Fadeev's theorem on entropy

Notes by: Arthur Parzygnat

March 29, 2016

Abstract

In this expository note, we state and prove Fadeev's theorem on a set of axioms characterizing the entropy function on finite probability spaces based on Rényi's simplification [2]. We use the outline of the proof mentioned in [1]. There is absolutely nothing new here and is merely meant to make sure I understand the proof. The proof is elementary in the sense that it only use combinatorics, number theory, and a small amount of analysis.

Lemma 1. *Let $\phi : \mathbb{Z}^+ \to \mathbb{R}$ be a function satisfying*

$$\phi(nm) = \phi(n) + \phi(m) \quad \forall\ n, m \in \mathbb{Z}^+ \tag{2}$$

and

$$\lim_{n\to\infty} \Big(\phi(n+1) - \phi(n)\Big) = 0. \tag{3}$$

Then there exists a constant $c \in \mathbb{R}$ such that

$$\phi(n) = c \log_2 n \quad \forall\ n \in \mathbb{Z}^+. \tag{4}$$

The following proof follows [2] very closely.

Proof. For every $N \in \{2, 3, 4, \dots\}$, define $g_N : \mathbb{Z}^+ \to \mathbb{R}$ by

$$\mathbb{Z}^+ \ni n \mapsto g_N(n) := \phi(n) - \frac{\phi(N)\log_2(n)}{\log_2(N)}. \tag{5}$$

The idea of the proof will be to show that

$$\lim_{n\to\infty} \frac{g_N(n)}{\log_2(n)} = 0 \tag{6}$$

which shows that $\frac{\phi(N)}{\log_2(N)}$ is a constant, c, independent of N. Note that

$$g_N(N) = \phi(N) - \frac{\phi(N)\log_2(N)}{\log_2(N)} = 0, \tag{7}$$

$$\begin{aligned} g_N(nm) &= \phi(nm) - \frac{\phi(N)\log_2(nm)}{\log_2(N)} \\ &= \phi(n) + \phi(m) - \frac{\phi(N)\big(\log_2(n) + \log_2(m)\big)}{\log_2(N)} \\ &= g_N(n) + g_N(m), \end{aligned} \tag{8}$$

and

$$\lim_{n\to\infty}\Big(g_N(n+1)-g_N(n)\Big)=\lim_{n\to\infty}\Big(\phi(n+1)-\phi(n)\Big)+\frac{\phi(N)}{\log_2(N)}\lim_{n\to\infty}\left(\log_2\left(\frac{n}{n-1}\right)\right)=0. \tag{9}$$

Setting $G_N : \{-1,0,1,2,\dots\}\to\mathbb{R}_{\geq 0}$ to be the function defined by

$$G_N(k) := \begin{cases} 0 & \text{if } k=-1 \\ \max\limits_{n\in[N^k,N^{k+1})\cap\mathbb{Z}} |g_N(n)| & \text{if } k\in\{0,1,2,\dots\} \end{cases} \tag{10}$$

and

$$\delta_N(k) := \max_{n\in[N^k,N^{k+1})\cap\mathbb{Z}} \big|g_N(n+1)-g(n)\big|, \tag{11}$$

(9) implies

$$\lim_{k\to\infty}\delta_N(k)=0. \tag{12}$$

For all $n\in[N^k,N^{k+1})\cap\mathbb{Z}$,

$$\begin{aligned} \frac{|g_N(n)|}{\log_2(n)} &\leq \frac{G_N(k)}{\log_2(n)} && \text{by (10)} \\ &\leq \frac{G_N(k)}{k\log_2(N)} && \text{since } \log_2(N^k)\leq\log_2(n). \end{aligned} \tag{13}$$

Hence,

$$\lim_{n\to\infty}\frac{|g_N(n)|}{\log_2(n)}\leq\lim_{k\to\infty}\frac{G_N(k)}{k\log_2(N)}=\frac{1}{\log_2(N)}\lim_{k\to\infty}\frac{G_N(k)}{k} \tag{14}$$

and thus to prove (6), it suffices to prove

$$\lim_{k\to\infty}\frac{G_N(k)}{k}=0. \tag{15}$$

Now fix $n\in\mathbb{N}$ and let $k\in\mathbb{Z}$ be such that $N^k\leq n<N^{k+1}$. Then since

$$0\leq n-N\left\lfloor\frac{n}{N}\right\rfloor<N, \tag{16}$$

where $\lfloor\cdot\rfloor:\mathbb{R}_{\geq 0}\to\mathbb{N}$ denotes the floor function (the largest integer less than its input),

$$g_N(n)=g_N(n)+\sum_{l=N\lfloor\frac{n}{N}\rfloor}^{n-1}\Big(\overbrace{g_N(l)-g_N(l)}^{0}\Big)=g_N\left(N\left\lfloor\frac{n}{N}\right\rfloor\right)+\sum_{l=N\lfloor\frac{n}{N}\rfloor}^{n-1}\Big(g_N(l+1)-g_N(l)\Big) \tag{17}$$

and so taking the absolute value gives

$$\begin{aligned} |g_N(n)| &\leq \left|g_N\left(N\left\lfloor\frac{n}{N}\right\rfloor\right)\right|+\sum_{l=N\lfloor\frac{n}{N}\rfloor}^{n-1}\Big|g_N(l+1)-g_N(l)\Big| \\ &\leq \left|g_N\left(N\left\lfloor\frac{n}{N}\right\rfloor\right)\right|+\left(n-1-N\left\lfloor\frac{n}{N}\right\rfloor\right)\delta_N(k) && \text{by def'n of } \delta_N(k) \text{ in (11)} \\ &\leq \left|g_N\left(N\left\lfloor\frac{n}{N}\right\rfloor\right)\right|+N\delta_N(k) && \text{by (16).} \end{aligned} \tag{18}$$

