we can use the GMD to find out how far a wire is from itself for the purpose

66 Maxwell **E&M**, Vol 2, Article [704], p342-343.

67 BS Sci. 169, p13, formula [11]. A greatly extended version is also given by Coffin, p14, formula [13].

68 See, for example: Gray 1921, pages 490 - 492.

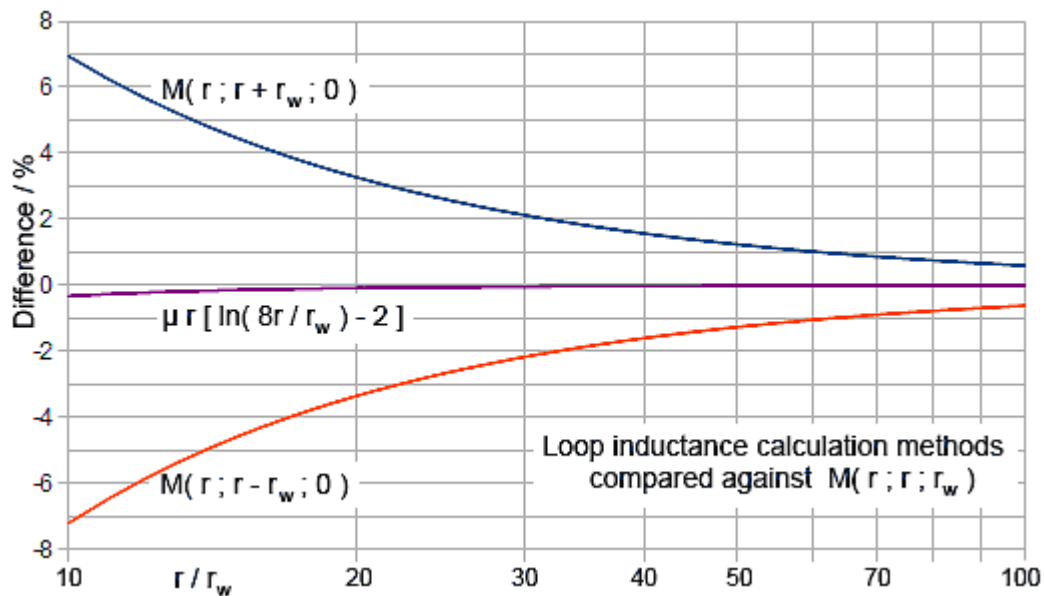
69 **Radio-Frequency Measurements by Bridge and Resonance Methods**, L. Hartshorn (Principal Scientific Officer, British National Physical Laboratory), Chapman & Hall, 1940 (Vol. X of "Monographs on Electrical Engineering", ed. H P Young). 3rd imp. 1942. [Hartshorn 1940]. Ch VIII, section 3, p146.

70 Rayleigh, **Scientific Papers Vol 2**, page 15, formula (13). Also given in BS Sci. 169, p111, formula (63). A version appearing in BS RP90, p166 formula (2), has a typographic error. The second ρ^2/a^2 (i.e.; r_w^2/r^2) term should be positive. Note that in all sources: $-7/2 = -2 + 1/4$. To remove the internal inductance component, change the term to -2 (i.e.; subtract $1/4$ from the first term in the power series). The formula requires correction for insulating space if used for multi-turn coils, but it is accurate to <1 part in 10^3 for 1-turn loops of ρ/a (i.e., r_w/r) ≤ 0.1 .

71 Maxwell **E&M**, Vol 2, Article [701], Second expression for M, page 340.

72 <http://electronbunker.ca/CalcMethods1b.html>

of self-inductance calculation, but we have a choice regarding how to arrange a pair of notional filaments separated by the GMD. Neglecting the effects of non-uniform current distribution, the average self-inductance is obtained by taking two filaments at the average coil radius and making the axial separation equal to g (i.e., the GMD). If we regard M as a function: $M(r_1; r_2; s)$, then this is the $M(r; r; g)$ choice. There are, of course, an infinite number of choices that give a distance of g between the filaments, but the extremes are: $M(r; r+g; 0)$ and $M(r; r-g; 0)$. The percentage difference between the extreme choices and the average choice is shown below plotted against r/r_w . Also shown for comparison, is a calculation using the approximate formula (10.3). Note that internal inductance is not considered in these calculations, and so $g = r_w$. The calculations are given in the spreadsheet [Loop_funcs.ods](#), sheet 1.



The difference between the upper and lower curves and the average is substantial (ca. $\pm 7\%$) when $r/r_w = 10$, but even at $r/r_w = 100$ it remains at $\pm 0.6\%$. Note also that most radio coils use relatively thick wire and so tend to fall in the middle of the range covered by this graph. Thus the choice of definition is critical for accurate results; and making an inappropriate choice is a potential source of systematic error.

Now recall that we apparently obtained the formula (10.2) from Maxwell's first elliptic integral expression (M701.1a) by placing the filaments at the average radius and the inner radius, with zero axial displacement. Not shown on the graph (but shown in the spreadsheet) is the fact that (10.2) evaluated by calling separate complete elliptic integral functions produces results that are exactly in agreement with $M(r; r-r_w; 0)$. This confirms that there are no approximations involved in the transformation⁷³ between Maxwell's two formulae (M701.1) and (M701.2).

The reason for performing these comparisons however, is that Ramo et al have effectively identified the ubiquitous formula (10.3) as originating from the $M(r; r-r_w; 0)$ case. Since the average choice $M(r; r; g)$ is the correct one, this would make (10.3) extremely suspect. We find however, that the simple formula gives a result very close to $M(r; r; r_w)$. In fact it is remarkably good, being asymptotically correct as $r/r_w \rightarrow \infty$, and exhibiting only -0.32% deviation from the $r_1 = r_2$ elliptic integral calculation at $r/r_w = 10$ and -1.5% at $r/r_w = 5$. This suggests that the derivation given by Ramo et al is an invention to fit the facts, rather than a legitimate route to the result. Certainly, it is extraordinary that the crude process of crossing-out small terms only serves to

⁷³ In **BS RP16**, on p489, Grover points-out that Landen's transformation is used to get from M701.1 to M701.2 and gives refs.. The transformation is also described in the article at: http://en.wikipedia.org/wiki/Landen's_transformation

improve the accuracy.

It is, incidentally, perfectly reasonable to use filaments with zero axial displacement when calculating the self-inductance of a loop; but not in the way that Ramo et al envisaged it. As has been pointed out by Bob Weaver⁷⁴; the correct approach when using concentric filaments is to define them so that the *geometric mean* of their radii is equal to the mean radius of the loop. Thus; if the filament radii are r_a and r_b and the loop radius is r :

$$r = \sqrt{r_a r_b}$$

The filament separation is, of course, the GMD; and so, if r_a is the radius of the inner filament:

$$g = r_b - r_a$$

Using the second expression to substitute for r_b in the first expression gives:

$$r^2 = r_a (r_a + g)$$

i.e.;

$$r_a^2 - r_a g - r^2 = 0$$

Applying the standard solution for quadratic equations we get:

$$r_a = \frac{-g \pm \sqrt{g^2 + 4r^2}}{2}$$

In this case, since the square root term is larger than g , the solution with the positive square root is the appropriate one. Hence:

$$r_a = \frac{\sqrt{g^2 + 4r^2} - g}{2}$$

and

$$r_b = r_a + g$$

When these expressions for the filament radii are used for the concentric filament case, the calculated loop mutual inductance is exactly the same as that obtained using filaments of equal radius and an axial displacement of g , i.e.;

$$M(r_a; r_b; 0) = M(r; r; g)$$

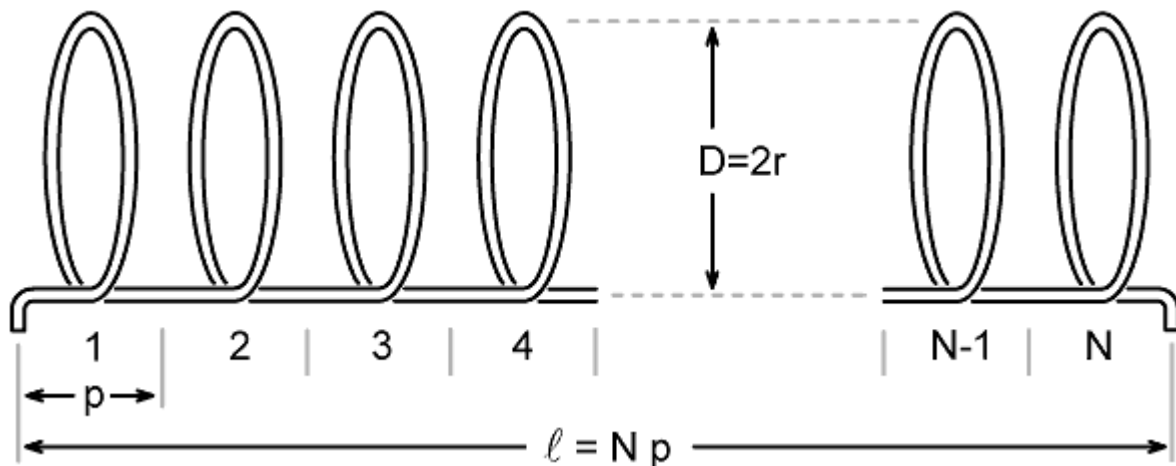
when r_a and r_b are as defined above.

74 GMD and wire loop inductance. Bob Weaver. Private e-mail communication, 25th Aug 2012.

10c. Solenoid inductance by the summation method

The summation method entails modelling the solenoid as a set of N wire rings (where N is the number of turns) and adding-up all of the ring self-inductances and inter-ring mutual inductances. For the purposes of the model, each ring lies in a plane perpendicular to the coil axis, is centred on the coil axis, and is positioned at the middle of the pitch interval for the turn to which it corresponds (so that the relationship $\ell = N p$ is preserved).

A structure that gives something approaching a physical realisation is depicted below. The actual model however is an abstraction. All rings carry the same radial current in the same direction, but the mechanism by which the current gets from one ring to the next is not considered. By joining the rings electrically in series, as has been done in the diagram, we can see that it is not possible to transfer the conduction current without creating additional inductance (due to a current in the axial direction); but in conventional modelling practice, that inductance (which amounts only to the partial inductance of a conductor the length of the coil) is not included. Note incidentally that the Lorenz current-sheet model also only allows conduction in the radial direction, and so suffers from the same limitation.



The ring self and mutual inductance contributions can be obtained using one of the Maxwell complete elliptic integral expressions for the mutual inductances of coaxial circles; i.e., (M701.1) or (M701.2) as discussed in section 10b. For the mutual inductances of round wires, the geometric mean distance (GMD) between loops is simply the axial distance from wire-centre to wire-centre. For the self-inductances, the GMD is as appropriate for the type of calculation required (e.g., $g = r_w$, the wire radius, for a round wire excluding internal inductance). Note that a pair of loops has two mutual inductances (M_{N_1,N_2} and M_{N_2,N_1}) and so makes two contributions to the total. The complete set of contributions is listed below:

Magnetic interaction	Contribution
N loop self-inductances	$\rightarrow N \times M(r; r; g)$
N-1 loop pairs spaced at distance p	$\rightarrow 2(N-1) \times M(r; r; p)$
N-2 loop pairs spaced at distance 2p	$\rightarrow 2(N-2) \times M(r; r; 2p)$
N-3 loop pairs spaced at distance 3p	$\rightarrow 2(N-3) \times M(r; r; 3p)$
...	...
2 loop pairs spaced at distance (N-2)p	$\rightarrow 4 \times M(r; r; \{N-2\}p)$
1 loop pair spaced at distance (N-1)p	$\rightarrow 2 \times M(r; r; \{N-1\}p)$

where $M(r_1; r_2; s)$ represents the mutual inductance between coaxial circular filaments of radii r_1 and r_2 and axial separation s .

Careful consideration of the list will show that the total number of inductance contributions in the set is N^2 , but multiple instances of identical contribution result in the need for only N separate mutual inductance calculations. Using standard notation, the inductance of a solenoid by the summation method is thus:

$$L = NM(r ; r ; g) + 2 \sum_{m=1}^{N-1} (N-m) M(r ; r ; mp)$$

Or, alternatively:

$$L = NM(r ; r ; g) + 2 \sum_{m=1}^{N-1} m M(r ; r ; (N-m)p)$$

Note that both of these expressions are made-up of two distinct parts: a sum of self-inductances (all identical), and a sum of mutual inductances (not all identical).

