Analytical approximations for the inverse Langevin function via linearization, error approximation, and iteration

Roy M. Howard

Abstract
This paper details an analytical framework, based on an intermediate function, which facilitates analytical approximations for the inverse Langevin function—a function without an explicit analytical form. The approximations have relative error bounds that are typically much lower than those reported in the literature and which can be made arbitrarily small. Results include convergent series expansions in terms of polynomials and sinusoids which have modest relative error bounds and convergence properties but are convergent over the domain of the inverse Langevin function. An important advance is to use error approximations, and then iterative relationships, which allow simple initial approximations for the inverse Langevin function, with modest relative errors, to generate approximations with arbitrarily low relative errors. One example is that of an initial approximating function, with a relative error bound of 0.00969, which yields relative error bounds of $2.77 \times 10^{-6}$ and $2.66 \times 10^{-16}$ after the use of first-order error approximation and then first-order iteration. Functions with much lower error bounds are possible and are detailed. First- and second-order Taylor series can be used to simplify the error- and iteration-based approximations.

Keywords
Inverse Langevin function · Inverse function · Linearization · Error approximation · Legendre basis set · Taylor series

Introduction
The Langevin function arises in diverse contexts including the classical model of paramagnetism (Langevin 1905) and the ideal freely jointed chain model (e.g., Fiasconaro and Falo 2018). In the ideal jointed chain model, as illustrated in Fig. 1, a chain comprising of $n$ identical rigid segments and, in an equilibrium energy environment, is such that independent movement of each segment is valid with each segment moving independently and with an angle consistent with a uniform distribution on the interval $(\pi, \pi]$. With an applied point force, the expected chain extension, normalized, is given by the Langevin function (e.g., Iliafar et al. 2013), which is defined according to

$$L(x) = \begin{cases} \coth(x) - \frac{1}{x} & x \in (0, \infty) \\ 0 & x = 0 \end{cases}$$

and whose graph is shown in Fig. 2 for the case of $x > 0$. In this equation, the non-descriptive variables of $x$ (representing the normalized force for the freely jointed chain) and $L$ (representing the normalized chain extension) are used. The inverse Langevin function, denoted $L^{-1}$, is shown in Fig. 2 for the case of $y > 0$, and is of interest as, for the freely jointed chain case, it defines the normalized force required for a given expected normalized chain extension. The freely jointed chain model, and chain-based network models, have applications, for example, in rubber elasticity (e.g., Ehret 2015), the micromodelling of polymers (e.g., Arruda and Boyce 1993, Boyce and Arruda 2000, Hessain and Steinmann 2012) and the biophysics of macromolecules (e.g., Holzapfel 2005).

As the Langevin function and the inverse Langevin function are antisymmetric, it is sufficient to consider these functions, respectively, over the intervals $[0, \infty)$ and $(0, 1)$. As is evident in Fig. 2, the inverse Langevin function has a singularity at the end point of the interval which complicates analysis and an explicit analytical expression for this function is an unsolved problem. Finding an approximate analytical expression has received attention over a considerable period of time, e.g., Kuhn and Grün (1942), Treloar (1954, 1975), and Cohen (1991), and improved approximations have been detailed in recent years: Itskov et al. (2012), Nguessong et al. (2014),
Darabi and Itskov (2015), Jedynak (2015), Kröger (2015), Marchi and Arruda (2015), Rickaby and Scott (2015), Petrosyan (2017), Jedynak (2018) and Marchi and Arruda (2019). These papers provide overviews of the approaches used and these include the use of Taylor series, several custom series, and the dominant approach of Padé approximants including the use of minimax optimization techniques. The Cohen approximation (Cohen 1991) has a maximum relative error of 0.0494 and improved approximations with maximum relative errors approaching $10^{-4}$ (Marchi and Arruda 2015, Jedynak 2018) have been reported with the recent work of Marchi and Arruda (2019) detailing approximations with maximum relative errors less than $10^{-6}$. Benitez and Montáns (2018) utilized discretization and interpolation via cubic splines to obtain highly accurate numerical results; a maximum relative error of the order of $10^{-11}$ can be achieved with the use of 10,000 points.

This paper details approaches for finding analytical approximations to the inverse Langevin function that can be made arbitrarily accurate. The approach taken is distinctly different from those reported in the literature and is based on determining an intermediate function that linearizes the Langevin function. This approach was initially motivated by the potential of higher order spline functions for function approximation (Howard 2019). The use of an intermediate function facilitates, first, the definition of convergent series for the inverse Langevin function over its complete domain—an unsolved problem. Second, the use of an intermediate function facilitates error approximation and functional iteration. When combined, these lead to approximations for the inverse Langevin function which can be made arbitrarily accurate over its complete domain. For example, a first-order iteration based on a simple approximating function has relative error bounds of better than $6 \times 10^{-8}$, $3 \times 10^{-16}$, and $3 \times 10^{-30}$ depending on the order of error approximation used. Higher-order error approximation and/or higher levels of iteration lead to significantly lower relative error bounds. The error-based approximations can be simplified by using first- and second-order Taylor series.

The structure of the paper is illustrated in Fig. 3. Section “Approximation via an intermediate function” details the theory underpinning the use of an intermediate function for finding approximations to the inverse Langevin function. This section includes a direct linearization approach which results in a non-convergent series. Section “Convergent series for inverse Langevin function” details convergent series for the inverse Langevin function which are based on utilizing suitable basis sets to approximate the error arising from linearization. Section “Improved approximation via error approximation” details how error approximation, of different orders, can be used to yield improved approximations. Section “Improved approximation via iteration” details approximations, with lower relative error bounds, that arise from function iteration. Associated results are detailed in Section “Results: Error approximation and function iteration”.

Section “Simplified approximation via Taylor series” details how the error-based approximations can be simplified by using the first- and second-order Taylor series. Conclusions are provided in Section “Conclusion”.

**Notation**

For a function $f$ defined over the interval $[\alpha, \beta]$, an approximating function $f_A$ has a relative error, at a point $x_1$, defined according to

$$\text{re}(x_1) = 1 - \frac{f_A(x_1)}{f(x_1)} \quad (2)$$

![Graph of the Langevin and inverse Langevin functions](image)
The relative error bound for the approximating function over the interval \([\alpha, \beta]\) is defined according to
\[
\text{re} = \max\{\text{re}(x_1) : x_1 \in [\alpha, \beta]\} \tag{3}
\]

The notation \(x \in \{0^+, 1^−\}\) is used with the meaning, respectively, of the limits as \(x\) approaches zero from above and one from below. The notation \(f^{(k)}(x) = \frac{df}{dx} f(x)\) is used in some instances.

Mathematica has been used to facilitate analysis and to obtain numerical results. In general, relative error results for approximations to the inverse Langevin function have been obtained by sampling the interval \((0, 1)\) with a resolution of 0.001.

**Approximation via an intermediate function**

The Langevin function is defined over the interval \([0, \infty)\) and the determination of an approximation to the inverse of this function is facilitated if a monotonically increasing intermediate function, denoted \(f\), can be defined which, as illustrated in Figure 4, changes from zero to infinity as its argument changes from zero to one and creates an approximately linear function \(L[f(x_1)]\) with respect to \(x_1\).

**Theorem 1 Approximation via an intermediate function**

Consider the case of a monotonically increasing intermediate function \(f\) which, as illustrated in Fig. 4, is such that \(L(x) = L[f(x_1)]\) for \(x_1 = [0, 1), x \in [0, \infty)\) and where \(L[f(x_1)]\) is close to being linear with a slope close to unity for \(x_1 \in [0, 1)\). When the linearity is such that
\[
L[f(x_1)] = x_1 + \varepsilon(x_1) \quad x_1 \in [0, 1), |\varepsilon(x_1)| \ll 1
\]

then the inverse Langevin function can be approximated according to
\[
\begin{align*}
\text{or } L[f(x_1)] &= x_1 \left[1 + \varepsilon(x_1)\right] \quad x_1 \in [0, 1), |\varepsilon(x_1)| \ll 1
\end{align*}
\]

For both cases, a simple as possible approximation for the inverse Langevin function is...
\[ x = L^{-1}(y) = f(y) \quad y \in [0, 1). \]  

(6)

As the inverse Langevin function is antisymmetric

\[ x = L^{-1}(y) = -f(-y) \quad y \in (-1, 0) \]  

(7)

**Proof**

The relationships

\[ y = L(x) = L[f(x_1)] = x_1 + \varepsilon(x_1) \]  

(8)

imply \( x_1 = y - \varepsilon(x_1) \) and

\[ x = L^{-1}(y) = f(x_1) = f[y - \varepsilon(x_1)] \]  

(9)

With the assumption of \( \varepsilon(x_1) \ll x_1 \), it follows that \( y \approx x_1 \) and the required relationship

\[ x = L^{-1}(y) = f[y - \varepsilon(y)] \]  

(10)

then follows. The alternative relationships follow in a similar manner with the first order approximation

\[ x = L^{-1}(y) = f(x_1) = f \left[ \frac{y}{1 + \varepsilon(x_1)} \right] \approx f[y(1 - \varepsilon(x_1))] \]  

(11)

being made.

\[ L[f(x_1)] \approx \begin{cases} 
0 & \text{if } \frac{1}{f(x_1)} \to 1 \\
\frac{1}{2} f(x_1) - \frac{1}{3!} f^2(x_1) + \frac{1}{4!} f^3(x_1) - \frac{1}{5!} f^4(x_1) + \cdots & \text{if } x_1 \to 0 
\end{cases} \]  

(12)

**Implication**

These results imply that to find approximations to the inverse Langevin function, it is sufficient to find functions \( f \) which effectively linearize the Langevin function such that Eq. 4 holds.

**Determining intermediate function**

As is clear from Theorem 1, the better the intermediate function \( f \) is in linearizing \( L[f(x_1)] \), the better the approximation \( L^{-1}(y) \approx f(y) \) is for the inverse Langevin function. To determine appropriate intermediate functions, a useful starting point is to consider the rates of change of \( L[f(x_1)] \) at the points \( x_1 \in \{0^+, 1^-\} \). The following theorem facilitates this.

**Theorem 2 Approximations for transformed Langevin function at 0,1**

For the case where \( f(x_1) \) changes monotonically between zero and infinity as \( x_1 \) changes between zero and one, the following are valid approximations for the transformed Langevin function \( L[f(x_1)] \) for, respectively, the right neighbourhood of the point zero and the left neighbourhood of the point one:

**Proof**

The proof is detailed in Appendix 1.

