

Diagrammatics of information

Mee Seong Im, Clement Kam, and Caden Pici

ABSTRACT. We introduce a diagrammatic perspective for Shannon entropy created by the first author and Mikhail Khovanov and connect it to information theory and mutual information. We also give two complete proofs that the 5-term dilogarithm deforms to the 4-term infinitesimal dilogarithm.

1. Introduction

Mathematics and theoretical physics have emerged in the rapid development of modern technology in recent decades. In particular, information theory [18, 9, 24, 1, 21, 4] plays a pivotal role in communication, quantification, storage, and transfer of knowledge. One of the key measures in information theory is entropy, which quantifies the uncertainty in a random variable or the outcome of a random process. In geometric information theory [22, 23, 26, 27, 25], spaces and manifolds of probability distributions are studied from a geometric and analysis point-of-view.

Applications and interpretations of one-dimensional cobordisms have recently emerged, predominantly explored by M.S. Im and her collaborators, e.g., see [6, 7, 8, 10, 11, 15, 12, 13, 14, 16, 17]. Im and M. Khovanov introduce in [12] a diagrammatical perspective of entropy and cocycles. In particular, they view entropy as certain cobordisms (see, e.g., [14, 19]) where information about the network only depends on the boundary. This coincides with the fact that in physics, the rich information about black holes also only depends on the geometric and topological structure of their boundaries.

Furthermore, this paper has the potential to influence multiple fields. In information theory, the diagrammatic approach could lead to new ways of visualizing and analyzing entropy, particularly in complex systems. In topological quantum field theories (TQFT), the connection to information-theoretic concepts may open avenues for applying diagrammatic techniques to quantum information problems. Our proposed future directions of extending the framework to other entropy types and higher dimensions suggest that this work could serve as a foundation for a broader research program, amplifying its impact.

In this paper, we recall diagrammatic interpretation of Shannon entropy in [12], and reinterpret them in terms of information theory. We also provide two complete

2020 *Mathematics Subject Classification.* Primary: 57K16, 18M30, 28D20; Secondary: 37A35, 68P30, 94A15, 94A40.

Key words and phrases. Topological quantum field theory, TQFT, diagrammatic algebras, Shannon entropy, information theory, mutual information.

proofs that certain 5-term dilogarithm deforms into the 4-term infinitesimal dilogarithm.

Acknowledgments. The first author would like to thank Mikhail Khovanov and the Research Scientists at the Naval Research Laboratory for extensive discussions. The first author would like to thank the Office of Naval Research (ONR) and United States Naval Research Laboratory (NRL) in Washington, DC for their support.

2. Background

Let \mathbf{k} be a field of characteristic 0.

2.1. Cathelineau's \mathbf{k} -vector space. J.-L. Cathelineau constructed a \mathbf{k} -vector space $J(\mathbf{k})$ in [3], with its spanning set $\langle a, b \rangle$, where $a, b \in \mathbf{k}$, satisfying the relations:

- (1) (symmetry) $\langle a, b \rangle = \langle b, a \rangle$,
- (2) (scaling) $\langle ca, cb \rangle = c\langle a, b \rangle$, $c \in \mathbf{k}$,
- (3) (2-cocycle relation) $\langle a, b+c \rangle + \langle b, c \rangle = \langle a+b, c \rangle + \langle a, b \rangle$.

For relation (3), the field $(\mathbf{k}, +)$ under addition is viewed as a group. Furthermore, the symbols $\langle a, b \rangle$ are reminiscent of values of 2-cocycles.

First, we state a few preliminary lemmas:

LEMMA 2.1. *In the vector space $J(\mathbf{k})$, we have $\langle a, 0 \rangle = \langle 0, a \rangle = 0$, and $\langle a, -a \rangle = 0$.*

PROOF. We will first prove $\langle a, 0 \rangle = 0$. Using (2), $\langle 0, 0 \rangle = \langle 0a, 0b \rangle = 0\langle a, b \rangle = 0$. So we obtain $\langle a, 0+0 \rangle + \langle 0, 0 \rangle = \langle a+0, 0 \rangle + \langle a, 0 \rangle$ by setting $b = c = 0$ in (3). This tells us that $\langle a, 0 \rangle = 0$. Using symmetry in (1), $\langle 0, a \rangle = \langle a, 0 \rangle = 0$. Now, to prove the last equality, $\langle a, -a \rangle = a\langle 1, -1 \rangle$. So $\langle 1, -1 \rangle = \langle -(-1), -1 \rangle = -\langle -1, 1 \rangle$. We move the symbols to one side and use symmetry to obtain $0 = \langle 1, -1 \rangle + \langle -1, 1 \rangle = \langle 1, -1 \rangle + \langle 1, -1 \rangle = 2\langle 1, -1 \rangle$. Since \mathbf{k} is a field of characteristic 0, $\langle 1, -1 \rangle = 0$. Thus $\langle a, -a \rangle = 0$. \square

LEMMA 2.2. *We have $\langle 1-a, -1 \rangle = \langle a, 1-a \rangle$.*

PROOF. Applying (3) and letting $a = a$, $b = 1-a$, and $c = -1$, we have $\langle a, 1-a \rangle + \langle a+1-a, -1 \rangle = \langle a, 1-a-1 \rangle + \langle 1-a, -1 \rangle$. So $\langle a, 1-a \rangle + \langle 1, -1 \rangle = \langle a, -a \rangle + \langle 1-a, -1 \rangle$ since $\langle a, -a \rangle = 0$. We thus obtain the equality. \square

If \mathbf{k} is a field of characteristic $p \neq 0$, then

$$(2.1) \quad \sum_{n=1}^p \langle 1, n1 \rangle = 0.$$

The vector space $J(\mathbf{k})$ is known as the space of infinitesimal dilogarithms, which is often infinite-dimensional over \mathbf{k} .

In the case when $\mathbf{k} = \mathbb{R}$, $J(\mathbf{k})$ is isomorphic to the entropy of finite random variables.

2.2. Vector space isomorphic to Shannon's entropy. The \mathbf{k} -vector space $\beta(\mathbf{k})$ has a spanning set $[a]$, $a \in \mathbf{k}^*$ is invertible, of vectors and relations:

- (1) $[1] = 0$,
- (2) $[a] - [b] + a\left[\frac{b}{a}\right] + (1-a)\left[\frac{1-b}{1-a}\right] = 0$, $a \in \mathbf{k} \setminus \{0, 1\}$, $b \in \mathbf{k}^*$.

This vector space also satisfies $[a] = [1-a]$.

LEMMA 2.3. *The vector space $\beta(\mathbf{k})$ satisfies $\left[\frac{1}{a}\right] = -\frac{1}{a}[a]$.*

PROOF. Let $b = 1$ in (2) to obtain

$$[a] - \cancel{[1]} + a \left[\frac{1}{a} \right] + (1-a) \cancel{\left[\frac{1-1}{1-a} \right]} = 0.$$

This completes the proof. \square

If \mathbf{k} is a field of characteristic $p \neq 0, 2$, then

$$\sum_{n=2}^{p-1} [n1] = 0.$$

The space $\beta(\mathbf{k})$ generalizes entropy, which we explain in Section 3. This 4-term equation (2) is a limit of the 5-term equation for the dilogarithm (see [Zagier]), hence the name infinitesimal dilogarithm. For completeness, we provide a proof of this derivation in Section 10.

2.3. Isomorphism of $J(\mathbf{k})$ and $\beta(\mathbf{k})$. Cathelineau proves in [2] that the two vector spaces $J(\mathbf{k})$ and $\beta(\mathbf{k})$ are isomorphic by sending $[a] \mapsto \langle a, 1-a \rangle$. Conversely, $\langle a, -a \rangle \mapsto [0] = 0$, and $\langle a, b \rangle \mapsto (a+b) \left[\frac{a}{a+b} \right]$ whenever $a+b \neq 0$. Furthermore, $\langle a, b \rangle = \langle b, a \rangle$ if and only if $[a] = [1-a]$.