A Basic routine that performs the calculation for coils made from cylindrically-symmetric wire is shown below:

```
Function Ljcm(ByVal N as long, r as double, x as double, g as double) as double
' Calc L of round-wire solenoid in uH using summation method. Version 2.20
' Calls function Mpar for mutual inductances. N = no. of turns.
' r=coil radius /mm, x=coil length /mm, g = conductor self gmd /mm.
If N < 1 then
  Ljcm = 0
elseif g <= 0 then
  Ljcm = 1E30
else
  Dim p as double, L as double, m as long
  ' calculate the N loop self inductances using the conductor gmd:
  L = N*Mpar( r , g )
  ' calculate the N(N-1) mutual inductances:
  p =x/N
  for m = 1 to N-1
    L = L + 2*m*Mpar( r , (N-m)*p )
  next
  Ljcm = L/1000
endif
end function
```

Referring to the program code: $Mpar(r ; g)$ is a routine that calculates the mutual inductance of parallel coaxial filaments, in nH, using Maxwell's formula (**M701.2**) and Bob Weaver's AGM algorithm. (see the macro library of the accompanying spreadsheet [Rosa_v_summation.ods](#)).

The input argument g is used to control whether the function calculates external inductance only ($g = r_w$) or includes DC internal inductance [$g = r_w \exp(-1/4)$]. Including internal inductance in this way is not recommended because the actual length of the wire in a coil is slightly longer than $n\pi D$, but the facility is useful for comparing the results against other methods. Note that setting g to zero results in infinite inductance and thus constitutes an invalid input. The condition is therefore

trapped before it can be passed to `Mpar()`, and the function returns $L/\mu H = 10^{30}$.

Notice also that the GMD for mutual inductance calculation is taken to be an integer multiple of the pitch, i.e., the GMD is taken from wire-centre to wire-centre. This is correct for points external to round wires and cylindrical tubes, but not for current-sheet segments or any other conductor cross-section that lacks infinite rotational symmetry. Consequently, the program cannot be used to calculate the inductance of coils having non-cylindrical conductors (although, of course, it can be modified fairly easily to include a GMD factor in the call to `Mpar`).

An idea of the relative computational efficiency of the summation method (for solenoids) can be gleaned by assuming (for the sake of argument) that we will use complete elliptic integrals for those parts of any calculation that have a major effect on accuracy. We will allow however, that when several elliptic integral terms have the same value of modulus (κ), then they can all be calculated in the same program loop; so the speed of computation is not so much linked to the number of terms, but to the number of different modulus values required.

When the Rosa method is used, the initial step of current-sheet inductance calculation usually gets the solenoid inductance to well within 3%. The precision requirement for the Rosa corrections is therefore much less stringent than for the determination of Nagaoka's coefficient. Thus we might argue that Rosa's method requires only one call upon an elliptic integral calculation routine (ignoring the fact that we have approximation methods good enough to eliminate even that).

The summation method, in contrast, always requires N calls to a routine for calculating K-E. Thus it is very roughly N -times slower than the Rosa method for a given level of precision. This incidentally, is not particularly important for spot calculations, or for graph-plotting if the maximum number of turns is reasonably small; but for extensive repetitive calculations it can prove to be an encumbrance (especially when running programs via an interpreter).

The main drawback of the summation method however is not necessarily obvious upon initial examination. It is that the mutual-inductance part of the summation is strictly limited to integer values of N . This makes the method completely unsuitable for any kind of simulation that seeks to return a value of N on a continuum; i.e.; it cannot be used for the purpose of coil optimisation except in a very restricted sense. There is a great deal of difference between setting up a model that gives values of N requiring truncation to a whole number, and setting up a model that can only use whole numbers internally for some variables.

Thus, if there is need to supply a compelling modern-day reason for using the Rosa method, it is that it can free us from the integer-turns restriction. It cannot do so directly, incidentally, because the mutual inductance correction is derived from the integer mathematics of the summation method; but it leads to continuous approximate functions of high precision, and these provide the necessary facility for interpolation.

A further general point that should be kept in mind, is that the summation method is not perfect, even within the limitations already discussed. The reason is that, although the Maxwell elliptic integral formulae for mutual inductance are exact for coaxial circular filaments; the calculation of self-inductance on the basis that it is the same as the mutual inductance of a pair of filaments separated by the internal GMD is an approximation that is true only when the GMD is much smaller than the loop radius⁷⁵ (i.e., the curvature of the loop is assumed to be negligible). Also, of course, it must be remembered that the turns of a helical coil are not rings.

⁷⁵ See **BS Sci. 31**, p173. But note that Wein's formula mentioned there is wrong. See **BS Sci 169**, footnote on p111.

10d. Rosa's self-inductance correction

Rosa's correction coefficient for self-inductance (k_s here, A or G elsewhere) is derived from the difference in the inductance of a single-turn loop of round wire compared to that of a single-turn current sheet.

The inductance of a relatively-short current sheet can be obtained to a good approximation by using the Rayleigh-Niven formula given in [section 8b](#). For the special case of a single turn of a multi-turn coil (i.e., a very short current sheet), Rosa considered the truncated form (just the first term in the ℓ/D power series) to be sufficient, i.e.;

$$k_{L1} = \frac{2}{\pi}(\ell/D) \left[\ln(4D/\ell) - \frac{1}{2} \right]$$

Putting this into equation [\(9.1\)](#) with $N=1$ we get:

$$L_{1s} = \mu_{(x)} r \left[\ln(4D/\ell) - \frac{1}{2} \right]$$

where $\mu_{(x)}$ is the external permeability (i.e., the permeability of the medium outside the conductor).

For the current-sheet equivalent to one-turn of a coil, Rosa defined the problem in such a way that the length of the cylinder corresponded to the pitch p of the inductor to which the correction is to be applied. Hence, substituting $2r = D$ and $p = \ell$, we get:

$$L_{1s} = \mu_{(x)} r \left[\ln\left(\frac{8r}{p}\right) - \frac{1}{2} \right]$$

Now turning to the loop inductance part: as we saw in [section 10b](#), the external inductance of a single turn of round wire is given to a good approximation by the formula [\(10.3\)](#):

$$L_{1x} = \mu r \left[\ln\left(\frac{8r}{(d/2)}\right) - 2 \right]$$

Rosa however, also included the DC internal inductance in his expression. This was given earlier (in [section 6](#)) as:

$$L_{i(dc)} = \ell_w \frac{\mu_{(i)}}{8\pi}$$

In this case, the length of the wire is $2\pi r$ (the circumference of the loop), and so:

$$L_{i(dc)} = \mu_{(i)} \frac{r}{4}$$

For the case where the wire is non magnetic, and the coil is not in proximity to any magnetic materials:

$$\mu_{(x)} = \mu_{(i)} = \mu_0$$

Hence, the total inductance of a loop of non-magnetic round wire, at low frequencies, in the absence of a magnetic core is:

$$L_{1(\text{dc})} = \mu_0 r \left[\ln \left(\frac{8r}{(d/2)} \right) - 2 + \frac{1}{4} \right]$$

The correction for self inductance given by Rosa is obtained by subtracting the inductance of a single-turn current-sheet of length p from the inductance of a single-turn wire loop and multiplying by the number of turns (N); i.e., the per turn correction is applied N times and we have:

$$N(L_{1(\text{dc})} - L_{1s}) = \mu_0 r N \left[\ln \left(\frac{8r}{(d/2)} \right) - 2 + \frac{1}{4} - \ln \left(\frac{8r}{p} \right) + \frac{1}{2} \right]$$

Observe here that subtracting logarithms is the same as performing a division of the numbers within. Hence:

$$(L_{1(\text{dc})} - L_{1s}) = \mu_0 r \left[\ln \left(\frac{2p}{d} \right) - \frac{3}{2} + \frac{1}{4} \right]$$

The $\frac{1}{4}$ being the internal inductance term. Note however, that Rosa defined his correction in equation (10.1) as negative. Thus:

$$k_{s(\text{dc})} = - \frac{(L_{1(\text{dc})} - L_{1s})}{\mu_0 r}$$

Hence, Rosa's original self-inductance correction coefficient, $k_{s(\text{dc})}$ (also known as A or G), depends only on the wire-pitch / wire-diameter ratio and is given by:

$k_{s(\text{dc})} = \frac{5}{4} - \ln \left(\frac{2p}{d} \right)$	10.4
--	-------------

This can also be written:

$k_{s(\text{dc})} = \ln \left(1.7542 \frac{d}{p} \right)$	10.4a
--	--------------

where $1.7542 = \frac{\exp(5/4)}{2}$

p/d is always > 1 , for reasons set out in [section 5a](#), and because close-spaced coils must be wound using insulated wire. When $p/d = \exp(5/4)/2$, $k_s = 0$, and so coils with a gap between turns of about $\frac{3}{4}$ of the wire diameter require no self-inductance correction at low frequencies.

Separation of internal inductance

Rosa's self-inductance correction, given as above in various textbooks, was, of course, never intended to be used for high-frequency calculations. Having preserved the identity of the internal inductance component in the derivation above however, we are in a position to modify the formula

to improve its versatility. The most obvious approach is to make a complete separation between the calculation of the internal and external inductances. By so doing, it becomes possible to have different external and internal permeabilities, allowing the model to deal with magnetic cores (or even magnetic wires); and the full frequency dependence of internal inductance can be included if required.

The self inductance correction, for *external inductance only*, is obtained by adding $\frac{1}{4}$ to equation (10.4):

$k_{s(x)} = \frac{3}{2} - \ln\left(\frac{2p}{d}\right)$	10.5
---	-------------

Using this expression, the inductance of a round-wire solenoid becomes:

$L = L_s - \mu_{(x)} r N (k_{s(x)} + k_m) + L_i$	[Henrys]	10.6
--	----------	-------------

The calculation of the internal inductance, L_i , has already been discussed in [section 6](#).

The separate calculation of internal and external inductance is slightly more accurate than Rosa's method, even at low frequencies. By comparing (10.4) and (10.5) we can see that the internal inductance term included by Rosa is:

$$L_{i(dc)} = \frac{\mu_{(i)} r N}{4}$$

this expression being obtained by estimating the length of the wire in the coil as N times the circumference, i.e.;

$$\ell_w = 2\pi r N$$

This estimate is only accurate when the pitch is small relative to the coil diameter. When p is relatively large, as it sometimes is in radio coils, it is better to use an exact expression for the wire length, i.e., (5.1) or (5.2):

$$\ell_w = \frac{\sqrt{(2\pi r N)^2 + \ell^2}}{\cos\psi} = \frac{2\pi r N}{\cos\psi}$$

where ℓ is the solenoid length and ψ is the pitch angle.

Note that, even when the internal inductance component is removed from the correction coefficient, there is still a mechanism by which $k_{s(x)}$ can vary with frequency. This is the repulsion between current streams in adjacent conductors (another manifestation of the proximity effect) that will cause the effective pitch / diameter ratio to increase. Noting the sign conventions used in equations (10.5) and (10.6), such repulsion will give rise to a small *increase* in the self inductance of a turn, but this will be offset by the error incurred by neglecting the proximity-induced reduction in the internal inductance (see [section 6](#)).

Precision and applicability of Rosa's self-inductance correction

The Rosa self-inductance correction per turn is obtained by subtracting two formulae, one being the wire loop inductance formula (10.3) and the other being the truncated Rayleigh-Niven current-sheet formula with $N=1$ and the length of the sheet defined as the axial length of a turn.

As we have seen in section 10b, the simple wire loop formula is rather good, but it will become inaccurate for coils wound with very thick wire. We might address that problem by calculating the correction term using the elliptic integral expression for $M(r; r; g)$ instead of inferior approximations. Remembering that Rosa's corrections are second-order adjustments to a much larger current-sheet inductance; the practical circumstances under which it will become necessary to do that are likely to be somewhat rare, but it suggests a definitive calculation procedure against which to evaluate less computationally intensive approaches.

The use of the truncated Rayleigh-Niven sheet formula is perhaps more contentious, but not greatly so. Using pitch p in place of coil length ℓ , this formula is good up to about $p/D = 0.1$, i.e., $\psi = \arctan(p/\pi D) = 1.82^\circ$. This constitutes a fairly wide-spaced coil, but wider spacing is occasionally encountered in VHF or UHF radio practice.

Grover addressed the Rosa-correction issue in BS Research Paper 90⁷⁶, in response to some bungled attempts to show that the method is flawed (it isn't). He was able to show that the Rosa-Nagaoka approach is practically equivalent to the summation method, for which it was intended to provide a computationally-efficient alternative. There is an issue however in that both the summation and the Rosa methods start to become inaccurate as the pitch-angle increases. This shared pathology, which we will investigate later, is of course nothing to do with the Rosa corrections; but in discussing the matter, Grover pointed to a more precise version of the self-inductance correction. He did so only to show that it would make no practical difference; but we might now take the view that if we are to correct for the shortcomings of these methods and thereby extend the applicable range of pitch, we must at least make sure that there are no defects in the basic assumptions.

The extended correction formula written by Grover is obtained by combining the complete Rayleigh-Niven current sheet formula with the wire loop formula. Note that this will not cure the limitations of the underlying theory, but it will extend the precision of the method (as opposed to the accuracy) to $p/D = 0.7$, i.e.; $\psi = \text{Arctan}(0.7/\pi) = 12.6^\circ$.