**Transformed Langevin function with unity rate of change at interval end points**

The goal is for the transformed function \( L[f(x_1)] \) to be linear with unity slope over the interval \([0, 1]\). A starting requirement is for \( L[f(x_1)] \) to have unity slope at \( x \in \{0^+, 1^-\} \).

**Theorem 3 Transformed Langevin function with unity rate of change**

For the case where \( f(x_1) \) changes monotonically between zero and infinity as \( x_1 \) changes between zero and one, and in a manner such that the approximations stated in Theorem 2 are valid, it follows that \( L[f(x_1)] \) has unity slope at \( x_1 = 0^+ \), and at \( x_1 = 1^- \), when

\[ f^{(1)}(x_1)|_{x_1=0^+} = 3 \quad f^{(1)}(x_1)|_{x_1=1^-} = 0 \]  

(13)

**Proof**

The proof is detailed in Appendix 2.

**Notes**

These constraints allow the coefficients in a valid function form for \( f(x_1) \) to be simultaneously solved (see Kröger 2015 for the Padé approximant case). Petrosyan (2017) solves for a valid function approximation via combining two function forms which satisfy, separately, the two asymptotic constraints. The result that the derivative of an approximation for the inverse Langevin function should have a rate of change of 3 at the origin is well known, widely used (e.g., Darabi and Itskov 2015), and consistent with a Taylor series expansion for this function, e.g., Itskov et al. 2012.
Low-order forms for intermediate function

There are many choices for a function that goes from zero to infinity as its argument goes from zero to one, for example \( f(x_1) = \ln \left[ \frac{1}{x_1}, (1 + \alpha_1 x_1 + \ldots) \right] \). However, this form does not satisfy the constraints specified in Theorem 3. The following polynomial form

\[
f(x_1) = \frac{3 x_1 (1 + k_1 x_1 + k_2 x_1^2 + k_3 x_1^3 + \ldots)}{1-x_1}
\]

(14)

has potential with the first-, second-, and third-order expressions being solved consistent with the constraints specified in Theorem 3 to yield:

\[
f(x_1) = \frac{3 x_1 (1-2 x_1/3)}{1-x_1} \quad \text{1st order}
\]

\[
f(x_1) = \frac{3 x_1 (1 + k_1 x_1 + k_2 x_1^2)}{1-x_1} \quad k_2 = -k_1 - \frac{2}{3} \quad \text{2nd order}
\]

(16)

\[
f(x_1) = \frac{3 x_1 (1 + k_1 x_1 + k_2 x_1^2 + k_3 x_1^3)}{1-x_1} \quad k_3 = -k_1 - k_2 - \frac{2}{3} \quad \text{3rd order}
\]

(17)

Examples of second- and third-order functions that are close to optimum in the sense of minimizing the magnitude of the maximum relative error, over the interval [0, 1) in the approximation \( L^{-1}(y) \approx f(y) \) for the inverse Langevin function, are

\[
f(x_1) = \frac{3 x_1}{1-x_1} \left[ 1 - \frac{24 x_1}{25} + \frac{22 x_1^2}{75} \right]
\]

(18)

\[
f(x_1) = \frac{3 x_1}{1-x_1} \left[ 1 - \frac{459 x_1}{500} + \frac{47 x_1^2}{250} + \frac{19 x_1^3}{300} \right].
\]

(19)

These two functions lead to maximum relative error magnitudes in the approximation \( L^{-1}(y) \approx f(y) \), respectively, of 0.00969 and 0.00583.

Higher-order forms for intermediate function

It is the case that the functions \( f \) defined by Eqs. 15, 18, and 19 have the required properties for the intermediate function, namely, being monotonically increasing between zero and infinity as its argument changes between zero and one and being such that \( L(f(x_1)) \) has unity rate of change as the points of zero and one are approached. These functions are low-order approximations and higher-order approximations are possible as stated in the following theorem:

**Theorem 4 Higher order forms for intermediate function**

The coefficients in the general form for the intermediate function \( f \), defined by Eq. 14, can be solved based on imposing the constraints of zero rate of change for higher-order derivatives at the points zero and one according to

\[
\frac{d}{dx} L[f(x_1)] = 1 \quad \text{for } x_1 \in \{0^+, 1^-\}
\]

(20)

\[
\frac{d^n}{dx^n} L[f(x_1)] = 0 \quad \text{for } x_1 \in \{0^+, 1^-\}, k \in \{2, 3, \ldots\}
\]

(21)

\[
to \text{yield, potentially, increasingly linear forms for } L[f(x_1)].
\]

The results for first-, second-, and sixth-order approximations, respectively, are:

\[
f(x_1) = \frac{3 x_1}{1-x_1} \left[ 1 - \frac{2 x_1}{3} \right] \quad \frac{d}{dx} L[f(x_1)] = 1, \quad x_1 \in \{0^+, 1^-\}
\]

(22)

\[
\frac{d^2}{dx^2} L[f(x_1)] = 0, \quad x_1 \in \{0^+, 1^-\}
\]

\[
f(x_1) = \frac{3 x_1}{1-x_1} \left[ 1 - \frac{2 x_1}{3} \right] \quad \frac{d}{dx} L[f(x_1)] = 1, \quad x_1 \in \{0^+, 1^-\}
\]

(21)

\[
f(x_1) = \frac{3 x_1}{1-x_1} \left[ 1 - \frac{x_1}{75} + \frac{2}{5} - \frac{2 x_1}{3} \right] \quad \frac{d}{dx} L[f(x_1)] = 1, \quad x_1 \in \{0^+, 1^-\}
\]

(22)

\[
\frac{d^2}{dx^2} L[f(x_1)] = 0, \quad x_1 \in \{0^+, 1^-\}
\]

\[
\frac{d^3}{dx^3} L[f(x_1)] = 0, \quad x_1 \in \{0^+, 1^-\}
\]

(23)

The third-, fourth-, fifth-, seventh-, and eighth-order approximations are detailed in Appendix 3.

**Proof**

Mathematica was used to solve for the coefficients by using the approximations stated in Theorem 2.

**Results**

The linearity of the function \( L[f(x_1)] \) is illustrated in Fig. 5 for the first-order case as specified by Eq. 21. Higher orders closely approximate the linear line with unity slope. The relative errors in the various orders of approximation for the inverse Langevin function, based on \( L^{-1}(y) \approx f(y) \), where \( f(y) \) is specified by Eqs. 21 to 23 and Eq. 122 to 126, are shown in Fig. 6. The relative error bounds, respectively, are 0.13, 0.0264, 0.0137, 0.0106, 6.06 \times 10^{-3}, 2.61 \times 10^{-3}, 3.15 \times 10^{-3}, and 6.18 \times 10^{-3} for the first- to eighth-order approximations. These results indicate that the series does not converge and with the sixth-order approximation yielding the smallest relative error bound of 2.61 \times 10^{-3}.

**Notes**

The polynomial-based approximations, as specified by Eqs. 21 to 23 and Eq. 122 to 126, and resulting from the linearization constraints specified by Eq. 20, show modest convergence and then divergence with respect to providing an
approximation $L^{-1}(y) \approx f(y)$ for the inverse Langevin function. The relative error bound of $2.61 \times 10^{-3}$ for the optimum sixth-order series, as specified by Eq. 23, is between that of the inverse Langevin approximations proposed by Kröger 2015, with a bound of $2.8 \times 10^{-3}$, and the non-linear function proposed by Petrosyan 2017, with a bound of $1.8 \times 10^{-3}$. Optimized Padé approximants, e.g., Marchi and Arruda (2019), yield better convergence with a $[4/4]$ expression having a relative error bound of $1.8 \times 10^{-4}$.

Simulation results indicate that the second-, third-, and fourth-order approximations, as specified by Eqs. 22, 122, and 123, represent lower bounds for the inverse Langevin function over the interval $[0, 1)$.

An unsolved problem is a convergent series for the inverse Langevin function. The use of an intermediate function allows convergent series to be defined for the inverse Langevin function and this is detailed in the following section.

**Convergent series for inverse Langevin function**

Consider the approximation for the inverse Langevin function $L^{-1}(y) \approx f(y)$, which is based on an intermediate function $f$ that satisfies $L[f(x_1)]=x_1 + \varepsilon(x_1)$, $|\varepsilon(x_1)| \ll x_1$. The relative error in the approximation is

$$\varepsilon(y) = 1 - \frac{f(y)}{L^{-1}(y)}$$

(24)

Consider an associated relative error function

$$\varepsilon_1(y) = \frac{L^{-1}(y)}{f(y)} - 1 = \frac{\varepsilon(y)}{1 - \varepsilon(y)} \approx \varepsilon(y)$$

(25)

where the approximation is valid when $|\varepsilon(y)| \ll 1$. When the approximation $L^{-1}(y) \approx f(y)$ is good over the domain of $[0, 1)$, the associated relative error is well defined and can be approximated arbitrarily accurately using a suitable basis set. This is the basis for a convergent series for the inverse Langevin function as detailed in the following theorem.

**Theorem 5 Convergent series for inverse Langevin function**

Consider an intermediate function $f$, which leads to the error function $\varepsilon_1(y) = \frac{L^{-1}(y)}{f(y)} - 1$ being a smooth-bounded integrable function on the interval $[0, 1)$. For this case, an orthonormal basis set $\{b_0, b_1, \ldots\}$ for the interval $[0,1]$ leads to the convergent series

$$L^{-1}(y) = f(y) \left[ 1 + c_0 b_0(y) + c_1 b_1(y) + \ldots \right] \quad y \in [0, 1)$$

(26)
where \( b_i \) \( i \in \{0, 1, 2, \ldots\} \), is the \( i \)th orthonormal basis function and the \( i \)th coefficient, \( c_i \), is defined according to

\[
c_i = \int_0^1 \left[ \frac{L^{-1}(y)}{f(y)} - 1 \right] b_i^*(y) dy \quad i \in \{0, 1, \ldots\} \tag{27}
\]

Here \( b_i^* \) is the conjugate of \( b_i \).

Proof

Convergence is guaranteed because, \( \varepsilon_1(y) \) has been assumed to be a smooth, integrable function and the set of functions \( b_i \), \( i \in \{0, 1, 2, \ldots\} \), has been assumed to be an orthonormal basis set for the interval [0, 1). A general reference is Debnath and Mikusinski 1999.