Via the isomorphism $J(\mathbf{k}) \cong \beta(\mathbf{k})$, the 2-cocycle condition in (3) for $J(\mathbf{k})$ is equivalent to

$$(a+b+c) \left[\frac{a}{a+b+c} \right] + (b+c) \left[\frac{b}{b+c} \right] = (a+b+c) \left[\frac{a+b}{a+b+c} \right] + (a+b) \left[\frac{a}{a+b} \right].$$

Cathelineau shows that there is an isomorphism between the second homology and his construction $H_2(\text{Aff}_1(\mathbf{k}), \mathbf{k}_r) \simeq \beta(\mathbf{k}) \simeq J(\mathbf{k})$, where

$$\text{Aff}_1(\mathbf{k}) = \left\{ \begin{pmatrix} c & a \\ 0 & 1 \end{pmatrix} : c \in \mathbf{k}^*, a \in \mathbf{k} \right\}$$

is the group of affine symmetries of a \mathbf{k} -line and \mathbf{k}_r is a suitable right $\text{Aff}_1(\mathbf{k})$ -module. We refer to [12] for further details.

3. Shannon entropy and Cathelineau's vector space

3.1. Shannon entropy. Let $X = \{x_1, \dots, x_n\}$, a finite set. Shannon entropy of a finite probability distribution p_X on X that associates probabilities p_1, \dots, p_n , where $\sum_{i=1}^n p_i = 1$, $0 < p_i < 1$ to its points x_i , respectively, is given by

$$(3.1) \quad H(p_X) = - \sum_{i=1}^n p_i \log p_i.$$

We can also think of $H(p_X) = -E[\log p_i] = E\left[\log \frac{1}{p_i}\right]$, where $E[\log p_i]$ is the expected value of $\log p_i$. When $n = 2$, the function $H(p_X)$ becomes a function of a single probability $p := p_1$. Then the entropy is

$$(3.2) \quad H(p) = -p \log p - (1-p) \log(1-p).$$

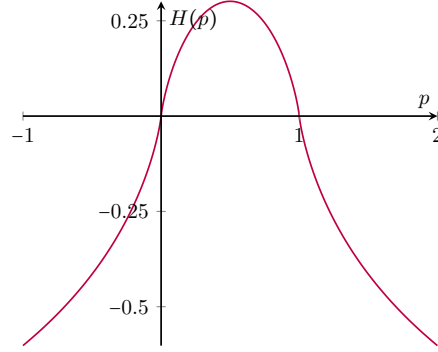


FIGURE 3.1.1. Entropy function $H(p) = -p \log |p| - (1-p) \log |1-p|$ for $-1 \leq p \leq 2$.

It is natural to extend H from the open interval $(0, 1)$ to all real numbers by defining $H(0) = H(1) = 0$ and

$$(3.3) \quad H(p) := -p \log |p| - (1-p) \log |1-p|.$$

Then H is defined and continuous on all of \mathbb{R} . The graph of H is given in Figure 3.1.1.

LEMMA 3.1. *We have $H(a) = H(1-a)$ and $H\left(\frac{1}{a}\right) = -\frac{1}{a}H(a)$.*

PROOF. For the first equality, we have

$$H(a) = -a \log |a| - (1-a) \log |1-a| = -(1-a) \log |1-a| - (1-(1-a)) \log |1-(1-a)| = H(1-a).$$

For the second equality, we have

$$\begin{aligned} H\left(\frac{1}{a}\right) &= -\frac{1}{a} \log \left|\frac{1}{a}\right| - \left(1 - \frac{1}{a}\right) \log \left|1 - \frac{1}{a}\right| \\ &= \frac{1}{a} \log |a| - \left(1 - \frac{1}{a}\right) \log \left|\frac{a-1}{a}\right| \\ &= \frac{1}{a} \log |a| - \left(1 - \frac{1}{a}\right) (\log |1-a| - \log |a|) \\ &= \log |a| - \left(1 - \frac{1}{a}\right) \log |1-a| \\ &= -\frac{1}{a} (-a \log |a| - (1-a) \log |1-a|) \\ &= -\frac{1}{a} H(a). \end{aligned}$$

This concludes the proof. \square

The entropy function satisfies a four-term functional equation, see [20],

$$(3.4) \quad H(p) - H(q) + p H\left(\frac{q}{p}\right) + (1-p) H\left(\frac{1-q}{1-p}\right) = 0.$$

Equation (3.4) together with $H(1-p) = H(p)$ and continuity property uniquely determine H as a nonzero function $\mathbb{R} \rightarrow \mathbb{R}$. See [5, 20]. Also see [28, Equation (38)].

PROPOSITION 3.2. *Let $H : \mathbb{R} \rightarrow \mathbb{R}$, where H is (3.3) satisfying conditions $H(0) = H(1) = 0$. Then H satisfies the 4-term relation (3.4).*

Since the proof of Proposition 3.2 is not explicitly stated anywhere in the literature, we will provide its proof.

PROOF. We have the following set of equalities:

$$\begin{aligned}
 -H(p) &= p \log |p| + (1-p) \log |1-p|, \\
 -H(q) &= q \log |q| + (1-q) \log |1-q|, \\
 -H\left(\frac{q}{p}\right) &= \frac{q}{p} \log \left| \frac{q}{p} \right| + \left(1 - \frac{q}{p}\right) \log \left| 1 - \frac{q}{p} \right| \\
 (3.5) \quad &= \frac{q}{p} (\log |q| - \log |p|) + \frac{p-q}{p} (\log |p-q| - \log |p|), \\
 -H\left(\frac{1-q}{1-p}\right) &= \frac{1-q}{1-p} \log \left| \frac{1-q}{1-p} \right| + \frac{q-p}{1-p} \log \left| \frac{q-p}{1-p} \right| \\
 &= \frac{1-q}{1-p} (\log |1-q| - \log |1-p|) + \frac{q-p}{1-p} (\log |q-p| - \log |1-p|).
 \end{aligned}$$

So

$$\begin{aligned}
 -pH\left(\frac{q}{p}\right) - (1-p)H\left(\frac{1-q}{1-p}\right) &= q (\log |q| - \log |p|) + (p-q) (\log |p-q| - \log |p|) \\
 &\quad + (1-q) (\log |1-q| - \log |1-p|) + (q-p) (\log |q-p| - \log |1-p|) \\
 &= q \log |q| - q \log |p| + \cancel{(p-q) \log |p-q|} - (p-q) \log |p| \\
 &\quad + (1-q) \log |1-q| - (1-q) \log |1-p| + \cancel{(q-p) \log |q-p|} - (q-p) \log |1-p| \\
 &= q \log |q| + (1-q) \log |1-q| - \cancel{q \log |p|} + \cancel{q \log |p|} - p \log |p| \\
 &\quad - \log |1-p| + \cancel{q \log |1-p|} - \cancel{q \log |1-p|} + p \log |1-p| \\
 &= q \log |q| + (1-q) \log |1-q| - (p \log |p| + (1-p) \log |1-p|) \\
 &= -H(q) + H(p).
 \end{aligned}$$

Therefore, $H(p) - H(q) + pH\left(\frac{q}{p}\right) + (1-p)H\left(\frac{1-q}{1-p}\right) = 0$ holds. \square

We conclude this section by mentioning an important relationship relating Cathelineau's vector space [2, 3] and Kontsevich's observation [20], which is

$$(3.6) \quad \langle p_1, p_2 \rangle = \begin{cases} (p_1 + p_2) H\left(\frac{p_1}{p_1 + p_2}\right) & \text{if } p_1 + p_2 \neq 0, \\ 0 & \text{if } p_1 + p_2 = 0, \end{cases}$$

relating Cathelineau's vector space and entropy.

