Article

MMS

Mathematics and Mechanics of Solids
17(7): 693–701

DOI: 10.1177/1081286511429886
mms.sagepub.com
SAGE

Taylor expansion of the inverse function with application to the Langevin function

Mikhail Itskov, Roozbeh Dargazany and Karl Hörnes
Department of Continuum Mechanics, RWTH Aachen University, Germany

Received 13 July 2011; accepted 21 October 2011

Abstract
A Taylor power series is a powerful mathematical tool, which can be used to express an inverse function especially if it is given in an implicit form. This is for example the case for the inverse Langevin function, which is an indispensable ingredient of full-network rubber models. In the present paper, we propose a simple recurrence procedure for calculating Taylor series coefficients of the inverse function. This procedure is based on the Taylor series expansion of the original function and results in a simple recurrence formula. This formula is further applied to the inverse Langevin function. The convergence radius of the resulting series is evaluated. Within this convergence radius the obtained approximation of the inverse Langevin function demonstrates better agreement with the exact solution in comparison to different Padé approximants.

Keywords
higher-order coefficients, inverse function, Langevin function, non-Gaussian statistics, Taylor expansion

1. Introduction

Let $y = f(x)$ be a one-to-one function defined on a range of real numbers X. The function f^{-1} is called the inverse of f if

$$x = f^{-1}(f(x)) \quad \forall x \in X. \tag{1}$$

Often the function f^{-1} cannot be expressed in an explicit form. In this case, a Taylor power series (provided it converges) is almost the only representation that can be used for the inverse function. The coefficients of this series can generally be expressed in terms of the corresponding coefficients of the original function by means of direct insertion (see, e.g. [1]). However, only a small number of the first coefficients can be found by this procedure.

The classical approach to the Taylor expansion of f^{-1} is due to Lagrange. Accordingly, if $f(0) = 0$ and $f'(0) \neq 0$ then

$$f^{-1}(y) = \sum_{n=1}^{\infty} \frac{y^n}{n!} \left[\frac{d^{n-1}}{dx^{n-1}} \left(\frac{x}{f(x)} \right)^n \right]_{x=0}, \tag{2}$$

which is referred to as Lagrange's inversion formula [2]. It requires, however, multiple explicit differentiations of the quotient $\left(\frac{x}{f(x)}\right)^n$ and is thus not easy for practical applications. This is also the case for the Faà di Bruno

Corresponding author:
Mikhail Itskov, Department of Continuum Mechanics, RWTH Aachen University, Eilfschornsteinstr. 18, 52064 Aachen, Germany.
Email: itskov@km.rwth-aachen.de

formula [3], which gives an explicit expression for the nth derivative of the composition of two functions in terms of their derivatives.

Chernoff [4] introduced a movable strip method in order to evaluate coefficients of the inverse series. This method did not, however, give a direct result for the derivatives of inverse functions. Traub [5] used a theorem of Jabotinsky [6] (representation of the inverse function in matrices) in order to prove a formula by Ostrowski [7]. This formula is based on the parameters β_i, which have to satisfy a set of additional equalities and inequalities. This set has to be solved for every order n, which becomes numerically very expensive especially for larger n.

Apostol [8] formulated a recursive polynomial, which gives the nth derivative of $x = x(y)$ as a function of all previous derivatives. This approach was later simplified by introducing a combinatorial argument for the representation of the nth derivative [9]. Although successive differentiation is no longer required, the computation of higher-order terms remains numerically very expensive.