Suitable basis sets and results

Suitable basis sets for the interval [0,1] include the Legendre and the standard sinusoidal basis set and these are detailed in Appendix 4.

To illustrate the potential for a convergent series, consider the function \( f \) defined by Eq. 18:

\[
f(y) = \frac{3y}{1-y} \left[ 1 - \frac{24y}{25} + \frac{22y^2}{75} \right]. \tag{28}
\]

The associated error function \( \varepsilon_1 \) is shown in Fig. 7.

Results: Legendre basis set

The Legendre basis set leads to a polynomial series approximation for \( L^{-1}(y) \) according to

\[
L^{-1}(y) = \frac{3y}{1-y} \left[ 1 - \frac{24y}{25} + \frac{22y^2}{75} \right] \left[ 1 + c_0 b_0(y) + c_1 b_1(y) + c_2 b_2(y) + \ldots \right] \tag{29}
\]

where \( c_0 = 9.65679 \times 10^{-4} \), \( c_1 = -2.18413 \times 10^{-3} \), \( c_2 = -4.19473 \times 10^{-3} \), \( c_3 = 8.87659 \times 10^{-4} \), \( c_4 = 3.24574 \times 10^{-3} \), \( c_5 = 7.98418 \times 10^{-4} \), \( c_6 = -2.72786 \times 10^{-4} \), \( c_7 = -3.56513 \times 10^{-4} \), \( c_8 = -1.38877 \times 10^{-4} \), \( c_9 = 1.33846 \times 10^{-5} \) and \( c_{10} = 5.20035 \times 10^{-5} \), etc. Simplification leads to the standard polynomial forms. For example, the approximations of orders 4 and 10 are:

\[
L^{-1}(y) = \frac{3y}{1-y} \left[ 1 - \frac{24y}{25} + \frac{22y^2}{75} \right] \left[ 1 + a_0 + a_1 y + a_2 y^2 + \ldots + a_{10} y^{10} \right]
\]

\[
a_0 = 1.114252 \times 10^{-4} \quad a_1 = -5.41586 \times 10^{-3} \quad a_2 = 0.695079 \quad a_3 = -5.789972
\]

\[
a_4 = 36.861391 \quad a_5 = -143.92813 \quad a_6 = 343.76325 \quad a_7 = -507.39216
\]

\[
a_8 = 449.12483 \quad a_9 = -217.30958 \quad a_{10} = 44.029235
\]

Orders zero to three are not of sufficiently high order to approximate the error function \( \varepsilon_1(y) \). The relative error bounds for the fourth- to tenth-order approximations are \( 2.8 \times 10^{-3} \), \( 2.6 \times 10^{-3} \), \( 1.6 \times 10^{-3} \), \( 5.1 \times 10^{-4} \), \( 4.0 \times 10^{-4} \), \( 3.5 \times 10^{-4} \), and \( 1.2 \times 10^{-4} \). The relative error bound for the twentieth-order series is \( 3.2 \times 10^{-6} \). This is an error level that is better than a [6/6] Padé approximation specified by Marchi and Arruda (2019) (relative error bound of \( 6.37 \times 10^{-6} \)). The graphs of selected relative errors are shown in Fig. 8.

Results: sinusoidal basis set

The sinusoidal basis set leads to the series for \( L^{-1}(y) \) according to

\[
L^{-1}(y) = \frac{3y}{1-y} \left[ 1 - \frac{24y}{25} + \frac{22y^2}{75} \right] \left[ 1 + c_1 b_1(y) + c_2 b_2(y) + \ldots \right]
\]

where \( c_1 = 3.91964 \times 10^{-3} \), \( c_2 = 3.07689 \times 10^{-3} \), \( c_3 = -6.46390 \times 10^{-3} \), \( c_4 = 1.78090 \times 10^{-3} \), \( c_5 = -1.07345 \times 10^{-3} \), \( c_6 = 5.46701 \times 10^{-4} \), \( c_7 = -1.24337 \times 10^{-4} \), \( c_8 = -7.75576 \times 10^{-5} \), \( c_9 = -4.42517 \times 10^{-5} \) and \( c_{10} = -2.43545 \times 10^{-5} \) etc. Thus:
Orders zero to two are not of sufficiently high order to approximate $\varepsilon_1(y)$. The relative error bounds for the third to tenth order approximations are $2.9 \times 10^{-3}$, $1.2 \times 10^{-3}$, $3.0 \times 10^{-4}$, $3.4 \times 10^{-4}$, $2.4 \times 10^{-4}$, $1.8 \times 10^{-4}$, $1.5 \times 10^{-4}$, and $1.3 \times 10^{-4}$. The relative error bound for a twentieth-order series is $3.3 \times 10^{-5}$. Selected relative errors are shown in Fig. 9.

Notes

The rate of convergence with both the Legendre and sinusoidal basis sets is modest. In general, a starting approximation for the inverse Langevin function, with a lower relative error bound, is likely to have a more oscillatory error function that needs to be approximated and this leads to more terms in the series approximation before a better approximation for the inverse Langevin function arises. For example, a tenth-order Legendre series approximation based on the initial approximating function given by Kröger (2015)

$$L^{-1}(y) \approx \frac{3y}{1-y^2} \left[ \frac{24y}{25} \left( 1 - \frac{22y^2}{75} \right) \right] \times \left[ 1 + 3.91964 \times 10^{-3} \sin(\pi y) + 3.07689 \times 10^{-3} \sin(2\pi y) - 6.46390 \times 10^{-3} \sin(3\pi y) + 1.78090 \times 10^{-3} \sin(4\pi y) - 1.07345 \times 10^{-3} \sin(5\pi y) + 5.46701 \times 10^{-3} \sin(6\pi y) + \ldots \right]. \quad (33)$$

leads to an improved approximation with a relative error bound of $1.2 \times 10^{-4}$ for a tenth-order series. This relative error bound is close to that achieved by the approximation specified by Eq. 31, which is based on an initial starting approximation with a higher relative error.

The coefficients in the convergent series for the inverse Langevin function can be computed with arbitrarily high
accuracy. Importantly, the series converges over the domain \([0, 1]\). This is in contrast with a Taylor series, e.g., Itskov et al. (2012), where convergence breaks down around 0.904. Dargazany (2013) details an algorithm that facilitates the efficient evaluation of the higher-order derivatives involved in a Taylor series approximation.

The following section details approaches with significantly lower relative error bounds.

**Improved approximation via error approximation**

Consider the result, consistent with Theorem 1, for an intermediate function \(f\) which is such that \(L[f(x_1)]\) is close to being linear:

\[
y = L(x) = L[f(x_1)] = x_1 + \epsilon(x_1)
\]

\[
x = L^{-1}(y) = f[y - \epsilon(x_1)] \quad x_1, y \in [0, 1).
\]

If an approximation to the error function \(\epsilon\) can be made, then an improved estimate for the inverse Langevin function results. Specifically, for \(y\) fixed, an approximation to the error \(\epsilon\) at the point \(x_1\), which is in terms of \(y\), is required.

Consider the general case, and the illustration shown in Fig. 10, for a function \(q\) which is close to being linear with a slope close to unity such that \(q(x_1) = x_1 + \epsilon(x_1)\) is an appropriate model. The goal is to find approximations to the error \(\epsilon(x_1) = q(x_1) - x_1\), at the point \(x_1\), which are in terms of \(q(y_1)\) and \(y_1\) where \(y_1 = q(x_1), x_1 = q^{-1}(y_1)\).

**Theorem 6 Zero-, first-, and second-order error approximations**

For a function \(q\), as illustrated in Fig. 10, which is such that the model \(q(x_1) = x_1 + \epsilon(x_1), |\epsilon(x_1)| < x_1\), is valid, approximations for \(x_1 = q^{-1}(y_1)\) and \(\epsilon(x_1) = q(x_1) - x_1\) are, first, for a zero-order error approximation:

\[
x_1 \approx q^{-1}(y_1) - q(y_1) \approx q(x_1) - y_1.
\]

Second, for a first-order error approximation:

\[
x_1 \approx x_1 = y_1 - \frac{q(y_1) - y_1}{q'(y_1)} \quad \epsilon(x_1) \approx \frac{q(y_1) - y_1}{q'(y_1)}.
\]

Third, for a second-order error approximation:

\[
x_1 \approx x_1 = y_1 - \frac{q(y_1) - y_1}{q'(y_1)} - \frac{q(y_1) - y_1}{q'(y_1)} \cdot \frac{q(y_1) - y_1}{q'(y_1)}
\]

where the latter approximations assume \(\frac{2q(y_1)q'(y_1)}{[q'(y_1)]^2} \ll 1\).

**Proof**

The proofs for these approximations are detailed in Appendix 5.

**Utilizing zero-, first-, and second-order error approximations**

For the case where \(L[f(x_1)]\) is approximately linear according to \(L[f(x_1)] = x_1 + \epsilon(x_1)\), it is the case, according to Theorem 1, that \(L^{-1}(y) = [y - \epsilon(x_1)]\). The approximations for \(\epsilon(x_1)\), detailed in Theorem 6, lead to the following approximations, with improved accuracy, for the inverse Langevin function:
Theorem 7 Zero-, first-, and second-order error-based approximations

Consider an initial approximation to the inverse Langevin function of $L^{-1}(y) \approx f(y)$ which is based on an intermediate function $f$ that linearizes the Langevin function such that the model $L[f(x_1)] = x_1 + \varepsilon(x_1)$ is valid with $|\varepsilon(x_1)| \ll x_1$. Approximations with a lower relative error bound are defined by

$$L^{-1}(y) \approx f_0(y)$$

where

$$f_0(y) = f[y - \varepsilon_k(y)] \quad k \in \{0, 1, 2\}$$

and with the zero-, first-, and second-order error approximations defined according to:

$$
\varepsilon_0(y) = L[f(y)] - y \Rightarrow f_0(y) = f[2y - L[f(y)]]
$$

$$
\varepsilon_1(y) = \frac{d}{dy} L[f(y)] \Rightarrow f_0(y) = f[y - \frac{d}{dy} L[f(y)]]
$$

$$
\varepsilon_2(y) = \frac{d^2}{dy^2} L[f(y)] \Rightarrow f_0(y) = f[y - \varepsilon_2(y)]
$$

The second-order error function can be approximated leading to

$$
\varepsilon_{2A}(y) = \frac{d}{dy} L[f(y)] \quad 0 \leq y \leq 1
$$

In these functions:

$$
\frac{d}{dy} L[f(y)] = \frac{d}{dy} f(y) \cdot \frac{d}{df} L[f(y)]
$$

$$
\frac{d^2}{dy^2} L[f(y)] = \frac{d^2}{dy^2} f(y) \cdot \frac{d}{df} L[f(y)] + \left[ \frac{d}{dy} f(y) \right]^2 \frac{d^2}{dy^2} L[f(y)]
$$

and

$$
\frac{d^2}{df} L(f) = \frac{1}{f^2} - \text{csch} (f)^2
$$

$$
\frac{d^2}{df} L(f) = \frac{-2}{f^3} + 2 \coth (f) \text{csch} (f)^2
$$

Proof

Using the result, consistent with Theorem 1, of $L^{-1}(y) = f[y - \varepsilon(x_1)] \quad y \in \{0, 1\}$, these results follow from the approximations for $\varepsilon(x_1)$ specified in Theorem 7.

