Generalized entropies and logarithms and their duality relations

Rudolf Hanel\textsuperscript{a}, Stefan Thurner\textsuperscript{a,b}, and Murray Gell-Mann\textsuperscript{b,1}

\textsuperscript{a}Section for Science of Complex Systems, Medical University of Vienna, 1090 Vienna, Austria; and \textsuperscript{b}Santa Fe Institute, Santa Fe, NM 87501

Contributed by Murray Gell-Mann, September 28, 2012 (sent for review August 6, 2012)

For statistical systems that violate one of the four Shannon–Khinchin axioms, entropy takes a more general form than the Boltzmann–Gibbs entropy. The framework of superstatistics allows one to formulate a maximum entropy principle with these generalized entropies, making them useful for understanding distribution functions of non-Markovian or nonergodic complex systems. For such systems where the composability axiom is violated there exist only two ways to implement the maximum entropy principle, one using escort probabilities, the other not. The two ways are connected through a duality. Here we show that this duality fixes a unique escort probability, which allows us to derive a complete theory of the generalized logarithms that naturally arise from the violation of this axiom. We then show how the functional forms of these generalized logarithms are related to the asymptotic scaling behavior of the entropy.

\[ S_{cd} \propto \sum \Gamma(1 + d, 1 - c \log p_i), \]

where \( \Gamma \) is the incomplete gamma function and \((c, d)\) are constants that are uniquely determined by the scaling properties of the statistical system in its thermodynamic limit. In previous work (7) we were able to show that for systems where the first three SK axioms hold, there exist only two ways to formulate a consistent maximum entropy principle. Starting with an entropy of “trace form,”

\[ S[p] = \sum_{i=1}^{W} s(p_i), \]

the maximization condition becomes \( \delta \Phi = 0 \), with

\[ \Phi[p] = S[p] - \alpha \left( \sum p_i \log p_i \right) - \beta \left( \sum \Lambda_i[p, s] e_i - U \right), \]

where the last two terms are the constraints. The first of the two possible approaches [Hanel–Thurner (HT) approach] (8, 9) uses a generalized entropy and the usual form of the constraint,

\[ Q_{HT}[p] = p_i. \]

The other approach, suggested in Tsallis and Souza (10) (TS approach), uses a generalized entropy and a more general way to impose constraints:

\[ Q_{TS}[p, s] = \frac{p_i + \nu s(p_i)}{\sum p_i + \nu s(p_i)}. \]

\( P_i \) is a so-called escort probability and \( \nu \) is a real number. Though in the HT case the constraint has the usual interpretation as an energy constraint, we do not attempt to give a physical interpretation of the escort probabilities. The two approaches have been shown to be connected by a duality map \( \cdot \colon S_{HT} \leftrightarrow S_{TS} \), with “” (meaning applying “” twice) being the identity (7). A special case of this duality has been observed in Ferri et al. (11).

Entropies can be conveniently formulated using their associated generalized logarithms. We first specify the space \( \mathcal{L} \) of proper generalized logarithms \( \Lambda \in \mathcal{L} \). We consider a generalized logarithm to be proper if the following properties hold: (i) \( \Lambda \) is a differentiable function \( \Lambda: \mathbb{R}^+ \rightarrow \mathbb{R} \). This is necessary for a finite second derivative of the entropy; (ii) \( \Lambda \) is monotonically increasing, which is a consequence of the second SK axiom; (iii) \( \Lambda(1) = 0 \), which captures the requirement that the entropy of single-state systems is 0; and (iv) \( \Lambda(1) = 1 \), is needed to fix the units of entropy.

In both approaches (HT and TS) there exist proper generalized logarithms \( \Lambda_{HT} \) and \( \Lambda_{TS} \) such that

\[ s_{HT}(p_i) = -k \int_0^{p_i} dx \Lambda_{HT}(x/x_0), \]

and

\[ s_{TS}(p_i) = -k \int_0^{p_i} dx \Lambda_{TS}(x/x_0), \]

with \( x_0 \) a constant. If both approaches predict the same distribution function \( p = \{ p_i \}_{i=1}^{W} \) as a result of the maximization of Eq. 3, then it can be shown that the two entropic functions \( s_{HT} \) and \( s_{TS} \) are one-to-one related by

\[ \frac{1}{\Lambda_{TS}(x)} - \frac{1}{\Lambda_{HT}(x)} = kU. \]
In the following, we set $k = 1$; this can be achieved either by choosing physical units accordingly, or by simply absorbing $k$ into $\nu$, so that $\nu$ becomes a dimensionless parameter.