In this paper, we derive higher-order derivatives of inverse functions using the Taylor series and taking advantage of power series raised to powers. This approach does not require successive differentiations or special calculus knowledge and results in a straightforward recurrent procedure. This procedure is applied for example to the Langevin function defined by

$$y = \mathcal{L}(x) = \coth(x) - \frac{1}{x}. \tag{3}$$

The inverse Langevin function $\mathcal{L}^{-1}(y)$ results from the non-Gaussian statistical theory of rubber elasticity (see e.g. [10]) as the entropic force developed by polymer chains. For a chain with an end-to-end distance r consisting of N segments each of length l, this entropic force per unit referential area is written as

$$f(N, r) = \frac{KT}{l}\mathcal{L}^{-1}\left(\frac{r}{Nl}\right), \tag{4}$$

where T denotes the absolute temperature and K is the Boltzmann constant. As the chain end-to-end distance r approaches its maximal value Nl corresponding to a fully stretched, straight chain, the force tends to infinity. This implies asymptotic behavior of $\mathcal{L}^{-1}(y)$ in the vicinity of $y = 1$.

The chain force (4) described by the inverse Langevin function is an important ingredient of full-network rubber models (see e.g. [11,12]). They are obtained by integration of single-chain strain energies over all spatial directions by taking into account the chain distribution function. There are other material models designed to account for non-Gaussian effects, such as for example the Gent model [13]. Horgan and Saccomandi [14] demonstrated that the Gent model is even able to reflect the maximal chain extensibility resulting from the non-Gaussian chain statistics. The crucial advantage of the full-network model utilizing the inverse Langevin function is that it is generally anisotropic. The anisotropy of the virgin material can result from a non-uniform spatial chain distribution and can be taken into account by the above mentioned spatial distribution function [11, 12]. Alternatively, the anisotropy can be strain induced and caused by damage evolution, such as for example of the Mullins type (see e.g. [15, 16]).

The inverse Langevin function cannot be expressed in an explicit form. For this reason, it is either approximated by rational functions [17–19] or by means of the Taylor expansion [20, 21]. The rational functions below mostly represent or are based on a Padé approximation and can describe the asymptotic behavior near $y = 1$ relatively precisely:

- Cohen [17]: $\mathcal{L}^{-1}(y) \approx y\frac{3-\frac{36}{35}y^2}{1-\frac{33}{35}y^2}$, which is further rounded to $\mathcal{L}^{-1}(y) \approx y\frac{3-y^2}{1-y^2}$,
- Puso [18]: $\mathcal{L}^{-1}(y) \approx \frac{3y}{1-y^3}$,
- Treloar [19]: $\mathcal{L}^{-1}(y) \approx \frac{3y}{1-0.6y^2-0.2y^4-0.2y^6}$.

The Taylor expansion, on the other hand, benefits from its higher accuracy (except in the neighborhood of the asymptotic point $y = 1$). Nevertheless, only the first 20 coefficients for the Taylor expansion of the inverse Langevin function are known so far [20]. This is due to the fact that the calculation time of Taylor coefficients grows very fast with their order.

This calculation can be facilitated by means of our recurrent procedure for higher-order derivatives of the inverse function. For the inverse Langevin function this procedure yields coefficients of the Taylor series expressed in terms of Bernoulli numbers. By this means, more than a hundred coefficients have been easily calculated using the numerical algorithm presented below. Finally, the accuracy of the resulting Taylor expansion is evaluated in comparison with the existing approximation functions.

2. Taylor series of an inverse function

We consider a function $y(x)$, which can be expanded into a Taylor power series around a point x_0 by

$$y(x_0 + \xi) = \sum_{n=0}^{\infty} \frac{\xi^n}{n!} y_n(x_0), \tag{5}$$

where

$$y_n(x_0) = \left.\frac{d^n y}{dx^n}\right|_{x_0}, \quad n = 1, 2, \dots \tag{6}$$