Results

Consider the relatively simple intermediate function $f$, defined by Eq. 18, and the approximation $L^{-1}(y) \approx f(y)$ which has a modest relative error bound of $9.69 \times 10^{-3}$. The magnitude of the relative errors in approximations to the inverse Langevin equation, based on this function, and for the cases of zero-, first-, and second-order error approximations, as specified by Theorem 7, are shown in Fig. 11. The relative error bounds for these cases are detailed in Table 1 and also for the case of a function, with a lower relative error bound, that is specified by Eq. 34.

Notes

These results show the significant improvement in utilizing accurate approximations for the error function.
Table 1 The relative error bounds in approximations to the inverse Langevin function defined by \(L^{-1}(y) = f(y) - \varepsilon_0(y)\) where \(\varepsilon_0(y)\) is defined by the zero-, first-, and second-order error approximations specified in Theorem 7.

| Error approx. | \(f(y): E q. 18\) | \(f(y): E q. 34\) |
|---------------|-----------------|-----------------|
| No error approx. | 0.00969 | 0.00275 |
| order 0: \(\varepsilon_0(y)\) | 1.31 × 10^{-4} | 1.45 × 10^{-5} |
| order 1: \(\varepsilon_1(y)\) | 2.77 × 10^{-6} | 9.22 × 10^{-8} |
| order 2: \(\varepsilon_2(y)\) | 1.61 × 10^{-8} | 1.54 × 10^{-10} |
| order 2: \(\varepsilon_3(y)\) | 1.53 × 10^{-8} | 1.58 × 10^{-10} |

The approximation to the second-order error function, as defined by Eq. 45, yields results that are consistent with the precise form as specified by Eq. 44. For example, for the function \(f\) defined by Eq. 18, the approximation specified by Eq. 44 yields a relative error bound of 1.61 × 10^{-8} while the approximation specified by Eq. 45 yields a slightly better error bound of 1.53 × 10^{-8}.

**Higher-order error approximation**

Given the significant improvement in approximations for the inverse Langevin function that can be obtained by zero-, first-, and second-order error approximations, it is of interest if higher-order approximations for the error function can be specified. Such approximations are detailed in the following theorem.

**Theorem 8 Higher order approximations for error function**

Consider an initial approximation to the inverse Langevin function of \(L^{-1}(y) \approx f_0(y)\). Improved approximations, with a lower maximum relative error bound, are defined by \(L^{-1}(y) = f_0(y)\) where \(f_0(y) = f[y - \varepsilon_k(y)]\) and \(\varepsilon_k, k \in \{1, 2, \ldots\}\), is a \(k\)th order approximation for the error function. First-, second-, and third-order approximations are defined according to:

\[
\varepsilon_1(y) = \frac{-a_0}{a_1} = \frac{L[f(y)] - y}{dy/dy} L[f(y)]
\]

\[
\varepsilon_2(y) = \frac{-a_0}{a_1} \left[1 + \frac{a_0 a_2}{a_1^2 - 2a_0 a_2}\right]
\]

\[
\varepsilon_3(y) = \frac{-a_0}{a_1} \delta_2, \quad \delta_2 = 1 + \frac{a_0 a_2}{a_1^2 - 2a_0 a_2}
\]

Higher-order approximations can be specified by iteration and according to

\[
\varepsilon_k = \varepsilon_k^{(1)} (1 + \Delta_k) \quad \Delta_k = \frac{-p_k(\varepsilon_k^{(1)})}{\varepsilon_k^{(1)} - p_k^{(1)}(\varepsilon_k^{(1)})}
\]

where

\[
p_k(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_k z^k
\]

\[
p_k^{(1)}(z) = a_1 + 2a_2 z + \cdots + k a_k z^{k-1}
\]

**Proof**

The proofs for these results are detailed in Appendix 6.

**Results**

Results, for the functions \(f\) defined by Eqs. 18 and 34, are tabulated in Table 2 and, as expected, show the significant improvement in accuracy with higher-order error approximations and the improved accuracy by starting with an initial function with a lower relative error bound. The usual trade-off of complexity for accuracy applies.

**Improved approximation via iteration**

An initial approximating function \(f\) for the inverse Langevin function, i.e., \(L^{-1}(y) \approx f(y)\), leads to a better approximation of \(L^{-1}(y) \approx f_0(y)\) where \(f_0(y) = f[y - \varepsilon(y)]\) and \(\varepsilon(y)\) is one of the error approximations specified in Theorem 7 and Theorem 8. Iteration, to yield lower error bounds, is possible as stated in the following theorem.

Table 2 Relative error bounds in approximations to the inverse Langevin function, defined by \(L^{-1}(y) = f[y - \varepsilon_k(y)]\), for the case of the error function approximations, \(\varepsilon_k(y)\), specified in Theorem 8.

| Error approx. | \(f(y): E q. 18\) | \(f(y): E q. 34\) |
|---------------|-----------------|-----------------|
| 0: no error approx. | 0.00969 | 0.00275 |
| order 1: \(\varepsilon_1(y)\) | 2.77 × 10^{-6} | 9.22 × 10^{-8} |
| order 2: \(\varepsilon_2(y)\) | 1.61 × 10^{-8} | 1.54 × 10^{-10} |
| order 3: \(\varepsilon_3(y)\) | 1.98 × 10^{-10} | 1.04 × 10^{-12} |
| order 4: \(\varepsilon_4(y)\) | 1.37 × 10^{-12} | 2.34 × 10^{-15} |
| order 5: \(\varepsilon_5(y)\) | 4.83 × 10^{-15} | 5.26 × 10^{-18} |
Theorem 9 Inverse Langevin approximation: first-order iteration

Consider any of the approximations to the inverse Langevin function as specified by $L^{-1}(y) \approx f_0(y)$ where $f_0$ has one of the forms specified in Theorem 7 or Theorem 8. With iteration, each of these functions yields an improved approximation to the inverse Langevin function according to

$$L^{-1}(y) \approx f_1(y)$$  \hspace{1cm} (55)

where

$$f_1(y) = f_0[y - \varepsilon_k(y)]$$  \hspace{1cm} (56)

and $\varepsilon_k(y)$ has one of the forms specified by Theorem 7 or Theorem 8. Examples include:

$$\varepsilon_0(y) = L[f_0(y)] - y \Rightarrow f_1(y) = f_0[2y - L[f_0(y)]]$$  \hspace{1cm} (57)

$$\varepsilon_1(y) = \frac{d}{dy}L[f_0(y)] \Rightarrow f_1(y) = f_0[y - \frac{L[f_0(y)] - y}{\frac{d}{dy}L[f_0(y)]}]$$  \hspace{1cm} (58)

$$\varepsilon_{2A}(y) = L[f_0(y)] - y + \frac{d^2}{dy^2}L[f_0(y)] \cdot L[f_0(y)] - y^2$$  \hspace{1cm} (59)

$$\Rightarrow f_1(y) = f_0[y - \varepsilon_{2A}(y)]$$  \hspace{1cm} (60)

where $\varepsilon_{2A}$ is defined by Eqs. 51 and 52 with $f(y)$ replaced by $f_0(y)$.

Proof

These results follow directly from Theorem 7 and Theorem 8.

Approximation via second-order iteration

Consider any of the approximations to the inverse Langevin function as specified by $L^{-1}(y) \approx f_1(y)$ where $f_1$ has one of the forms specified in Theorem 9. Iteration by using the results stated in Theorem 7 or Theorem 8 leads to improved approximations for the inverse Langevin function:

Theorem 10 Inverse Langevin approximation: second-order iteration

Each of the approximations for the inverse Langevin function, as specified in Theorem 9, can be used as a basis for an improved approximation according to

$$L^{-1}(y) \approx f_2(y)$$  \hspace{1cm} (61)

where

$$f_2(y) = f_1[y - \varepsilon_k(y)]$$  \hspace{1cm} (62)

and $\varepsilon_k(y)$ has one of the forms specified by Theorem 7 or Theorem 8. Examples include:

$$\varepsilon_0(y) = L[f_1(y)] - y \Rightarrow f_2(y) = f_1[2y - L[f_1(y)]]$$  \hspace{1cm} (63)

$$\varepsilon_1(y) = \frac{d}{dy}L[f_1(y)] \Rightarrow f_2(y) = f_1[y - \varepsilon_1(y)] = f_1[y - \frac{L[f_1(y)] - y}{\frac{d}{dy}L[f_1(y)]}]$$  \hspace{1cm} (64)

$$\varepsilon_{2A}(y) = L[f_1(y)] - y + \frac{d^2}{dy^2}L[f_1(y)] \cdot [L[f_1(y)] - y]^2$$  \hspace{1cm} (65)

$$\Rightarrow f_2(y) = f_1[y - \varepsilon_{2A}(y)]$$

Proof

These results follow directly from Theorem 7 and Theorem 9.

Approximation via higher-order iteration

By using the same order of error approximation at each iteration, general expressions for approximations to the inverse Langevin function, based on iteration, can be specified:

Theorem 11 Iteration with set order of error approximation

Direct iteration, with a set error approximation type at each stage, leads to the following general result:

$$L^{-1}(y) \approx f_i(y) = f_{i-1}[y - \varepsilon_k(y)]$$  \hspace{1cm} (66)

where, for example:

$$f_i(y) = f_{i-1}[y - \varepsilon_0(y)] \quad \varepsilon_0(y) = L[f_{i-1}(y)] - y$$  \hspace{1cm} (67)

$$f_0(y) = f_0(2y - L[f(y)])$$

$$f_i(y) = f_{i-1}[y - \varepsilon_1(y)] \quad \varepsilon_1(y) = \frac{d}{dy}L[f_{i-1}(y)]$$  \hspace{1cm} (68)

$$f_0(y) = f_0 \left[ y - \frac{L[f(y)] - y}{\frac{d}{dy}L[f(y)]} \right]$$

$\copyright$ Springer
Alternative forms for iterative approximations

The following theorem details two alternative forms for iterative approximations to the inverse Langevin function. These functions illustrate the complex nature of iterative approximate expressions for the inverse Langevin function.

