

# A Chain Rule for the Randomness Deficiency Function

Samuel Epstein\*

October 18, 2023

## Abstract

This paper is an exposition of the addition equality theorem for algorithmic entropy in [G01], applied to the Cantor space. This application implies that randomness deficiency of infinite sequences obeys the chain rule, analogous to the finite Kolmogorov complexity case. This is a generalization of van Lambalgen's Theorem. It is unclear whether this result is folklore, but in any case, this paper presents a dedicated proof of the equality. In addition, a dual integration trick shortens the proof.

## 1 Introduction

Prefix free Kolmogorov complexity,  $\mathbf{K}$ , obeys the chain rule, with for  $x, y \in \{0, 1\}^*$ ,

$$\mathbf{K}(x, y) =^+ \mathbf{K}(x) + \mathbf{K}(y|x, \mathbf{K}(x)).$$

In this paper, we apply the addition equality theorem for algorithmic entropy in [G01] to the specific case of infinite sequences. We also shorten the proof using an integration trick. The consequence to this is a result about randomness deficiency  $\mathbf{D}$ , where for computable probability  $\mu$ , for infinite sequences,  $\mathbf{D}(\alpha|\mu, x) = \sup_n -\log \mu(\alpha[0..n]) - \mathbf{K}(\alpha[0..n]|x)$ . The randomness deficiency over the space  $\{0, 1\}^\infty \times \{0, 1\}^\infty$ , is  $\mathbf{D}(\alpha, \beta|\mu, \nu) = \sup_n -\log \mu(\alpha[0..n]) - \log \nu(\beta[0..n]) - \mathbf{K}(\alpha[0..n]|\beta[0..n])$ . The discrete case for  $\mathbf{d}(x|p) = -\log p(x) - \mathbf{K}(x)$  is trivial. The result detailed in this paper is as follows.

**Theorem.** ([G01]) *Relativized to probabilities  $\mu$  and  $\nu$  over  $\{0, 1\}^\infty$ ,*

$$\mathbf{D}(\alpha, \beta|\mu, \nu) =^+ \mathbf{D}(\alpha|\mu) + \mathbf{D}(\beta|\nu, (\alpha, \lceil \mathbf{D}(\alpha|\mu) \rceil)).$$

This is a generalization of van Lambalgen's Theorem, which states  $(\alpha, \beta)$  is ML random iff  $\alpha$  is ML random and  $\beta$  is ML random with respect to  $\alpha$ . If one were to take the complexities of the probabilities  $\mu$  and  $\nu$  into account (that is, they are no longer  $O(1)$ ) then the theorem statement and proof become more nuanced. This generalization can be seen in [G01]. An open question is whether  $\mathbf{D}$  follows the linear inequalities that parallel Shannon entropy  $\mathcal{H}$ , as Kolmogorov complexity was shown to do [HRSV00]:

**Conjecture.** *Given  $\{\alpha_1 \dots \alpha_n\} \in \{0, 1\}^{\infty n}$ , and random variables  $\{\beta_1, \dots, \beta_n\}$  is it the case that over all  $W \subseteq \{1, \dots, n\}$ , and  $\lambda_W \in \mathbb{R}$ ,*

$$\sum_W \lambda_W \mathbf{D}(\alpha^W) \leq 0 \Leftrightarrow \sum_W \lambda_W \mathcal{H}(\beta^W) \geq 0?$$

---

\*JP Theory Group. samepst@jpththeorygroup.org

## 2 Results

For the nonnegative real function  $f$ , we use  $<^+ f$ ,  $>^+ f$ , and  $=^+ f$  to denote  $< f + O(1)$ ,  $> f - O(1)$ , and  $= f \pm O(1)$ . As shown in [G01],  $2^{\mathbf{D}(\alpha|\mu)} \stackrel{*}{=} \mathbf{t}_\mu(\alpha)$  where  $\mathbf{t}_\mu$  is a universal lower computable  $\mu$ -test. Furthermore, a modification of the proof Theorem 2.3.4 in [G01] to the  $\{0, 1\}^\infty \times \{0, 1\}^\infty$  space can be used to show that  $2^{\mathbf{D}(\alpha, \beta|\mu, \nu)} \stackrel{*}{=} \mathbf{t}_{\mu, \nu}(\alpha, \beta)$ , where  $\mathbf{t}_{\mu, \nu}$  is a universal lower computable test over  $\{0, 1\}^\infty \times \{0, 1\}^\infty$ . For measure  $\mu$  and lower continuous function  $f$  over  $\{0, 1\}^\infty$ , we use the notation  $\mu^x f(x) = \int_{x \in \{0, 1\}^\infty} f(x) d\mu(x)$ . Throughout this section, the universal Turing machine is assumed to be relativized to probabilities  $\mu$  and  $\nu$  over  $\{0, 1\}^\infty$ . This means that there is an  $O(1)$  sized program that can compute  $\mu(x\{0, 1\}^\infty)$  uniformly in  $x \in \{0, 1\}^*$ , and similarly for  $\nu$ .