Using this and the fact that

$$\begin{aligned} g_N\left(N\left\lfloor \frac{n}{N}\right\rfloor\right) &= g_N(N) + g_N\left(\left\lfloor \frac{n}{N}\right\rfloor\right) \quad \text{by (8)} \\ &= g_N\left(\left\lfloor \frac{n}{N}\right\rfloor\right) \qquad\qquad \text{by (7)} \end{aligned} \tag{19}$$

results in

$$|g_N(n)| - \left|g_N\left(\left\lfloor \frac{n}{N}\right\rfloor\right)\right| \leq N\delta_N(k). \tag{20}$$

As n varies over $[N^k, N^{k+1}) \cap \mathbb{Z}$, $\left\lfloor \frac{n}{N}\right\rfloor$ varies over $[N^{k-1}, N^k)$. Hence,

$$\begin{aligned} G_N(k) - G_N(k-1) &= \max_{n\in[N^k,N^{k+1})\cap\mathbb{Z}} |g_N(n)| - \max_{n\in[N^k,N^{k+1})\cap\mathbb{Z}} \left|g_N\left(\left\lfloor \frac{n}{N}\right\rfloor\right)\right| \\ &\leq \max_{n\in[N^k,N^{k+1})\cap\mathbb{Z}} \left(|g_N(n)| - \left|g_N\left(\left\lfloor \frac{n}{N}\right\rfloor\right)\right|\right) \\ &\leq \max_{n\in[N^k,N^{k+1})\cap\mathbb{Z}} N\delta_N(k) \qquad \text{by (20)} \\ &= N\delta_N(k). \end{aligned} \tag{21}$$

Summing this last inequality over all $k \in \{0, 1, \ldots, m\}$ gives (for large m)

$$G_N(m) = G_N(m) - \overbrace{G_N(-1)}^{0} = \sum_{k=0}^{m}\Big(G_N(k) - G_N(k-1)\Big) \leq N\sum_{k=0}^{m}\delta_N(k). \tag{22}$$

Dividing by m gives

$$\frac{G_N(m)}{m} \leq N\frac{\sum_{k=0}^{m}\delta_N(k)}{m}. \tag{23}$$

As $m \to \infty$, since $\lim_{k\to\infty}\delta_N(k) = 0$, for any $\epsilon > 0$, there exists some m' so that $\delta_N(m) \leq \frac{\epsilon}{N}$ for all $m > m'$. Hence,

$$\begin{aligned} \lim_{m\to\infty}\frac{G_N(m)}{m} &\leq N\lim_{m\to\infty}\frac{\sum_{k=0}^{m}\delta_N(k)}{m} \\ &\leq N\lim_{m\to\infty}\left(\frac{\sum_{k=0}^{m'}\delta_N(k)}{m} + \frac{\sum_{k=m'+1}^{m}\delta_N(k)}{m}\right) \\ &\leq N\sum_{k=0}^{m'}\delta_N(k)\lim_{m\to\infty}\left(\frac{1}{m}\right) + N\lim_{m\to\infty}\left(\left(\frac{m-m'-1}{m}\right)\frac{\epsilon}{N}\right) \\ &\leq \epsilon. \end{aligned} \tag{24}$$

Since ϵ can be chosen arbitrarily small, this proves (15). By the outline at the beginning of this proof, this shows that

$$\phi(N) = c\log_2(N) \quad \text{for all } N \in \{2, 3, \ldots\}. \tag{25}$$

By the assumption (2), $\phi(1) = \phi(1) + \phi(1)$ which proves $\phi(1) = 0$. Hence,

$$\phi(n) = c\log_2(n) \quad \text{for all } n \in \mathbb{Z}^+. \tag{26}$$

□

We take the following definition for classical entropy.

Definition 27. Let Δ^{n-1} denote the $(n-1)$-simplex, i.e.

$$\Delta^{n-1} := \left\{ (p_1, \ldots, p_n) \in \mathbb{R}^n \;\middle|\; \sum_{i=1}^{n} p_i = 1 \text{ and } p_i \geq 0 \; \forall \; i = 1, \ldots, n \right\}, \tag{28}$$

and

$$\Delta := \coprod_{n=1} \Delta^{n-1}, \tag{29}$$

the disjoint union of all n-simplices. An element of Δ^{n-1} is called a *probability distribution on n events*. For each $n \in \mathbb{Z}^+$, the function $H_n : \Delta^{n-1} \to \mathbb{R}$ defined by

$$H_n(p_1, \ldots, p_n) := -\sum_{i=1}^{n} p_i \log_2 p_i. \tag{30}$$

is called the *entropy of the probability distribution* $(p_1, \ldots, p_n) \in \Delta^{n-1}$. The function

$$H := \coprod_{n=1} H_n, \tag{31}$$

defined by

$$H(p_1, \ldots, p_n) := H_n(p_1, \ldots, p_n), \tag{32}$$

is called the *Shannon entropy* function.