**Theorem 12 Alternative forms for iterative approximations**

First, consider the function $f_1$ defined by a first-order iteration, as specified by Eq. 57, where iteration is based on a zero-order error approximation (Eq. 42). An expanded expression is

$$L^{-1}(y) ≈ f_1(y) = f[h_1(y)]$$

where, respectively, for zero- and first-order error approximation:

$$h(y) = 2y - L[f(y)] \quad h_1(y) = 2y - L[f[h(y)]]$$

Second, consider the function $f_2$ defined by a second-order iteration, as specified by Eq. 63, where iteration is based on a zero-order error approximation (Eq. 42). An expanded expression is

$$L^{-1}(y) ≈ f_2(y) = f[h_1(y)]$$

Higher order iteration forms follow in a consistent manner.

**Proof**

The iteration results follow directly from Theorem 7, Theorem 9, and Theorem 10. The proof for the specific first- and second-order iteration forms is detailed in Appendix 7.

Evaluation of some of the expressions in these equations is facilitated by the use of Eqs. 46, 47 and the results:
Table 3  Relative error bounds for approximations to the inverse Langevin function. Iteration based on functions defined by a zero-, first-, second-, and third-order error approximations and with a base function $f$ defined by Eq. 80

| Iteration order/approximating function | Zero-order error: $\varepsilon_0(y)$ | First order error: $\varepsilon_1(y)$ | Second-order error: $\varepsilon_2(y)$ | Second-order error: $\varepsilon_2A(y)$ | Third-order error: $\varepsilon_3(y)$ |
|----------------------------------------|---------------------------------------|--------------------------------------|-------------------------------------|-------------------------------------|-----------------------------|
| 0: $f_0(y)$ (Theorem 7)                | $1.31 \times 10^{-4}$                 | $2.77 \times 10^{-6}$                 | $1.61 \times 10^{-6}$                 | $1.53 \times 10^{-8}$                 | $1.98 \times 10^{-10}$               |
| 1: $f_1(y)$ (Theorem 9)               | $5.08 \times 10^{-8}$                 | $2.66 \times 10^{-16}$                 | $2.07 \times 10^{-30}$                 | $1.75 \times 10^{-30}$                 |                                           |
| 2: $f_2(y)$ (Theorem 10)              | $1.37 \times 10^{-14}$                 | $1.03 \times 10^{-45}$                 |                                           |                                           |                                           |
| 3: $f_3(y)$ (Theorem 11)              | $2.05 \times 10^{-27}$                 |                                           |                                           |                                           |                                           |

The proofs for these results are detailed in Appendix 8.

**Results: error approximation and function iteration**

Relatively simple approximations for the inverse Langevin function, with modest relative error bounds—bounds greater than $10^{-4}$, include:

$$L^{-1}(y) \approx g_1(y) = \frac{3y}{1-y} \left[ 1 - \frac{24y}{25} + \frac{22y^2}{75} \right]$$

this paper $|\text{re}| < 9.69 \times 10^{-3}$ \hfill (80)

$$L^{-1}(y) \approx g_2(y) = \frac{3y}{1-y} \left[ 1 - \frac{459y}{500} + \frac{47y^2}{250} + \frac{19y^3}{300} \right]$$

this paper $|\text{re}| < 5.83 \times 10^{-3}$ \hfill (81)

$$L^{-1}(y) \approx g_3(y) = \frac{3y - \frac{y^2}{5} (6y^2 + y^4 - 2y^6)}{1-y^2}$$

Kroger (2015) $|\text{re}| < 2.8 \times 10^{-3}$ \hfill (82)

$$L^{-1}(y) \approx g_4(y) = \frac{3y}{1-y} \left[ 1 - \frac{3y^2}{5} - \frac{3y^3}{5} + \frac{3y^3}{175} - \frac{99y^4}{175} + \frac{123y^5}{35} - \frac{612y^6}{105} - \frac{3974y^7}{175} + \frac{3604y^8}{175} - \frac{146y^9}{35} \right]$$

\hfill (83)

This paper: 6th-order approx. (Eq. 23): $|\text{re}| < 2.61 \times 10^{-3}$

$$L^{-1}(y) \approx g_5(y) = 3y + \frac{y^2}{5} \sin \left( \frac{2y}{3} \right) + \frac{y^3}{1-y^2}$$

Petrosyan (2017) $|\text{re}| < 1.8 \times 10^{-3}$ \hfill (84)

$$L^{-1}(y) \approx g_6(y) = \frac{y(3-y^2)}{1-y^2} - \frac{y^{10/3}}{2} + 3y^4 \left( \frac{y - 76}{100} \right) \left( y - 1 \right)$$

Nguessong (2014) $|\text{re}| < 7.2 \times 10^{-4}$ \hfill (85)
Table 4  Relative error bounds for approximations to the inverse Langevin function. Iteration based on functions defined by a zero-, first-, and second-order error approximation and with a base function $f$ defined by Eq. 85

| Iteration order/approximating function | Zero-order error: $\varepsilon_0(y)$ | First-order error: $\varepsilon_1(y)$ | Second-order error: $\varepsilon_2(y)$ | Second-order error: $\varepsilon_2(y)$ | Third-order error: $\varepsilon_3(y)$ |
|----------------------------------------|-------------------------------------|-------------------------------|---------------------------------|-------------------------------|---------------------------------|
| $0$: $f_0(y)$ (Theorem 7)              | $1.14 \times 10^{-6}$               | $2.42 \times 10^{-6}$         | $2.26 \times 10^{-12}$          | $2.25 \times 10^{-12}$          | $3.30 \times 10^{-15}$          |
| $1$: $f_1(y)$ (Theorem 9)              | $4.96 \times 10^{-12}$              | $4.31 \times 10^{-15}$        | $2.80 \times 10^{-45}$          | $2.78 \times 10^{-45}$          |                                  |
| $2$: $f_2(y)$ (Theorem 10)             | $1.71 \times 10^{-22}$              | $1.1 \times 10^{-71}$         |                                |                               |                                  |
| $3$: $f_3(y)$ (Theorem 11)             | $4.49 \times 10^{-43}$              |                                |                                |                               |                                  |

$L^{-1}(y) = g_7(y) = \frac{3y + a_2y^2 + a_3y^3 + a_4y^4}{1 - y + b_2(y - y^2)} + b_3(y^2 - y^3) + b_4(y^3 - y^4)$

$a_2 = -6.98408968$, $a_3 = 5.69026957$, $a_4 = -1.35415696$

$b_2 = -1.33411915$, $b_3 = 0.0391556$, $b_4 = 0.64694651$

$L^{-1}(y) = g_6(y) = f(h(y))$

$f(y) = \frac{3y}{1-y} \left[ \frac{1-24y^2}{25} + \frac{22y^3}{75} \right]$

$h(y) = 2y - L[f(y)]$  \tag{87}

This paper: zero-order error approx. (Eq. 42): $|re| < 1.31 \times 10^{-4}$.

$L^{-1}(y) = g_9(y) = \frac{3y}{1-y} \left[ \frac{1-24y^2}{25} + \frac{22y^3}{75} \right]$

$[1 + a_0 + a_1y + a_2y^2 + \cdots + a_{10}y^{10}]$  \tag{88}

This paper: Legendre series (Eq. 31): $|re| < 1.2 \times 10^{-4}$.

Lower error bounds via iterative error approximation

Much lower relative error bounds are possible when error approximation and iteration are utilized. Indicative results are detailed in Table 3 for the case of the function specified by Eq. 80 which has the relative error bound of $9.69 \times 10^{-3}$.

Lower relative error bounds arise by using an initial function approximation with a lower error bound. For example, the function specified by Nguessong (Eq. 85), which has the relative error bound of $7.2 \times 10^{-4}$, yields the results specified in Table 4.

Lower-error bounds via better base functions

By utilizing a base function with a lower relative error bound, improved approximations for the inverse Langevin function are possible with error approximation and iteration. Results are detailed in Table 5.

Notes

The variation, with different base functions, in the relative error bound of approximations for the inverse Langevin function are illustrated in Fig. 12 for the case of a first iteration based on zero order error approximation (see Eq. 57). In general, the upper bound on the magnitude of the relative error occurs in the mid-band region of the interval $[0, 1]$—the exception being the case of $g_7(y)$ defined by Eq. 86. For this function, the coefficients defining the approximation are not of sufficient accuracy to yield high-order approximations with iteration and a floor in the relative error is evident in Table 5. This problem can be simply overcome by entering the coefficient numbers with higher precision.

As is evident in the results tabulated in Tables 3, 4, and 5, the improvement in minimizing the relative error bound is dramatic with error approximation and iteration. The function specified by Eq. 80 approximates the inverse Langevin function with a relative error bound of 0.00969. As detailed in Table 3, this relative error bound can be reduced to $5.08 \times 10^{-6}$, $2.66 \times 10^{-10}$, and $2.06 \times 10^{-30}$ by a first-order iteration based, respectively, on zero-, first-, and second-order error approximations. A comparison of the results between Tables 3 and 4 shows the natural improvement with error approximation and iteration when a base function with a lower relative error bound is used.

