1

Some refelections about the Lambert W Function as inverse of $x \cdot \log(x)$

Ulrich Tamm

***Abstract*— The Lambert W function fulfills $W(y) \cdot e^{W(y)} = y$. With the choice $y = \log(x)$ it can hence be applied to invert the function $f(x) = x \cdot \log(x)$, which is of some interest in the problems discussed. Further applications of the Lambert W function in information theory are briefly surveyed.**

***Index Terms*— Lambert W function, take away games, binary entropy**

I. Introduction

Some reflections will be presented about situations in which the function $f(x) = x \cdot \log(x)$ has to be inverted. Actually, this was motivated a long time ago by the analysis of a certain take–away game, a generalization of the well–known Fibonacci Nim. In the formula for the winning positions a recursion $a(n) = a(n-1) + a(n-k)$ occured where k is about $c \cdot \ln(c)$ for some $c > 1$.

At the Workshop "Information Theory in Mathematics", Balatonlelle, Hungary, 2000, the author presented a short lecture about the problem and asked if somebody in the audience might have some idea about how to invert $x \cdot \log(x)$. In the coffee break Benjamin Weiss pointed out the (those days) recent paper by Corless, Gonnet, Hare, Jeffrey and Donald Knuth [9] about the Lambert W function.

The Lambert W function, actually, is the inverse of $y \cdot e^y$, but with the choice of $y = \ln(x)$ also $x \cdot \ln(x)$ can be inverted. Several formulae and applications are listed in [9]. For instance, also tower functions of the form $x^{x^{x^{\cdots}}}$ might be inverted via Lambert W.

Since $x \cdot e^x$ is not an injective function, the inverse function splits into two branches, $W_0(x)$ for $x \geq \frac{1}{e}$ and $W_{-1}(x)$ for $x < \frac{1}{e}$. W_0 has the nice Taylor expansion (related to the enumeration of trees)

$$W_0(x) = \sum_{n=0}^{\infty} \frac{(-n)^{n-1}}{n!} x^n$$

Unfortunately, application of the Lambert W function – more exactly, its branch W_0 – did not yield much insight into the recursion under discussion for the take–away game. However, some interesting problems remain, which are stated in the next section.

The reason to ask information theorists about the inverse of $x \cdot \log x$ was, of course, the occurrence of this term in the entropy function. In Section 2 it is briefly shown that it can be used to approximately determine the probability p for a given binary entropy if p is small. Here, the branch W_{-1} comes into play. However, the approach does not seem to be very useful because an expression of the Lambert function in terms of elementary functions does not seem to be at hand.

Finally, in Section 3, the development of applications of the Lambert W function in information theory since 1996 is briefly surveyed.

II. Inversion of $x \cdot \log x$

Schwenk [26] found the winning positions of a certain take-away game via the recursion

$$a(n) = a(n-1) + \min\left\{a(i) : a(i) \geq \frac{a(n-1)}{c}\right\}$$

Here, $c > 1$ is a real number determining the rules of the game: in each turn the player is allowed to remove at most c times as many pieces from the pile as his opponent in his turn before.

In [26] it was further shown that finally

$$a(n) = a(n-1) + a(n-k)$$

for a fixed number $k = k(c)$. The game is also discussed by Berlekamp, Conway, and Guy in volume 3 of their famous work "Winning Ways for Your Mathematical Plays", where they ask for the values of k, when c is an integer number. For instance for $c = 2$ the Fibonacci numbers arise, hence $k(2) = 1$. For small k exact values are known, in general, $k(c)$ is about $c \cdot \ln(c)$, e.g., [33], where "about" means that

$$c \cdot \ln(c) - \ln(c) < k(c) < c \cdot \ln(c) + \ln(c)$$

Our computer observations suggest that $k(c)$ is indeed very close to $c \cdot \ln(c)$.

For this reason the interest in the inverse function of $c \cdot \ln(c)$ arose. Actually, application of the Lambert W function did not yield the desired insight. However, two interesting questions arose:

Question 1: Are there further applications of sequences $x_1, x_2, ...$ with almost constant differences $x_{i+1} \cdot \ln(x_{i+1}) - x_i \cdot \ln(x_i)$?

Question 2: A number approximately 0.7322... seems to involved. What is this number? It might be related to the gamma function.

Another situation where the inversion of $c \cdot \ln(c)$ might come into play is the conclusion from a given value of the binary entropy to the underlying probabilities.