Remark 33. The symmetric group on n letters acts on each Δ^{n-1} as follows. Let $\sigma \in S_n$ be a permutation on n distinct elements. Then

$$\sigma \cdot (p_1, \ldots, p_n) := (p_{\sigma(1)}, \ldots, p_{\sigma(n)}). \tag{34}$$

Lemma 35. *The entropy function satisfies the following properties.*

(a) For each n, the diagram

$$\begin{array}{ccc} \Delta^{n-1} & \xrightarrow{\sigma} & \Delta^{n-1} \\ & {\scriptstyle H}\searrow \quad \swarrow{\scriptstyle H} & \\ & \mathbb{R} & \end{array} \tag{36}$$

commutes for all $\sigma \in S_n$.

(b) The function $H_1 \equiv H|_{\Delta^1} : \Delta^1 \to \mathbb{R}$ is continuous.

(c) $H(\frac{1}{2}, \frac{1}{2}) = 1$.

(d) For each $n \in \mathbb{Z}^+$,

$$H\big(tp_1, (1-t)p_1, p_2, \ldots, p_n\big) = H(p_1, \ldots, p_n) + p_1 H(t, 1-t) \tag{37}$$

for all $t \in [0, 1]$.

Proof. (a) follows from commutativity of addition. (b) follows from the fact that $x \ln x$ is continuous on $[0,1]$ with the assumption that $0 \ln 0 = 0$. (c) follows from

$$2\left(-\frac{1}{2}\log_2\left(\frac{1}{2}\right)\right) = \log_2 2 = 1. \tag{38}$$

(d) follows from

$$\begin{aligned} H\big(tp_1, (1-t)p_1, p_2, \dots, p_n\big) &= -tp_1\log_2(tp_1) - (1-t)p_1\log_2\big((1-t)p_1\big) - \sum_{k=2}^{n} p_2\log_2 p_2 \\ &= -tp_1\log_2(t) - tp_1\log_2(p_1) - (1-t)p_1\log_2(1-t) \\ &\quad - (1-t)p_1\log_2(p_1) - \sum_{k=2}^{n} p_2\log_2 p_2 \\ &= p_1 H(t, 1-t) - p_1\log_2(p_1) - \sum_{k=2}^{n} p_2\log_2 p_2 \\ &= H(p_1,\dots,p_n) + p_1 H(t,1-t). \end{aligned} \tag{39}$$

□

Theorem 40. *Let $J : \Delta \to \mathbb{R}$ be any function satisfying conditions (a)—(d) from Lemma 35 (with H replaced by J). Then $J = H$.*

Proof. We will prove this theorem in a series of steps which consist of (1) using J to construct a function $\phi : \mathbb{Z}^+ \to \mathbb{R}$ and showing it satisfies the conditions of Lemma 1, (2) extending the results to rational numbers, and (3) extending them to real numbers.

Step 1. We follow the suggestion of [1] (which is probably due to Fadeev) and define $\phi : \mathbb{Z}^+ \to \mathbb{R}$ by

$$\mathbb{Z}^+ \ni n \mapsto \phi(n) := \begin{cases} 0 & \text{if } n = 1 \\ J\big|_{\Delta^{n-1}}\left(\frac{1}{n},\dots,\frac{1}{n}\right) & \text{otherwise.} \end{cases} \tag{41}$$

We will prove the formula

$$\phi(nm) = \phi(n) + \phi(m) \tag{42}$$

by induction on n. The base case is for $n = 2$ and results in

$$\begin{aligned} \phi(2m) &= J\big|_{\Delta^{2m-1}}\left(\frac{1}{2m},\dots,\frac{1}{2m}\right) \\ &= J\big|_{\Delta^{2m-1}}\left(\left[\frac{1}{2}\right]\left[\frac{1}{m}\right], \left[\frac{1}{2}\right]\left[\frac{1}{m}\right], \frac{1}{2m},\dots,\frac{1}{2m}\right) \\ &\overset{(37)}{=} J\big|_{\Delta^{2m-2}}\left(\frac{1}{m}, \frac{1}{2m},\dots,\frac{1}{2m}\right) + \frac{1}{m} J\big|_{\Delta^1}\left(\frac{1}{2},\frac{1}{2}\right) \\ &\overset{(36)}{=} J\big|_{\Delta^{2m-2}}\left(\frac{1}{2m},\dots,\frac{1}{2m},\frac{1}{m}\right) + \frac{1}{m} J\big|_{\Delta^1}\left(\frac{1}{2},\frac{1}{2}\right) \\ &\ \ \vdots \quad \text{repeat } m-1 \text{ more times} \\ &= J\big|_{\Delta^{m-1}}\left(\frac{1}{m},\dots,\frac{1}{m}\right) + \frac{m}{m} J\big|_{\Delta^1}\left(\frac{1}{2},\frac{1}{2}\right) \\ &= \phi(m) + \phi(2). \end{aligned} \tag{43}$$