Approximation with high accuracy and modest complexity

The results detailed in Table 5 indicate that approximations for the inverse Langevin function, based on a first-order iteration...
chosen base function $f$ that is higher than that required for most applications. For averse Langevin function

Simplified approximation via Taylor series

Relative error bounds, over the interval $[0, 1)$, for approximations to the inverse Langevin function based on the functions $g_1, \ldots, g_7$, defined by Eqs. 80 to 86

| Base function | No error approx. $\epsilon_0(y)$ Eq. 42 | No iteration: $\epsilon_1(y)$ Eq. 43 | No iteration: $\epsilon_2(y)$ Eq. 44 | First iteration: $\epsilon_3(y)$ Eq. 57 | First iteration: $\epsilon_4(y)$ Eq. 58 | First iteration: $\epsilon_5(y)$ Eq. 59 | Second iteration: $\epsilon_6(y)$ Eq. 63 |
|---------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| $g_1(y)$     | $9.69 \times 10^{-3}$           | $1.31 \times 10^{-4}$           | $2.77 \times 10^{-6}$           | $1.61 \times 10^{-8}$           | $5.08 \times 10^{-10}$          | $2.26 \times 10^{-12}$          | $1.75 \times 10^{-14}$           |
| $g_2(y)$     | $5.83 \times 10^{-3}$           | $5.98 \times 10^{-5}$           | $9.67 \times 10^{-7}$           | $3.06 \times 10^{-9}$           | $1.18 \times 10^{-11}$          | $1.53 \times 10^{-13}$          | $4.92 \times 10^{-16}$           |
| $g_3(y)$     | $2.75 \times 10^{-3}$           | $1.45 \times 10^{-5}$           | $9.22 \times 10^{-7}$           | $1.54 \times 10^{-9}$           | $5.80 \times 10^{-12}$          | $1.40 \times 10^{-14}$          | $5.68 \times 10^{-17}$           |
| $g_4(y)$     | $2.61 \times 10^{-3}$           | $1.24 \times 10^{-5}$           | $9.72 \times 10^{-7}$           | $2.20 \times 10^{-9}$           | $4.28 \times 10^{-12}$          | $2.17 \times 10^{-14}$          | $2.45 \times 10^{-17}$           |
| $g_5(y)$     | $1.79 \times 10^{-3}$           | $3.88 \times 10^{-6}$           | $1.52 \times 10^{-7}$           | $1.86 \times 10^{-9}$           | $2.91 \times 10^{-11}$          | $5.83 \times 10^{-13}$          | $8.52 \times 10^{-16}$           |
| $g_6(y)$     | $7.22 \times 10^{-4}$           | $1.14 \times 10^{-7}$           | $2.42 \times 10^{-9}$           | $2.26 \times 10^{-11}$          | $4.96 \times 10^{-13}$          | $4.31 \times 10^{-15}$          | $2.78 \times 10^{-18}$           |
| $g_7(y)$     | $1.80 \times 10^{-4}$           | $1.11 \times 10^{-7}$           | $1.11 \times 10^{-10}$          | $5.54 \times 10^{-12}$          | $9.26 \times 10^{-14}$          | $5.11 \times 10^{-16}$          | $5.11 \times 10^{-19}$           |

and a first-order error approximation, yield relative error bounds of the order of $10^{-16}$ or better—a level of accuracy that is higher than that required for most applications. For a chosen base function $f$, the approximation is

$$L^{-1}(y) \approx f_0(y) = f(y - L[f(y)])$$

(89)

$$f_0'(y) = \left[ f'(y) \frac{df}{df} L[f(y)] \right] + \frac{1}{f'(y)} \left[ L[f(y)] - y \right] \left[ f''(y) \frac{df}{df} L[f(y)] + \left( f'(y) \frac{df}{df} L[f(y)] \right)^2 \frac{d^2f}{df^2} L[f(y)] \right]$$

(90)

For $y$ fixed, the evaluation of $L^{-1}(y)$ requires the determination of $f(y), f_0'(y), f_0''(y), L[f(y)], \frac{df}{df} L[f(y)], \frac{d^2f}{df^2} L[f(y)]$ etc.

**Simplified approximation via Taylor series**

Consider, the initial error-based approximations for the inverse Langevin function

Fig. 12 Graphs of the magnitude of the relative error in approximations to $L^{-1}(y)$ based on a first iteration and on a zero-order error approximation (Eq. 57). The base functions, $f$, are defined by Eqs. 80 to 86.

$$L^{-1}(y) \approx f_0(y) = f(y - \epsilon_k(y)) \quad k \in \{0, 1, 2\}$$

(92)

as specified by Eqs. 42 to 45, where it is expected that the error term is small, i.e., $|\epsilon_k(y)| < y$. This justifies the simplified approximations for the inverse Langevin function as stated in the following theorem:
Theorem 13 Taylor series-based approximations

A first- and second-order Taylor series yield the simplified approximations for the inverse Langevin function:

\[ L^{-1}(y) \approx f_0(y) \quad (93) \]

where

\[
\begin{align*}
\hat{f}_0(y) &= f(y) - \varepsilon_k(y) f^{(1)}(y) & \text{1st order Taylor} \\
\tilde{f}_0(y) &= f(y) - \varepsilon_k(y) f^{(1)}(y) + \frac{\varepsilon_k^2(y)}{2} f^{(2)}(y) & \text{2nd order Taylor} \\
\end{align*}
\]

(94)

For the zero-order error case (Eq. 42) the respective first- and second-order Taylor series based approximations are:

\[
\begin{align*}
\hat{f}_0(y) &= f(y) - [L[f(y)] - y] f^{(1)}(y) \\
\tilde{f}_0(y) &= f(y) - [L[f(y)] - y] f^{(1)}(y) + \frac{[L[f(y)] - y]^2}{2} f^{(2)}(y)
\end{align*}
\]

(95)

For a first-order error-based approximation (Eq. 43) the respective first- and second-order Taylor series based approximations are:

\[
\begin{align*}
\hat{f}_0(y) &= f(y) - \frac{d}{df} L[f(y)] \\
\tilde{f}_0(y) &= f(y) - \frac{d}{df} L[f(y)] + \frac{1}{2} \left( \frac{d}{df} L[f(y)] \right)^2 f^{(2)}(y)
\end{align*}
\]

(96)

Higher-order error approximations follow in an analogous manner.

Proof
These results arise from the first- and second-order Taylor series based on the point \( y \), i.e.

\[
\begin{align*}
f(y + \Delta) &\approx f(y) + \Delta f^{(1)}(y) \\
f(y + \Delta) &\approx f(y) + \Delta f^{(1)}(y) + \frac{\Delta^2}{2} f^{(2)}(y)
\end{align*}
\]

(97)

where \( \Delta = - \varepsilon_k(y) \).

Explicit error-based approximations

Theorem 14 Explicit error-based approximations

The error-based approximations for the inverse Langevin function, as specified in Theorem 13, have the following form:

\[ L^{-1}(y) \approx f_0(y) \quad (98) \]

where, first, for a zero-order error-based approximation and respectively, for a first- and second-order Taylor series approximation:

\[
\begin{align*}
f_0(y) &= f(y) - \left[ \coth [f(y)] - \frac{1}{f(y)} \right] f^{(1)}(y) \\
f_0(y) &= f(y) - \left[ \coth [f(y)] - \frac{1}{f(y)} \right] f^{(1)}(y) + \frac{1}{2} \left[ \coth [f(y)] - \frac{1}{f(y)} \right]^2 f^{(2)}(y)
\end{align*}
\]

(99)

Second, for a first order error-based approximation:

\[
\begin{align*}
f_0(y) &= \frac{2f(y) + yf^2(y) \sinh[f(y)]^2 - f^3(y) - f^2(y) \sinh[2f(y)]}{2 \sinh[f(y)]^2 - f^2(y)} \\
f_0(y) &= \frac{2f(y) + yf^2(y) \sinh[f(y)]^2 - f^3(y) - f^2(y) \sinh[2f(y)]}{2 \sinh[f(y)]^2 - f^2(y)}
\end{align*}
\]

(100)

Results for higher-order error approximations follow in a natural manner.

Proof
These results arise from Theorem 13 and the definition of the Langevin function and its derivative.

First- and second-order iteration

The approximations detailed in Theorem 14 can be used to specify first-, second-, and higher-order iteration-based approximations for the inverse Langevin function:

Theorem 15 First- and second-order iteration: first-order error and first-order Taylor

Consider the case of a first-order error approximation with either a first- or second-order Taylor series approximation (Eq. 96). A first-order iteration approximation to the inverse Langevin function is defined according to

\[ L^{-1}(y) \approx f_1(y) \quad (102) \]

where

\[ f_1(y) = \frac{2f_0(y) + yf^2_0(y) \sinh[f_0(y)]^2 - f^3_0(y) - f^2_0(y) \sinh[2f_0(y)]}{2 \sinh[f_0(y)]^2 - f^2_0(y)} \]

(103)

for the case of a first-order Taylor series approximation. Here, \( f_0 \) is defined by Eq. 100. For a second-order, Taylor series approximation
Here, series approximation:

\[ f_1(y) = \left[ \frac{2f_0(y)}{f_0(y)/y} + yf_1(y) \right] \sinh[f_1(y)]y - f_0(y) \sinh[2f_0(y)/]2 \]

and based on the Taylor series approximations defined by Eqs. 100, 101, 102, and 103. Each group of results represents a distinct order of functional approximation defined by Eq. 80 and with the use of first- and second-order Taylor series approximations

Table 6 The relative error bounds, over the interval (0,1), for approximations to the inverse Langevin function based on the function \( f \) defined by Eq. 80

| Taylor series approximation | \( f(y) \) | \( f_0(y), \varepsilon_0 (y) \) | \( f_1(y), \varepsilon_1 (y) \) | \( f_2(y), \varepsilon_0 (y) \) | \( f_1(y), \varepsilon_1 (y) \) | \( f_2(y), \varepsilon_0 (y) \) | \( f_1(y), \varepsilon_1 (y) \) |
|-----------------------------|-------------|----------------|----------------|----------------|----------------|----------------|----------------|
| None: exact case            | 9.69 × 10^{-3} | 1.31 × 10^{-6} | 5.08 × 10^{-8} | 2.66 × 10^{-16} | 1.37 × 10^{-14} | 1.03 × 10^{-45} |
| First-order                 | 1.16 × 10^{-4} | 9.39 × 10^{-5} | 3.17 × 10^{-8} | 8.80 × 10^{-9} | 1.58 × 10^{-11} | 7.74 × 10^{-17} |
| Second-order                | 1.31 × 10^{-4} | 2.43 × 10^{-6} | 5.07 × 10^{-8} | 1.77 × 10^{-14} | 1.37 × 10^{-14} | 3.10 × 10^{-46} |

\[ f_0(y) = f(y) - \frac{L[f(y)] - y}{y} \]

It then follows from Eq. 58 that

\[ L^{-1}(y) \simeq f_1(y) = f_0(y) - \frac{L[f_0(y)] - y}{f_0(y)} \]

The first-order Taylor series approximation, as specified by Eq. 96, yields the required result of

\[ L^{-1}(y) \simeq f_1(y) - \frac{L[f_0(y)] - y}{f_0(y)} \]

The other results follow in an analogous manner.