Recall that the complete solenoid (partial) inductance can be expressed in the form:

$$L = L_s - \mu_{(x)} r N (k_{s(x)} + k_m) + L_i$$

Hence the complete self-inductance correction is:

$$\Delta L_{\text{self}} = - \mu_{(x)} r N k_{s(x)} + L_i$$

Internal inductance is given by (6.5) as:

$$L_i = \frac{\mu_{(i)} r N}{\cos \psi} \frac{\Theta}{4}$$

Where Θ (Theta) is a factor that varies between 1 at low frequencies and 0 as $f \rightarrow \infty$; and the factor $1/\cos\psi$ corrects for the difference between the circumference of a loop perpendicular to the axis and the conductor-length of a helical turn. Hence:

⁷⁶ A comparison of formulas for the inductance of coils and spirals wound with wire of large cross section. F W Grover, BS J. Research. Vol 3. 1929. [BS RP90], Equivalence of the summation method and the Rosa method. p165-170, esp. formula (7), p167.

$$\Delta L_{\text{self}} = N \left[-\mu_{(x)} r k_{s(x)} + \frac{\mu_{(i)} r \Theta}{4 \cos \psi} \right]$$

This contracts into Rosa's original form when we put $\mu_{(i)} = \mu_{(x)} = \mu$, set $\Theta = 1$, and neglect the pitch correction, i.e.;

$$\Delta L_{\text{self}} = \mu r N (-k_{s(x)} + 1/4) = -\mu r N k_{s(dc)}$$

Here, avoiding the unfortunate practice of making approximations at the definition stage; we can dispense with the proliferation of permeabilities by expressing the internal permeability relative to the external permeability; i.e.;

$$\mu_{r(i)} = \frac{\mu_{(i)}}{\mu_{(x)}}$$

This is convenient for air-cored coils, because in that case, assuming copper, silver or some other non-magnetic wire-making material, $\mu_{r(i)} = 1$ to within a few parts per 1000. Also, we can model an external permeability greater than μ_0 , i.e.; $\mu = \mu_{r(x)} \mu_0$, by making $\mu_{r(i)} = 1/\mu_{r(x)}$. Hence:

$$\Delta L_{\text{self}} = \mu r N \left[-k_{s(x)} + \frac{\mu_{r(i)} \Theta}{4 \cos \psi} \right]$$

We can also reduce distracting detail still further by including the internal relative permeability as part of the internal inductance factor; i.e.: $\Theta' = \mu_{r(i)} \Theta$. Thus any derivation of the correction coefficient can start with the rigorously defined form:

$$\frac{-\Delta L_{\text{self}}}{\mu r N} = k_{s(x)} - \frac{\Theta'}{4 \cos \psi}$$

Also note that, since the correction per turn needs to be applied N times:

$$\Delta L_{\text{self}} = N (L_1 - L_{1s})$$

where L_{1s} is the inductance of a 1-turn current sheet, and L_1 is the complete inductance of a 1-turn wire loop. Thus:

$k_s = k_{s(x)} - \frac{\Theta'}{4 \cos \psi} = \frac{L_{1s} - L_1}{\mu r}$ <p>Where $\Theta' = \mu_{r(i)} \Theta$</p>	<p>Formal definition for Rosa's self-inductance correction coefficient.</p>
---	---

Also, if we exclude internal inductance we get:

$$k_{s(x)} = \frac{L_{1s} - L_{1x}}{\mu r}$$

Where L_{1x} is the external inductance of a 1-turn loop.

Using the complete Rayleigh-Niven sheet formula (**section 8b**) with $N=1$ and $\ell = p$ gives the inductance of a 1-turn of the current sheet as:

$$L_{1s} = \mu r \left[\ln\left(\frac{8r}{p}\right) - \frac{1}{2} + \frac{1}{32}\left(\frac{p}{r}\right)^2 \left[\ln\left(\frac{8r}{p}\right) + \frac{1}{4} \right] \right]$$

The loop external inductance can be accurately approximated using (**10.3**):

$$L_{1x} = \mu r \left[\ln\left(\frac{8r}{r_w}\right) - 2 \right]$$

Hence, subtracting the two formulae:

$$k_{s(x)} = \ln\left(\frac{8r}{p}\right) - \frac{1}{2} + \frac{1}{32}\left(\frac{p}{r}\right)^2 \left[\ln\left(\frac{8r}{p}\right) + \frac{1}{4} \right] - \ln\left(\frac{8r}{r_w}\right) + 2$$

Thus, noting that $2r_w = d$ (the wire diameter) and that inverting the argument of a logarithm changes its overall sign:

$$k_{s(x)} = \frac{3}{2} - \ln\left(\frac{2p}{d}\right) + \frac{1}{32}\left(\frac{p}{r}\right)^2 \ln\left[\left(\frac{8r}{p}\right) + \frac{1}{4}\right]$$

10.7

Comparing this with the simple version of Rosa's coefficient (**10.5**), we may note that, although the extended formula for k_s remains dimensionless, it now depends not only on the pitch to wire-diameter ratio p/d , but also on the ratio of pitch to coil radius or diameter, i.e., p/r or p/D . Expressed using the coil diameter it becomes:

$$k_{s(x)} = \frac{3}{2} - \ln\left(\frac{2p}{d}\right) + \frac{1}{8}\left(\frac{p}{D}\right)^2 \ln\left[\left(\frac{4D}{p}\right) + \frac{1}{4}\right]$$

$$p \leq 0.7 D \quad , \quad D \geq 10 d \quad .$$

10.7a

To include internal inductance, we simply subtract $\frac{1}{4}$ to be consistent with NBS documents, or more generally, we subtract $\Theta'/(4 \cos\psi)$. Note also, as mentioned in **section 5**:

$$\frac{1}{\cos\psi} = \sec\psi = \sqrt{1 + \left(\frac{p}{2\pi r}\right)^2}$$

Thus:

$$k_s = \frac{3}{2} - \ln\left(\frac{2p}{d}\right) + \frac{1}{32}\left(\frac{p}{r}\right)^2 \ln\left[\left(\frac{8r}{p}\right) + \frac{1}{4}\right] - \frac{\Theta'}{4} \sqrt{1 + \left(\frac{p}{2\pi r}\right)^2}$$

10.8

Where: $\Theta' = \frac{\mu_{(i)}}{\mu_{(x)}} \Theta = \mu_{r(i)} \Theta$

Alternatively, using the coil diameter:

$$k_s = \frac{3}{2} - \ln\left(\frac{2p}{d}\right) + \frac{1}{8}\left(\frac{p}{D}\right)^2 \ln\left[\left(\frac{4D}{p}\right) + \frac{1}{4}\right] - \frac{\Theta'}{4} \sqrt{1 + \left(\frac{p}{2\pi r}\right)^2}$$
10.8a

Notice here incidentally, that we can force the formula to calculate $k_{s(x)}$ by setting $\Theta' = 0$.

A Basic macro function that performs the calculation is shown below:

```
Function Rosaksx(byval pdw as double, pD as double, fi as double) as double
' Rosa's self inductance correction, extended series version for p/D <= 0.7 (psi <= 13 deg.)
' pdw = pitch / wire diam , pD = pitch / coil diam,
' fi = internal inductance factor, zero for none, 1 for LF.
Rosaksx = 1.5 - log(2*pdw) + pD*pD*( log(4/pD) + 0.25 )/8 - fi*sqr(1 +(pD/pi)*(pD/pi) )/4
end function
```

We can also apply a correction to the formula (**10.8a**) to extend its range to $p/D = 1$, i.e., $\psi = \arctan(1/\pi) = 17.7^\circ$. To do that, we simply add extra terms to convert the Rayleigh-Niven formula into Coffin's formula (see **section 8b**). The extra terms (using p instead of ℓ) are:

$$\delta k_s = -\frac{1}{64}\left(\frac{p}{D}\right)^4 \left[\ln\left(\frac{4D}{p}\right) - \frac{2}{3} \right] + \frac{5}{1024}\left(\frac{p}{D}\right)^6 \left[\ln\left(\frac{4D}{p}\right) - \frac{109}{120} \right] - \frac{35}{16384}\left(\frac{p}{D}\right)^8 \left[\ln\left(\frac{4D}{p}\right) - \frac{431}{420} \right]$$

A Basic function extended in this way is:

```
Function Rosaks18(byval pdw as double, pD as double, fi as double) as double
' Rosa's self inductance correction, extended series version for p/D <= 1 (psi <= 18deg.)
' pdw = pitch / wire diam , pD = pitch / coil diam,
' fi = internal inductance factor, zero for none, 1 for LF.
Dim pd2 as double, pd4 as double, lge as double, dk as double
pd2 = pD*pD
pd4 = pd2*pd2
lge = log(4/pD)
dk = -pd4*(lge - 2/3)/64 + 5*pd2*pd4*(lge - 109/120)/1024 - 35*pd4*pd4*(lge - 431/420)/16384
Rosaks18 = 1.5 - log(2*pdw) + pD2*(lge + 0.25)/8 - fi*sqr(1 +(pD/pi)*(pD/pi) )/4 + dk
end function
```

If a closed-form expression that remains precise for all possible pitch angles is required however, it is better to use one of the continuous empirical formulae for Nagaoka's coefficient to provide the current-sheet terms. The best of these is Bob Weaver's modified version of Wheeler 1982-7 (**W82-7W**) as discussed in **section 8e**. That formula, with p substituted for ℓ gives:

$$\frac{L_{Is}}{\mu r} = \ln \left[1 + \frac{\pi D}{2 p} \right] + \frac{1}{P_{nom} \left(\frac{p}{D} \right)}$$

P_{nom} being a correcting polynomial defined as:

$$P_{nom} \left(\frac{p}{D} \right) = \frac{1}{\left[\ln \left(\frac{8}{\pi} \right) - \frac{1}{2} \right]} + 3.437 \left(\frac{p}{D} \right) + \frac{24}{(3\pi^2 - 16)} \left(\frac{p}{D} \right)^2 - \frac{0.47}{\left(0.755 + \frac{D}{p} \right)^{1.44}}$$

The four empirical constants shown give a precision of ± 20.5 ppM, which makes the evaluation of 1-turn sheet inductance practically exact in comparison to the wire loop formula to be subtracted from it. The first term in the series is best evaluated as as given (rather than entered as a constant in truncated form) because it ensures that the formula is asymptotic in the short-coil limit, this being an important requirement for the modelling of single turns.

Thus, using the polynomial-corrected formula, and also improving the accuracy of the round-wire loop part by using the complete Rayleigh-Niven short coil external inductance formula (10.3a) we get:

$$k_s = \ln \left[1 + \frac{\pi D}{2 p} \right] + \frac{1}{P_{nom} \left(\frac{p}{D} \right)} - \ln \left(\frac{8D}{d} \right) + 2 - \frac{1}{8} \left(\frac{d}{D} \right)^2 \left[\ln \left(\frac{8D}{d} \right) + \frac{1}{3} \right] - \left(\frac{\Theta'}{4} \right) \sqrt{1 + \left(\frac{p}{\pi D} \right)^2} \quad \mathbf{10.9}$$

A Basic function that performs this calculation is shown below. The 90 in the function name implies that it remains precise for a pitch angle of 90° (albeit without the ability to compensate for the theoretical limitations of the Lorenz and Maxwell methods).

```
Function Rosaks90(byval pdw as double, pD as double, fi as double) as double
'Rosa's self-inductance correction, continuous empirical version 2.00.
' Uses Bob Weaver's modified version of Wheeler 82-7 for the current sheet part
' and the complete Rayleigh-Niven (13) formula for the round-wire part.
' pdw = pitch / wire diam , pD = pitch / coil diam,
' fi = internal inductance factor, zero for none, 1 for LF.
Dim Ddw as double, pn as double, k0 as double, k2 as double, w as double, lg as double
Ddw = (1/pD)/(1/pdw)
k0 = 1/(log(8/pi)-0.5)
k2 = 24/(3*pi*pi -16)
w = -0.47/(0.755 + 1/pD)^1.44
pn = k0 + 3.437*pD + k2*pD*pD + w
lg = log(8*Ddw)
Rosaks90 = log(1+pi/(2*pD))+1/pn-lg+2-(lg+1/3)/(8*Ddw*Ddw) -fi*sqr(1+(pD/pi)*(pD/pi))/4
end function
```


Precise calculation of Rosa's self-inductance correction

The ultimate version of Rosa's correction is, of course, one that has no approximations within the limitations of the model. Defining the external self inductance correction coefficient as before:

$$k_{s(x)} = \frac{L_{1s} - L_{1x}}{\mu r}$$

we note that the one-turn current sheet inductance can be obtained by calling a function for the exact calculation of Nagaoka's coefficient with D/p as the argument instead of D/ℓ . The complete expression for the inductance of 1-turn of the sheet is then, using (9.1):

$$L_{1s} = \mu r \frac{\pi}{2} \left(\frac{D}{p} \right) \text{Nagaoka} \left(\frac{D}{p} \right)$$