If $y_1(x_0) \neq 0$, there exists a unique inverse function $x(y)$ whose Taylor series around $y_0 = y(x_0)$ can be given by

$$x(y_0 + \eta) = \sum_{n=0}^{\infty} \frac{\eta^n}{n!} x_n(y_0). \tag{7}$$

The higher-order derivatives of the inverse function

$$x_n(y_0) = \left.\frac{d^n x}{dy^n}\right|_{y_0}, \quad n = 1, 2, \dots, \tag{8}$$

represent an indispensable ingredient of the series (7). In the following we are going to express these derivatives in explicit form in terms of y_n, $n = 1, 2, \dots$, which are assumed to be known. To this end, we apply and further develop a procedure proposed by Liptaj [22]. Accordingly, we first write

$$x_0 + \xi = x(y(x_0 + \xi)) = x(y(x_0) + \eta) = x_0 + \sum_{j=1}^{n} \frac{\eta^j}{j!} x_j + O\left(\eta^{n+1}\right), \tag{9}$$

where in comparison to [22] the series for

$$\eta = \sum_{i=1}^{\infty} \frac{\xi^i}{i!} y_i \tag{10}$$

is not truncated. Setting in (9) $n = 1$ we obtain

$$\xi = x_1 \sum_{i=1}^{\infty} \frac{\xi^i}{i!} y_i + O\left(\eta^2\right).$$

Enforcing $\xi \to 0$ leads to the well-known result

$$x_1 = \frac{1}{y_1}. \tag{11}$$

Let us now assume that x_i, $i = 1, 2, \dots, n-1$, are known. In order to express x_n we rewrite (9) as follows

$$\xi = \sum_{j=1}^{n-1} \frac{\eta^j}{j!} x_j + \frac{\eta^n}{n!} x_n + O\left(\eta^{n+1}\right).$$

Hence,

$$x_n = n! \frac{\xi - \sum_{j=1}^{n-1} \frac{\eta^j}{j!} x_j - O\left(\eta^{n+1}\right)}{\eta^n}. \tag{12}$$

Using a representation for a power series raised to powers (see e.g. [23], p. 17) by taking (10) into account we can write

$$\eta^j = \left(\sum_{i=1}^{\infty} \frac{\xi^i}{i!} y_i \right)^j = \xi^j \left(\sum_{i=0}^{\infty} \frac{\xi^i}{(i+1)!} y_{i+1} \right)^j = \sum_{k=j}^{\infty} P_{j,k} \xi^k, \quad j = 1, 2, \ldots, \tag{13}$$

where

$$P_{j,j} = y_1^j, \quad j = 1, 2, \ldots,$$
$$P_{j,k} = \frac{1}{(k-j)\, y_1} \sum_{l=1}^{k-j} (lj - k + j + l) \frac{y_{l+1}}{(l+1)!} P_{j,k-l}, \quad k = j+1, j+2, \ldots \tag{14}$$

Inserting representation (13) into (12) and enforcing again $\xi \to 0$ we obtain

$$x_n = \lim_{\xi \to 0} n! \frac{\xi - \sum_{j=1}^{n-1} \frac{x_j}{j!} \sum_{k=j}^{\infty} P_{j,k} \xi^k}{\sum_{k=n}^{\infty} P_{n,k} \xi^k}. \tag{15}$$

The above limit exists and is final if

$$x_1 P_{1,1} = 1, \quad \sum_{j=1}^{k} \frac{x_j}{j!} P_{j,k} = 0, \quad k = 2, 3 \ldots, n-1, \tag{16}$$

which leads to

$$x_n = -\frac{n!}{y_1^n} \sum_{j=1}^{n-1} \frac{x_j}{j!} P_{j,n}, \quad n = 2, 3, \ldots \tag{17}$$

It is seen that this solution along with (11) satisfies the conditions (16).

3. Application to the Langevin function

The Langevin function (3) has a removable singularity at $x_0 = 0$ and can be expanded as a Taylor series (5) at this point. Indeed, using the result (see e.g. [1])

$$\coth(x) = \frac{1}{x} + \sum_{n=1}^{\infty} \frac{2^{n+1} B_{n+1}}{(n+1)!} x^n, \tag{18}$$

where B_n, $n = 1, 2, \ldots$, denote Bernoulli numbers, we can write

$$\mathcal{L}(x) = \sum_{n=1}^{\infty} \frac{2^{n+1} B_{n+1}}{(n+1)!} x^n. \tag{19}$$

According to the definition of the Bernoulli numbers $B_n = 0$ for odd n except for $n = 1$ (see e.g. [1]). The values of the first few Bernoulli numbers are given in Table 1. Thus, comparison with (5) yields

$$y_n(0) = \frac{2^{n+1}}{n+1} B_{n+1}, \quad n = 1, 2, 3 \ldots \tag{20}$$

Table 1. The first Bernoulli numbers.