U. Tamm is with the Department of Economy, University of Applied Sciences, Bielefeld, Germany on leave from the German Language Department of Business Informatics, Marmara University, Istanbul (email tamm@ieee.org)

Of course, $h(p) = -p \cdot \log_2(p) - (1-p) \cdot \log_2(1-p)$ is the sum of two terms but for small p the first summand is strongly dominating. Further, then the approximation $\ln(1-p) \approx -p$ is rather accurate. Thus

$$h(p) = -\frac{1}{\ln 2}\big(p \cdot \ln(p) + (1-p) \cdot \ln(1-p)\big)$$
$$\approx -\frac{1}{\ln 2}\big(p \cdot \ln(p) + (1-p) \cdot (-p)\big)$$
$$\approx -\frac{1}{\ln 2}\big(p \cdot \ln(p) - p\big) = -\frac{1}{\ln 2} \cdot p \cdot (\ln(p) - 1)$$
$$= -\frac{1}{\ln 2} \cdot p \cdot \ln(\tfrac{p}{e}) = -\frac{1}{\ln 2} \cdot e \cdot \frac{p}{e} \cdot \ln(\tfrac{p}{e}).$$

Hence, given a value $-\ln(p) - \ln(1-p) = x$ a small p may be approximated by $\overline{p} = e^{W_{-1}(\frac{x}{e}+1)}$.

If one is only interested in the numerical values this might be quite okay. For instance, for $x = 0.19851524$ it is $p = 0.05$ and $\overline{p} = 0.049576$, and for $x = 0.056002$ it is $p = 0.01$ and $\overline{p} = 0.09989$.

On the other hand, e.g., via Newton iteration the exact values can be obtained.

More interesting might be a further simplification via an analytical approximation of the Lambert W expression. Unfortunately, there are two obstacles here.

First, the W function returns the logarithm of p, which for small p has a large absolute value outside the convergence radius of the Taylor series.

Secondly, the existing approximation formulae, e.g., [2], [7], [31] would yield a polynomial with $\ln(|\ln x|)$ as argument which seems to be difficult to handle.

III. Further applications in information theory

After Corless et al. had published their paper [9] the Lambert W function attracted a lot of interest. It is implemented in the standard mathematical software packages as Maple, Mathematica, Matlab, etc. There is even a discussion in the mathematical community that it might deserve the status of a special function (as, for instance, the gamma function). This is mainly due to the many applications in between found in physics, differential equations, combinatorics and graph theory, etc. (cf. [14], [5])

The development with most impact for electrical engineering seems to be the description of the relation between voltage current and resistance in a diode in terms of the Lambert W function [1].

Soon after the publication of [9] the Lambert W function entered the literature in various areas of information and coding theory, e.g., [29], [19], [10], [17], [18], [24], [25], [30] and related fields as signal processing [6], networks [4], [8], compression [23] or image coding [11], [32]

The list is by far not complete but it shows that the applications seem to fall into the following categories:

1) Asymptotic analysis of codes using the Taylor series of the Lambert W function, e.g., [29], [15],

2) Optimization problems in which an expression of type $w \cdot e^w$ or $y \cdot \log y$ arises, e.g., [10], [19], [25], [30],

3) New entropy concepts obtained by modifying the entropy axioms - mostly discussed by researchers in physics, e.g., [16], [27],

4) Analysis of Kullback-Leibler divergence, e.g, [11], [20], [28].

5) Related statistics, for instance, analysis of heavy tailed data [12], [13], [21], [22]

References

[1] T.C. Banwell, "Bipolar transistor circuit analysis using the Lambert function", *IEEE Trans. Circuits and Systems - I. Fundamental Theory and Applications*, vol. 47(6), pp. 1621-1633, 2000.

[2] D.A. Barry, J.Y. Parlange, L. Li, H. Prommer, C.J. Cunningham and F. Stagnitti, "Analytical approximations for real values of the Lambert W-function", *Mathematics and Computers in Simulation*, vol. 53, pp. 95-103, 2000.

[3] E.R. Berlekamp, J.H. Conway and R.K. Guy, *Winning Ways for Your Mathematical Plays*, Academic Press, 1982.

[4] F. Brah, A. Zaidi, J. Louveaux and L. Vandendorpe, "On the Lambert-W function for constrained resource allocation in cooperative networks", *EURASIP Journal on Wireless Communications and Networking*, pp. 1-13, 2011.

[5] P. Brito, F. Fabiao and A. Staubyn, "Euler, Lambert, and the Lambert W function today", *The Mathematical Scientist*, vol. 33(2), 2008.

[6] F. Chapeau-Blondeau and A. Monir, "Numerical evaluation of the Lambert W-function and application to generation of generalized Gaussian noise with exponent 1/2", *IEEE Trans. Signal Processing*, vol. 50(9), pp. 2160-2165, 2002.

[7] I. Chatzigeorgiou, "Bounds on the Lambert function and their application to the outage analysis of user cooperation", *IEEE Communications Letters*, vol. 17(8), pp. 1505-1508, 2013.

[8] D. Chen, M. Haenggi and J.N. Laneman, "Distributed spectrum efficient routing algorithms in wireless networks", *IEEE Trans. Wireless Communications*, vol. 7(12) , pp. 5297-5305 , 2008.