LEMMA 3.3. *Equation (3.6) satisfies symmetry, scaling, and the 2-cocycle relation.*

PROOF. We will first prove symmetry:

$$\begin{aligned}
 \langle p_1, p_2 \rangle &= (p_1 + p_2) H\left(\frac{p_1}{p_1 + p_2}\right) \\
 &= (p_1 + p_2) \left(-\frac{p_1}{p_1 + p_2} \log \left| \frac{p_1}{p_1 + p_2} \right| - \left(1 - \frac{p_1}{p_1 + p_2}\right) \log \left| 1 - \frac{p_1}{p_1 + p_2} \right| \right)
 \end{aligned}$$

$$\begin{aligned}
&= (p_1 + p_2) \left(-\frac{p_1}{p_1 + p_2} \log \left| \frac{p_1}{p_1 + p_2} \right| - \frac{p_2}{p_1 + p_2} \log \left| \frac{p_2}{p_1 + p_2} \right| \right) \\
&= (p_1 + p_2) \left(-\frac{p_2}{p_1 + p_2} \log \left| \frac{p_2}{p_1 + p_2} \right| - \frac{p_1}{p_1 + p_2} \log \left| \frac{p_1}{p_1 + p_2} \right| \right) \\
&= (p_1 + p_2) \left(-\frac{p_2}{p_1 + p_2} \log \left| \frac{p_2}{p_1 + p_2} \right| - \left(1 - \frac{p_2}{p_1 + p_2} \right) \log \left| 1 - \frac{p_2}{p_1 + p_2} \right| \right) \\
&= (p_1 + p_2) H \left(\frac{p_2}{p_1 + p_2} \right) \\
&= \langle p_2, p_1 \rangle.
\end{aligned}$$

Next, we will prove scaling:

$$\begin{aligned}
\langle c p_1, c p_2 \rangle &= (c p_1 + c p_2) H \left(\frac{c p_1}{c p_1 + c p_2} \right) \\
&= (c p_1 + c p_2) \left(-\frac{c p_1}{c p_1 + c p_2} \log \left| \frac{c p_1}{c p_1 + c p_2} \right| - \left(1 - \frac{c p_1}{c p_1 + c p_2} \right) \log \left| 1 - \frac{c p_1}{c p_1 + c p_2} \right| \right) \\
&= c(p_1 + p_2) \left(-\frac{\cancel{c} p_1}{\cancel{c} p_1 + \cancel{c} p_2} \log \left| \frac{\cancel{c} p_1}{\cancel{c} p_1 + \cancel{c} p_2} \right| - \frac{\cancel{c} p_2}{\cancel{c} p_1 + \cancel{c} p_2} \log \left| \frac{\cancel{c} p_2}{\cancel{c} p_1 + \cancel{c} p_2} \right| \right) \\
&= c(p_1 + p_2) \left(-\frac{p_1}{p_1 + p_2} \log \left| \frac{p_1}{p_1 + p_2} \right| - \frac{p_2}{p_1 + p_2} \log \left| \frac{p_2}{p_1 + p_2} \right| \right) \\
&= c(p_1 + p_2) H \left(\frac{p_1}{p_1 + p_2} \right) \\
&= c \langle p_1, p_2 \rangle.
\end{aligned}$$

Finally, we will prove the 2-cocycle condition:

$$\begin{aligned}
\langle p_1, p_2 \rangle &= (p_1 + p_2) H \left(\frac{p_1}{p_1 + p_2} \right) \\
&= \cancel{(p_1 + p_2)} \left(-\frac{p_1}{\cancel{p_1 + p_2}} \log \left| \frac{p_1}{p_1 + p_2} \right| - \frac{p_2}{\cancel{p_1 + p_2}} \log \left| \frac{p_2}{p_1 + p_2} \right| \right) \\
&= -p_1 \log \left| \frac{p_1}{p_1 + p_2} \right| - p_2 \log \left| \frac{p_2}{p_1 + p_2} \right| \\
&= -p_1 \log |p_1| - p_2 \log |p_2| + (p_1 + p_2) \log |p_1 + p_2|
\end{aligned}$$

while

$$\begin{aligned}
\langle p_1 + p_2, p_3 \rangle &= (p_1 + p_2 + p_3) H \left(\frac{p_1 + p_2}{p_1 + p_2 + p_3} \right) \\
&= H(p_1 + p_2) \quad \text{since } p_1 + p_2 + p_3 = 1 \\
&= -(p_1 + p_2) \log |p_1 + p_2| - (1 - (p_1 + p_2)) \log |1 - (p_1 + p_2)| \\
&= -(p_1 + p_2) \log |p_1 + p_2| - p_3 \log |p_3|.
\end{aligned}$$

So the sum is:

$$\begin{aligned}
\langle p_1, p_2 \rangle + \langle p_1 + p_2, p_3 \rangle &= -p_1 \log |p_1| - p_2 \log |p_2| + \cancel{(p_1 + p_2) \log |p_1 + p_2|} \\
&\quad - \cancel{(p_1 + p_2) \log |p_1 + p_2|} - p_3 \log |p_3| \\
&= -p_1 \log |p_1| - p_2 \log |p_2| - p_3 \log |p_3| \\
&= H(p_X).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
 \langle p_2, p_3 \rangle &= (p_2 + p_3) H\left(\frac{p_2}{p_2 + p_3}\right) \\
 &= \cancel{(p_2 + p_3)} \left(-\frac{p_2}{\cancel{p_2 + p_3}} \log \left| \frac{p_2}{p_2 + p_3} \right| - \frac{p_3}{\cancel{p_2 + p_3}} \log \left| \frac{p_3}{p_2 + p_3} \right| \right) \\
 &= -p_2 \log \left| \frac{p_2}{p_2 + p_3} \right| - p_3 \log \left| \frac{p_3}{p_2 + p_3} \right| \\
 &= -p_2 \log |p_2| - p_3 \log |p_3| + (p_2 + p_3) \log |p_2 + p_3|
 \end{aligned}$$

while

$$\begin{aligned}
 \langle p_1, p_2 + p_3 \rangle &= (p_1 + p_2 + p_3) H\left(\frac{p_1}{p_1 + p_2 + p_3}\right) \\
 &= H(p_1) \quad \text{since } p_1 + p_2 + p_3 = 1 \\
 &= -p_1 \log |p_1| - (1 - p_1) \log |1 - p_1| \\
 &= -p_1 \log |p_1| - (p_2 + p_3) \log |p_2 + p_3|.
 \end{aligned}$$

So the sum is:

$$\begin{aligned}
 \langle p_2, p_3 \rangle + \langle p_1, p_2 + p_3 \rangle &= -p_2 \log |p_2| - p_3 \log |p_3| + \cancel{(p_2 + p_3) \log |p_2 + p_3|} \\
 &\quad - p_1 \log |p_1| - \cancel{(p_2 + p_3) \log |p_2 + p_3|} \\
 &= -p_1 \log |p_1| - p_2 \log |p_2| - p_3 \log |p_3| \\
 &= H(p_X).
 \end{aligned}$$

This concludes the proof. \square

LEMMA 3.4. *If $p_1 + p_2 = 1$, we have*

$$(3.7) \quad H(p_1) = H(p_2).$$

PROOF. This follows from Lemma 3.3. \square

Let $\rho : H(\mathbf{k}) \rightarrow J(\mathbf{k})$, where $\rho(H(a)) = \langle a, 1 - a \rangle$.

THEOREM 3.5. *The map ρ is an isomorphism of vector spaces.*

PROOF. We have

$$\rho(H(a)) = \langle a, 1 - a \rangle = \langle 1 - a, a \rangle = \langle 1 - a, 1 - (1 - a) \rangle = \rho(H(1 - a))$$

and

$$\begin{aligned}
 \rho\left(H\left(\frac{1}{a}\right)\right) &= \left\langle \frac{1}{a}, 1 - \frac{1}{a} \right\rangle = -\frac{1}{a} \langle -1, -a + 1 \rangle = -\frac{1}{a} \langle -1, 1 - a \rangle \\
 &= -\frac{1}{a} \langle 1 - a, -1 \rangle \stackrel{\dagger}{=} -\frac{1}{a} \langle a, 1 - a \rangle = -\frac{1}{a} \rho(H(a)),
 \end{aligned}$$

where \dagger holds by Lemma 2.2.