The full implication of Eq. 7, which is related to the essence of this paper, can be summarized as follows. The statistical properties of a physical system—for instance, a superstatistical system as discussed in Hanel et al. (7)—uniquely determine the entropy $S_{HT}$. A priori, there exists a spectrum of TS entropies, $S_{TS,\nu}$, whose boundaries are determined by the properties of the generalized logarithm associated with $S_{HT}$. Moreover, these properties determine a particular value $\nu^*$, so that $S_{TS,\nu}$ and $S_{HT}$ become a pair of dual entropies. This unique duality allows us to derive a complete theory of generalized logarithms naturally arising as a consequence of the fourth SK axiom being violated. We present a full understanding of how the TS and HT approaches are interrelated and derive the most general form of families of generalized logarithms that are compatible with a maximum entropy principle and the first three SK axioms. Finally, we demonstrate how these logarithms can be classified according to their asymptotic scaling properties, following the results presented in Hanel and Thurner (6).

### Duality

In contrast to the images of generalized logarithms, which need not span $\mathbb{R}$ completely and can differ from one another, the domain of generalized logarithms themselves is always all of $\mathbb{R}$. For these reasons, one may classify generalized logarithms according to the minimum and maximum values of their images and consider the group $G$ of order-preserving automorphisms on $\mathbb{R}$, that keep an infinitesimal neighborhood of $1 \in \mathbb{R}$, invariant, as the means to generate these classes. In the following, we call the elements of $G$ of this automorphism group scale transformations. More precisely, $g \in G$ is a scale transformation if it is differentiable and maps $\mathbb{R}^+$ to $\mathbb{R}^+$, one-to-one, $g(0) > 0$, $g(1) = 1$, and $g'(1) = 1$. From these properties it follows that $g(0) = 0$ and $\lim_{x \to \infty} g(x) = \infty$. Finally, we use the notation $f \circ g(x) = f(g(x))$.

Scale transformations leave the image of a generalized logarithm invariant, which allows us to parameterize classes in the following way. Given a proper generalized logarithm $\Lambda \in \mathcal{L}$, we write for its maximum and minimum values

$$\Lambda \equiv \min\{\Lambda(x) | x \in \mathbb{R}^+\}, \quad \Lambda^\prime \equiv \max\{\Lambda(x) | x \in \mathbb{R}^+\},$$

and define two functions

$$\nu_-[\Lambda] \equiv -\frac{1}{\Lambda} \leq 0, \quad \nu_+[\Lambda] \equiv -\frac{1}{\Lambda} \geq 0,$$

which associate numbers $\nu_-$ and $\nu_+$ to any $\Lambda$. For their sum we write $\nu^* = \nu_+ + \nu_-$. Next, we define sets of proper generalized logarithms,

$$\mathcal{L}_{\nu^*} \equiv \{ \Lambda \in \mathcal{L} \mid \nu_-[\Lambda] = \nu_+[\Lambda] \text{ and } \nu_- = \nu_+[\Lambda] \}.$$  

Members of $\mathcal{L}_{\nu^*}$ all have the same maximum and minimum values. In fact, the $\mathcal{L}_{\nu^*}$ are exactly the equivalence classes in $\mathcal{L}$ generated by $G$; two generalized logarithms $\Lambda^{(A)}$ and $\Lambda^{(B)}$ are considered equivalent if there exists a scale transformation $g \in G$ such that $\Lambda^{(B)} = \Lambda^{(A)} \circ g$. The space of generalized logarithms can be written as the union of these sets, $\mathcal{L} = \cup_{\nu^*} \mathcal{L}_{\nu^*}$.