Assume that $\phi(km) = \phi(k) + \phi(m)$ for $k = 2, \dots, n-1$. Let

$$n = q_1^{\alpha_1} \cdots q_l^{\alpha_l} \tag{44}$$

be a prime factor decomposition of n with

$$q_1 < \cdots < q_l \tag{45}$$

and all $\alpha_i \in \mathbb{Z}^+$. Then

$$\begin{aligned}
\phi(nm) &= J\Big|_{\Delta^{nm-1}} \left(\frac{1}{nm}, \dots, \frac{1}{nm}\right) \\
&= J\Big|_{\Delta^{nm-1}} \left(\left[\frac{1}{2}\right]\left[\frac{2}{nm}\right], \left[\frac{1}{2}\right]\left[\frac{2}{nm}\right], \frac{1}{nm}, \dots, \frac{1}{nm}\right) \\
&\overset{(37)}{=} J\Big|_{\Delta^{nm-2}} \left(\frac{2}{nm}, \frac{1}{nm}, \dots, \frac{1}{nm}\right) + \frac{2}{nm} J\Big|_{\Delta^1}\left(\frac{1}{2}, \frac{1}{2}\right) \\
&= J\Big|_{\Delta^{nm-2}} \left(\left[\frac{2}{3}\right]\left[\frac{3}{nm}\right], \left[\frac{1}{3}\right]\left[\frac{3}{nm}\right], \frac{1}{nm}, \dots, \frac{1}{nm}\right) + \frac{2}{nm} J\Big|_{\Delta^1}\left(\frac{1}{2}, \frac{1}{2}\right) \\
&\overset{(37)}{=} J\Big|_{\Delta^{nm-3}} \left(\frac{3}{nm}, \frac{1}{nm}, \dots, \frac{1}{nm}\right) + \frac{3}{nm} J\Big|_{\Delta^1}\left(\frac{2}{3}, \frac{1}{3}\right) + \frac{2}{nm} J\Big|_{\Delta^1}\left(\frac{1}{2}, \frac{1}{2}\right) \\
&\ \ \vdots \quad \text{repeat } q_1 - 3 \text{ more times} \\
&= J\Big|_{\Delta^{nm-q_1}} \Bigg(\frac{q_1}{nm}, \underbrace{\frac{1}{nm}, \dots, \frac{1}{nm}}_{nm-q_1 \text{ terms}}\Bigg) + \underbrace{\sum_{k=2}^{q_1} \frac{k}{nm} J\Big|_{\Delta^1}\left(\frac{k-1}{k}, \frac{1}{k}\right)}_{\chi}
\end{aligned} \tag{46}$$

Using (36), and then repeating this on the first term gives

$$\begin{aligned}
J\Big|_{\Delta^{nm-q_1}} \left(\frac{q_1}{nm}, \frac{1}{nm}, \dots, \frac{1}{nm}\right) &= J\Big|_{\Delta^{nm-(q_1-1)-1}} \left(\frac{1}{nm}, \dots, \frac{1}{nm}, \frac{q_1}{nm}\right) \\
&= J\Big|_{\Delta^{nm-2(q_1-1)-1}} \Bigg(\underbrace{\frac{1}{nm}, \dots, \frac{1}{nm}}_{nm-2(q_1-1)-2 \text{ terms}}, \frac{q_1}{nm}, \frac{q_1}{nm}\Bigg) + \chi \\
&\ \ \vdots \quad \text{repeat } l - 2 \text{ more times} \\
&= J\Big|_{\Delta^{nm-l(q_1-1)-1}} \Bigg(\underbrace{\frac{1}{nm}, \dots, \frac{1}{nm}}_{nm-l(q_1-1)-l \text{ terms}}, \underbrace{\frac{q_1}{nm}, \dots, \frac{q_1}{nm}}_{l \text{ terms}}\Bigg) + (l-1)\chi.
\end{aligned} \tag{47}$$

This procedure terminates at $l = \frac{nm}{q_1}$ since $nm - \frac{nm}{q_1}(q_1 - 1) - \frac{nm}{q_1} = 0$ so that there are no more $\frac{1}{nm}$ terms and all the terms are of the form $\frac{q_1}{nm}$. Then $nm - \frac{nm}{q_1}(q_1 - 1) - 1 = \frac{nm}{q_1} - 1$ and we get

$$\begin{aligned}
\phi(nm) &= \overbrace{J\Big|_{\Delta^{\frac{nm}{q_1}-1}} \Bigg(\underbrace{\frac{q_1}{nm}, \dots, \frac{q_1}{nm}}_{\frac{nm}{q_1} \text{ terms}}\Bigg)}^{\phi\left(\frac{n}{q_1}m\right)} + \frac{nm}{q_1} \sum_{k=2}^{q_1} \frac{k}{nm} J\Big|_{\Delta^1}\left(\frac{k-1}{k}, \frac{1}{k}\right) \\
&= \phi\left(\frac{n}{q_1}\right) + \phi(m) + \sum_{k=2}^{q_1} \frac{k}{q_1} J\Big|_{\Delta^1}\left(\frac{k-1}{k}, \frac{1}{k}\right)
\end{aligned} \tag{48}$$