B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9	B_{10}
$-\frac{1}{2}$	$\frac{1}{6}$	0	$-\frac{1}{30}$	0	$\frac{1}{42}$	0	$-\frac{1}{30}$	0	$\frac{5}{66}$

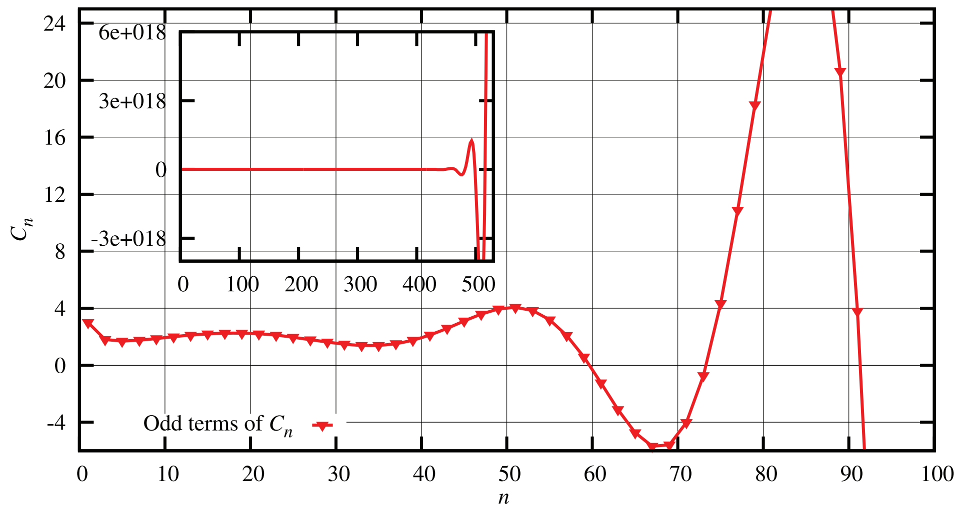

Figure 1. Magnitude of the odd coefficients C_n (21).

In the following we consider the Langevin function on the interval $x \in [0, \infty]$, where it is one-to-one and invertible. Thus, we can write

$$x = \mathcal{L}^{-1}(y) = \sum_{i=0}^{\infty} C_i y^i, \quad C_n = \frac{x_n}{n!}, \quad n = 1, 2, \ldots \tag{21}$$

Taking (17) into account we further obtain

$$C_0 = x_0 = 0, \quad C_1 = \frac{1}{y_1} = 3, \quad C_n = -3^n \sum_{j=1}^{n-1} C_j P_{j,n}, \quad n = 2, 3, \ldots, \tag{22}$$

where in view of (14) and (20)

$$P_{j,j} = \frac{1}{3^j}, \quad j = 1, 2, \ldots,$$
$$P_{j,k} = \frac{3}{(k-j)} \sum_{l=1}^{k-j} (lj - k + j + l) \frac{2^{l+2} B_{l+2}}{(l+2)!} P_{j,k-l}, \quad k = j+1, j+2, \ldots \tag{23}$$

Thus, the coefficients C_n result from the recurrent relations (22) and (23). Accordingly, all even coefficients are zero. The first thirty, odd coefficients obtained by (22) and (23) are given in Table 2.