With these definitions we now analyze the relation between the HT and TS approaches. Assuming that $\Lambda_{HT}$ is given, Eq. 7 implies

$$\Lambda_{TS,x}(x) = T_x \circ \Lambda_{HT}(x), \quad T_x(x) = \frac{1}{\frac{1}{x} + \nu}$$

$T_x$ is a shift operator with the property $T_x \circ T_y = T_{x+y}$. We have of course $\Lambda_{TS,x} = \Lambda_{HT}$. The fact that $\Lambda_{HT}$ is a proper generalized logarithm does not imply that $\Lambda_{TS,x}$ is also proper for all choices of $\nu$.

In fact, given that $\Lambda \in \mathcal{L}$, it can be shown (SI Materials and Methods) that $T_x \circ \Lambda \in \mathcal{L}$ if and only if $\nu_-[\Lambda] \leq \nu \leq \nu_+[\Lambda]$. Moreover, for $\Lambda \in \mathcal{L}_{\nu^*}$, and for $T_x \circ \Lambda$ being a proper generalized logarithm, it follows that $T_x \circ \Lambda \in \mathcal{L}_{\nu^*}$. As a consequence $\Lambda_{TS,x}(x) = T_x \circ \Lambda_{HT}(x)$ is proper only for $\nu_-[\Lambda_{HT}] \leq \nu \leq \nu_+[\Lambda_{HT}]$, and

$$\Lambda_{TS,x} \in \mathcal{L}_{\nu^*} \iff \Lambda_{HT} \in \mathcal{L}_{\nu^*}.$$  

This equation does not uniquely determine a duality relation * on $\mathcal{L}$, yet by imposing the condition that * commute with scale transformations $g \in G$, it can be shown (SI Materials and Methods) that * is given by

$$\Lambda^* = T_{\nu^*} \circ \Lambda \quad \text{for} \quad \Lambda \in \mathcal{L}_{\nu^*},$$

with the property

$$\Lambda \in \mathcal{L}_{\nu^*} \iff \Lambda^* \in \mathcal{L}_{\nu^*}.$$  

Thus, for each $\Lambda_{HT}$ there exists a unique value $\nu^* = \nu_+[\Lambda_{HT}] + \nu_-[\Lambda_{HT} + \nu_-[\Lambda_{HT}]]$ such that $\Lambda_{TS,x}$ is a proper generalized logarithm. The duality map * gives $\Lambda_{TS,x} = \Lambda_{HT}$. Furthermore, because $*$ and $g$ commute ($(g \circ \Lambda)^* = \Lambda^* \circ g$), any proper generalized logarithm $\Lambda$ can be decomposed into a specific representative $\Lambda_{\nu^*} \in \mathcal{L}_{\nu^*}$, and a scale transformation $g$, so that

$$\Lambda = \Lambda_{\nu^*} \circ g,$$

which implies that any $\Lambda_{HT}$ or $\Lambda_{TS,x}$ can be decomposed in this way, and that the dual logarithms $\Lambda_{HT}$ and $\Lambda^*_{HT} = \Lambda_{TS,x}$ transform identically under scale transformations.

### Functional Form of the Generalized Logarithms

Eq. 14 implies the existence of transformations that map members of $\mathcal{L}_{\nu^*}$ to members of $\mathcal{L}_{\nu^*}$. These maps can be used to represent the duality * on specific families $\Lambda_{\nu^*} \in \mathcal{L}_{\nu^*}$. $\Lambda(x) \to -\Lambda(1/x)$ is exactly such a map, because $\max\{-\Lambda(1/x)\mid x \in \mathbb{R}^+\} = \min\{\Lambda(x)\mid x \in \mathbb{R}^+\} = -\Lambda$. The same holds for $\min\{-\Lambda(1/x)\mid x \in \mathbb{R}^+\} = -\Lambda$, which allows us to construct $\Lambda_{\nu^*}$ with the properties

$$\Lambda_{\nu^*}(x) = -\Lambda_{\nu^*}(1/x).$$

By using Eq. 13 and inserting $\Lambda_{\nu^*}(x) = -\Lambda_{\nu^*}(1/x)$ into Eq. 7, we get

$$\frac{1}{\Lambda_{\nu^*}(1/x)} + \frac{1}{\Lambda_{\nu^*}(x)} = -\left(\nu_+ + \nu_-\right) = -\nu^*.$$  