**Proposition 1**  $\log \nu^y 2^{\mathbf{D}(x, y|\mu, \nu)} <^+ \mathbf{D}(x|\mu)$ .

**Proof.** Let  $f(x, \mu, \nu) = \log \nu^y 2^{\mathbf{D}(x, y|\mu, \nu)}$ . The function  $f$  is lower computable and has  $\mu^x 2^{f(x, \mu, \nu)} \leq 1$ . The proposition follows from the universal properties of  $\mathbf{t}_\mu$ , where  $2^f \stackrel{*}{<} \mathbf{t}_\mu$ .  $\square$

**Proposition 2** If  $i < j$ , then

$$i + \mathbf{D}(x|\mu, j) <^+ j + \mathbf{D}(x|\mu, i).$$

**Proof.** By the properties of  $\mathbf{D}$ , we have

$$\mathbf{D}(x|\mu, j) <^+ \mathbf{D}(x|\mu, i) + \mathbf{K}(j - i) <^+ \mathbf{D}(x|\mu, i) + j - i.$$

$\square$

**Definition 1** Let  $F : \{0, 1\}^\infty \rightarrow \mathbb{Z} \cup \{-\infty, \infty\}$  be a lower semicomputable function. An  $(\mu, F)$ -test is a function  $t : \{0, 1\}^\infty \times \{0, 1\}^\infty \rightarrow \mathbb{Z} \cup \{-\infty, \infty\}$  that is lower semicomputable and  $\mu^x t(x, y) \leq 2^{F(y)}$ . There exists a maximal  $(\mu, F)$  test,  $\mathbf{t}_{(\mu, F)}$ , such that  $t \stackrel{*}{<} \mathbf{t}_{(\mu, F)}$ .

**Proposition 3** Let  $F : \{0, 1\}^\infty \rightarrow \mathbb{Z} \cup \{-\infty, \infty\}$  be an upper semicomputable function,. For all  $x$  and with  $\mathbf{t}_{(\nu, F)}(y) > -\infty$ ,

$$\mathbf{t}_{(\nu, F)}(x, y) \stackrel{*}{=} 2^{F(y)} \mathbf{t}_\nu(x|y, -F(y)).$$

**Proof.** To prove the inequality  $\stackrel{*}{>}$ , let  $g(x, y, m) = \max_{i \geq m} 2^{-i} \mathbf{t}_\nu(x|y, i)$ . This function is lower computable, and decreasing in  $m$ . The function  $g(x, y) = g_\nu(x, y, -F(y))$  is lower semicomputable since  $-F$  is upper semi-computable. The multiplicative form of Proposition 2 implies

$$\begin{aligned} g(x, y, m) &\stackrel{*}{=} 2^{-m} \mathbf{t}_\nu(x|y, m) \\ g(x, y) &\stackrel{*}{=} 2^{F(y)} \mathbf{t}_\nu(x|y, -F(y)). \end{aligned}$$

Since  $\mathbf{t}_\nu$  is a test:

$$\begin{aligned} \nu^x 2^{-m} \mathbf{t}_\nu(x|y, m) &\leq 2^{-m} \\ \nu^x g(x, y) &\stackrel{*}{<} 2^{F(y)}, \end{aligned}$$

which implies  $g(x, y) \stackrel{*}{<} \mathbf{t}_{(\nu, F)}(x, y)$  by the optimality of  $\mathbf{t}_{(\nu, F)}$ . We now consider the upper bound. Since, given fixed  $y$ ,  $2^{-F(y)} \mathbf{t}_{(\nu, F)}(x, y)$  is an  $x$ -test conditional on  $y$  and  $-F(y)$ , we have

$$2^{-F(y)} \mathbf{t}_{(\nu, F)}(x, y) \stackrel{*}{<} \mathbf{t}(x|y, -F(y)).$$

□

The following Theorem is a specific case of Theorem 4.5.2 in [G01], to the Cantor space and with  $O(1)$  complexities for the probabilities. The proof is shortened by noting that  $f$  is a universal  $F$ -test.

**Theorem 1** *Relativized to probabilities  $\mu$  and  $\nu$  over  $\{0, 1\}^\infty$ ,*

$$\mathbf{D}(x, y|\mu, \nu) =^+ \mathbf{D}(x|\mu) + \mathbf{D}(y|\nu, (x, \lceil \mathbf{D}(x|\mu) \rceil)).$$

**Proof.** Let  $f(x, y) = 2^{\mathbf{D}(x, y|\mu, \nu)}$ . Proposition 1 implies there exists  $c \in \mathbb{N}$  with  $\nu^y f(x, y) \leq 2^{\mathbf{D}(x|\mu)+c}$ . Let  $F(x, \mu) = \lceil \mathbf{D}(x|\mu) \rceil$ . Note that if  $h$  is a lower computable function such that  $\nu^y h(x, y) \stackrel{*}{<} 2^{\mathbf{D}(x|\mu)}$ , then  $\mu^x \nu^y h(x, y) \stackrel{*}{<} \mu^x \mathbf{t}_\mu(x) \stackrel{*}{<} 1$ , so  $h \stackrel{*}{<} f$ , so  $f$  is a universal  $F$ -test. Proposition 3 (swapping  $x$  and  $y$ ) gives

$$\begin{aligned} \mathbf{D}(x, y|\mu, \nu) &= \log f(x, y) =^+ F(x) + \mathbf{D}(y|\nu, (x, -F(x))) \\ \mathbf{D}(x, y|\mu, \nu) &=^+ \mathbf{D}(x|\mu) + \mathbf{D}(y|\nu, (x, \lceil \mathbf{D}(x|\mu) \rceil)). \end{aligned}$$

□

## References

- [G01] P. Gács. Quantum Algorithmic Entropy. *Journal of Physics A Mathematical General*, 34(35), 2001.
- [HRSV00] D. Hammer, A. Romashchenko, A. Shen, and N. Vereshchagin. Inequalities for shannon entropy and kolmogorov complexity. *Journal of Computer and System Sciences*, 60(2):442–464, 2000.