Of special importance is the convergence of the series (21). Since the coefficients of this series are not monotonic (see Fig. 1), the convergence radius could hardly be defined by conventional approaches, as for example by the d'Alembert criterion. For this reason, we apply a procedure proposed by Mercer and Roberts [24]. To this end, the series of the inverse Langevin function (21) is first rewritten as

$$x = \sum_{i=1}^{\infty} C_{2i-1} y^{2i-1} = \frac{1}{\sqrt{Y}} \sum_{i=1}^{\infty} D_i Y^i, \quad D_j = C_{2j-1}, \quad j = 1, 2, \ldots, \tag{24}$$

Table 2. Taylor series coefficients of the inverse Langevin function.

n	C_n
1	3
3	$\frac{9}{5}$
5	$\frac{297}{175}$
7	$\frac{1539}{875}$
9	$\frac{126117}{67375}$
11	$\frac{43733439}{21896875}$
13	$\frac{231321177}{109484375}$
15	$\frac{20495009043}{9306171875}$
17	$\frac{1073585186448381}{476522530859375}$
19	$\frac{4387445039583}{1944989921875}$
21	$\frac{1000263375846831627}{453346207767578125}$
23	$\frac{280865021365240713}{133337119931640625}$
25	$\frac{148014872740758343473}{75350125192138671875}$
27	$\frac{137372931237386537808993}{76480377070020751953125}$
29	$\frac{41722474198742657618857192737}{25674386102028896409912109375}$
31	$\frac{12348948373636682700768301125723}{8344175483159391333221435546875}$
33	$\frac{5001286000741585238340074032091091}{3590449627018291035442047119140625}$
35	$\frac{185364329915163811141785118512534489}{132846636199676768311355743408203125}$
37	$\frac{6292216384025878939310787532157558451}{4160197291516193533960877227783203125}$
39	$\frac{299869254759556271677902570230858640837}{170568088952163934892395966339111328125}$
41	$\frac{316689568216860631885141475537178451746044283}{148810301691076651811854156590061187744140625}$
43	$\frac{670194310437429598283653289122392937145371137}{258140319260030926612400067554187774658203125}$
45	$\frac{19697015384373759058671314622426656486031717197919}{6332687088594936951180328352888571262359619140625}$
47	$\frac{178793788985653424246012689916144867915861856840849}{49756827124674504616416865629838774204254150390625}$
49	$\frac{323844166067349737493036492206152479344269351967043143667}{82039106319990447886052744467343800746748447418212890625}$
51	$\frac{200808116689754604893460969866238617668631975356485302537199}{49409916306357883385918130190559334540655314922332763671875}$
53	$\frac{27506481209689719715275759452624078040221544551995885750037973}{7164437864421893090958128877631103508395020663738250732421875}$
55	$\frac{16356939619211770477227805130221533318985185730316048281126247721}{5122573073061653560035062147506239008502439774572849273681640625}$
57	$\frac{30976199222209837888906735596203249520053107250807132769871859115868101}{14801698741072505317694781203956860833766632412399612367153167724609375}$
59	$\frac{519588001407316958447129785511020819131555326399179970047767492196701159}{902903623205422824379381653441368510859764577156376354396343231201171875}$