Dividing this expression by μr gives the coefficient as:

$$k_{s(x)} = \frac{\pi}{2} \left(\frac{D}{p} \right) \text{Nagaoka} \left(\frac{D}{p} \right) - \frac{L_{1x}}{\mu r}$$

Now, for the loop inductance, we can use Maxwell's second expression for M in complete elliptic integrals (M701.2) as discussed in section 10b. The reason for going back to basics in this way is that the correction coefficient is dimensionless, and can therefore be calculated by a function having dimensionless arguments. Putting $r_1 = r_2 = r$ gives:

$$L_1 = \mu r \frac{2}{\sqrt{\kappa_1}} [K(\kappa_1) - E(\kappa_1)] = M(r; r; g)$$

which gives gives the Rosa coefficient as:

$$k_s = \frac{\pi}{2} \left(\frac{D}{p} \right) \text{Nagaoka} \left(\frac{D}{p} \right) - \frac{2}{\sqrt{\kappa_1}} [K(\kappa_1) - E(\kappa_1)]$$

Note that the subscript x (for external) has here been dropped because the inclusion of internal inductance is now dependent on the choice of GMD as used in the (dimensionless) arguments of the elliptic integrals. The argument common to both elliptic integrals was given earlier as:

$$\kappa_1 = \frac{s_1 - s_2}{s_1 + s_2} \quad \text{where} \quad s_1 = \sqrt{(r_1 + r_2)^2 + s^2} \quad \text{and} \quad s_2 = \sqrt{(r_1 - r_2)^2 + s^2}$$

With $r = r_1 = r_2$, and the filament separation s made equal to the GMD, we get:

$$s_1 = \sqrt{4r^2 + g^2} \quad \text{and} \quad s_2 = g$$

i.e.:

$$\kappa_1 = \frac{\sqrt{(4r^2 + g^2)} - g}{\sqrt{(4r^2 + g^2)} + g}$$

Factoring out g and cancelling gives:

$$\kappa_1 = \frac{\sqrt{(2r/g)^2 + 1} - 1}{\sqrt{(2r/g)^2 + 1} + 1}$$

The ratio g/r is defined as:

$$\frac{g}{r} = \frac{r_w}{r} \exp(-\gamma) = \frac{d}{D} \exp(-\gamma)$$

where $\gamma = 0$ if only the external inductance is to be calculated, and $\gamma = 1/4$ if the DC value of the internal inductance is to be included for comparison with old methods and data. Note incidentally, that if we include internal inductance in this way, then the total conductor length is taken to be $N\pi D$ (i.e., the $\sec\psi$ factor is neglected), and the relative permeability of the wire is forced to be the same as that of the external medium. Hence, the feature is useful for testing, but it is not recommended for actual modelling. This is, incidentally, in contrast to the extended and continuous formulae discussed previously, where full provision for internal inductance was included.

So now we have a function that needs the calling arguments D/p (for the current sheet part) and d/D (for the wire loop part); and also, optionally, the exponent argument $-\gamma$ for the GMD factor. For consistency with the Basic functions for the closed-form formula (given above) however, the two unique arguments in the Basic function shown below have been chosen as p/d and p/D . In that case d/D is given by:

$$\frac{d}{D} = \frac{p/D}{p/d}$$

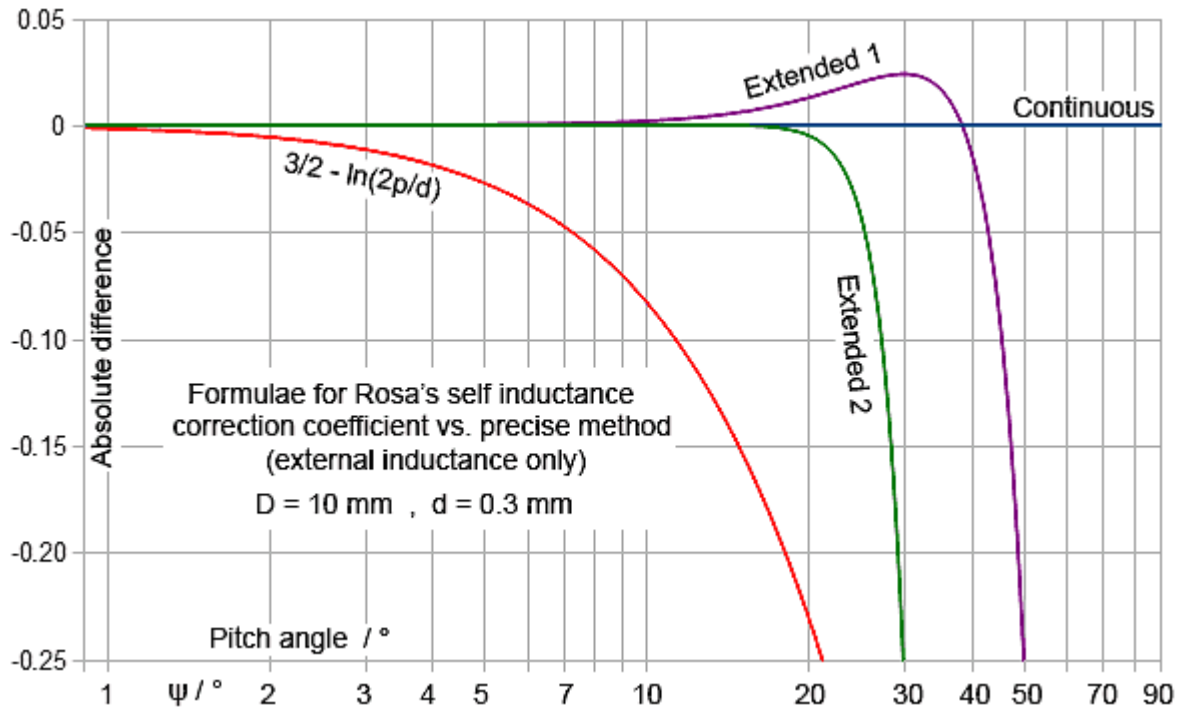
```
Function Rosaksp(byval pdw as double, pD as double, ga as double)
' Precise calculation of Rosa self inductance correction ks.
' Calls functions Nagaoka and KminusE
' pdw = pitch / wire diam. pD = pitch / coil diam.
' ga = gmd exponent, 0 for no int. inductance, -0.25 for LF.
Dim kL1s as double, kL1 as double, k1 as double, sr as double
sr = exp(ga)*pD/pdw
k1 = (sqr(4 + sr*sr) - sr)/(sqr(4 + sr*sr) + sr)
kL1 = 2*sqr(1/k1)*KminusE(k1)
kL1s = (pi/2)*Nagaoka(1/pD)/pD
Rosaksp = kL1s - kL1
end function
```

The function `KminusE` is a fast routine for calculating the complete elliptic integral combination $K(\kappa) - E(\kappa)$ using Bob Weaver's AGM algorithm (see the macro library in the spreadsheet [Loop_funcs.ods](#), version 1.00 onwards).

Note that further simplification might be possible by writing the expression for Nagaoka's coefficient in complete elliptic integrals and subtracting the loop inductance component from it. That operation however, does not appear to reduce the number of function calls for elliptic integral calculation, and so is unlikely to give significant improvement in computational efficiency.

Comparison of formulae for k_s

An evaluation of the closed formulae discussed above for the Rosa self-inductance correction coefficient is shown in the graph below. The baseline for comparison is the precise calculation using complete elliptic integrals as just described.



The calculations used to produce the curves are given in the spreadsheet [Loop_funcs.ods](#) , sheet 2.

- The curve marked " $3/2 - \ln(2p/d)$ " is produced using the simple formula (10.5) (acceptable for $\psi \leq 1.8^\circ$).
- "Extended 1" is the version (10.8a) using the full Rayleigh-Niven current-sheet formula (acceptable to about 13°).
- "Extended 2" is the version using Coffin's formula' (very precise up to 18°).
- "Continuous" is the version (10.9) based on Bob Weaver's modified form of Wheeler 1982-7 and the non-truncated Rayleigh-Niven round-wire external inductance formula (10.3a).

Regarding the matter of which of these calculation methods is best for routine use; it should be noted that, although it cannot be seen on the graph, the formula "Extended 2", based on the Coffin series is the most severely wrong once out of its range of application. Hence, like online inductance calculators in general, it harbours nasty surprises for those who encounter it divorced from its caveats. The others are more friendly and may be chosen according to practical considerations; but given the availability of a programming environment, it is hard to think of a reason for not using the continuous version (10.9).

The precise method is perhaps too computationally intensive for routine use; but even the interpreted Basic implementation is not painfully slow for the plotting of a few hundred graph points.

10e. Rosa's mutual inductance correction

The mutual inductance correction coefficient k_m (also known as B or H) was introduced in [section 10](#) as part of the general Rosa expression for solenoid inductance ([10.1](#)). The correction is required because pairs of round-wire loops and pairs of current-sheet loops separated by the same average distance do not have the same geometric mean distance (GMD) for mutual-inductance calculation. For conductors having cylindrical symmetry, the external GMD is the distance from centre-to-centre of the conductor cross sections (the cross sections being taken in a plane perpendicular to the conductor axes). For current-sheet segments however, the GMD is slightly less than the average (i.e., the arithmetic mean) separation

GMD for pairs of current-sheet segments

For the purpose of mutual inductance calculation, the GMD of a pair of current sheet segments is the same as that of a pair of straight lines of finite length both lying on the same straight line. The general solution for this was given by Maxwell⁷⁷, but here we are only interested in the special case in which the segments both have the same length p and the average separation between segments is an integer multiple of p . Of course, p corresponds to the pitch of the coil; and we will use m for the integer multiplier because it will turn out to correspond to the summation-index for the mutual-inductance part of the summation-method inductance calculation.

The required formula is easily obtained from the general solution and was given explicitly by Rosa as the starting point for the derivation of his correction coefficient⁷⁸. The expression, for \log_e of the GMD (with a change from Rosa's notation), is:

$$\ln(g) = \frac{(m+1)^2}{2} \ln[(m+1)p] - m^2 \ln(mp) + \frac{(m-1)^2}{2} \ln[(m-1)p] - \frac{3}{2} \quad (10.10)$$

From here on however, we will deviate from the narrative of Rosa's paper in order to clarify an important logical step. Firstly, in keeping with earlier discussion, we will represent the GMD as the average distance between segments multiplied by an exponential:

$$g = m p \exp(-\gamma)$$

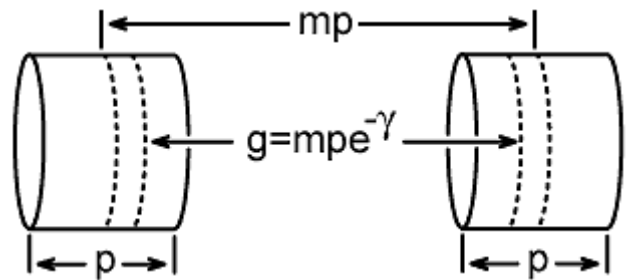
This means, noting the general logarithmic relationships $\ln(ab) = \ln(a) + \ln(b)$, and $\ln(e^x) = x$, that:

$$\ln(g) = \ln(m) + \ln(p) - \gamma$$

Now, inspect equation ([10.10](#)) and note that:

$$\frac{(m+1)^2}{2} + \frac{(m-1)^2}{2} - m^2 = 1$$

This means that if the factor p within each logarithm is moved into a separate logarithm term, there will be obtained a single $\ln(p)$ term on the right that will cancel the $\ln(p)$ component of the



77 On the GMD of two figures in a plane. J C Maxwell. Trans. Roy. Soc. Edinburgh. Vol. 26. 1872. p280 - 285. See page 282, formula (3).

78 Rosa, 1906, BS Sci. 31, pages 167 - 169.

logarithm of the GMD. Thus:

$$\ln(m) - \gamma = \frac{(m+1)^2}{2} \ln(m+1) - m^2 \ln(m) + \frac{(m-1)^2}{2} \ln(m-1) - \frac{3}{2} \quad (10.11)$$

What Rosa did was to set $p = 1$ to make it disappear from (10.10), a step that caused the definition of the GMD to change from that point onwards. That operation appears to make no sense unless we realise that what he continues to refer to as the GMD no longer has dimensions of length. In fact, all we are interested in is the factor by which the GMD differs from the average (arithmetic mean) distance. That can be obtained from the quantity γ (the argument of the GMD-factor exponent), which is now given by rearrangement of (10.11):

$$\gamma_m = (m^2 + 1) \ln(m) + \frac{3}{2} - \frac{(m+1)^2}{2} \ln(m+1) - \frac{(m-1)^2}{2} \ln(m-1) \quad (10.12)$$

Notice, incidentally, that although the definition of γ has not changed from that given by (10.11), we have by this point established that it is purely a function of m (but only for the special case of equal segments separated by an integer number of segment lengths). Hence it has been given the subscript m as a reminder and to distinguish it from instances related to other GMD problems.