[9] R.M. Corless, G.H.Gonnet, D.E.G. Hare, D.J. Jeffrey and D.H. Knuth, "On the Lambert W function", *Adv. Comput. Math.*, vol. 5, pp. 329-359, 1996.

[10] T. Cui, T. Ho and J. Kliewer, "Memoryless relay strategies for two-way relay channels: performance analysis and optimization", *IEEE Trans. Communications*, vol. 57(10), pp. 3132-3143, 2009.

[11] M. El Gheche, J.C. Pesquet and J. Farah, "A proximal approach for optimization problems involving Kullback divergences", *Proceedings IEEE ICASSP*, Vancouver, pp. 26-31, 2013.

[12] G.M. Goerg, "Lambert W random variables - a new family of generalized skewed distributions with applications to risk estimations", *The Annals of Applied Statistics*, vol. 5(3), pp. 2197-2230, 2011.

[13] G.M. Goerg, "The Lambert way to Gaussianize skewed, heavy tailed datawith the inverse of Tukey's h transformation as a special case", *arXiv: 1010.2265v5 [math.ST] 30. Dec. 2012*, 2012.

[14] B. Hayes, "Why W?", *American Scientist*, vol. 93, pp. 104-108, 2005.

[15] H.K. Hwang, "Uniform asymptotics of some Abel sums arising in coding theory", *Theoretical Computer Science*, vol. 263, pp. 145-158, 2001.

[16] P. Jizba and T. Arimitsu, "Towards information theory for q-nonextensive statistics without q-deformed distributions", *Physica A*, vol. 365, pp. 76-84, 2006.

[17] C. Koller, M. Haenggi, J. Kliewer and D. Costello Jr., "Joint channel/network coding for star networks", *Proceedings ISIT 2013*, Istanbul, 2013.

[18] Y. Liang, A. Goldsmith and M. Effros, "Distortion metrics of composite channels with receiver side information", *Proceedings IEEE ITW07*, Lake Tahoe, pp. 559-564, 2007.

[19] P.W. Moo and D.L. Neuhoff, "Uniform polar quantization revisited", *Proceedings IEEE ISIT 1998*, MIT, p. 100, 1998.

[20] F. Nielsen, "On the symetrical Kullback–Leibler Jeffreys centroids", *IEEE Signal Processing Letters*, vol. 20(7), pp. 657-660, 2013.

[21] A.G. Pakes, "Lambert's W, infinite divisibility and Poisson mixtures", *J. Math. Analysis Appl.*, vol. 378, pp. 480-492, 2011.

[22] , "Applications of the Lambert W function", *Int. J. Appl. Math. Statitics*, vol. 23(D11), 2011.

[23] G. Retvari, J. Tapolcai, A. Körösi, A. Majdan and Z. Heszberger, "Compressing IP forwarding tables: towards entropy bounds and beyond", *Proceedings SIGCOMM 13*, Hong Kong, pp. 111-122, 2013.

[24] I.Sason and H.D. Pfister, "Recent results on capacity achieving codes for the erasure channel with bounded complexity", *Proceedings IEEE 24th Convention of Electrical and Electronics Engineers in Israel*, Eilat, pp. 339-343, 2006.

[25] I. Sason and R. Urbanke, "Parity-check density versus performance of binary linear block codes over memoryless symmetric channels", *IEEE Trans. Information Theory*, vol. 49(7), pp. 1611-1635, 2003.

[26] A. Schwenk, ”Take away games”, *The Fibonacci Quarterly*, vol. 8(3), pp. 225-234, 1970.
[27] F. Shafee, ”Lambert function and a new non-extensive form of entropy”, *IMA J. Applied Mathematics*, vol. 72 , pp. 785-800, 2007.
[28] , ”On some exact densities in an exponential family”, *Robust 2000 (J. Antoch and G. Dohnal, eds.)*, pp. 298-307, Praha, 2001.
[29] W. Szpankowski, ”On asymptotics of certain recurrences arising in universal coding”,*Probl. Pered. Inf.*, vol. 34(2), pp. 56-61, 1998.
[30] S. Toumpis, R. Müller and J. Sayir, ”On the transport capacity of a multiple access Gaussian channel”, *Proceedings Int. Workshop on Wireless Ad Hoc Networks*, Oulu, Finnland, 2004.
[31] D. Veberic, ”Having fun with the Lambert W(x) function”, *arXiv:1003.1628v1 [cs.Ms] March 8*, 2010.
[32] J. Wright and Li Ma, ”Dense error correction via l1 minimization”, *Proceedings IEEE ICASSP*, 2009.
[33] M. Zieve, ”Take away games”, pp. 351-361 in *Games of No Chance*, *MSRI Publications*, vol. 29, Cambridge University Press, 1996.