Furthermore, equation (3.4) under ρ becomes

$$\begin{aligned}
 \rho(H(0)) &= \rho(H(a)) - \rho(H(b)) + a\rho\left(H\left(\frac{b}{a}\right)\right) + (1 - a)\rho\left(H\left(\frac{1 - b}{1 - a}\right)\right) \\
 &= \langle a, 1 - a \rangle - \langle b, 1 - b \rangle + a\left\langle \frac{b}{a}, 1 - \frac{b}{a} \right\rangle + (1 - a)\left\langle \frac{1 - b}{1 - a}, 1 - \frac{1 - b}{1 - a} \right\rangle \\
 &= \langle a, 1 - a \rangle - \langle b, 1 - b \rangle + \langle b, a - b \rangle + \langle 1 - b, 1 - a - (1 - b) \rangle
 \end{aligned}$$

$$\stackrel{\ddagger}{=} \langle a, 1-a \rangle - \langle b, 1-b \rangle - \langle a, b-a \rangle + \langle 1-b, b-a \rangle,$$

where \ddagger holds due to

$$\langle b, a-b \rangle = b \left\langle 1, \frac{a}{b} - 1 \right\rangle = -b \left\langle 1 - \frac{a}{b}, -1 \right\rangle = -b \left\langle \frac{a}{b}, 1 - \frac{a}{b} \right\rangle = -\langle a, b-a \rangle$$

via Lemma 2.2. Now, replacing a with a , b with $b-a$, and $c = 1-b$ in the 2-cocycle relation (3), we see that

$$\begin{aligned} & \langle a, b-a+1-b \rangle + \langle b-a, 1-b \rangle - \langle a+b-a, 1-b \rangle - \langle a, b-a \rangle \\ &= \langle a, 1-a \rangle + \langle 1-b, b-a \rangle - \langle a, b-a \rangle - \langle b, 1-b \rangle = 0. \end{aligned}$$

So

$$\begin{aligned} \rho(0) &= \rho(H(0)) = \rho \left(H(a) - H(b) + aH\left(\frac{b}{a}\right) + (1-a)H\left(\frac{1-b}{1-a}\right) \right) \\ &= \langle a, 1-a \rangle - \langle b, 1-b \rangle - \langle a, b-a \rangle + \langle 1-b, b-a \rangle = 0. \end{aligned}$$

Conversely, let $\lambda : J(\mathbf{k}) \rightarrow H(\mathbf{k})$, where $\lambda(\langle a, b \rangle) = (a+b)H\left(\frac{a}{a+b}\right)$ if $a+b \neq 0$, and $\lambda(\langle a, b \rangle) = 0$ if $a+b = 0$. So we have

$$\begin{aligned} \lambda(\langle a, b \rangle) &= (a+b)H\left(\frac{a}{a+b}\right) \\ &= (a+b)H\left(1 - \frac{a}{a+b}\right) \text{ by the first equality in Lemma 3.1} \\ &= (a+b)H\left(\frac{b}{a+b}\right) \\ &= (b+a)H\left(\frac{b}{a+b}\right) \\ &= \lambda(\langle b, a \rangle). \end{aligned}$$

Secondly, $\lambda(\langle ca, cb \rangle) = c\lambda(\langle a, b \rangle)$ holds by scaling in Lemma 3.3.

Lastly, we have

$$\begin{aligned} 0 &= \lambda(0) \\ &= \lambda(\langle a, b+c \rangle + \langle b, c \rangle - \langle a+b, c \rangle - \langle a, b \rangle) \\ &= \lambda(\langle a, b+c \rangle) + \lambda(\langle b, c \rangle) - \lambda(\langle a+b, c \rangle) - \lambda(\langle a, b \rangle) \\ &= (a+b+c)H\left(\frac{a}{a+b+c}\right) + (b+c)H\left(\frac{b}{b+c}\right) - (a+b+c)H\left(\frac{a+b}{a+b+c}\right) - (a+b)H\left(\frac{a}{a+b}\right) \\ &= (a+b+c)H\left(\frac{a}{a+b+c}\right) + (b+c)H\left(\frac{b}{b+c}\right) - (a+b+c)H\left(\frac{a+b}{a+b+c}\right) - (a+b)H\left(\frac{b}{a+b}\right) \end{aligned}$$

since $H\left(\frac{a}{a+b}\right) = H\left(\frac{b}{a+b}\right)$ by the first equality in Lemma 3.1.

Let

$$u = \frac{a}{a+b+c} \quad \text{and} \quad v = \frac{c}{a+b+c}.$$

Then

$$\begin{aligned} & (a+b+c)H\left(\frac{a}{a+b+c}\right) + (b+c)H\left(\frac{b}{b+c}\right) - (a+b+c)H\left(\frac{a+b}{a+b+c}\right) - (a+b)H\left(\frac{b}{a+b}\right) \\ &= (a+b+c) \left(H(u) + (1-u)H\left(\frac{1-u-v}{1-u}\right) - H(1-v) - (1-v)H\left(\frac{1-u-v}{1-v}\right) \right) \\ &= (a+b+c) \left(H(u) + (1-u)H\left(\frac{v}{1-u}\right) - H(1-v) - (1-v)H\left(\frac{u}{1-v}\right) \right) \end{aligned}$$

by the first equality in Lemma 3.1. Now let $p = u$ and $q = 1 - v$. Then

$$\begin{aligned}
 & H(u) + (1 - u)H\left(\frac{v}{1 - u}\right) - H(1 - v) - (1 - v)H\left(\frac{u}{1 - v}\right) \\
 &= H(p) + (1 - p)H\left(\frac{1 - q}{1 - p}\right) - H(q) - qH\left(\frac{p}{q}\right) \\
 &= H(p) + (1 - p)H\left(\frac{1 - q}{1 - p}\right) - H(q) + q\frac{p}{q}H\left(\frac{q}{p}\right) \\
 &= H(p) + (1 - p)H\left(\frac{1 - q}{1 - p}\right) - H(q) + pH\left(\frac{q}{p}\right) \\
 &= 0
 \end{aligned}$$

by the second equality in Lemma 3.1 and (3.4). \square

In characteristic $p > 0$, we have

$$\begin{aligned}
 \rho\left(\sum_{n=1}^p H(n)\right) &= \sum_{n=1}^p \rho(H(n)) = \sum_{n=1}^p \langle n, 1 - n \rangle = \sum_{n=1}^p \langle 1 - n, -1 \rangle \\
 &= -\sum_{n=1}^p \langle n - 1, 1 \rangle = -\sum_{n=1}^p \langle 1, n - 1 \rangle = -\sum_{n=1}^p \langle 1, n \rangle = 0,
 \end{aligned}$$

where the last equality holds by (2.1).

Entropy is related to the pair of symbols $\langle \cdot, \cdot \rangle$ via the following:

PROPOSITION 3.6. *Let p_1, \dots, p_n be a finite probability distribution on a finite set $X = \{x_1, \dots, x_n\}$. Then for $n \geq 3$,*

$$(3.8) \quad H(p_X) = H(p_1, \dots, p_n) = \sum_{j=2}^n \left\langle \sum_{i=1}^{j-1} p_i, p_j \right\rangle.$$

Expanding out (3.8), we have

$$H(p_1, \dots, p_n) = \langle p_1, p_2 \rangle + \langle p_1 + p_2, p_3 \rangle + \langle p_1 + p_2 + p_3, p_4 \rangle + \dots + \langle p_1 + p_2 + \dots + p_{n-1}, p_n \rangle.$$