Inserting $m=1$ into equation (10.12) we find that $\gamma_1 = (3/2) - \ln(4) = 0.1137\dots$. Hence the factor by which the average distance between two adjacent segments needs to be multiplied in order to obtain the GMD is $\exp(-0.1137\dots) = 0.8925\dots$. Thus the GMD is considerably less than the average distance, and so the mutual inductance is significantly greater than for a pair of round-wire turns at the same average separation. This however is the extreme case. As the current-sheet segments get further and further apart, the GMD tends gradually towards the arithmetic mean distance.

A formula for γ_m is needed, of course, not only for the purpose of enumerating Rosa's coefficient, but also for calculating the mutual inductance component of current-sheet inductance by the summation method (for determining the effect of approximations if nothing else). With that in mind, note by inspection, that as m becomes large, evaluating (10.12) becomes a matter of subtracting terms that become very similar in value. This means that the expression is prone to roundoff error. Rosa's solution was to take the series expansions of the logarithms and recombine the terms to give γ_m as an infinite series.

$$\gamma_m = \frac{1}{12m^2} + \frac{1}{60m^4} + \frac{1}{168m^6} + \frac{1}{360m^8} + \frac{1}{660m^{10}} + \dots$$

The generating function is:

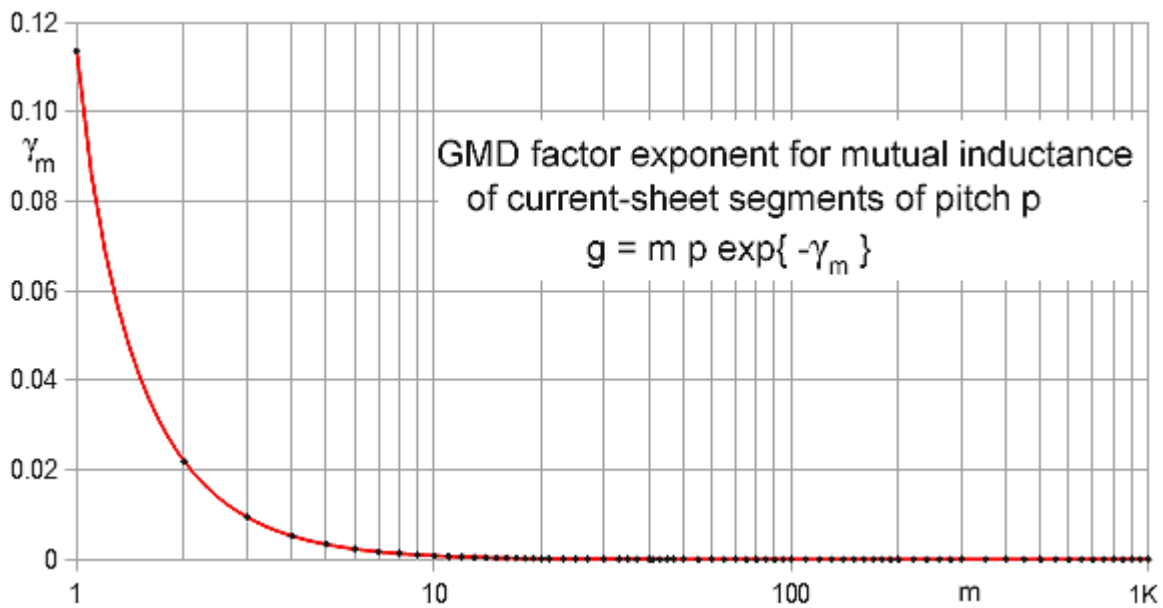
$$\gamma_m = \sum_{i=1}^{\infty} \left[\frac{1}{2i} + \frac{1}{2i+2} - \frac{2}{2i-1} \right] \frac{1}{m^{2i}} \quad (10.13)$$

Bob Weaver has also found an alternative version⁷⁹ that gives the integer denominators of the series coefficients directly:

⁷⁹ Investigation of Rosa's round wire mutual inductance correction. Robert Weaver, July 2008. [Weaver 2008] [http://electronbunker.ca/DLpublic/Rosa_Derivation.pdf also g3ynh.info/zdocs/magnetics/]

$$\gamma_m = \sum_{i=1}^{\infty} \frac{1}{(2i+1)(i+1)2i} \frac{1}{m^{2i}} \quad (10.13a)$$

Comparison of the closed form and series form expressions (10.12) and (10.13a) is given in the spreadsheet [Rosa_km.ods](#) (sheet 4). The calculations given there show that, in double-precision, in the Open Office Basic programming environment, the roundoff error in (10.12) has no effect beyond the 9th decimal place for $m < 1000$. The closed-form expression also has the advantage of being continuous. That can be seen in the graph below, which shows the smooth form of (10.12) with points from the integer-only (10.13a) superimposed.



Bear in mind however, that the GMD function is used for calculating current-sheet inductance, and also Rosa's round-wire corrections, by the summation method. This means that any roundoff error will accumulate in the mutual-inductance summation, actually reducing the overall precision to about 7 decimal places for coils approaching 1000 turns. That problem is circumvented by changing to the series summation for $m \geq 100$.

A Basic routine that uses the closed form expression for γ_m when $m < 100$ and then switches to the series form (10.13a) is shown in the box below.

Note, on examining the code; that in obtaining the quantity m^{2i} the exponentiation operator is not used. In general, exponentiation in programs should be avoided for large integer powers with non-optimising compilers and interpreters, because it involves taking logarithms and thereby introduces roundoff error. The specific reason for avoiding it in this case is that the exponentiation routine built into Open Office Basic becomes noticeably inaccurate for powers greater than about 10^6 (and, for such an extreme case, the same limitation will be present in other programming languages). The solution is to multiply the last-used value of m^{2i} by m^2 at the end of each summation loop, thereby obtaining the value required for the next term.

If the crossover point for using the summation rather than the continuous formula is changed to $m \geq 1000$; then rounding m to an integer from that point onwards has no effect in the 9th decimal place, and so the function becomes effectively continuous. The main use of the function however is in creating data for the development of empirical fitting functions; in which case, the $m \geq 100$ crossover is the choice that has no significant impact on overall precision.

```

Function gsm(byval m as double) as double
'GMD factor exp argument for M of current sheet segs. Ver. 2.0, 9th Aug. 2012.
' Uses continuous log formula for m < 100. Uses series summation for m >= 100.
' Based on the function dgmd by Bob Weaver.
if m <=1 then
  gsm = 1.5 - log(4)
elseif m < 100 then
  gsm = (m*m+1)*log(m) +1.5 -0.5*(m+1)*(m+1)*log(m+1) -0.5*(m-1)*(m-1)*log(m-1)
else
  Dim mi as long, mm as double, mmi as double, i as integer, sum as double, term as double
  mi = m
  mm = mi*mi
  mmi = mm
  for i = 1 to 18
    term = 1/((2*i+1)*(i+1)*2*i*mmi)
    sum = sum + term
    if term < 1E-36 then exit for
    mmi = mmi*mm
  next
  gsm = sum
endif
end function

```

Definition of Rosa's mutual inductance coefficient

For the summation-method calculation procedure, the inductance of a solenoid has the form:

$$L = L_{\text{self}} + L_{\text{mutual}}$$

where, as we saw in [section 10c](#), the mutual inductance part can be written:

$$L_{\text{mutual}} = 2 \sum_{m=1}^{N-1} (N-m) M(r; mp)$$

$M(r; mp)$ being, in principle, any function that can return the mutual inductance of coaxial circular filaments of equal radius r and axial separation mp . Recall also, that when the radii are equal, the axial filament separation is equal to the GMD for the type of conductor under consideration (it is written as mp above because the expression is for cylindrically symmetric conductors, in which case the GMD is the same as the average separation).

For the Rosa-method calculation procedure, the solenoid inductance has the form given earlier as equation (10.1):

$$L = L_s - \mu r N (k_s + k_m) \quad [\text{Henry}]$$

Where L_s is the equivalent current-sheet inductance, and k_s and k_m are the round-wire correction

coefficients (k_s on this occasion being assumed to include an internal-inductance component). With a formula for the GMDs of pairs of current sheet segments we can, of course, now envisage calculating the current-sheet inductance using the summation method. Thus:

$$L = L_{\text{self}} + L_{\text{mutual}} = L_{s,\text{self}} + L_{s,\text{mutual}} - \mu r N k_s - \mu r N k_m$$

The self-inductance part was dealt with in [section 10d](#), leaving us with the mutual inductance part:

$$L_{\text{mutual}} = L_{s,\text{mutual}} - \mu r N k_m$$

i.e.:

$k_m = \frac{L_{s,\text{mutual}} - L_{\text{mutual}}}{\mu r N}$	Formal definition for Rosa's mutual-inductance correction coefficient.
---	--

which, noting that the GMD for current sheet segments is $m p \exp(-\gamma_m)$, can be written explicitly as:

$$k_m = \frac{2}{\mu r N} \sum_{m=1}^{N-1} (N-m) \left[M(r; m p \exp(-\gamma_m)) - M(r; m p) \right] \quad (10.14)$$

We could, of course, now adopt the practice of calculating k_m using one of Maxwell's elliptic integral formulae and the function for γ_m described earlier; but there would be no computational advantage (indeed, there would be a disadvantage) in doing so. The expression above is however, general, and is therefore suitable both for evaluating approximations and as a starting point for finding approximations.

Rosa's formula for k_m (aka B or H)

In the early part of the 20th Century; the preferred approach to the problem of calculating the mutual inductance of circular filaments was not to use Maxwell's elliptic integral formulae, but to use one of the various available series expansions (with a view to truncating it after the minimum possible number of terms). For the derivation of his coefficient, Rosa used the Maxwell general series formula for coaxial circles⁸⁰, which, as befits the present problem, can be simplified by making both radii equal and identifying the axial separation as the GMD. The resulting expression is:

$$M(r; g) = \mu r \left[\left(1 + \frac{3}{16} \left(\frac{g}{r} \right)^2 \right) \ln \left(\frac{8r}{g} \right) - 2 - \frac{1}{16} \left(\frac{g}{r} \right)^2 \right] \quad (10.15)$$

Rosa however, chose to neglect terms in $(g/r)^2$ and higher, to obtain the simple formula:

$$M(r; g) = \mu r \left[\ln \left(\frac{8r}{g} \right) - 2 \right] \quad (10.15a)$$

This is the loop formula (10.3) with r_w replaced by g . We therefore know, from the discussion in

⁸⁰ Maxwell E&M, Vol 2, Article [705], p345. Also BS Sci. 169, p13, formulae [10] and [12].

section 10b, that the expression is fairly accurate for $r/g > 10$. This is fine for calculating the self-inductance of wire loops; but for the mutual inductances of turns on a solenoid, the distance can easily exceed the coil radius. The choice therefore seems inappropriate on first consideration. It is in the nature of this formula however, to under-predict (eventually reporting negative inductance when $r/g < e^2/8 = 0.9236..$). We also know that the difference between the mutual inductances for pairs of current-sheet loops and pairs of wire loops disappears when the separation is large. Hence the truncated formula should (*perhaps*) provide a fair approximation when used solely for the purpose of calculating differences.

Substituting **(10.15a)** into **(10.14)** gives an astonishing simplification:

$$k_m = \frac{2}{N} \sum_{m=1}^{N-1} (N-m) \gamma_m \quad (10.16)$$

This seemingly innocuous formula however, has previously caused the greatest difficulty in the matter of adapting the Rosa method for use with electronic computers. Recall that γ_m is given (according to Rosa's preference at least) by the series expression **(10.13)**; which means that **(10.16)** is actually a double summation. The calculation is so laborious by manual methods that is doubtful that either Rosa or Grover ever imagined that users of the NBS approach would want to do it. Great effort was therefore put into the tabulation of k_m (B or H), and little thought was given to the matter of explaining or refining the generating function. Readers of Rosa's original 1906 paper will note (for example) that not only does he discard the pitch from the GMD formula apparently arbitrarily, he also uses the same symbol for the summation index and for turns in the coil. This is easily done in private notebooks, since the person performing the derivation knows which symbol is which, but it should not have ended up in print. The subject is taken up again by Grover in BS Research Paper 90, and although this results in a new formula that we will examine shortly, the basic derivation is still far from clear. Hopefully, the logical development given above will help to unravel the matter.

Practical methods for calculating k_m

An empirical formula⁸¹ for the calculation of k_m was first given by this author (DWK) in 2006. It was obtained by fitting the four-figure data given in Grover's 1946 monograph⁸², but it was also obvious from the fitting residuals that there were some minor errors in the table. At the time however, the only NBS data available to the author were in Grover's book and in the 1911 first edition of what was later to become BS Science Paper 169 (the 1911 data being even less accurate). Hence the work appeared to be the best that could be done with the resources available.