where $Y = y^2$. Then, we introduce the parameter

$$Z_n = \sqrt{\frac{D_{n+1}D_{n-1} - D_n^2}{D_n D_{n-2} - D_{n-1}^2}} \tag{25}$$

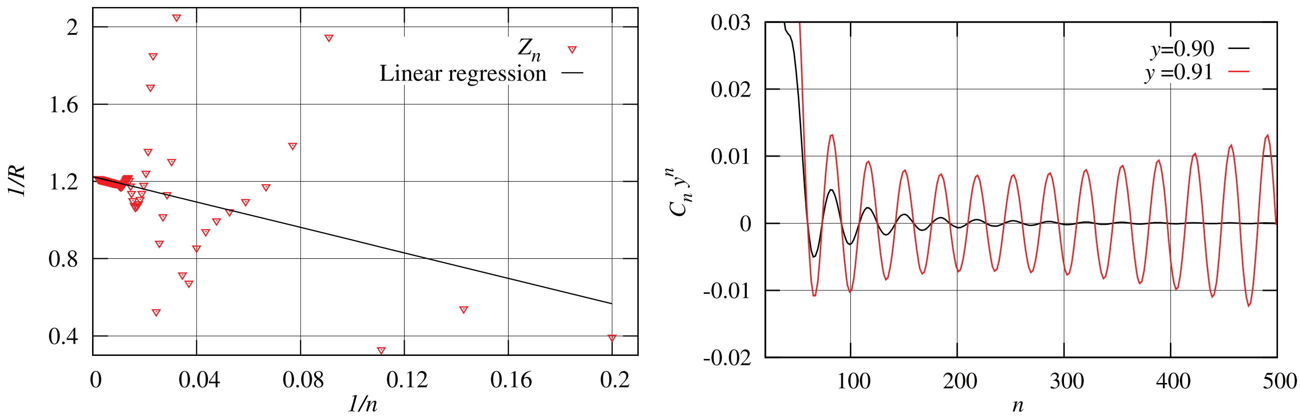

Figure 2. (left) Evaluation of the convergence radius on the basis of the procedure by [24], (right) convergence test of the Taylor expansion near the convergence radius. The even terms of the series are not included.

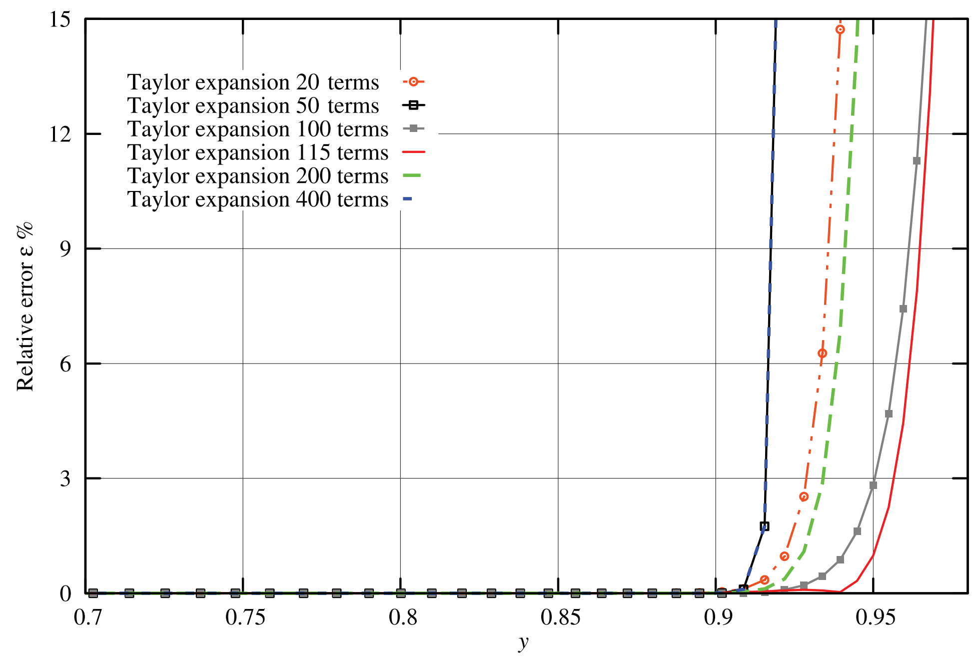

Figure 3. Comparison of the inverse Langevin function series based on different numbers of terms.

and plot it versus $\frac{1}{n}$. The convergence radius R of the series in Y (24) results from

$$\frac{1}{R} = \lim_{\frac{1}{n} \to 0} Z_n. \tag{26}$$

This limit can be evaluated by a simple linear regression illustrated in Fig. 2 (left). On the basis of 500 terms it yields $R \approx 0.816$. Accordingly, the convergence of the initial series in y is calculated by $r = \sqrt{R} \approx 0.904$. This result is illustrated in Fig. 2 (right), where the odd series terms $C_n y^n$ are plotted for two values of y below and above the convergence radius. We observe a crucial difference in the series behavior in spite of the fact that these values of y are close to each other.