**Results**

A first-order Taylor series approximation, as defined in Eq. 94, in general, leads to significantly higher relative errors in approximations to the inverse Langevin function. In contrast, a second-order Taylor series approximation leads to relative errors comparable with the non-approximated case. Illustrative results are shown in Table 6 where the base function defined by Eq. 80 has been used.
Example

Consider the relatively simple expression, as defined by Eq. 100, which is for a first-order Taylor approximation to a first-order error-based approximation, i.e.,

\[ L^{-1}(y) = f_0(y) \]

\[ = \frac{[2f(y) + yf^2(y)] \sinh [f(y)]^2 - f^3(y) - f^2(y) \sinh [2f(y)]}{\sinh [f(y)]^2 - f^2(y)} \quad (111) \]

When the function \( f \) is that defined by Petrosyan (Eq. 84 with a relative error bound of \( 1.8 \times 10^{-3} \)), or Ngessong (Eq. 85 with a relative error bound of \( 7.2 \times 10^{-4} \)), the relative error bounds, respectively, are \( 3.20 \times 10^{-6} \) and \( 3.81 \times 10^{-7} \). These bounds are comparable with the [7/7] approximation given in Marchi and Arruda (2019) which has a relative error bound of \( 6.92 \times 10^{-7} \). If this function is used iteratively, according to

\[ L^{-1}(y) = f_0(y) \]

\[ = \frac{[2f_0(y) + yf^2_0(y)] \sinh [f_0(y)]^2 - f^3_0(y) - f^2_0(y) \sinh [2f_0(y)]}{\sinh [f_0(y)]^2 - f^2_0(y)} \quad (112) \]

then the relative error bounds reduce, respectively, to \( 1.03 \times 10^{-11} \) and \( 1.09 \times 10^{-13} \). As is clear from the results detailed in Table 6, these error bounds can be significantly improved upon by utilizing a second-order Taylor series approximation.

Computational complexity

For Padé-based approximations, where a measure of the computational complexity involved in the evaluation of the inverse Langevin function for a set argument can be clearly defined, a graph of the relative error bound versus computational complexity can readily be generated, e.g., Kröger 2015 and Jedynak 2018. The results shown in Fig. 13 are indicative of the relative error improvement that is possible with an increase in functional and computational complexity. An unsolved problem is the determination of the computational complexity, in terms of the number of basic operations (addition, subtraction, multiplication, division, ...), for the approximations detailed in the paper, in particular for the Taylor series approximations detailed in Eqs. 100, 101, 103, 104, 106, and 107. There is potential, e.g., Muller (2006) and Brent (2018), for the computational efficiency of specific functions to be enhanced by innovative approaches. Further research is warranted.

Conclusion

In this paper, an analytical framework has been detailed which underpins, first, convergent series approximations for the inverse Langevin function and, second, analytical approximations for this function, with potentially arbitrarily small maximum relative error magnitudes. The basis for both approaches is the definition of an intermediate function \( f \) that leads to \( L[f(x)] \) being a closely linear function with a slope that is close to one over the interval \([0, 1]\). This function allows, first, the definition of an error function for the inverse Langevin function which can be approximated via a basis set decomposition. Second, it allows error approximation and then function iteration which leads, potentially, to arbitrarily low relative error bounds in approximations.

Basis set decomposition for the defined error function underpins convergent series approximations for the inverse Langevin function. A tenth-order Legendre basis set, based on an initial approximating function as specified by Eq. 80, leads to a series approximation for the inverse Langevin equation with a relative error bound of \( 1.2 \times 10^{-4} \). A twentieth-order series has a relative error bound of \( 3.2 \times 10^{-6} \).

A modest approximating function (Eq. 80), with a relative error bound of 0.00969, leads to relative error bounds of \( 1.31 \times 10^{-4}, 2.77 \times 10^{-5} \), and \( 1.61 \times 10^{-8} \) with zero-, first-, and second-order error approximation. First-order iteration, based on a first-order error approximation, leads to a relative error bound of \( 2.66 \times 10^{-16} \). Significantly lower relative error bounds can be obtained by higher-order iteration and by using a second-, or higher-, order error approximation. These results represent significant improvements on published analytical approximations for the inverse Langevin function.

First- and second-order Taylor series for the error-based approximations to the inverse Langevin function can be used to obtain simplified function forms. Whilst the first-order Taylor series leads to results with a much lower accuracy, the second-order Taylor series results in expressions without accuracy compromise. As usual, there is a trade-off between complexity and accuracy.

Acknowledgments The support of Prof. A. Zoubir, SPG, Technische Universität Darmstadt, Germany, who hosted a visit where the research underpinning this paper was completed, is gratefully acknowledged. The attention to detail by the reviewers is acknowledged.

Funding Open Access funding enabled and organized by Projekt DEAL.

Appendix 1: Proof of Approximations for transformed Langevin function at 0,1

Consider the alternative form for the Langevin function

\[ L(x) = \frac{e^x + e^{-x}}{e^x - e^{-x}} - 1 = \frac{(x-1) + (x+1)e^{-2x}}{x(1-e^{-2x})} \quad (113) \]
which implies

\[
L[f(x_1)] = \frac{f(x_1) - 1 + [f(x_1) + 1]e^{-2f(x_1)}}{f(x_1)[1-e^{-2f(x_1)}]} \tag{114}
\]

The basis for an approximation for \(L[f(x_1)]\), close to the origin, is the standard Taylor series expansion for the exponential function. Use of such a series leads to:

\[
L[f(x_1)] = \frac{f(x_1) - 1 + [f(x_1) + 1] - 1 + [f(x_1) + 1]e^{-2f(x_1)}}{f(x_1)[1-e^{-2f(x_1)}]} - \frac{2f(x_1) - 2f(x_1)}{2f(x_1)[1-e^{-2f(x_1)}]} + \cdots
\]

Simplification leads to the first required result:

\[
L[f(x_1)] = \frac{2f(x_1) - 2f(x_1) + 2f^2(x_1) + 1 - 1}{f(x_1)[1-e^{-2f(x_1)}]} = \frac{f(x_1)[1-2f(x_1) + f^2(x_1) + f(x_1) - 1]}{1-f(x_1) + 2f(x_1) - 2f(x_1) + \cdots} \tag{115}
\]

For \(x_1 \to 1\) from below, and with \(f(x_1) \to \infty\), an appropriate approximation for \(L[f(x_1)]\) is:

\[
L[f(x_1)] = \frac{f(x_1) - 1 + [f(x_1) + 1]e^{-2f(x_1)}}{f(x_1)[1-e^{-2f(x_1)}]} - \frac{f(x_1) - 1}{f(x_1)} = 1 - \frac{1}{f(x_1)} \tag{117}
\]

This result is consistent with \(\coth(x)\) converging to unity much faster than \(1/x\) converges to zero as \(x \to \infty\).

**Appendix 2: Proof of Transformed Langevin function with unity rate of change**

Utilizing the first approximation stated in Theorem 2, which is valid for \(x_1 = 0^+\), it follows that

\[
\frac{d}{dx_1} L[f(x_1)] = \frac{f^{(1)}(x_1) - 1}{f(x_1)} - \frac{f^{(1)}(x_1) - 1}{f(x_1)} = 1 - \frac{1}{f(x_1)} \tag{118}
\]

As \(f(x_1) \to 0\) and \(L[f(x_1)] \to 0\) as \(x_1 \to 0\) from above, it follows that

\[
\frac{d}{dx_1} L[f(x_1)] = \frac{f^{(1)}(x_1)}{f(x_1)[1-e^{-2f(x_1)}]} = \frac{1}{3} \quad x_1 = 0^+ \tag{119}
\]

The second approximation stated in Theorem 2 implies

\[
\frac{d}{dx_1} L[f(x_1)] = \frac{f^{(1)}(x_1)}{f^2(x_1)} = 1 - \frac{1}{x_1} \tag{120}
\]

For unity rate of change at \(x_1 = 0^+\) and \(x_1 = 1\), these two equations imply:

\[
f^{(1)}(x_1)|_{x_1=0^+} = 3 \quad f^{(1)}(x_1)|_{x_1=1^-} = 0 \tag{121}
\]

**Appendix 3: Additional Results: Higher order forms for intermediate function**

The third, fourth, fifth, seventh and eighth order approximations, respectively, are:

\[
f(x_1) = \frac{3x_1}{1-x_1} \left[1 - x_1 + \frac{3x_1^2}{5} - \frac{4x_1^3}{5} + \frac{4x_1^4}{5} - \frac{4x_1^5}{15}\right] \tag{122}
\]

\[
f(x_1) = \frac{3x_1}{1-x_1} \left[1 - x_1 + \frac{3x_1^2}{5} - \frac{3x_1^3}{5} + \frac{x_1^4}{15} + \frac{11x_1^5}{15} - \frac{3x_1^6}{5} + \frac{2x_1^7}{15}\right] \tag{123}
\]

\[
f(x_1) = \frac{3x_1}{1-x_1} \left[1 - x_1 + \frac{3x_1^2}{5} - \frac{3x_1^3}{5} + \frac{99x_1^4}{175} - \frac{29x_1^5}{35} + \frac{23x_1^6}{35} + \frac{26x_1^7}{21} + \frac{76x_1^8}{175}\right] \tag{124}
\]

\[
f(x_1) = \frac{3x_1}{1-x_1} \left[1 - x_1 + \frac{3x_1^2}{5} - \frac{3x_1^3}{5} + \frac{99x_1^4}{175} - \frac{29x_1^5}{35} + \frac{23x_1^6}{35} + \frac{26x_1^7}{21} + \frac{76x_1^8}{175}\right] \tag{125}
\]
For such a basis set, the coefficients are defined according to, e.g., Lebedev (1972):

\[ b_1(x) = \sqrt{2i + 1}L_i(2x-1) \quad i \in \{0, 1, \ldots\} \]  \hspace{1cm} (127)

Assuming \( \epsilon_i(y) \), as defined by Eq. 25, is an odd function, and defined on the interval \([-1, 1]\), the orthonormal sinusoidal basis set is defined by

\[ b_i(x) = \sin(2\pi f_i x) \quad f_0 = 1/2 \quad i \in \{0, 1, \ldots\} \]  \hspace{1cm} (130)

For such a basis set, the coefficients are defined according

\[ c_i = 2f_0^1 \left[ \frac{L^{-1}(y)}{f(y)} \right]^{-1} b_i(y) \, dy \quad i \in \{0, 1, \ldots\} \]  \hspace{1cm} (131)

where the factor of two arises from use of the interval \([-1, 1]\) and the fact that on this interval the basis functions are odd functions.