PROOF. Since

$$\begin{aligned}
 & \langle p_1, p_2 \rangle + \langle p_1 + p_2, p_3 \rangle + \langle p_1 + p_2 + p_3, p_4 \rangle + \dots + \langle p_1 + p_2 + \dots + p_{n-1}, p_n \rangle \\
 &= (p_1 + p_2)H\left(\frac{p_1}{p_1 + p_2}\right) + (p_1 + p_2 + p_3)H\left(\frac{p_1 + p_2}{p_1 + p_2 + p_3}\right) + \dots + \\
 &\quad + (p_1 + p_2 + p_3 + \dots + p_n)H\left(\frac{p_1 + p_2 + \dots + p_{n-1}}{p_1 + p_2 + p_3 + \dots + p_n}\right) \\
 &= (p_1 + p_2) \left(-\frac{p_1}{p_1 + p_2} \log \left| \frac{p_1}{p_1 + p_2} \right| - \left(1 - \frac{p_1}{p_1 + p_2} \right) \log \left| 1 - \frac{p_1}{p_1 + p_2} \right| \right) \\
 &\quad + (p_1 + p_2 + p_3) \left(-\frac{p_1 + p_2}{p_1 + p_2 + p_3} \log \left| \frac{p_1 + p_2}{p_1 + p_2 + p_3} \right| \right. \\
 &\quad \left. - \left(1 - \frac{p_1 + p_2}{p_1 + p_2 + p_3} \right) \log \left| 1 - \frac{p_1 + p_2}{p_1 + p_2 + p_3} \right| \right) + \dots + \\
 &\quad + (p_1 + p_2 + p_3 + \dots + p_n) \left(-\frac{p_1 + p_2 + \dots + p_{n-1}}{p_1 + p_2 + p_3 + \dots + p_n} \log \left| \frac{p_1 + p_2 + \dots + p_{n-1}}{p_1 + p_2 + p_3 + \dots + p_n} \right| \right. \\
 &\quad \left. - \left(1 - \frac{p_1 + p_2 + \dots + p_{n-1}}{p_1 + p_2 + p_3 + \dots + p_n} \right) \log \left| 1 - \frac{p_1 + p_2 + \dots + p_{n-1}}{p_1 + p_2 + p_3 + \dots + p_n} \right| \right)
 \end{aligned}$$

$$\begin{aligned}
&= \cancel{(p_1 + p_2)} \left(-\frac{p_1}{\cancel{p_1 + p_2}} \log \left| \frac{p_1}{p_1 + p_2} \right| - \frac{p_2}{\cancel{p_1 + p_2}} \log \left| \frac{p_2}{p_1 + p_2} \right| \right) \\
&\quad + \cancel{(p_1 + p_2 + p_3)} \left(-\frac{p_1 + p_2}{\cancel{p_1 + p_2 + p_3}} \log \left| \frac{p_1 + p_2}{p_1 + p_2 + p_3} \right| - \frac{p_3}{\cancel{p_1 + p_2 + p_3}} \log \left| \frac{p_3}{p_1 + p_2 + p_3} \right| \right) \\
&\quad + \dots + \cancel{(p_1 + p_2 + \dots + p_n)} \left(-\frac{p_1 + p_2 + \dots + p_{n-1}}{\cancel{p_1 + p_2 + p_3 + \dots + p_n}} \log \left| \frac{p_1 + p_2 + \dots + p_{n-1}}{p_1 + p_2 + p_3 + \dots + p_n} \right| \right. \\
&\quad \left. - \frac{p_n}{\cancel{p_1 + p_2 + p_3 + \dots + p_n}} \log \left| \frac{p_n}{p_1 + p_2 + p_3 + \dots + p_n} \right| \right) \\
&= -p_1 \log \left| \frac{p_1}{p_1 + p_2} \right| - p_2 \log \left| \frac{p_2}{p_1 + p_2} \right| - (p_1 + p_2) \log \left| \frac{p_1 + p_2}{p_1 + p_2 + p_3} \right| - p_3 \log \left| \frac{p_3}{p_1 + p_2 + p_3} \right| \\
&\quad - \dots - (p_1 + p_2 + \dots + p_{n-1}) \log \left| \frac{p_1 + p_2 + \dots + p_{n-1}}{p_1 + p_2 + p_3 + \dots + p_n} \right| \\
&\quad - p_n \log \left| \frac{p_n}{p_1 + p_2 + p_3 + \dots + p_n} \right| \\
&= -p_1 \log |p_1| - p_2 \log |p_2| + \cancel{(p_1 + p_2) \log |p_1 + p_2|} - \cancel{(p_1 + p_2) \log |p_1 + p_2|} - p_3 \log |p_3| \\
&\quad + \cancel{(p_1 + p_2 + p_3) \log |p_1 + p_2 + p_3|} - \cancel{(p_1 + p_2 + p_3) \log |p_1 + p_2 + p_3|} + \dots + \\
&\quad + \cancel{(p_1 + p_2 + \dots + p_{n-1}) \log |p_1 + p_2 + \dots + p_{n-1}|} \\
&\quad - \cancel{(p_1 + p_2 + \dots + p_{n-1}) \log |p_1 + p_2 + \dots + p_{n-1}|} - p_n \log |p_n| \\
&\quad + \underbrace{(p_1 + p_2 + p_3 + \dots + p_n) \log |p_1 + p_2 + p_3 + \dots + p_n|}_1 \\
&= -p_1 \log |p_1| - p_2 \log |p_2| - p_3 \log |p_3| - \dots - p_n \log |p_n| = H(p_X).
\end{aligned}$$

This concludes the proof. \square

4. Joint entropy

4.1. Joint entropy of 2 random variables. Let X be a discrete random variable associated with probabilities p_1, \dots, p_n . Recall Shannon entropy in (3.1). Let Y be a discrete random variable associated with probabilities q_1, \dots, q_k . Then the joint entropy of X and Y is defined to be

$$(4.1) \quad H(p_X, p_Y) = - \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log p_{ij},$$

where $p_{ij} = P(X = p_i, Y = q_j)$, the probability that the random variables $X = p_i$ and $Y = q_j$. Then similar to (3.6), we have

$$(4.2) \quad \langle p_{ij}, p_{\ell k} \rangle = \begin{cases} (p_{ij} + p_{\ell k}) H \left(\frac{p_{ij}}{p_{ij} + p_{\ell k}} \right) & \text{if } p_{ij} + p_{\ell k} \neq 0, \\ 0 & \text{if } p_{ij} + p_{\ell k} = 0. \end{cases}$$

PROPOSITION 4.1. Equation (4.2) satisfies symmetry, scaling, and the 2-cocycle relation for the vector space $J(\mathbf{k})$.

The proof of Proposition 4.1 is similar to the proof of Lemma 3.3.

Similar in spirit to Proposition 3.6 and in order to provide a motivation for Theorem 4.3, we have the following:

PROPOSITION 4.2. *Let p_1, p_2 and q_1, q_2 be finite probability distributions on sets $X = \{x_1, x_2\}$ and $Y = \{y_1, y_2\}$, respectively. Let p_{ij} be the probability of obtaining p_i and q_j . Then a relation between entropy and Cathelineau's vector space is*

$$(4.3) \quad H(p_X, p_Y) = \langle p_{11}, p_{21} \rangle + \langle p_{12}, p_{22} \rangle + \langle p_{11} + p_{21}, p_{12} + p_{22} \rangle.$$

PROOF. Using (4.2), we have:

$$\begin{aligned} \langle p_{11}, p_{21} \rangle &= (p_{11} + p_{21}) H\left(\frac{p_{11}}{p_{11} + p_{21}}\right), \\ \langle p_{12}, p_{22} \rangle &= (p_{12} + p_{22}) H\left(\frac{p_{12}}{p_{12} + p_{22}}\right), \\ \langle p_{11} + p_{21}, p_{12} + p_{22} \rangle &= (p_{11} + p_{21} + p_{12} + p_{22}) H\left(\frac{p_{11} + p_{21}}{p_{11} + p_{21} + p_{12} + p_{22}}\right). \end{aligned}$$