by the induction hypothesis. Now, the leftover term becomes

$$\begin{aligned}\sum_{k=2}^{q_1}\frac{k}{q_1}J\Big|_{\Delta^1}\left(\frac{k-1}{k},\frac{1}{k}\right) &\overset{(36)}{=} \sum_{k=2}^{q_1}\frac{k}{q_1}J\Big|_{\Delta^1}\left(\frac{1}{k},\frac{k-1}{k}\right)\\ &\overset{(37)}{=} \sum_{k=2}^{q_1-1}\left[J\Big|_{\Delta^{q_1-k+1}}\left(\frac{1}{q_1},\frac{k-1}{q_1},\underbrace{\frac{1}{q_1},\dots,\frac{1}{q_1}}_{q_1-k\text{ terms}}\right)-J\Big|_{\Delta^{q_1-k}}\left(\frac{k}{q_1},\underbrace{\frac{1}{q_1},\dots,\frac{1}{q_1}}_{q_1-k\text{ terms}}\right)\right]\\ &\quad + J\Big|_{\Delta^1}\left(\frac{1}{q_1},\frac{q_1-1}{q_1}\right)\\ &\overset{(36)}{=} J\Big|_{\Delta^{q_1-1}}\left(\frac{1}{q_1},\dots,\frac{1}{q_1}\right)\\ &= \phi(q_1)\end{aligned} \tag{49}$$

since all the other terms cancel. In the second equality, we set $p_1=\frac{k}{q_1}$ and $t=\frac{1}{k}$ and used (37). Therefore,

$$\phi(nm)=\phi\left(\frac{n}{q_1}\right)+\phi(m)+\phi(q_1)=\phi\left(\frac{n}{q_1}q_1\right)+\phi(m)=\phi(n)+\phi(m) \tag{50}$$

again by the induction hypothesis. Next we prove

$$\lim_{n\to\infty}\Big(\phi(n+1)-\phi(n)\Big)=0. \tag{51}$$

As a preliminary, we first prove that $J|_{\Delta^1}(0,1)=0$. This follows from the following calculation

$$\begin{aligned}J|_{\Delta^1}(0,1) &\overset{(37)}{=} J|_{\Delta^2}\big([t][0],[1-t][0],1\big)+0J|_{\Delta^1}(t,1-t)\\ &\overset{(36)}{=} J|_{\Delta^2}(1,0,0)\\ &= J|_{\Delta^2}\big([1][1],[1-1][1],0\big)\\ &\overset{(37)}{=} J|_{\Delta^1}(1,0)+1J|_{\Delta^1}(1,0)\end{aligned} \tag{52}$$

which by (36) gives

$$J|_{\Delta^1}(0,1)=2J|_{\Delta^1}(0,1) \qquad \Rightarrow \qquad J|_{\Delta^1}(0,1)=0 \tag{53}$$

as needed. Now, to prove (51), we first rewrite the terms as

$$\begin{aligned}\phi(n+1)-\phi(n) &\overset{(49)}{=} \sum_{k=2}^{n+1}\frac{k}{n+1}J\Big|_{\Delta^1}\left(\frac{k-1}{k},\frac{1}{k}\right)-\sum_{k=2}^{n}\frac{k}{n}J\Big|_{\Delta^1}\left(\frac{k-1}{k},\frac{1}{k}\right)\\ &= J\Big|_{\Delta^1}\left(\frac{n}{n+1},\frac{1}{n+1}\right)-\frac{1}{n(n+1)}\sum_{k=2}^{n}kJ\Big|_{\Delta^1}\left(\frac{k-1}{k},\frac{1}{k}\right)\end{aligned} \tag{54}$$

using our earlier calculation. By condition (b) of Lemma 35, i.e. continuity of $J|_{\Delta^1}$, and the fact that $J|_{\Delta^1}(1,0)=J|_{\Delta^1}(0,1)=0$, for any $\epsilon>0$, there exists an N_1 such that

$$\left|J\Big|_{\Delta^1}\left(\frac{n}{n+1},\frac{1}{n+1}\right)\right|<\frac{\epsilon}{3} \qquad \forall\; n>N_1. \tag{55}$$

Furthermore, there exists an $N > N_1$ such that

$$\frac{1}{n+1}\left|\sum_{k=2}^{N_1} J\Big|_{\Delta^1}\left(\frac{k-1}{k}, \frac{1}{k}\right)\right| < \frac{\epsilon}{3} \qquad \forall\ n > N \tag{56}$$

since the term in the absolute value is finite. Using these two bounds, we obtain