The empirical formula was made public via the author's website, accompanied by a comment to the effect that the original generating function was 'not well described' in the sources used. There the matter rested until June 2008, when Bob Weaver wrote⁸³ to say that he had managed to deduce the original function (or its equivalent) using the current-sheet mutual GMD series formula given by Grover⁸⁴ and what was essentially a process of reverse-engineering. The expression he gave was identical to **(10.16)** in all but notation, and the derivation given above was greatly assisted by prior knowledge of that result. The GMD series formula given by Grover moreover is in truncated form with no explanation of how the coefficients come about; and by studying the pattern of numbers

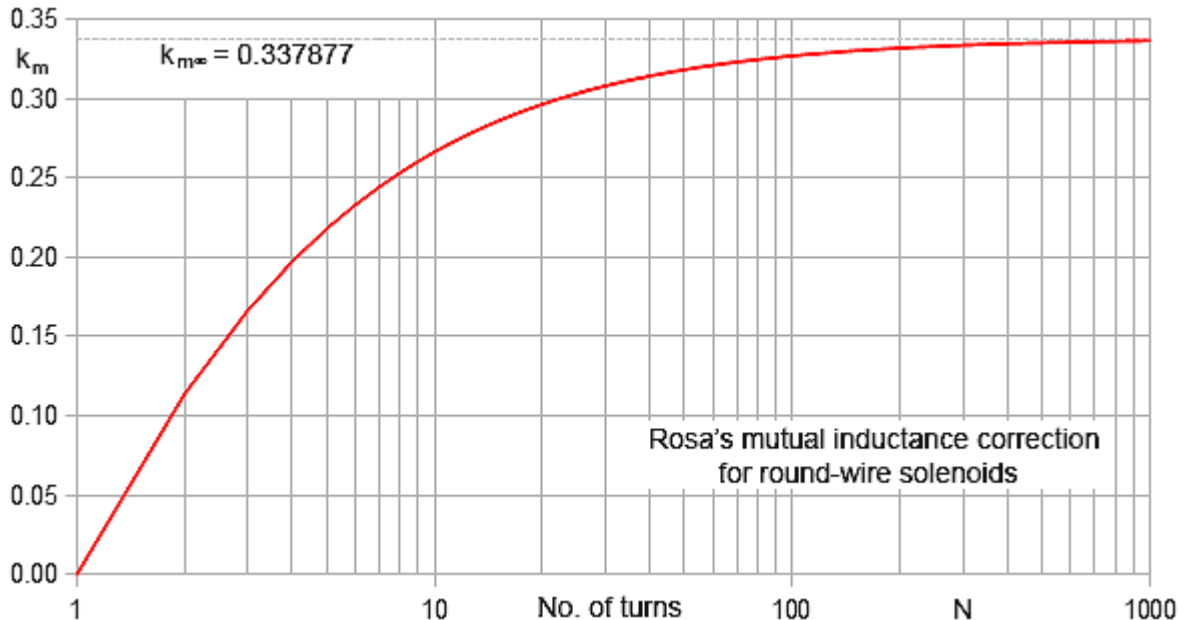
81 See the spreadsheet **Rosa_km.ods**, sheet 3.

82 **Grover 1946**, Table 39, p150.

83 Rosa's round wire corrections for mutual inductance. Bob Weaver, private e-mail communication, 16th June 2008.

84 **Grover 1946**, page 20, formula (e).

Bob was able to deduce the formula (10.13a), which is actually more convenient than Rosa's original. The Rosa coefficient k_m , as given by (10.16) using a program routine written by Bob⁸⁵, is shown plotted below. Note that it is purely a function of N .



The following analytical values of k_m are also useful when dealing with special cases:

$N = 1$	$k_{m1} = 0$
$N = 2$	$k_{m2} = \ln(1/4) + 3/2 = 0.113\ 705\ 638 = \gamma_1$
$N \rightarrow \infty$	$k_{m\infty} = \ln(2\pi) - 3/2 = 0.337\ 877\ 066$

With the help of Bob Weaver's investigative work, we can be confident of the definition of the Rosa mutual inductance correction and the method by which the tables were calculated. All findings are, incidentally, also confirmed in BS Research Paper 90⁸⁶. It should be understood however, that recalculation of the data for large N takes many hours using interpreted Basic and a fast personal computer, and so the summation formula (10.16) is not a good basis for a general-purpose solenoid-modelling subroutine. A further limitation of the formula, of course, is that it is for integer values of N only.

In view of the computationally-intensive nature of the calculation, and the need to provide an inherently continuous function, the empirical formula first obtained in 2006 was revisited. This time however, the raw data were machine-calculated, free from illegitimate errors, and precise to at least 10 decimal places. The development of the fitting function and its subsequent updating is discussed in a separate article⁸⁷. Details of the fit can be examined in the spreadsheet [Rosa_km.ods](#) (sheet 2). The best optimised version is:

85 <http://electronbunker.ca/CalcMethods2a.html>. See also the article cited earlier as **Weaver 2008**. All functions discussed in this section will be found in the Basic macro library of the spreadsheet [Rosa_km.ods](#).

86 **BS RP90**, Section V, page 174, formula (28).

87 **Rosa's mutual inductance correction for the round-wire solenoid**. David Knight, April 2010. [**Knight 2010**] g3ynh.inf/zdocs/magnetics/. A Basic routine is given in the macro library of the spreadsheet: [Rosa_km.ods](#).

$$k_m = \left[\ln(2\pi) - \frac{3}{2} \right] \left[1 - \frac{1 - 0.01711}{N - 0.01711} \right] + \ln(N) \left[\frac{-0.16641}{N} + \frac{0.00479}{N^2} + \frac{0.001772}{N^3} \right] \quad (10.17)$$

The maximum difference between this function and (10.16) is $\pm 0.000\,001\,04$ (i.e., $< 1.1 \times 10^{-6}$).

A further approximation formula, more precise than the one above, was obtained from information given in BS Research Paper 90. In that paper, Grover sets out to defend the Rosa method by comparing it with a modified version of the summation-method as developed by Strasser. Strasser's approach replaces the mutual inductance part of the summation-method with a formula involving a tabulated coefficient. That gives an improvement in the tractability of the calculation procedure for moderate N , but the coefficient is difficult to calculate for large N by straightforward summation. Asymptotic formulae for the calculation of Strasser's coefficient have however been developed; and Grover goes on to show that these are related to the Rosa coefficient and can therefore be used to obtain an expression for it. The investigation led him to the following formula⁸⁸:

$$k_m = \ln(2\pi) - \frac{3}{2} - \frac{\ln(N)}{6N} - \frac{0.330842}{N} - \frac{1}{120N^3} + \frac{1}{504N^5} \quad \text{for } N \geq 4$$

This expression is extremely precise for large N , and so enabled Grover to update the existing 1916 NBS table (which, given the difficulty in obtaining the coefficient by summation, is only reliable to two decimal places for large N). A difficulty with the Grover formula however, is that it is not accurate for $N < 4$, and does not give zero for $N=1$. Hence it cannot be used as it stands in functions and computer programs. The problem is easily solved however; by adding more terms and by finishing the series with a closing term calculated to give the correct result⁸⁹ when $N=1$. This, with some adjustment of Grover's empirical coefficient to minimise runout, gives the optimised formula:

$k_m = \ln(2\pi) - \frac{3}{2} - \frac{\ln(N)}{6N} - \frac{0.33084236}{N} - \frac{1}{120N^3} + \frac{1}{504N^5} - \frac{0.0011925}{N^7} + \frac{c_9}{N^9}$	10.18
--	--------------

Where the coefficient c_9 is calculated from the other terms in order to force k_m to zero when $N=1$; i.e.:

$$c_9 = - \left[\ln(2\pi) - \frac{3}{2} - 33084236 - \frac{1}{120} + \frac{1}{504} - 0.0011925 \right] = 0.000507000$$

The formula (10.18) gives a maximum absolute error of $\pm 0.000\,000\,013$ (i.e., 1.3×10^{-8}) for integer N . A reduction in precision is to be expected for non-integer values of $N < 5$, due to undulation of the error curve in that region, but the error is unlikely to extend into the 6th decimal place.

Note that, although the formula given above will return k_m to at least 6 decimal places; such extreme precision is not necessary in practice. Four decimal places is more than adequate, one reason being that the self, mutual, and internal inductance corrections for realistic wire will generally affect the total inductance at the level of a few % at most. Thus the determination of small correction terms even to a modest accuracy of about 1% moves the overall effect on precision to the level of a few parts in 10 000. Also, it must be understood that the formula (10.16)

⁸⁸ BS RP90, page 176, formula (31).

⁸⁹ Knight 2010. Optimisation of Grover's 1929 series formula. See also spreadsheet Rosa_km.ods, sheet 1.

from which the k_m correction originates is, in the first place, an approximation.

Accurate calculation of the mutual inductance coefficient

The methods given above for the calculation of Rosa's coefficient are extremely precise, which is to say that they return a value to a large number of decimal places; but precision does not imply accuracy (i.e. truth). The issue here is that the mutual inductance formula (10.15a) from which the correction is derived is actually a fairly crude approximation when the distance between filaments is a substantial proportion of the loop radius. Thus the only appraisal of Rosa's calculation procedure that we can give with any certainty at this point is that it will be exact for solenoids having zero pitch-angle. Such coils are, of course, physically impossible. We do know, however, that Rosa's original method is very good in practice for calculating the inductance of fine-pitch solenoids; a provenance albeit somewhat dependent on the fact that the corrections are second-order on an initial current-sheet calculation.

Prior to the development of the electronic computer, no one was in a position to improve matters. Even the precise evaluation of Rosa's original formula was impractical for large N until Grover carried out his investigation of Strasser's method (as described in Research Paper 90 and discussed above). There the matter was left moreover, perhaps because the explanations given were difficult to follow and the built-in approximations were forgotten. Having re-worked the derivation however, we can now see the problem. The mutual inductance correction is dependent on the pitch-to-diameter ratio (p/D). As the turns of the coil get further and further away from each other, the differences in the GMDs for mutual inductance calculation for current-sheet segments and wire loops become smaller and smaller. Hence the Rosa correction should diminish as the coil pitch increases, becoming zero when the pitch-angle reaches 90° (at which point, the turns are infinitely far apart). This means (bearing in mind that the corrections are subtracted from the current-sheet inductance) that the Rosa method always underestimates the mutual inductance contribution to the total inductance. Obviously, this error must be small for coils of reasonably fine pitch, but so far we have no idea of its magnitude.

This gap in our knowledge can of course be filled by developing a program to calculate the correction exactly. That will not lead directly to a practical calculation method because the program will be computationally intensive; but it will provide a source of data for the development of other methods.

A potentially exact expression for k_m was given earlier as (10.14).

$$k_{m\psi} = \frac{2}{\mu r N} \sum_{m=1}^{N-1} (N-m) \left[M(r; mp \exp(-\gamma_m)) - M(r; mp) \right]$$

Note however, that an additional subscript ψ has been appended as a reminder that the function we will be working with from now on is dependent on the pitch angle. Also note that since:

$$\frac{p}{D} = \pi \tan \psi$$

(see section 5) saying that a function is dependent on ψ is the same as saying that it is dependent on the coil pitch-to-diameter ratio.

For the exact calculation of the mutual inductances we can use Maxwell's elliptic integral formula (**M701.2**) (see **section 10b**) with $r_1 = r_2$. This gives:

$$M(r; g) = \mu r \frac{2}{\sqrt{\kappa_1}} [K(\kappa_1) - E(\kappa_1)]$$

with modulus:

$$\kappa_1 = \frac{\sqrt{1+(D/g)^2} - 1}{\sqrt{1+(D/g)^2} + 1}, \quad (\text{where } D = 2r).$$

The current sheet and round wire parts of the problem, or course, require different moduli because they have different geometric mean distances. Hence we will define:

For the round-wire coil: $g_w = m p$

and for the current sheet: $g_s = m p \exp(-\gamma_m) = g_w \exp(-\gamma_m)$

Thence:

$$\frac{g_w}{D} = \frac{mp}{D}, \quad \frac{g_s}{D} = \frac{g_w}{D} \exp(-\gamma_m),$$

$$\kappa_s = \frac{\sqrt{1+(D/g_s)^2} - 1}{\sqrt{1+(D/g_s)^2} + 1} \quad \text{and} \quad \kappa_w = \frac{\sqrt{1+(D/g_w)^2} - 1}{\sqrt{1+(D/g_w)^2} + 1}.$$

and we calculate $k_{m\psi}$ from the following expression:

$$k_{m\psi} = \frac{4}{N} \sum_{m=1}^{N-1} (N-m) \left[\frac{K(\kappa_s) - E(\kappa_s)}{\sqrt{\kappa_s}} - \frac{K(\kappa_w) - E(\kappa_w)}{\sqrt{\kappa_w}} \right] \quad (10.19)$$

A Basic routine that performs the calculation is given below. It will be found in the macro library of the spreadsheet **Rosa_km.ods** (versions after 9th August 2012).