Outside the convergence radius, an optimization between the number of coefficients and the accuracy of the calculation is required. In Fig. 3 the series truncated after different numbers of terms are compared in the range $[0, 0.98]$. It is seen that the solution based on 115 series terms exhibits the smallest relative error especially outside the convergence radius.

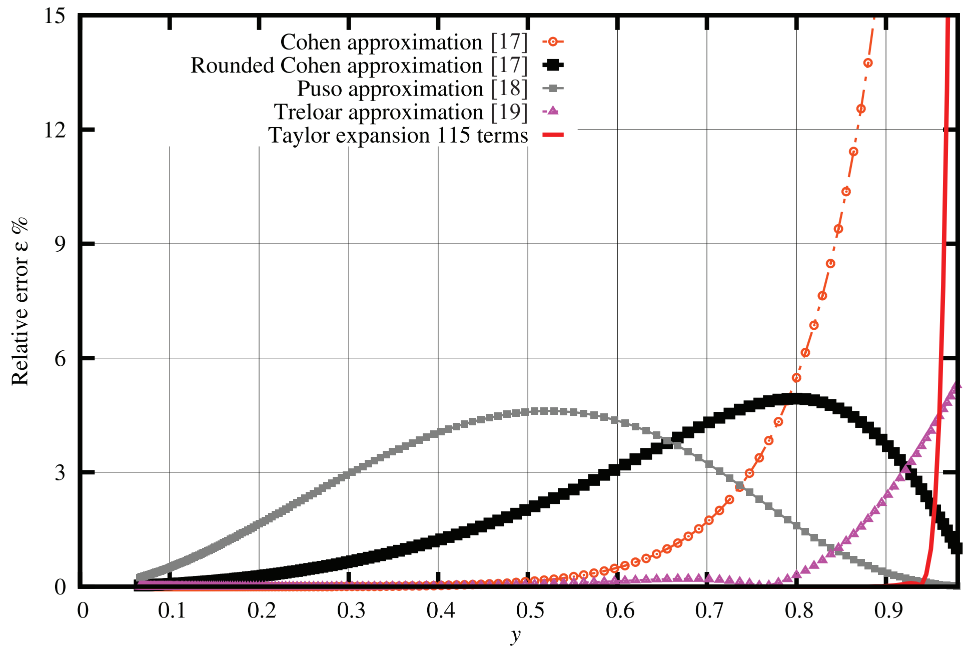

Figure 4. Comparison of different approximations for the inverse Langevin function with respect to the relative error.

In Fig. 4 this solution is compared with other approximations of the inverse Langevin function. One can observe that the solution based on 115 series terms shows the best accuracy within the region $[0, 0.95]$.

4. Conclusion

We have derived a simple recurrent formula for the Taylor series coefficients of an inverse function. The presented procedure only requires the Taylor series coefficients of the initial function to be known. The formula is applied for example to the inverse Langevin function resulting from the statistical theory of rubber elasticity as the chain force. The inverse Langevin function cannot be represented in an explicit form and requires either a series or rational function representation, as for example by a Padé approximation. The resulting series expressed in terms of Bernoulli numbers converges within a relatively wide range of chain stretches except for the area close to the value 1 corresponding to the fully stretched chain. Within this convergence radius the approximation obtained demonstrates better agreement with the exact solution in comparison to different Padé approximants.

As well as the full-network rubber model using the inverse Langevin function discussed above, there are certainly many other examples in mechanics where the proposed solution can be implemented. For instance, Steward et al. [25] presented an analytical formulation of the Cauchy integral applied to boundary value problems with curved boundaries. One indispensable ingredient of this formulation is the Taylor series representation of the inverse function.

Funding

This work was supported by Deutsche Forschungsgemeinschaft (DFG) [grant number IT 67/6-1].

Conflict of interest

None declared.