$$\begin{aligned}
\Big|\phi(n+1) - \phi(n)\Big| &= \left|J\Big|_{\Delta^1}\left(\frac{n}{n+1}, \frac{1}{n+1}\right) - \frac{1}{n(n+1)}\sum_{k=2}^{n} kJ\Big|_{\Delta^1}\left(\frac{k-1}{k}, \frac{1}{k}\right)\right| \\
&\le \left|J\Big|_{\Delta^1}\left(\frac{n}{n+1}, \frac{1}{n+1}\right)\right| + \frac{1}{n(n+1)}\left|\sum_{k=2}^{N_1} kJ\Big|_{\Delta^1}\left(\frac{k-1}{k}, \frac{1}{k}\right)\right| \\
&\quad + \frac{1}{n(n+1)}\sum_{k=N_1+1}^{n} k\left|J\Big|_{\Delta^1}\left(\frac{k-1}{k}, \frac{1}{k}\right)\right| \\
&\le \frac{\epsilon}{3} + \left(\frac{N_1}{n}\right)\left(\frac{1}{n+1}\right)\left|\sum_{k=2}^{N_1} J\Big|_{\Delta^1}\left(\frac{k-1}{k}, \frac{1}{k}\right)\right| + \left(\frac{1}{n(n+1)}\right)\left(\frac{\epsilon}{3}\right)\left|\sum_{k=N_1+1}^{n} k\right| \qquad (57) \\
&\le \frac{\epsilon}{3} + \left(\frac{N_1}{n}\right)\left(\frac{\epsilon}{3}\right) + \left(\frac{1}{n(n+1)}\right)\left(\frac{\epsilon}{3}\right)\left|\frac{n(n+1)}{2} - \frac{N_1(N_1+1)}{2}\right| \\
&\le \frac{\epsilon}{3} + \left(\frac{N_1}{n}\right)\left(\frac{\epsilon}{3}\right) + \frac{\epsilon}{6} + \left(\frac{N_1(N_1+1)}{n(n+1)}\right)\left(\frac{\epsilon}{6}\right) \\
&\le \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{6} + \frac{\epsilon}{6} \\
&= \epsilon \qquad \forall\ n > N,
\end{aligned}$$

which proves (51). Thus by Lemma 1, there exists a constant $c \in \mathbb{R}$ such that

$$\phi(n) \equiv J\Big|_{\Delta^{n-1}}\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = c\log_2(n) \quad \text{for } n \in \mathbb{Z}^+. \tag{58}$$

By condition (c) of Lemma 35, $c = 1$.

Step 2. We first explicitly calculate $J\big|_{\Delta^1}\left(\frac{n-1}{n}, \frac{1}{n}\right)$ for n a positive integer greater than 2. Then, we calculate $J\big|_{\Delta^1}\left(\frac{m}{n}, \frac{m-n}{n}\right)$ for $m, n \in \mathbb{Z}^+$ with $n > m > 2$. Finally, we calculate $J\big|_{\Delta^{n-1}}\left(\frac{w_1}{z_1}, \dots, \frac{w_n}{z_n}\right)$ for $z_i > w_i > 2$ all integers for all $i = 1, \dots, n$. The first claim is

$$J\Big|_{\Delta^1}\left(\frac{n-1}{n}, \frac{1}{n}\right) = \phi(n) - \frac{n-1}{n}\phi(n-1). \tag{59}$$

The proof is done by induction. The base case of $n = 2$ is obvious. Assume the formula holds for all integers up to n. Then

$$
\begin{aligned}
J\Big|_{\Delta^1}\left(\frac{n}{n+1},\frac{1}{n+1}\right) &\overset{(37)}{=} J\Big|_{\Delta^2}\left(\left[\frac{1}{n}\right]\left[\frac{n}{n+1}\right],\left[1-\frac{1}{n}\right]\left[\frac{n}{n+1}\right],\frac{1}{n+1}\right) \\
&\qquad - \frac{n}{n+1}J\Big|_{\Delta^1}\left(\frac{1}{n},1-\frac{1}{n}\right) \\
&\overset{(36)}{=} J\Big|_{\Delta^2}\left(\frac{n-1}{n+1},\frac{1}{n+1},\frac{1}{n+1}\right) - \frac{n}{n+1}J\Big|_{\Delta^1}\left(\frac{n-1}{n},\frac{1}{n}\right) \\
&\vdots \quad \text{repeat } n-2 \text{ more times} \\
&= \phi(n+1) - \sum_{k=1}^{n-1}\frac{n+1-k}{n+1}J\Big|_{\Delta^1}\left(\frac{n-k}{n+1-k},\frac{1}{n+1-k}\right) \qquad (60) \\
&= \phi(n+1) - \sum_{k=1}^{n-1}\frac{n+1-k}{n+1}\left(\phi(n+1-k) - \frac{n-k}{n+1-k}\phi(n-k)\right) \\
&= \phi(n+1) + \frac{1}{n+1}\sum_{k=1}^{n-1}\Big((n-k)\phi(n-k) - (n+1-k)\phi(n+1-k)\Big) \\
&= \phi(n+1) - \frac{n}{n+1}\phi(n),
\end{aligned}
$$

where we have used the induction hypothesis in the fourth equality. Note that in the last equality all terms but the $k=1$ term on the right cancel. We next prove the formula

$$
J\Big|_{\Delta^1}\left(\frac{m}{n},\frac{n-m}{n}\right) = \phi(n) - \frac{m}{n}\phi(m) - \frac{n-m}{n}\phi(n-m) \qquad (61)
$$