Note that $p/D = 0$ is an illegal argument for the elliptic integral formula; and so in that case the program evaluates the function using expression (10.16). The current sheet GMDs are obtained using the function $gsm(m)$ (discussed above), and the K-E complete elliptic integral combination is obtained using the function $KminusE(k)$.

If experimenting with this routine in a spreadsheet, be sure to cut the results from the cells after calculation and put them back using 'paste special / numbers only'. If the function call is left active, the cell will be recalculated upon reopening the spreadsheet. If there are hundreds of instances, particularly of calculations for small pitch angle, that can take many hours; and there is no means of escape save for killing the spreadsheet program via the computer operating system.

```

Function kmref(ByVal N as long, pD as double) as double
' Reference function for Rosa-method mutual inductance coeff. Version 3.0, 9th Aug. 2012.
' Calls functions gsm and KminusE. N = no. of turns. pD = coil pitch / diam.
Dim m as long, term as double, sum as double
if N <= 1 then
  kmref = 0
elseif pD = 0 then
' Routine for p/D = 0 (use analytical value if N=2)
  if N = 2 then
    kmref = log(0.25) + 1.5
  else
    for m = 1 to N-1
      term = (N-m)*gsm(m)
      sum = sum + term
    next m
    kmref = 2*sum/N
  endif
else
' Routine for p/D > 0
Dim gf as double, Dss as double, Dsw as double
for m = 1 to N-1
  gf = exp( -gsm(m) )
  Dsw = 1/(m*pD)
  Dss = Dsw/gf
  ks = ( sqrt(Dss*Dss+1)-1 )/( sqrt(Dss*Dss+1)+1 )
  kw = ( sqrt(Dsw*Dsw+1)-1 )/( sqrt(Dsw*Dsw+1)+1 )
  term = 4*(1-m/N)*( sqrt(1/ks)*KminusE(ks) - sqrt(1/kw)*KminusE(kw) )
  sum = sum + term
  if term < 1E-15 then exit for
next
kmref = sum
endif
end function

```

In the routine for $p/D > 0$ (the second part of the program), a test is included in the summation loop. The procedure calculates only differences. Hence, with the summation performed in order of increasing GMD, it is permissible to jump out of the loop if the term being calculated turns out to be very small. Note however, that the exit criterion used above, which is that the summation can be aborted when terms become smaller than 10^{-15} , does not imply that the precision of the result will be to 15 decimal places.

Say, for example, that we abort a summation because the last term was almost negligibly less than 10^{-15} , but there are still a million terms to go. If the rate of convergence is very slow (which it most definitely is) then the termination criterion used only strictly guarantees an accuracy of 9 decimal places for $N \leq 10^6$ (because if the neglected 1 million terms are all nearly the same, they would have added-up to enough to roll-over the digit in the 9th decimal place). In practice however, convergence will always be somewhat faster than the most pessimistic assumption. It was found, for example, that adjusting the convergence criterion to be 3 orders of magnitude smaller than 10^{-15} caused no changes greater than 1 in the 11th decimal place for $N = 10^7$.

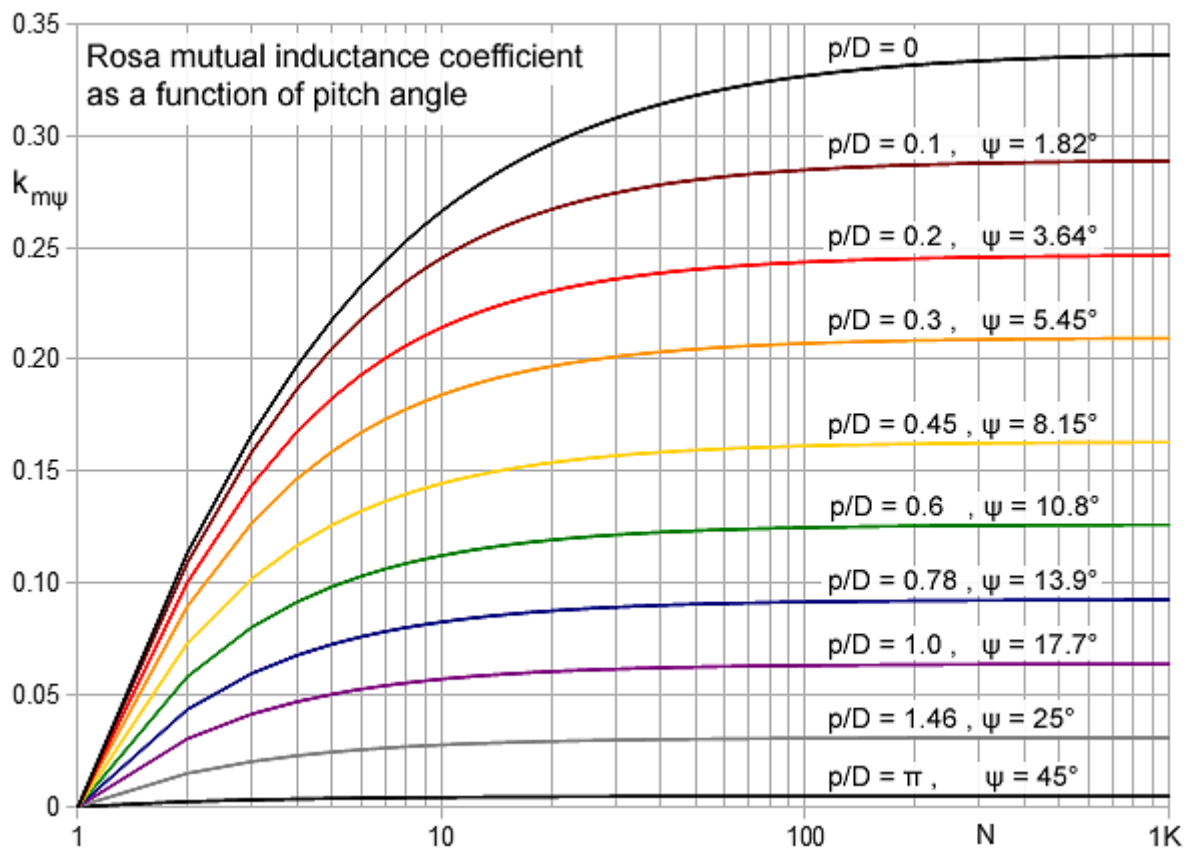
There is, of course, no practical requirement for the modelling coils of more than a million turns;

but the point is to find the asymptotic value for $k_{m\psi}$ at any pitch angle to a sensible precision (say, 6 decimal places). Also, we know when the summation has converged with the asymptotic value to 6 decimal places for $p/D = 0$, because we have the analytical value for $k_{m\infty}$ in that case (and calculations for small pitch angle are by far the slowest to converge). It transpires that $N = 10^7$ is more than enough to produce the asymptotic value to 6 places for $p/D = 0$. Hence that number of turns is comfortably adequate to give the asymptotic value to 6 decimal places in all instances. With the convergence criterion given, a calculation for $p/D = 0.0001$ and $N = 10^7$ using an Intel 2.66 GHz quad-core computer took 75 seconds. Calculations for larger pitch angles take less time than this; but calculations for $p/D = 0$ with large N take much longer.

Notice that the routine for $p/D = 0$ (the first part of the program) does not have a convergence criterion. That is because, for a hypothetical coil with all of the turns coincident (which is what that limit implies), the summation does not converge. All we can do to obtain the solution is carry on to the bitter end. Spot calculations, using a somewhat ancient workstation (as above) took 108 seconds per million turns (18 minutes for $N = 10^7$).

The $p/D = 0$ routine, incidentally, duplicates program functionality previously provided by Bob Weaver⁹⁰. The replication enabled the various coding methods to be checked against Bob's earlier calculations.

The results of a few hours of computer time are plotted below. The numbers can be inspected by viewing the spreadsheet [Rosa_km.ods](#) (sheet 5).



It has been mentioned already that, until we devise a correction for the axial current contribution to the total inductance; we are pushing the Rosa method beyond its range of applicability if we use it to calculate inductance for coils having p/D greater than about 0.1 ($\psi > 1.8^\circ$). The

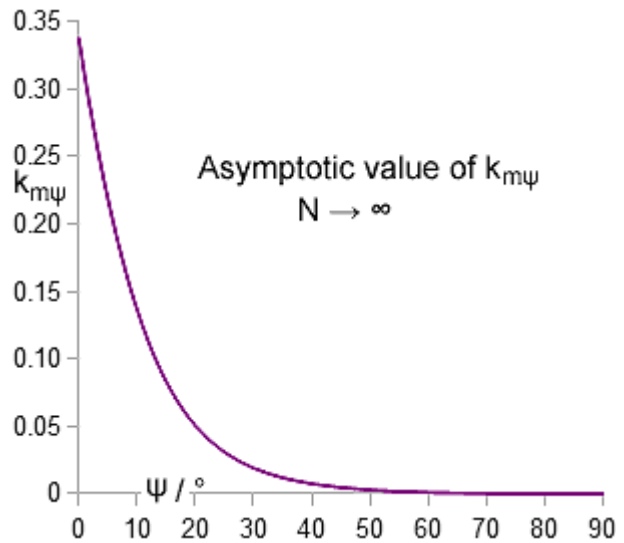
⁹⁰ See: <http://electronbunker.ca/CalcMethods2a.html>

uncomfortable discovery is therefore that the accuracy of the original Rosa coefficient is not particularly good for $0 < p/D < 0.1$.

The breakdown of Rosa's approximation can be seen in the graph on the right, which is a curve obtained by plotting $k_{m\psi}$ vs. ψ with $N = 10^7$. This shows that the asymptotic value of $k_{m\psi}$ decays (approximately) exponentially as the pitch increases. The decay law is given to a first approximation by the expression.:

$$k_{m\psi\infty} = \left[\ln(2\pi) - \frac{3}{2} \right] \exp\left(-1.6257 \frac{p}{D}\right) \pm 0.0028$$

This means that the rate of change $\partial k_{m\psi} / \partial \psi$ has its greatest magnitude in the range associated with practical coils.



An empirical function for the accurate mutual inductance coefficient

In order to calculate the Rosa mutual inductance coefficient in a manner that is accurate for all pitch angles, it is first necessary to have a way of calculating accurate values of $k_{m\psi\infty}$. One method for that is to add a correcting polynomial to the expression given above, and subject both the polynomial coefficients and the single empirical coefficient in the exponential bracket to simultaneous machine optimisation. Two examples are given below, becoming more complicated as the accuracy is improved.

The Basic function below, which was obtained by fitting the numbers used to produce the graph above, calculates $k_{m\psi\infty}$ to within ± 0.00008 , and is asymptotically correct as $p/D \rightarrow 0$ and $p/D \rightarrow \infty$.

```
Function kmpinf(byval pD as double)
' calculates asymptotic value of Rosa B correction as a function of p/D.
' Version 1.0 (+/- 0.00008), 9th Aug. 2012. pD = pitch / coil Diam.
Dim kminf as double
kminf = log(2*pi) - 1.5
if pD <= 0 then
  kmpinf = kminf
else
  Dim poly as double, x as double
  x = 1/(1+pD)
  Poly = log(1+pD)*(-0.0905*x^3 + 1.7565*x^4 - 2.1277*x^5 + 1.0967*x^6 - 0.0664*x^7 -
  -0.4186*x^8 - 0.363*x^9 + 0.499*x^10)
  kmpinf = kminf*exp(-2.3736*pD) + poly
endif
end function
```

Note that the correcting polynomial has the argument:

$$x = \frac{1}{1+p/D}$$

This is chosen because it is valid for any value of p/D . The factor $\ln(1 + p/D)$ that operates on the whole polynomial is zero when $p/D = 0$ and so forces the correction term to zero at that point. The expression used will actually return the correct value when $p/D = 0$, but the condition is trapped in the process of disallowing negative values of p/D .