This equation may have many solutions $\Lambda_{\nu^*}$, but we can restrict ourselves to finding a particular one; all of the others can be obtained by scale transformations, which is seen as follows: Suppose $\Lambda^{(A)}_{\nu^*}$ and $\Lambda^{(B)}_{\nu^*}$ are both solutions of Eq. 17; then according to Eq. 15 for any pair $(\nu_+, \nu_-)$ there exists a scale transformation $g_{\nu^*}$ such that $\Lambda^{(B)}_{\nu^*} = \Lambda^{(A)}_{\nu^*} \circ g_{\nu^*}$. Because $g_{\nu^*}$ must leave $\Lambda_{\nu^*}$ invariant (this is not the case for arbitrary scale transformations $g \in G$), these scale transformations have two properties. The first property is $g_{\nu^*}^{-1}(x)g_{\nu^*}(1/x) = 1$, which makes them members of a subgroup $g \in G$ of all possible scale transformations $g \in G$. The second property is $g_{\nu^*}^{-1} = g_{-\nu^*}$, and follows from the fact that * commutes with scale transformations.
A particular solution of Eq. 17 is given by

$$\Lambda_{\nu, \nu_c}(x) = \left( \frac{1}{\nu_c x^2 h\left( \frac{x}{2} \log(x) \right)} - \frac{\nu_c + \nu}{2} \right)^{-1},$$  \[18\]

with $h: \mathbb{R} \to [-1, 1]$ a continuous, monotonically increasing, odd function, with $\lim_{x \to \infty} h(x) = 1$ and $h'(0) = 1$. It can easily be verified that this solution has all of the required properties: $\Lambda_{\nu, \nu_c}$ is a proper logarithm with $\Lambda_{\nu, \nu_c} \in E_{\nu, \nu_c}$ (correct minimum and maximum), $\Lambda_{\nu, \nu_c}(1/x) = \Lambda_{-\nu, -\nu_c}(x)$, $\Lambda_{\nu, \nu_c}(x) = h(\nu \log(x))/\nu$ is self-dual, and $\lim_{x \to 0} \lim_{x \to \infty} \Lambda_{\nu, \nu_c}(x) = \log(x)$.

The above argument means that we can generate a specific family of logarithms $\Lambda_{\nu, \nu_c}$, following Eq. 16, by choosing one particular function $h$ [e.g., $h(x) = \tanh(x)$] and then using scale transformations to reach all other possibilities. In particular, some family $\Lambda_{\nu, \nu_c}$ with the property $\Lambda_{\nu, \nu_c}(x) = -\Lambda_{-\nu, -\nu_c}(x)$ can be reached by a family of scale transformations $g_{\nu, \nu_c} = \mathcal{E}_{\nu, \nu_c} \circ \Lambda_{\nu, \nu_c} \in \mathcal{G}$, where $\mathcal{E}_{\nu, \nu_c} \equiv \Lambda_{\nu, \nu_c}$ are generalized exponential functions (inverse functions of logarithms). Moreover, if $\Lambda_{\nu, \nu_c}$ also follows Eq. 16, then $g_{\nu, \nu_c} \in \mathcal{G}$.

The family of dual logarithms discussed in Hanel et al. (7) is obtained in the framework presented here by setting either $\nu_c = 0$ or $\nu = 0$. These classes correspond to logarithms that are unbounded either from below or from above, whereas the duality maps $\mathcal{E}_{0,0} \leftrightarrow \mathcal{E}_{\nu, \nu_c}$. Moreover, in Hanel et al. (7), only pairs of dual logarithms have been considered such that $\Lambda^*(x) = -\Lambda(1/x)$, and the part that scale transformations play in the unique definition of $^*$ had not yet been described.

We are now in a position to understand all observable distribution functions emerging from the two approaches in terms of a single two-parameter family of generalized logarithms $\Lambda_{\nu, \nu_c}$ and a scale transformation. This result now raises the question of how $\Lambda_{\nu, \nu_c}$ is related to the two-parameter logarithms associated with the $(c, d)$ entropies in Eq. 1 (6), and will further clarify the role of the scale transformations.