for integers $n > m > 2$. First, we have

$$
\begin{aligned}
J\Big|_{\Delta^1}\left(\frac{m}{n},\frac{n-m}{n}\right) &\overset{(37)}{=} J\Big|_{\Delta^2}\left(\left[\frac{1}{m}\right]\left[\frac{m}{n}\right],\left[\frac{m-1}{m}\right]\left[\frac{m}{n}\right],\frac{n-m}{n}\right) - \frac{m}{n}J\Big|_{\Delta^1}\left(\frac{1}{m},\frac{m-1}{m}\right) \\
&\overset{(36)}{=} J\Big|_{\Delta^2}\left(\left[\frac{m-1}{n-1}\right]\left[\frac{n-1}{n}\right],\left[\frac{n-m}{n-1}\right]\left[\frac{n-1}{n}\right],\frac{1}{n}\right) - \frac{m}{n}J\Big|_{\Delta^1}\left(\frac{m-1}{m},\frac{1}{m}\right) \qquad (62) \\
&\overset{(37)}{=} J\Big|_{\Delta^1}\left(\frac{n-1}{n},\frac{1}{n}\right) - \frac{m}{n}J\Big|_{\Delta^1}\left(\frac{m-1}{m},\frac{1}{m}\right) + \frac{n-1}{n}J\Big|_{\Delta^1}\left(\frac{m-1}{n-1},\frac{n-m}{n-1}\right)
\end{aligned}
$$

This gives a recursion relation so that we can repeat until the term on the right-hand-side terminates, which happens after m iterations. Doing this, we get

$$\begin{aligned}
J\Big|_{\Delta^1}\left(\frac{m}{n},\frac{n-m}{n}\right) &= J\Big|_{\Delta^1}\left(\frac{n-1}{n},\frac{1}{n}\right) - \frac{m}{n}J\Big|_{\Delta^1}\left(\frac{m-1}{m},\frac{1}{m}\right) \\
&\quad + \frac{n-1}{n}\left(J\Big|_{\Delta^1}\left(\frac{n-2}{n-1},\frac{1}{n-1}\right) - \frac{m-1}{n-1}J\Big|_{\Delta^1}\left(\frac{m-2}{m-1},\frac{1}{m-1}\right)\right) \\
&\quad + \frac{n-1}{n}\left(\frac{n-2}{n-1}J\Big|_{\Delta^1}\left(\frac{m-2}{n-2},\frac{n-m}{n-2}\right)\right) \\
&= \frac{n-0}{n}J\Big|_{\Delta^1}\left(\frac{n-1}{n},\frac{1}{n}\right) + \frac{n-1}{n}J\Big|_{\Delta^1}\left(\frac{n-2}{n-1},\frac{1}{n-1}\right) \\
&\quad - \frac{m}{n}J\Big|_{\Delta^1}\left(\frac{m-1}{m},\frac{1}{m}\right) - \frac{m-1}{n}J\Big|_{\Delta^1}\left(\frac{m-2}{m-1},\frac{1}{m-1}\right) \\
&\quad + \frac{n-2}{n}J\Big|_{\Delta^1}\left(\frac{m-2}{n-2},\frac{n-m}{n-2}\right) \\
&\ \ \vdots \quad \text{repeat } m-3 \text{ more times} \\
&= \sum_{k=0}^{m-1}\left(\frac{n-k}{n}J\Big|_{\Delta^1}\left(\frac{n-1-k}{n-k},\frac{1}{n-k}\right) - \frac{m-k}{n}J\Big|_{\Delta^1}\left(\frac{m-1-k}{m-k},\frac{1}{m-k}\right)\right) \\
&\overset{(59)}{=} \frac{1}{n}\sum_{k=0}^{m-1}\Big((n-k)\phi(n-k) - (n-1-k)\phi(n-1-k)\Big) \\
&\quad + \frac{1}{n}\sum_{k=0}^{m-1}\Big((m-1-k)\phi(m-1-k) - (m-k)\phi(m-k)\Big)\Big) \\
&= \frac{1}{n}\Big(n\phi(n) - (n-m)\phi(n-m) - m\phi(m) + 0\Big) \\
&= \phi(n) - \frac{m}{n}\phi(m) - \frac{n-m}{n}\phi(n-m),
\end{aligned} \tag{63}$$

which is the formula we tried to prove. Note that by Step 1, this says

$$\begin{aligned}
J\Big|_{\Delta^1}\left(\frac{m}{n},\frac{n-m}{n}\right) &= \log_2 n - \frac{m}{n}\log_2 m - \frac{n-m}{n}\log_2(n-m) \\
&= -\frac{m}{n}\log_2\left(\frac{m}{n}\right) - \frac{n-m}{n}\log_2\left(\frac{n-m}{n}\right)
\end{aligned} \tag{64}$$

so we see that it is convenient to write this as

$$J\Big|_{\Delta^1}\left(\frac{m}{n},\frac{n-m}{n}\right) = \frac{m}{n}\Big(\phi(n) - \phi(m)\Big) + \frac{n-m}{n}\Big(\phi(n) - \phi(n-m)\Big). \tag{65}$$