The $k_{m\psi\infty}$ function shown below is accurate to within $\pm 0.000\ 000\ 75$

```

Function kmpinf(byval pD as double)
' Asymptotic value of Rosa B correction as a function of p/D.
' Version 2.1 (+/- 0.000 000 75 ). D W Knight, 10th Oct. 2012.
' pD = pitch / coil Diam.
Dim kminf as double
kminf = log(2*pi) - 1.5
if pD <= 0 then
  kmpinf = kminf
elseif pD > 72 then
  kmpinf = 0
else
  Dim x1 as double, x2 as double, x3 as double, x4 as double
  Dim x5 as double, x6 as double, x7 as double, poly as double, k as double
  x1 = pi*(1-1/sqr(1+2.185698*pD))
  x2 = pi*(1-1/sqr(1+2.557432*pD))
  x3 = pi*(1-1/sqr(1+1.995623*pD))
  x4 = pi*(1-1/sqr(1+2.117234*pD))
  x5 = pi*(1-1/sqr(1+2.82975*pD))
  x6 = pi*(1-1/sqr(1+2.320973*pD))
  x7 = pi*(1-1/sqr(1+1.315417*pD))
  poly = 0.02545714*sin(x1) + 0.021469*sin(2*x1) + 0.004551*sin(3*x2) _
  -0.00203236*sin(4*x3) -0.00067821*sin(5*x4) -0.00006356*sin(6*x4) _
  +0.00000823*sin(7*x5) +0.00009831*sin(8*x5) -0.00003428*sin(9*x6) _
  -0.00004050*sin(10*x6) -0.00000801*sin(11*x7) -0.00000511*sin(12*x7)
  k = kminf*exp(-2.295314*pD) +poly
  if k < 0 then k = 0
  kmpinf = k
endif
end function

```

The simple exponential decay function gives zero error as $p/D \rightarrow 0$ and as $p/D \rightarrow \infty$. Hence, it being sensible policy to use only fitting functions that preserve analytical values, a correcting polynomial that goes to zero at these limits is required. This is obtained by defining baseline functions (x_i) that vary between 0 and π as the pitch angle ψ varies between 0 and 90° . Hence any term defined as:

$$T_n = c_n \sin(n x_i)$$

will be zero both when $\psi = 0$ and when $\psi = 90^\circ$. The machine optimisation process tends to produce error curves that undulate. The baseline functions are chosen and adjusted so that the

undulations on a particular scale appear approximately sinusoidal. Multiplying the sine function that gives the same number of undulations by a suitable coefficient and then subtracting it from the error curve reduces the error. This minimisation process can be performed successively to reduce the error to any desired degree. In principle, each sine term can have its own baseline function. In practice, the baseline parameter is much less critical than the term coefficient, and so, in cases where the baseline parameters appear similar, several terms can be made to share the same baseline function. In practice also; the initial fitting formula is constructed by applying successive approximations, but the final formula is obtained by allowing all parameters to vary simultaneously.

>>>> work in progress

Empirical functions for $k_{m\psi}$

>>>>>

Some notes on the work in progress:

An empirical function for $k_{m\psi}$ is in development:

I have calculated a large array of data, with N increasing on going down a column, and pitch angle varying between 0 and 90 deg on going along a row. The contents of each column in turn have been fitted it to an extended version of my original 2006 empirical function for k_m (see spreadsheet [Rosa_kmp.ods](#)). The parameter sets for each fit were then put into an array in place of the column data.

I find that if I approach the fitting in the same way in each case, the parameters evolve smoothly. Hence I should be able to create empirical functions that produce the requisite parameter set for each input value of p/D .

The downside of this method is that the fitting needs to be very precise, because the errors in the parameters are transmitted into the final result. I could envisage doing a final fitting run on all of the data in one go, but since a good minimisation on 15 parameters and a couple of hundred data points can take a day, that would probably take weeks. Consequently, the best option appears to be to get the parameters right in the first round.

A note on the choice of fitting functions:

Legendre and Chebyshev polynomials provide a way of decorrelating the parameters to simplify fitting, but ultimately they can be decomposed into simple polynomials, and so don't reduce the number of parameters required, and certainly don't respect limiting values in the error function. I note however, that the Chebyshev polynomial terms can be represented as:

$$T_n(x) = \cos[n \arccos(x)]$$

and this has led me to a useful approach. For example: For the asymptote-value function $k_{minf}(\psi)$ (assuming that pitch angle is used as the argument), we have:

$$k_{\min f}(0) = \ln(2 \pi) - 3/2$$

and

$$k_{\min f}(\pi / 2) = 0$$

Then most of the fitting is done by the expression

$$k_{\min f}(\psi) = [\ln(2 \pi) - 3/2] \exp\{-c p/D\}$$

where

$$p / D = \pi \tan(\psi)$$

This leaves an error function that is analytically zero at both extremes. So I chose the form:

$$k_{\min f}(\psi) = [\ln(2 \pi) - 3/2] \exp\{-c p/D\} + \text{Polynom}(\psi)$$

with terms in the polynomial:

$$T_n(x) = C_n \sin(n \pi x_i)$$

where x_i is a function that varies between 0 and 1 as ψ varies between 0 and $\pi/2$. All such terms are, of course, anchored to zero at the ends, but the trick lies in the function x_i . The optimisation process tends to produce an error function that undulates. By choosing x_i to space the undulations evenly, x_i turns the error curve into an approximate sine wave on its scale. Hence, using an appropriate x_i function, and choosing the coefficient according to the height, the error curve can be successively minimised at each undulation frequency. In the most extreme cases, a different x_i function is required for every term, but the difference is sometimes quite small (or can be forced to be small) between adjacent terms and so sets of terms can share the same x_i to reduce the number of parameters.

The method works well on error curves that have a roughly sinusoidal appearance. In the event that the error dies off exponentially at the extremes, the polynomial can be multiplied by a Gaussian (this is reminiscent of the Hermite functions, which are obtained by multiplying Hermite polynomials by a Gaussian), or by a Lorentzian (or modified Lorentzian, of which much has been written in my article on internal impedance).

>>>>>>>

Comparison of Rosa and summation methods - error due to curvature.

>>>>>>>

11. Helicity

The formulae and calculation methods for static inductance outlined so far will produce results that are consistent with practically all traditional approaches to the problem of solenoid inductance calculation. There is however, a limitation on the method, which is almost invariably ignored.

We might say that, since we cannot measure inductance using DC, and since single-layer air-cored solenoids only start to become practical as inductors at radio frequencies; then we tend to test solenoid models against measurements made at relatively high frequencies. We then note that the actual inductance, after correction for leads and self capacitance, comes out a little lower than the value calculated using the average coil diameter. We can put this down to non-uniform current distribution; and then on the basis that the calculation is only a quasi-static approximation, satisfy ourselves that this is as good as can be obtained without introducing empirical adjustments.

In reality however, the decline of inductance with frequency is even greater than traditional modelling methods would have us believe, because those methods do not account for all of the static inductance. We might have suspected that from the curiously convenient fact that Nagaoka's coefficient depends only on ℓ/D . When something seems too good to be true, there is usually a reason.

>>>> work in progress

12. Combined static magnetic corrections

Having now collected and evaluated all of the necessary formulae and functions for the application of the modified Rosa-Nagaoka inductance calculation method; it is worth noting that it is possible to combine all of the solenoid static-field corrections into a single coefficient. This coefficient is useful when building lumped coil-inductance into general circuit models, and it is an important parameter in transmission-line inductor models.

We start by inserting the expression for current sheet inductance (7.1) into the modified form of Rosa's general inductance equation (10.4). This gives:

$$L = \left[\mu_{(x)} \frac{\pi r^2}{\ell} N^2 k_L \right] - \mu_{(x)} r N (k_{s(x)} + k_m) + L_i$$

L_i can be expressed in terms of the internal inductance factor Θ , which was given as part of the formula (6.1). Thus:

$$L_i = \ell_w \frac{\mu_{(i)}}{8\pi} \Theta$$

and substituting for ℓ_w using equation (5.2) gives:

$$L_i = \frac{2\pi r N}{\cos\psi} \frac{\mu_{(i)}}{8\pi} \Theta$$

Hence:

$$L = \left[\mu_{(x)} \frac{\pi r^2}{\ell} N^2 k_L \right] - \mu_{(x)} r N (k_{s(x)} + k_m) + \frac{\mu_{(i)} r N \Theta}{4 \cos\psi}$$

Now removing the factor $\mu_{(x)} \pi r^2 N^2 / \ell$ from the second and third terms gives:

$$L = \frac{\mu_{(x)} \pi r^2 N^2}{\ell} \left[k_L - \frac{\left(k_m + k_{s(x)} - \frac{(\mu_{(i)}/\mu_{(x)})\Theta}{4 \cos\psi} \right) \ell}{\pi r N} \right]$$

Using the substitution $\ell / r = 2 \ell / D$ then gives:

$L = \frac{\mu_{(x)} \pi r^2 N^2}{\ell} \left[k_L - \frac{2 \left(k_m + k_{s(x)} - \frac{(\mu_{(i)}/\mu_{(x)})\Theta}{4 \cos\psi} \right) \ell}{\pi D N} \right]$	[Henrys]	12.1
---	----------	------

Hence we can write a compact expression for lumped solenoid inductance as:

$L = \mu \frac{\pi r^2}{\ell} N^2 k_H$	[Henrys]	12.2
--	----------	------

where:

$k_H = k_L - \frac{2}{\pi}(\ell/D) \frac{\left(k_m + k_{s(x)} - \frac{(\mu_i/\mu_x)\Theta}{4 \cos \psi} \right)}{N}$	12.3
---	-------------

To make this expression more compact, we can also put:

$$z_k = \frac{2}{\pi}(\ell/D) \quad \text{and} \quad k_s = k_{s(x)} - \frac{(\mu_i/\mu_x)\Theta}{4 \cos \psi}$$

so that:

$$k_H = k_L - z_k \frac{(k_m + k_s)}{N} \tag{12.3a}$$

Note that equation (12.3) tells us that the round-wire corrections disappear when N is large or when ℓ/D is small; i.e., they are intermediate corrections and are not important for very short coils, and are only required for long coils if the number of turns is low in comparison to ℓ/D . The corrections are also generally small for coils of unexceptional design; the error incurred in neglecting them is usually less than 2%, and so it is often acceptable to ignore them in rough engineering calculations. When the round wire corrections are neglected, i.e., when it is assumed that $k_H = k_L$, this is known as the *current-sheet approximation*.

13. Apparent inductance and equivalent lumped inductance

At this point, we have collected all of the formulae needed to calculate the quasi-static magnetic contribution to the impedance seen when looking into the terminals of a solenoid inductor. In other words, we can calculate some inductance L (neglecting qualifying subscripts), and calculate a reactance $X_L = 2\pi fL$ from that. This reactance however, will only correspond to the actual or measured reactance at low frequencies.

As mentioned earlier, the static magnetic model does not allow for the finite velocity of electromagnetic waves travelling along the wire. This means that the impedance seen at the coil terminals will be different from that predicted by the method outlined so far (even after correction for parasitic contributions), and this can be accounted-for by evoking a hypothetical 'self-capacitance', C_L , in parallel with the coil. Hence the actual reactance (neglecting losses) will be:

$$X'_L = X_L // X_{CL} = \frac{X_L X_{CL}}{X_L + X_{CL}}$$

where: $X_{CL} = \frac{-1}{2\pi f C_L}$

The self-capacitance can be neglected at low frequencies because the magnitude of X_{CL} is large when f is small, in which case $(X_L + X_{CL}) \approx X_{CL}$ and so $X'_L \approx X_L$.

In early publications (before the widespread availability of electronic computers), it was generally assumed that the "true" inductance of a coil (after correction for self-capacitance) is constant. This is reasonable when working to an accuracy of a few percent, but it leads to an unfortunate and potentially confusing usage; which is that of referring to the pure inductance component as the "low-frequency inductance". Since we have added a number of high-frequency corrections to the magnetic model, the term 'low-frequency inductance' should now be taken to mean 'the inductance when the skin depth is greater than the wire radius', i.e., the inductance below the LF-HF transition frequency as discussed in [section 6a](#). A preferable term for the quasi-static magnetic component is; ***equivalent lumped inductance***: and the inductance obtained by dividing the measured reactance by $2\pi f$ should be referred to as the ***apparent inductance***.

14. Solenoid inductance calculation vs. measurement

>>> work in progress



99.

Revision History:

- Original HTML article started in 2005 and made public in 2007.
- 0.01 12th June 2012. Begin porting HTML into odt and pdf formats. Preface added. Editing done up to section 6d.
 - 0.02 18th June 2012. + Section 6e.
 - 0.03 19th June 2012. + Section 6f.
 - 0.04 19th June 2012. + Sections 7, 7a, + 7b started.
 - 0.05 23rd June 2012. + up to section 9. New content being added.
 - 0.06 23rd June 2012. Porting fairly complete.
 - 0.07 27th June 2012. Sections 5a, 5b added. Decision to split document into two.
 - 0.08 1st July 2012.
 - 0.09 4th July 2012. Foreword. Expanding work on Rosa corrections.
 - 0.10 5th July 2012. Rosa corrections.
 - 0.11 6th July 2012.
 - 0.12 6th July 2012. Summation method.
 - 0.13
 - 0.14
 - 0.15
 - 0.16
 - 0.17
 - 0.18 24th Oct. 2012. Notes on the state of the investigation prior to suspension of work. (DWK underwent surgery for replacement aortic valve on 2nd Nov. 2012).
 - 0.19 24th Jan. 2013. Minor corrections.
 - 0.20 4th Feb. 2016. Changes to the introduction. Reformatted equations using OO Math. Added section 5c. Updated section 6 for PACAML. Added program code for Lintern.