**Appendix 5: Proof of Zero, first and second order error approximations**

Consider a first-order Taylor series for the function \( q \) around the point \( y_1 = q(x_1) \):

\[ q(x) \approx q(y_1) + (x-y_1)q^{(1)}(y_1) \Rightarrow \epsilon(x_1) = y_1 - x_1 \approx \frac{q(y_1)-q(x_1)}{q^{(1)}(y_1)} = \frac{q(y_1)-y_1}{q^{(1)}(y_1)} \]  \hspace{1cm} (132)

\[ x_1 = y_1 - \epsilon(x_1) \approx y_1 - \frac{q(y_1)-y_1}{q^{(1)}(y_1)} \]

which are the results for the first-order error approximation case. For the case of constant slope, which is close to unity, such that the approximation \( q^{(1)}(y_1) = 1 \) is valid, the zero-order approximations arise:

\[ \epsilon(x_1) \approx q(y_1)-y_1 \quad x_1 \approx 2y_1 - q(y_1) \]  \hspace{1cm} (133)

To establish a second-order error approximation, consider a second-order Taylor series for the function \( q \) around the point \( y_1 = q(y_1) \):

\[ q(x) \approx q(y_1) + (x-y_1)q^{(1)}(y_1) + \frac{(x-y_1)^2}{2}q^{(2)}(y_1) \Rightarrow \epsilon(x_1) = y_1 - x_1 \approx \frac{q^{(1)}(y_1)}{q^{(2)}(y_1)} \left[ 1 \pm \sqrt{1 - \frac{2q^{(2)}(y_1)[q(y_1)-y_1]}{[q^{(1)}(y_1)]^2}} \right] \]  \hspace{1cm} (134)

Small errors are consistent with the negative sign prior to the square root. Hence:

\[ x_1 = y_1 - \epsilon(x_1) \approx y_1 - \frac{q^{(1)}(y_1)}{q^{(2)}(y_1)} \times \left[ 1 - \sqrt{1 - \frac{2q^{(2)}(y_1)[q(y_1)-y_1]}{[q^{(1)}(y_1)]^2}} \right] \]  \hspace{1cm} (135)

For the case where \( \left| \frac{2q^{(2)}(y_1)[q(y_1)-y_1]}{[q^{(1)}(y_1)]^2} \right| < 1 \), the approximation \( \sqrt{1 + \frac{q^{(2)}(y_1)}{q^{(1)}(y_1)^2}} \approx 1 \) leads to
It then follows that the error defined by
\[ \varepsilon(x_1) = y_1 - x_1 \] (138)
is the root of the polynomial equation in terms of \( x_1 - y_1 \):
\[
\frac{(-1)^n}{n!} q^n(y_1) + \cdots + \frac{(x-y_1)^2}{2} q^{(2)}(y_1) + (x-y_1) q^{(1)}(y_1) + \Delta_2 = 0
\] (139)
which, ideally, is a better approximation than that arising from a first-order Taylor series as specified in Eq. 132.

Consider the \( n \)th order polynomial
\[ p_n(z) = a_n z^n + \cdots + a_1 z + a_0 \] (140)
where \( z = \varepsilon(x_1) = y_1 - x_1 \) and
\[
a_n = \frac{(-1)^n}{n!} q^n(y_1) \quad \cdots \quad a_2 = \frac{1}{2} q^{(2)}(y_1) \\
a_1 = -q^{(1)}(y_1) \quad a_0 = q(y_1) - y_1
\] (141)

For a first-order polynomial the solution for \( p_1(z) = 0 \) is exact:
\[ z_1 = z_0(1 + \Delta_1) = -\frac{a_0}{a_1} \quad z_0 = -\frac{a_0}{a_1} , \Delta_1 = 0 \] (142)

Consider using this solution as the basis for an approximation to the second-order polynomial equation of \( p_2(z) = a_2 z^2 + a_1 z + a_0 = 0 \). With the solution to this equation denoted as \( z_2 \), and defined as \( z_2 = z_1(1 + \Delta_2) \), it follows, by using the approximation
\[ z_2^2 = z_1^2(1 + 2\Delta_2) \] (143)
that the solution for \( \Delta_2 \) is
\[ \Delta_2 = \frac{-p_2(z_1)}{z_1 p_2'(z_1)} \] (144)

iteration, in this manner, and by using the approximations \( z_k^m \approx z_{k-1}^m(1 + n \Delta_k) \) for \( n \in \{1, \ldots, k\} \), yields the general stated iteration formula.

### Appendix 7: Proof for results in Iteration with set order of error approximations

To prove the first-order iteration form specified by Eq. 70 and Eq. 71 consider:
\[
L^{-1}(y) = f_1(y) = f_0[2y - L[f_0(y)]] \\
f_0(y) = f[2y - L[f_0(y)]] = f[h(y)]
\] (145)
where \( h(y) = 2y - L[f_0(y)] \). This implies
\[
f_1(y) = f_0[2y - L[f(h(y))]] = f_0[h_1(y)] = f[h_1(y)]
\] (146)
where \( h_1(y) = 2y - L[f(h(y))] \). The form specified by Eqs. 70 and 72 follows in an analogous manner.

For the second-order iteration form specified by Eqs. 73 and 74 consider:
\[
L^{-1}(y) = f_2(y) = f_1[2y - L[f_1(y)]]
\] (147)
where, consistent with Eq. 146, \( f_1(y) = f[h_1(y)] \). With the definition of \( h_2(y) = 2y - L[f_1(y)] \) it follows that
\[
f_2(y) = f_1[h_2(y)] = f[h[h_1[y]]]
\] (148)
The form specified by Eq. 73 and 75 follows in an analogous manner.

### Appendix 8: Results for Alternative forms for iterative approximations

Consider the first-order iteration approximation based on a zero-order error approximation:
\[
L^{-1}(y) = f_1(y) = f_0[2y - L[f_0(y)]] \\
f_0(y) = f[2y - L[f(y)]]
\] (149)
Substitution yields

\[ f_1(y) = f_0 \left[ 2y - L \left[ \frac{2y - L[f(y)]}{2} \right] \right] \]
\[ = f \left[ 2 \left[ 2y - L \left[ \frac{2y - L[f(y)]}{2} \right] \right] \right] \]
\[ - L \left[ 2y - L \left[ \frac{2y - L[f(y)]}{2} \right] \right] \]
\[ = f \left[ 4y - 2L \left[ \frac{2y - L[f(y)]}{2} \right] \right] \]
\[ - L \left[ 2y - L \left[ \frac{2y - L[f(y)]}{2} \right] \right] \]
\[ = f \left[ 4y - 2L \left[ \frac{2y - L[f(y)]}{2} \right] \right] \]
\[ - L \left[ 2y - L \left[ \frac{2y - L[f(y)]}{2} \right] \right] \]  \hspace{1cm} (150)

This result can be written as

\[ f_1(y) = f[h(y)] \]  \hspace{1cm} (151)

where

\[ h(y) = 4y - 2L \left[ \frac{2y - L[f(y)]}{2} \right] - L \left[ \frac{2y - L[f(y)]}{2} \right] \]  \hspace{1cm} (152)

For the second-order iteration case, based on a zero-order error approximation, the approximation for the inverse Langevin function is

\[ L^{-1}(y) \approx f_2(y) = f_1[2y - L[f_1(y)]] \]  \hspace{1cm} (153)

where \( f_1(y) = f[h(y)] \). It then follows that

\[ f_2(y) = f_1[2y - L[f[h(y)]]] = f_1[h_1(y)] \]  \hspace{1cm} (154)

where

\[ h_1(y) = 2y - L[f(h(y))] \]
\[ = 2y - L \left[ 4y - 2L \left[ \frac{2y - L[f(y)]}{2} \right] - L \left[ \frac{2y - L[f(y)]}{2} \right] \right] \]  \hspace{1cm} (155)

The required result of

\[ f_2(y) = f \left[ 4h_1(y) - 2L \left[ \frac{2h_1(y) - L[f[h_1(y)]]}{2} \right] \right] \]
\[ - L \left[ 2h_1(y) - L \left[ \frac{2h_1(y) - L[f[h_1(y)]]}{2} \right] \right] \]  \hspace{1cm} (156)

then follows from Eq. 150.
Jedynak R (2015) Approximation of the inverse Langevin function revisited. Rheol Acta 54:29–39. https://doi.org/10.1007/s00397-014-0802-2

Jedynak R (2018) A comprehensive study of the mathematical methods used to approximate the inverse Langevin function. Math Mech Solids 24:1992–2016. https://doi.org/10.1177/1081286518811395

Kröger M (2015) Simple, admissible, and accurate approximants of the inverse Langevin and Brillouin functions, relevant for strong polymer deformations and flows. J Non-Newtonian Fluid Mech 223:77–87. https://doi.org/10.1016/j.jnnfm.2015.05.007

Kuhn W, Grün F (1942) Beziehungen zwischen elastischen Konstanten und Dehnungsdoppelbrechung hochelastischer Stoffe. Kolloid-Zeitschrift 101:248–271

Langevin P (1905) Sur la théorie du magnétisme. J Phys Theor Appl 4:678–693

Lebedev NN (1972) Special functions and their applications, Dover ch 4

Marchi BC, Arruda EM (2015) An error-minimizing approach to inverse Langevin approximations. Rheol Acta 54:887–902. https://doi.org/10.1007/s00397-015-0880-9

Marchi BC, Arruda EM (2019) Generalized error-minimizing, rational inverse Langevin approximations. Math Mech Solids 24:1630–1647. https://doi.org/10.1177/1081286517754131

Muller JM (2006) Elementary functions: algorithms and implementations, 2nd edn. Birkhauser

Nguessong AN, Beda T, Peyraut F (2014) A new based error approach to approximate the inverse Langevin function. Rheol Acta 53:585–591

Petrosyan R (2017) Improved approximations for some polymer extension models. Rheol Acta 56:21–26. https://doi.org/10.1007/s00397-016-0977-9

Rickaby SR, Scott NH (2015) A comparison of limited-stretch models of rubber elasticity. Int J Non-Linear Mech 68:71–86. https://doi.org/10.1016/j.ijnonlinmec.2014.06.009

Treloar LRG (1954) The photoelastic properties of short-chain molecular networks. Trans Faraday Soc 50:881–896

Treloar LRG (1975) The physics of rubber elasticity. Clarendon Press

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.