So

$$\begin{aligned} \langle p_{11}, p_{21} \rangle + \langle p_{12}, p_{22} \rangle + \langle p_{11} + p_{21}, p_{12} + p_{22} \rangle &= (p_{11} + p_{21}) H\left(\frac{p_{11}}{p_{11} + p_{21}}\right) \\ &\quad + (p_{12} + p_{22}) H\left(\frac{p_{12}}{p_{12} + p_{22}}\right) + (p_{11} + p_{21} + p_{12} + p_{22}) H\left(\frac{p_{11} + p_{21}}{p_{11} + p_{21} + p_{12} + p_{22}}\right) \\ &= (p_{11} + p_{21}) \left(-\frac{p_{11}}{p_{11} + p_{21}} \log\left(\frac{p_{11}}{p_{11} + p_{21}}\right) - \left(1 - \frac{p_{11}}{p_{11} + p_{21}}\right) \log\left(1 - \frac{p_{11}}{p_{11} + p_{21}}\right) \right) \\ &\quad + (p_{12} + p_{22}) \left(-\frac{p_{12}}{p_{12} + p_{22}} \log\left(\frac{p_{12}}{p_{12} + p_{22}}\right) - \left(1 - \frac{p_{12}}{p_{12} + p_{22}}\right) \log\left(1 - \frac{p_{12}}{p_{12} + p_{22}}\right) \right) \\ &\quad + (p_{11} + p_{21} + p_{12} + p_{22}) \left(-\frac{p_{11} + p_{21}}{p_{11} + p_{21} + p_{12} + p_{22}} \log\left(\frac{p_{11} + p_{21}}{p_{11} + p_{21} + p_{12} + p_{22}}\right) \right. \\ &\quad \left. - \left(1 - \frac{p_{11} + p_{21}}{p_{11} + p_{21} + p_{12} + p_{22}}\right) \log\left(1 - \frac{p_{11} + p_{21}}{p_{11} + p_{21} + p_{12} + p_{22}}\right) \right) \\ &= \cancel{(p_{11} + p_{21})} \left(-\frac{p_{11}}{\cancel{p_{11} + p_{21}}} \log\left(\frac{p_{11}}{p_{11} + p_{21}}\right) - \frac{p_{21}}{\cancel{p_{11} + p_{21}}} \log\left(\frac{p_{21}}{p_{11} + p_{21}}\right) \right) \\ &\quad + \cancel{(p_{12} + p_{22})} \left(-\frac{p_{12}}{\cancel{p_{12} + p_{22}}} \log\left(\frac{p_{12}}{p_{12} + p_{22}}\right) - \frac{p_{22}}{\cancel{p_{12} + p_{22}}} \log\left(\frac{p_{22}}{p_{12} + p_{22}}\right) \right) \\ &\quad + \cancel{(p_{11} + p_{21} + p_{12} + p_{22})} \left(-\frac{p_{11} + p_{21}}{\cancel{p_{11} + p_{21} + p_{12} + p_{22}}} \log\left(\frac{p_{11} + p_{21}}{p_{11} + p_{21} + p_{12} + p_{22}}\right) \right. \\ &\quad \left. - \frac{p_{12} + p_{22}}{\cancel{p_{11} + p_{21} + p_{12} + p_{22}}} \log\left(\frac{p_{12} + p_{22}}{p_{11} + p_{21} + p_{12} + p_{22}}\right) \right) \\ &= -p_{11} \log\left(\frac{p_{11}}{p_{11} + p_{21}}\right) - p_{21} \log\left(\frac{p_{21}}{p_{11} + p_{21}}\right) - p_{12} \log\left(\frac{p_{12}}{p_{12} + p_{22}}\right) - p_{22} \log\left(\frac{p_{22}}{p_{12} + p_{22}}\right) \\ &\quad - (p_{11} + p_{21}) \log\left(\frac{p_{11} + p_{21}}{p_{11} + p_{21} + p_{12} + p_{22}}\right) - (p_{12} + p_{22}) \log\left(\frac{p_{12} + p_{22}}{p_{11} + p_{21} + p_{12} + p_{22}}\right) \\ &= -p_{11} \log(p_{11}) - p_{21} \log(p_{21}) + \cancel{(p_{11} + p_{21}) \log(p_{11} + p_{21})} \\ &\quad - p_{12} \log(p_{12}) - p_{22} \log(p_{22}) + \cancel{(p_{12} + p_{22}) \log(p_{12} + p_{22})} \\ &\quad - \cancel{(p_{11} + p_{21}) \log(p_{11} + p_{21})} - \cancel{(p_{12} + p_{22}) \log(p_{12} + p_{22})} \\ &\quad + \underbrace{(p_{11} + p_{21} + p_{12} + p_{22}) \log(p_{11} + p_{21} + p_{12} + p_{22})}_1 \end{aligned}$$

$$= -p_{11} \log(p_{11}) - p_{21} \log(p_{21}) - p_{12} \log(p_{12}) - p_{22} \log(p_{22}) = H(p_X, p_Y).$$

□

More generally, we have the following:

THEOREM 4.3. *Let p_1, \dots, p_n and q_1, \dots, q_m be finite probability distributions on finite sets $X = \{x_1, \dots, x_n\}$ and $Y = \{y_1, \dots, y_m\}$, respectively. Let $p_{ij} = P(p_i, q_j)$, the probability of obtaining p_i and q_j . Then*

$$(4.4) \quad H(p_X, p_Y) = \sum_{i=1}^n \sum_{k=1}^{m-1} \left\langle \sum_{j=1}^k p_{ij}, p_{i,j+1} \right\rangle + \sum_{k=1}^{n-1} \left\langle \sum_{i=1}^k \sum_{j=1}^m p_{ij}, \sum_{\ell=1}^m p_{i\ell} \right\rangle.$$

The proof of Theorem 4.3 is a direct calculation, similar to the proof of Proposition 3.6 and Proposition 4.2. We give a diagrammatical calculus perspective of entropy in Section 5, which originated in [12]. Using this, one can easily prove Theorem 4.3.

4.2. Entropy for u random variables. One may extend (4.1) as follows. Consider the finite set $X_i = \{x_{i_1}, \dots, x_{i_{k_i}}\}$, where $1 \leq i \leq u$. Let p_{X_i} be a probability distribution on X_i that associates probabilities $p_{i_1}, \dots, p_{i_{k_i}}$ to the points $x_{i_1}, \dots, x_{i_{k_i}}$, respectively, where the sum $\sum_{j=1}^{k_i} p_{i_j} = 1$, $0 < p_{i_j} < 1$, $1 \leq i \leq u$.

Write

$$I_J := 1_{j_1} 2_{j_2} \cdots u_{j_u}, \quad J := (j_1, j_2, \dots, j_u), \quad \text{and} \quad K := (k_1, \dots, k_u).$$

Let the joint probability for random variables $\mathfrak{X}_1, \dots, \mathfrak{X}_u$ be

$$p_{I_J} = p_{1_{j_1} 2_{j_2} \cdots u_{j_u}} = P(\mathfrak{X}_1 = p_{1_{j_1}}, \mathfrak{X}_2 = p_{2_{j_2}}, \dots, \mathfrak{X}_u = p_{u_{j_u}}).$$

We obtain the joint entropy

$$(4.5) \quad H(p_{X_1}, \dots, p_{X_u}) = - \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} \cdots \sum_{j_u=1}^{k_u} p_{1_{j_1} 2_{j_2} \cdots u_{j_u}} \log(p_{1_{j_1} 2_{j_2} \cdots u_{j_u}}),$$

or more compactly, we write

$$(4.6) \quad H(p_{X_1}, \dots, p_{X_u}) = - \sum_{J=1}^K p_{I_J} \log(p_{I_J}).$$

PROPOSITION 4.4. *Under the conditions as above, we have*

$$(4.7) \quad \langle p_{I_J}, p_{I_{J'}} \rangle = (p_{I_J} + p_{I_{J'}}) H\left(\frac{p_{I_J}}{p_{I_J} + p_{I_{J'}}}\right),$$

which satisfies symmetry, scaling, and the 2-cocycle relation for $J(\mathbf{k})$ in Section 2.1.

The proof of Proposition 4.4 is similar to the proof of Lemma 3.3.

4.3. Rescaling a finite probability distribution by a scalar. Rescaling the pair of symbols $\langle cp_1, cp_2 \rangle = c \langle p_1, p_2 \rangle$ corresponds to rescaling entropy. In our diagrammatic story, it will correspond to a red wavy line, which is to the left of our graphical network.