References

[1] Bronstein, IN, Semendyayev, KA, Musiol, G, and Muehlig, H. *Handbook of Mathematics*. Berlin, Heidelberg, New York: Springer, 2004.

[2] Lagrange, JL. Nouvelle méthode pour résoudre les équations littérales par le moyen des séries. *Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Berlin* 1770; 24: 251–326.

[3] di Bruno, FF. *Théorie des formes binaire.* Brero succr. de P. Marietti, 1876.

[4] Chernoff, H. A note on the inversion of power series. *Mathematical Tables and Other Aids to Computation* 1947; 2: 331–335.

[5] Traub, J. On the nth derivative of the inverse function. *The American Mathematical Monthly* 1962; 69: 904–907.

[6] Jabotinsky, E. Representation of functions by matrices. Application to Faber polynomials. *Proc Amer Math Soc* 1953; 4: 546–553.

[7] Ostrowski, A. Le developpement de Taylor de la fonction inverse. *Comptes rendus de l'Académie des sciences Paris* 1957; 244: 429–430.

[8] Apostol, TM. Calculating higher derivatives of inverses. *The American Mathematical Monthly* 2000; 107: 738–741.

[9] Johnson, WP. Combinatorics of higher derivatives of inverses. *The American Mathematical Monthly* 2002; 109: 273–277.

[10] Kuhn, W, and Grün, F. Beziehungen zwischen elastischen Konstanten und Dehnungsdoppelbrechung hochelastischer Stoffe. *Colloid Polym Sci* 1942; 101: 248–271.

[11] Wu, PD, and van der Giessen, E. On improved 3-D non-Gaussian network models for rubber elasticity. *Mech Res Comm* 1992; 19: 427–433.

[12] Wu, PD, and van der Giessen, E. On improved network models for rubber elasticity and their applications to orientation hardening in glassy polymers. *J Mech Phys Solids* 1993; 41: 427–456.

[13] Gent, AN. A new constitutive relation for rubber. *Rubber Chem Technol* 1996; 69: 59–61.

[14] Horgan, CO, and Saccomandi, GA. A molecular-statistical basis for the Gent constitutive model of rubber elasticity. *J Elast* 2002; 68: 167–176.

[15] Dargazany, R, and Itskov, M. A network evolution model for the anisotropic Mullins effect in carbon black filled rubbers. *Int J Solids Struct* 2009; 46: 2967–2977.

[16] Göktepe, S, and Miehe, C. A micro–macro approach to rubber-like materials. Part III: The micro-sphere model of anisotropic Mullins-type damage. *J Mech Phys Solids* 2005; 53: 2259–2283.

[17] Cohen, A. A Padé approximant to the inverse Langevin function. *Rheologica Acta* 1991; 30: 270–273.

[18] Puso, M. *Mechanistic Constitutive Models for Rubber Elasticity and Viscoelasticity*. PhD thesis, University of California, Davis, 2003.

[19] Treloar, LRG. *The Physics of Rubber Elasticity*. Oxford University Press, 1975.

[20] Itskov, M, Ehret, AE, and Dargazany, R. A full-network rubber elasticity model based on analytical integration. *Math Mech Solids* 2010; 15: 655–671.

[21] Treloar, LRG. The photoelastic properties of short chain molecular networks. *Trans Faraday Soc* 1954; 50: 881–896.

[22] Liptaj, A. Higher order derivatives of the inverse function. http://www.scribd.com/doc/13699758, May 2007.

[23] Gradshteyn, I, and Ryzhik, I. *Table of Integrals, Series, and Products*. Elsevier, 2007.

[24] Mercer, GN, and Roberts, AJ. A centre manifold description of contaminant dispersion in channels with varying flow properties. *SIAM J Appl Math* 1990; 50: 1547–1565.

[25] Steward, DR, Le Grand, P, Jankovic, I, and Strack ODL. Analytic formulation of Cauchy integrals for boundaries with curvilinear geometry. *Proc Roy Soc Lond* 2008; 464: 223–248.