\section*{Logarithm and $(c, d)$ Entropy}

Generalized entropies can be classified with respect to their asymptotic scaling behavior in terms of two scaling exponents, $c$ and $d$, where $0 < c \leq 1$ and $d$ is a real number (6); they are obtained from the scaling relations

$$x^c = \lim_{x \to 0} \frac{s(\Lambda(x))}{s(x)} \quad \text{and} \quad (1 + a)^d = \lim_{x \to 0} \frac{s(x^{1+a})}{s(x)} \quad \text{for} \quad \Lambda \in \mathcal{E},$$  \[19\]

where $s$ is the summand in Eq. 2. Using Eq. 19, de l’Hôpital’s rule, and the fact that $s'(x) = -\Lambda(x)$, we find the exponents $(c, d)$ for a given $\Lambda \in \mathcal{E}$.

$$x^c = \lim_{x \to 0} \frac{\Lambda(\Lambda(x))}{\Lambda(x)} \quad \text{and} \quad (1 + a)^d = \lim_{x \to 0} \frac{\Lambda(x^{1+a})}{\Lambda(x)} \quad \text{for} \quad \Lambda \in \mathcal{E},$$  \[20\]

where we represent $\Lambda$ as $\Lambda = \Lambda_{\nu, \nu_c} \circ g$. In this way we get the dependence of $(c, d)$ as a function of $(\nu_c, \nu_c)$, $h$, and the scale transformation $g$. We first compute the asymptotic properties of $h$ and $g$, defining the exponents $c_{g,h}$ and $d_{g,h}$ by

$$x^c = \lim_{x \to 0} \frac{\phi_{g,h}(\Lambda(x))}{\phi_{g,h}(x)} \quad \text{and} \quad (1 + a)^d = \lim_{x \to 0} \frac{\phi_{g,h}(x^{1+a})}{\phi_{g,h}(x)} \quad \text{for} \quad \Lambda \in \mathcal{E},$$  \[21\]

where $\phi_{g,h} = 1 + h \circ \log \circ g(x)$. Note that $\log \circ g \in \mathcal{E}_{0,0}$. By defining $\Lambda_0 \equiv \log \circ g$, we compute its scaling exponents $c_0$ and $d_0$.

$$x^{c_0} = \lim_{x \to 0} \frac{\Lambda(\Lambda(x))}{\Lambda(x)} \quad \text{and} \quad (1 + a)^d_0 = \lim_{x \to 0} \frac{\Lambda(x^{1+a})}{\Lambda(x)} \quad \text{for} \quad \Lambda \in \mathcal{E},$$  \[22\]

With these preparations one can derive the results

$$c = \begin{cases} 1 & \text{for } \nu_c \neq 0, \\ 1 - c_{g,h} & \text{for } \nu_c = 0 \text{ and } c_0 \neq 1, \\ 1 - c_{g,h}(-\frac{c_0}{c_0 - 1}) & \text{for } \nu_c = 0 \text{ and } c_0 = 1, \\ 0 & \text{for } \nu_c = 0, \\ -d_{g,h} & \text{for } \nu_c = 0 \end{cases}$$  \[23\]

which demonstrate clearly that, given a fixed $h$, $c$ is controlled by $\nu_c$. (For $\nu_c = 0$ and $c_0 = 1$ and $d$ is determined by the scale transformation.

\section*{Examples}

\subsection*{Example 1. A simple choice for $h$}

For example, fix $h(x) = \tanh(x)$. From Eq. 18 we get for the generalized logarithm

$$\Lambda_{\nu, \nu_c}(x) = \frac{x^{\nu - \nu_c} - 1}{\nu - \nu_c x^{\nu - \nu_c}}.$$  \[24\]

The associated generalized exponential (inverse of the generalized logarithm) is

$$\mathcal{E}_{\nu, \nu_c}(x) = \left( \frac{1 + \nu + x}{1 + \nu + x} \right)^{-\nu - \nu_c}.$$  \[25\]

\subsection*{Example 2. Power laws}

By setting $h(x) = \tanh(x)$ and $\nu_c = 0$, we get from Eq. 24 the so-called $q$-logarithm, with $\Lambda_{0,q} = \log_q(x) = (1 - (1 - q)x)^{1/(1-q)}$, where $0 \leq q = 1 - \nu_c \leq 1$. The dual is $\Lambda_{q,0} = \log_{1/(1-q)}(x)$, and we recover the well-known duality for $q$-logarithms. It is also well known that $\log_q$ results from the use of escort distributions (12, 13, 10), whereas $\log_{2^{-q}}$ is a natural result of the HT approach (8, 9).