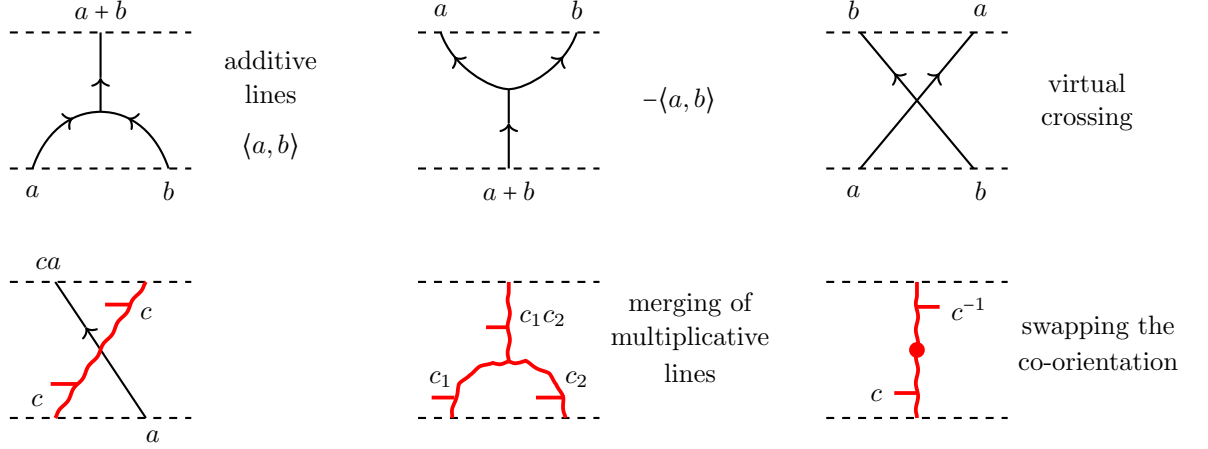


FIGURE 5.0.1. Upper left: Whenever two black additive lines merge, we evaluate $\langle a, b \rangle$ at the additive vertex. Upper middle: whenever two additive lines split, we evaluate $-\langle a, b \rangle$. Upper right: we are allowed to have virtual crossing whenever two additive lines cross but the intersection of these two lines is virtual, so there is no corresponding evaluation for these two additive lines. Bottom left: whenever a red line is to the left of a black line, then we rescale the value a of the additive line by $c \in \mathbf{k}^*$ of the multiplicative line. Bottom middle: multiplicative lines can merge, resulting in multiplication of their weights. Bottom right: we may have a red vertex on the multiplicative network, resulting in the swapping of co-orientations of the multiplicative line.

5. Diagrammatics of Shannon entropy

Let $\mathbf{k} = \mathbb{R}$ and let \mathbf{k}^* be the set of units in \mathbf{k} . In this section, we discuss a new perspective of entropy using diagrammatics, as introduced in [12].

We decorate the boundary of a network of lines from values in \mathbf{k} . Figure 5.0.1 gives us the properties of black and red (network) lines. Black lines are additive and take values in \mathbf{k} while red wavy lines are multiplicative and take values in \mathbf{k}^* . Whenever two additive (black) lines merge (at the intersections of additive lines), we evaluate at the additive vertex using the two symbols $\langle a, b \rangle$, see Figure 5.0.1 top left. Whenever two additive lines split, we evaluate at the additive vertex the pair of symbols $\langle a, b \rangle$, but with a negative sign: $-\langle a, b \rangle$, see Figure 5.0.1 top middle. We can also have a virtual crossing, where two lines cross but their intersection is said to be virtual, see Figure 5.0.1 top right. In the second row of Figure 5.0.1, we see that red wavy (multiplicative) lines have a conormal direction. In particular, whenever an additive line passes through a multiplicative line of weight $c \in \mathbf{k}^*$, we rescale the additive line from $a \mapsto ca$, see Figure 5.0.1 bottom left. Similar to the additive lines, whenever multiplicative lines merge, we multiply the two weights, Figure 5.0.1 bottom middle. Finally, we introduce a red dot to represent the switching of normal co-orientations, see Figure 5.0.1 bottom right.

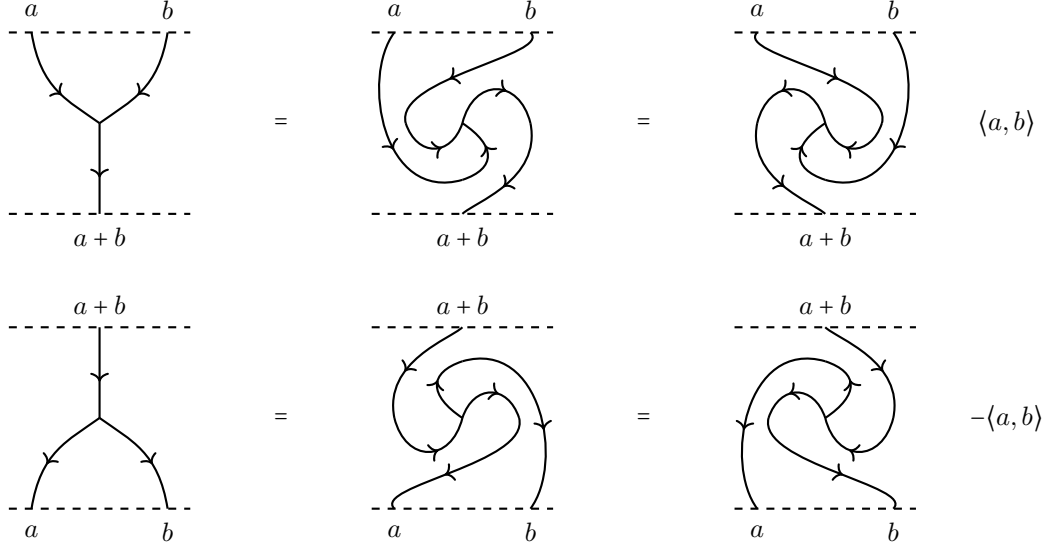


FIGURE 5.0.2. When additive lines are pointing downwards, we rotate the additive vertices in the network whilst fixing the boundary points so that the orientations at the additive vertices are upwards. Top row: the additive vertex in each of the figures gives the contribution of $\langle a, b \rangle$. Bottom row: the additive vertices give $-\langle a, b \rangle$.

When additive lines are oriented downwards, we keep the vertices to stay the same but rotate the additive vertex clockwise or counterclockwise, resulting in a contribution of $\langle a, b \rangle$ or $-\langle a, b \rangle$. See Figure 5.0.2.

Additive and multiplicative lines satisfy natural isotopy relations, as well as the Cathelineau's 2-cocycle relation in (3). The relation $\langle a, b+c \rangle + \langle b, c \rangle = \langle a+b, c \rangle + \langle a, b \rangle$ is satisfied by Figure 5.0.3 top left by summing over the additive vertices. Figure 5.0.3 top right satisfies the symmetry relation $\langle a, b \rangle = \langle b, a \rangle$. Figure 5.0.3 bottom left satisfies the scaling relation $\langle ca, cb \rangle = c\langle a, b \rangle$ since the additive vertex on the left-hand side of the diagram contributes $\langle ca, cb \rangle$ since the additive lines have weights ca and cb while the additive vertex on the right-hand side of the diagram contributes $\langle a, b \rangle$, which we then rescale by the scalar c . For Figure 5.0.3 bottom right, the diagrams show that we can pull apart additive and multiplicative lines that have crossed virtually.

Furthermore, in Figure 5.0.3 top left, if the values a, b, c sum to 1, i.e., $a+b+c = 1$, then we get Shannon entropy if all the lines in a diagram merge and we sum over all the additive vertex contributions. That is, consider Figure 5.0.4. Although Proposition 5.1 holds for n strands merging into one, we first prove the case for three additive lines.

PROPOSITION 5.1. *Each side in the equality in Figure 5.0.4 evaluates to Shannon entropy in (3.1).*

PROOF. Recall from [12] that a relation between diagrammatics and scaled entropy is (3.6). The additive vertex in the blue circle gives a contribution of