Finally, for the rational points of an arbitrary $(n-1)$-simplex, we will prove the formula

$$J\Big|_{\Delta^{n-1}}\left(\frac{w_1}{z_1},\ldots,\frac{w_n}{z_n}\right) = \sum_{k=1}^{n}\frac{w_k}{z_k}\Big(\phi(z_k) - \phi(w_k)\Big) \tag{66}$$

by induction. This formula holds for the base case $n = 2$ by our previous result. We assume it is true for n (for all rational probability measures) and show it holds for $n+1$. We get

$$
\begin{aligned}
&J\Big|_{\Delta^n}\left(\frac{w_1}{z_1},\dots,\frac{w_{n+1}}{z_{n+1}}\right) = J\Big|_{\Delta^n}\left(\frac{\frac{w_1}{z_1}}{\frac{w_1}{z_1}+\frac{w_2}{z_2}}\left[\frac{w_1}{z_1}+\frac{w_2}{z_2}\right],\frac{\frac{w_2}{z_2}}{\frac{w_1}{z_1}+\frac{w_2}{z_2}}\left[\frac{w_1}{z_1}+\frac{w_2}{z_2}\right],\frac{w_3}{z_3}\dots,\frac{w_{n+1}}{z_{n+1}}\right)\\
&\overset{(37)}{=} J\Big|_{\Delta^{n-1}}\left(\frac{w_1}{z_1}+\frac{w_2}{z_2},\frac{w_3}{z_3}\dots,\frac{w_{n+1}}{z_{n+1}}\right)+\left(\frac{w_1}{z_1}+\frac{w_2}{z_2}\right)J\Big|_{\Delta^1}\left(\frac{\frac{w_1}{z_1}}{\frac{w_1}{z_1}+\frac{w_2}{z_2}},\frac{\frac{w_2}{z_2}}{\frac{w_1}{z_1}+\frac{w_2}{z_2}}\right)\\
&=\left(\frac{w_1}{z_1}+\frac{w_2}{z_2}\right)\Big(\phi(z_1z_2)-\phi(w_1z_2+w_2z_1)\Big)+\sum_{k=3}^{n+1}\frac{w_k}{z_k}\Big(\phi(z_k)-\phi(w_k)\Big)\\
&+\left(\frac{w_1}{z_1}+\frac{w_2}{z_2}\right)\left(\frac{\frac{w_1}{z_1}}{\frac{w_1}{z_1}+\frac{w_2}{z_2}}\Big(\phi(w_1z_2+w_2z_1)-\phi(w_1z_2)\Big)+\frac{\frac{w_2}{z_2}}{\frac{w_1}{z_1}+\frac{w_2}{z_2}}\Big(\phi(w_1z_2+w_2z_1)-\phi(w_2z_1)\Big)\right)\\
&=\left(\frac{w_1}{z_1}+\frac{w_2}{z_2}\right)\Big(\phi(z_1)+\phi(z_2)-\phi(w_1z_2+w_2z_1)\Big)+\frac{w_1}{z_1}\Big(\phi(w_1z_2+w_2z_1)-\phi(w_1)-\phi(z_2)\Big)\\
&+\frac{w_2}{z_2}\Big(\phi(w_1z_2+w_2z_1)-\phi(w_2)-\phi(z_1)\Big)+\sum_{k=3}^{n+1}\frac{w_k}{z_k}\Big(\phi(z_k)-\phi(w_k)\Big)\\
&=\sum_{k=1}^{n+1}\frac{w_k}{z_k}\Big(\phi(z_k)-\phi(w_k)\Big)
\end{aligned}
\tag{67}
$$

which concludes the proof of the formula for the function for arbitrary rational points.
Step 3. Since

$$
J\Big|_{\Delta^{n-1}}\left(\frac{w_1}{z_1},\dots,\frac{w_n}{z_n}\right)=\sum_{k=1}^{n}\frac{w_k}{z_k}\Big(\phi(z_k)-\phi(w_k)\Big)=\sum_{k=1}^{n}\frac{w_k}{z_k}\log_2\left(\frac{z_k}{w_k}\right) \tag{68}
$$

can be written in terms of a finite sum of $J\big|_{\Delta^1}$ terms thanks to the earlier calculation (49), by continuity, the function is uniquely extended to arbitrary points of the $(n-1)$-simplices since the rationals are dense in the reals. Thus,

$$
J\Big|_{\Delta^{n-1}}(p_1,\dots,p_n)=\sum_{i=1}^{n}p_i\log_2\left(\frac{1}{p_i}\right)=-\sum_{i=1}^{n}p_i\log_2 p_i \tag{69}
$$

for all $(p_1,\dots,p_n)\in\Delta^{n-1}$ for all $n\in\mathbb{Z}^+$ greater than 1. □

Remark 70. If condition (c) in the statement of theorem 40 does not hold but is instead replaced by the condition that $0\le J(\frac{1}{2},\frac{1}{2})<\infty$, then

$$
J=J\left(\frac{1}{2},\frac{1}{2}\right)H. \tag{71}
$$

This follows from the proof by just modifying the argument at the end of Step 1.

References

[1] John C. Baez, Tobias Fritz, and Tom Leinster, *A Characterization of Entropy in Terms of Information Loss*, Entropy **13** (2011), 1945–1957. http://arxiv.org/abs/1106.1791.

[2] Alfréd Rényi, *On Measures of Entropy and Information*, Proc. Fourth Berkeley Symp. on Math. Statist. and Prob. **1** (1961), 547–561.