An example of a generalized logarithm that is not a power is obtained by taking $\nu_c = -\frac{1}{c_0}$ in Eq. 24. One obtains $\Lambda_{\nu, \nu_c}(x^{1+c_0} - 1)/(\nu_c (x + 1))$, with the dual $\Lambda_{-\nu, -\nu_c} = (x^{c_0} - 1)/(\nu_c (x^{1+c_0} - 1))$.

\subsection*{Example 3. Scale transformations}

Any generalized logarithm can be written as a composition of a representative logarithm from Eq. 18 and a scale transformation, $\Lambda = \Lambda_{\nu, \nu_c} \circ g$. For example, pick $\Lambda_{0,0}(x) = \log(x)$, and $g_{0,0}(x) = \exp[-(1-d \log(x)^2)]$, where $d > 0$ is a parameter of $g$. The generalized logarithm then becomes

$$\Lambda(x) = \Lambda_{0,0}(g(x)) = 1 - (1 - d \log(x)^2).$$  \[26\]

The associated generalized exponential is a stretched exponential, $\mathcal{E}_{0,0}(x) = \exp(-\frac{1}{d}(1-x)^d - 1)$, which is the known result for $(c, d)$ entropies with $c = 1$ and $d > 0$ (6, 14).

\subsection*{Example 4. Different choices for $h$}

Suppose that a physical situation demands a specific $\Lambda$ and two observers, $A$ and $B$, choose to represent $\Lambda$ differently. Observer $A$ chooses $h_A(x) = \frac{1}{2} \arctan\left(\frac{x}{2}\right)$ to represent $\Lambda$, so that $\Lambda_A = \Lambda_{A,\nu_c} \circ g_{A}$. Observer $B$ chooses $h_B(x) = \tan(x)$ to represent $\Lambda_B = \Lambda_{B,\nu_c} \circ g_{B}$. Then $\Lambda_A$ and $\Lambda_B$ can only differ by a scale transformation $g_{\pi} \in \mathcal{G}_0$ with
\[ \nu \equiv \frac{\nu - 1}{2}, \text{ and it follows that } g_\nu = \exp \left[ \frac{1}{2} h_A^{-1} \circ h_B(\frac{1}{2} \log(x)) \right]. \]

For the particular functions \( h_A \) and \( h_B \) we have chosen, we get
\[ g_\nu(x) = \exp \left[ \frac{1}{2} \tan \left( \frac{\nu - 1}{2} \right) \right]. \]

Discussion

By studying the two types of entropies that are related to the two possible ways to formulate a maximum entropy principle for systems that explicitly violate the fourth SK axiom, we find that there exists a unique duality that relates the two entropies. Consequently, thermodynamic properties derived from those two entropies will also be related through the duality. We show that the maximum and minimum of \( \Lambda_{HT} \) determine a unique value \( \nu^* \) for which \( \Lambda_{TS,\nu^*} \) is the dual of \( \Lambda_{HT} \). In this way it is possible for an object such as \( \Lambda_{HT} \), which does not explicitly carry an index \( \nu \), to become dual to an object that does, such as \( \Lambda_{TS,\nu^*} \).

Acknowledgments. R.H. and S.T. thank the Santa Fe Institute and M.G.-M. thanks the Aspen Center for Physics for their hospitality. Support for this work was provided in part by National Science Foundation Grant 1066293, Insight Venture Partners, and the Bryan J. and June B. Zwan Foundation (M.G.-M.).

Supporting Information

Hanel et al. 10.1073/pnas.1216885109

SI Materials and Methods

Theorem 1. Let $\Lambda \in \mathcal{L}$. Then $T_\nu \circ \Lambda \in \mathcal{L}$ if and only if $\nu \leq [\Lambda] \leq \nu_+ [\Lambda]$. Moreover, if $\Lambda \in \mathcal{L}_{\nu_-,\nu}$ and $T_\nu \circ \Lambda \in \mathcal{L}$ is a proper generalized logarithm, then $T_\nu \circ \Lambda \in \mathcal{L}_{\nu_-,\nu}$. Proof of Theorem 1. We recall that $T_\nu$ is defined as $T_\nu (\lambda) = (1/\nu + \nu)^{1/\nu}$. If there exists an $x > 0$ such that $\Lambda (x) = -1/\nu$, then $T_\nu \circ \Lambda (x)$ possesses a pole where $T_\nu \circ \Lambda (x)$ changes sign. In that case, $T_\nu \circ \Lambda$ is neither a continuous nor a monotonically increasing function and therefore not a proper generalized logarithm. Conversely, if no $x > 0$ exists such that $\Lambda (x) = -1/\nu$, then $T_\nu \circ \Lambda$ has no pole and is a continuous monotonically increasing function because $T_\nu (\lambda) = (1 + \nu \lambda)^{1/\nu} > 0$ for all $x$. Moreover, $T_\nu \circ \Lambda (1) = 1$ and $(T_\nu \circ \Lambda)' (1) = T_\nu '\circ (0) \Lambda' (1) = 1$. It follows that $T_\nu \circ \Lambda \in \mathcal{L}$. To find sufficient conditions for $T_\nu \circ \Lambda$ to have no pole, we first look at the case $\nu > 0$. In that case, no pole exists if $\lambda = \min \{ \lambda (x) | x \in \mathbb{R}^+ \} \geq 1 - 1/\nu$. In other words, $\nu \leq 1/\nu = \nu_+ [\Lambda]$. Now we turn to the case $\nu < 0$. Then no pole exists if $\lambda = \max \{ \lambda (x) | x \in \mathbb{R}^+ \} \leq 1 - 1/\nu$, which is to say $\nu \geq 1/\nu = \nu_- [\Lambda]$. Both cases together show that $T_\nu \circ \Lambda$ is a continuous function only if $\nu_- [\Lambda] \leq \nu \leq \nu_+ [\Lambda]$. Finally, if $T_\nu \circ \Lambda$ is continuous, then max$(T_\nu \circ \Lambda (x) | x \in \mathbb{R}^+) = T_\nu (\max \{ \lambda (x) | x \in \mathbb{R}^+ \})$, i.e., $T_\nu \circ \Lambda = T_\nu (\Lambda)$. An analogous relation holds for $\nu_+$, which completes the proof.

Theorem 2. Suppose a map $* \circ \sigma$ is given on $\mathcal{L}$ with the following properties: (i) $* \circ \sigma$ is the identity map; (ii) for each $\Lambda \in \mathcal{L}$ there exists a $\nu^* \circ \sigma$ such that $\Lambda^* = T_{\nu^*} \circ \Lambda$; and (iii) $* \circ \sigma$ commutes with scale transformations $g \in \mathcal{G}$, i.e., $(\Lambda \circ g)^* = \Lambda^* \circ g$. Then $* \circ \sigma$ is uniquely determined and $\nu^* \circ \sigma$ is given by

$$\nu^* \circ \sigma = \nu^+ + \nu_- \quad \text{for} \quad \Lambda \in \mathcal{L}_{\nu_-,\nu}. \quad \text{[S1]}$$

Furthermore, it follows from Theorem 1 that

$$\Lambda \in \mathcal{L}_{\nu_-,\nu} \Rightarrow \Lambda^* \in \mathcal{L}_{\nu_-,\nu} \quad \text{[S2]}$$

Proof of Theorem 2. The duality $* \circ \sigma$ on $\mathcal{L}$ can be constructed in the following way. From property (ii) stated in Theorem 2 we know there exists a functional $F : \mathcal{L} \to \mathbb{R}$ such that $\Lambda^* = T_{\nu^*} \circ \Lambda$. From properties (i) and (ii) we also know that $\Lambda = T_{\nu^*} \circ \Lambda^*$. Theorem 1 states that given $\Lambda \in \mathcal{L}_{\nu_-,\nu}$, the condition $\nu_i \geq [\Lambda] \geq \nu_-$ is necessary for $\Lambda^*$ to be a proper logarithm. As a consequence, we get $T_{\nu^*} \circ \Lambda \in \mathcal{L}_{\nu_-,\nu} = T_{\nu^*} \circ \Lambda^*$, which further implies $\nu_i \geq F(\Lambda) \geq \nu_-$ necessary is $\Lambda^* = F(\Lambda)$. Property (iii) implies that for any two logarithms $\Lambda_1$ and $\Lambda_2$ that are members of the same class $\mathcal{L}_{\nu_-,\nu}$, we get $F(\Lambda_1) = F(\Lambda_2)$. Therefore, $F$ can only be of the form

$$F(\Lambda) = f(\nu_+, -\nu_-) \quad \text{for} \quad \Lambda \in \mathcal{L}_{\nu_-,\nu}. \quad \text{[S3]}$$

where $f : \mathbb{R}_+^2 \to \mathbb{R}$. Using $T_\nu \circ T_\mu = T_{\nu+\mu}$ together with property (i) leads to

$$f(\nu_+, -\nu_-) = -f(\nu_+ - f(\nu_+, -\nu_-), - (\nu_+ - f(\nu_+ - \nu_-))). \quad \text{[S4]}$$

In other words, $f$ solves the two equations

(a) $f(x, y) = -f(x - f(x, y), y + f(x, y))$

(b) $x \geq f(x, y) \geq y$

for all $x, y \in \mathbb{R}_+$. Eq. [S5(b)] immediately tells us that $f(0, 0) = 0$. Consider a function $y(x, z)$ solving the implicit equation

$$f(x, y(x, z)) = z, \quad \text{[S6]}$$

and rewrite Eq. [S5(a)] as

$$x - f(x, y(x, z)) = -z. \quad \text{[S7]}$$

From Eqs. S6 and S7 one gets $y(x, z) = y(x, z) + z$. Also, $f(0, 0) = 0$ implies $y(0, 0) = 0$. By expanding $y(x, z) = \sum_{m,n=0}^{\infty} y_{m,n} x^m z^n$, one gets $y_{0,0} = 0$, and for the first order,

$$2y_{1,1} + y_{1,0} + 1 = 0. \quad \text{[S8]}$$

All coefficients $y_{m,n}$ of higher order follow equation

$$y_{m,n} = (-1)^n \sum_{k=0}^{n} \binom{m+k}{m} y_{m+k,n-k}. \quad \text{[S9]}$$

We also expand $f(x, y(x, z)) = \sum_{m,n=1}^{\infty} y_{m,n} x^m z^n$.

In other words, $f$ can only contribute to coefficients $y_{i,j}$ with indices $i \geq n$ or $j \geq n$. Comparing coefficients by order by order, one shows that only coefficients of the first order,

$$f(1,0)x + f(0,1)y(x, z) = z, \quad \text{[S10]}$$

contribute to solving Eq. S6. Thus, $y(x, z)$ can only be of the form $y(x, z) = y_0 x^0 + y_1 x^1$. This, together with Eqs. S8 and S7(b), implies $x \geq 2(y_1 x - y)/(1 + y_1, 1) \leq y$. Choosing $x = 0$ gives $1 \geq y_1, 1 \geq -1$. Setting $y = 0$ also implies $y_{1,0} \geq 1$, so that the only possible solution for $y_{1,0} = 1$. As a consequence of Eq. S8, $y_{1,0} = -1$. Therefore, we have $y(x, z) = x - z$ and $f$ has the unique solution $f(x, y) = x - y$. From this it follows that $\nu^* = f(\nu_+, -\nu_-) = \nu^+ + \nu_-$ for $\Lambda \in \mathcal{L}_{\nu_-,\nu}$, which means that $\nu^* \circ \sigma$ is uniquely defined. Because $\nu_- - \nu^* = -\nu_-$ and $\nu_+ - \nu^* = -\nu_+$, Theorem 1 implies that $\Lambda \in \mathcal{L}_{\nu_-,\nu} \Rightarrow \Lambda^* \in \mathcal{L}_{\nu_-,\nu}$, and this completes the proof.