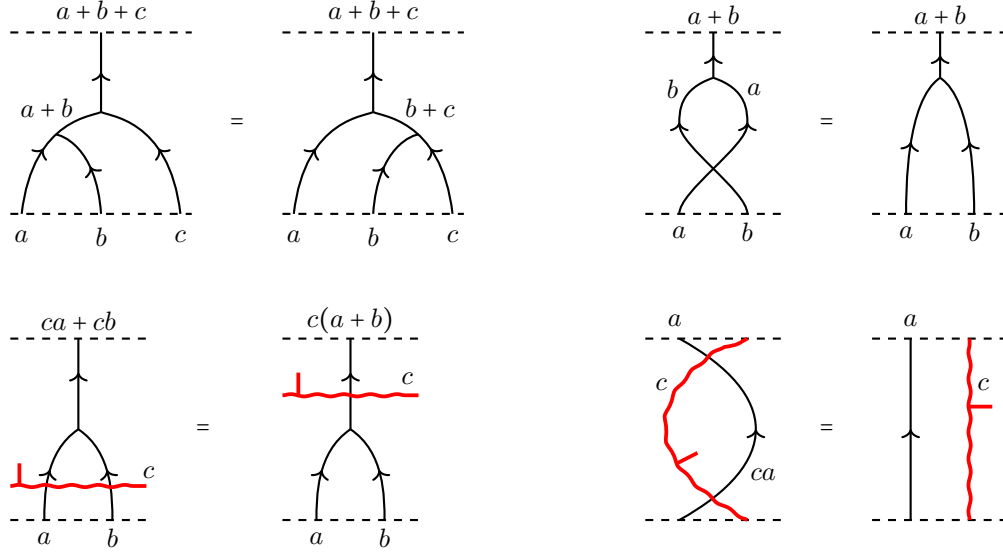


FIGURE 5.0.3. Top left: This diagram represents the 2-cocycle condition (3) in Cathelineau's \mathbf{k} -vector space. Top right: Two lines crossing is virtual, but the additive vertex contributes $\langle b, a \rangle = \langle a, b \rangle$. Bottom left: the additive vertex on the left hand side gives $\langle ac, bc \rangle$ while the additive vertex on the right hand side gives $c\langle a, b \rangle$. This cobordism implies that they are equal. Bottom right: These two isotopies imply that one may pull the red multiplicative line away from black additive line.

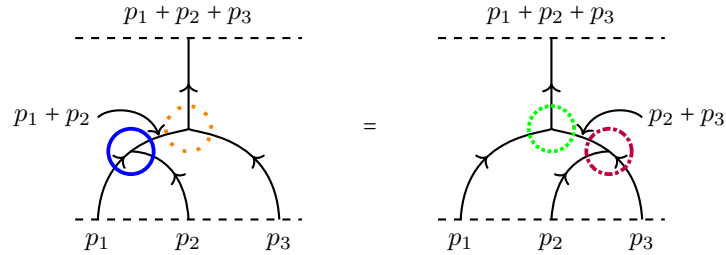


FIGURE 5.0.4. When $\mathbf{k} = \mathbb{R}$ and $p_1 + p_2 + p_3 = 1$, each diagram on the left and the right evaluates to Shannon entropy.

$\langle p_1, p_2 \rangle$. The additive vertex in the dotted orange circle gives a contribution of $\langle p_1 + p_2, p_3 \rangle$. So the sum of the two additive vertices gives $\langle p_1, p_2 \rangle + \langle p_1 + p_2, p_3 \rangle$.

On the other hand, the additive vertex in the purple dashed-dotted circle provides $\langle p_2, p_3 \rangle$. The additive vertex in the green dotted circle contributes $\langle p_1, p_2 + p_3 \rangle$. Combining the two additive vertices on the right hand side, we have $\langle p_2, p_3 \rangle + \langle p_1, p_2 + p_3 \rangle$. We see that

$$\langle p_1, p_2 \rangle + \langle p_1 + p_2, p_3 \rangle = \langle p_2, p_3 \rangle + \langle p_1, p_2 + p_3 \rangle,$$

which is the 2-cocycle condition for Cathelineau's vector space. So by the proof of Lemma 3.3, each side sums to $H(p_X)$. \square

Now consider Figure 5.0.5.

PROPOSITION 5.2. *Figure 5.0.5 corresponds to entropy $H(p_X) = -p_1 \log |p_1| - p_2 \log |p_2| - p_3 \log |p_3| - p_4 \log |p_4|$.*

PROOF. Here, we will compute only the left hand side in Figure 5.0.5, where $-p_1 \log p_1 - p_2 \log p_2 - p_3 \log(p_3) - p_4 \log p_4 = H(p_X)$. We will leave the other diagrams as an exercise for the reader. From bottom to top, the additive vertices contribute

$$\begin{aligned}\langle p_1, p_2 \rangle &= (p_1 + p_2) H\left(\frac{p_1}{p_1 + p_2}\right) \\ \langle p_1 + p_2, p_3 \rangle &= (p_1 + p_2 + p_3) H\left(\frac{p_1 + p_2}{p_1 + p_2 + p_3}\right) \\ \langle p_1 + p_2 + p_3, p_4 \rangle &= (p_1 + p_2 + p_3 + p_4) H\left(\frac{p_1 + p_2 + p_3}{p_1 + p_2 + p_3 + p_4}\right),\end{aligned}$$

respectively. Combining the three equations above gives

$$\begin{aligned}& \langle p_1, p_2 \rangle + \langle p_1 + p_2, p_3 \rangle + \langle p_1 + p_2 + p_3, p_4 \rangle \\&= (p_1 + p_2) H\left(\frac{p_1}{p_1 + p_2}\right) + (p_1 + p_2 + p_3) H\left(\frac{p_1 + p_2}{p_1 + p_2 + p_3}\right) \\&\quad + (p_1 + p_2 + p_3 + p_4) H\left(\frac{p_1 + p_2 + p_3}{p_1 + p_2 + p_3 + p_4}\right) \\&= (p_1 + p_2) \left(-\frac{p_1}{p_1 + p_2} \log \left| \frac{p_1}{p_1 + p_2} \right| - \left(1 - \frac{p_1}{p_1 + p_2} \right) \log \left| 1 - \frac{p_1}{p_1 + p_2} \right| \right) \\&\quad + (p_1 + p_2 + p_3) \left(-\frac{p_1 + p_2}{p_1 + p_2 + p_3} \log \left| \frac{p_1 + p_2}{p_1 + p_2 + p_3} \right| \right. \\&\quad \left. - \left(1 - \frac{p_1 + p_2}{p_1 + p_2 + p_3} \right) \log \left| 1 - \frac{p_1 + p_2}{p_1 + p_2 + p_3} \right| \right) \\&\quad + (p_1 + p_2 + p_3 + p_4) \left(-\frac{p_1 + p_2 + p_3}{p_1 + p_2 + p_3 + p_4} \log \left| \frac{p_1 + p_2 + p_3}{p_1 + p_2 + p_3 + p_4} \right| \right. \\&\quad \left. - \left(1 - \frac{p_1 + p_2 + p_3}{p_1 + p_2 + p_3 + p_4} \right) \log \left| 1 - \frac{p_1 + p_2 + p_3}{p_1 + p_2 + p_3 + p_4} \right| \right) \\&= \cancel{(p_1 + p_2)} \left(-\frac{p_1}{\cancel{p_1 + p_2}} \log \left| \frac{p_1}{p_1 + p_2} \right| - \frac{p_2}{\cancel{p_1 + p_2}} \log \left| \frac{p_2}{p_1 + p_2} \right| \right) \\&\quad + \cancel{(p_1 + p_2 + p_3)} \left(-\frac{p_1 + p_2}{\cancel{p_1 + p_2 + p_3}} \log \left| \frac{p_1 + p_2}{p_1 + p_2 + p_3} \right| \right. \\&\quad \left. - \frac{p_3}{\cancel{p_1 + p_2 + p_3}} \log \left| \frac{p_3}{p_1 + p_2 + p_3} \right| \right) \\&\quad + \cancel{(p_1 + p_2 + p_3 + p_4)} \left(-\frac{p_1 + p_2 + p_3}{\cancel{p_1 + p_2 + p_3 + p_4}} \log \left| \frac{p_1 + p_2 + p_3}{p_1 + p_2 + p_3 + p_4} \right| \right. \\&\quad \left. - \frac{p_4}{\cancel{p_1 + p_2 + p_3 + p_4}} \log \left| \frac{p_4}{p_1 + p_2 + p_3 + p_4} \right| \right) \\&= -p_1 \log \left| \frac{p_1}{p_1 + p_2} \right| - p_2 \log \left| \frac{p_2}{p_1 + p_2} \right| - (p_1 + p_2) \log \left| \frac{p_1 + p_2}{p_1 + p_2 + p_3} \right| - p_3 \log \left| \frac{p_3}{p_1 + p_2 + p_3} \right| \\&\quad - (p_1 + p_2 + p_3) \log \left| \frac{p_1 + p_2 + p_3}{p_1 + p_2 + p_3 + p_4} \right| - p_4 \log \left| \frac{p_4}{p_1 + p_2 + p_3 + p_4} \right|\end{aligned}$$

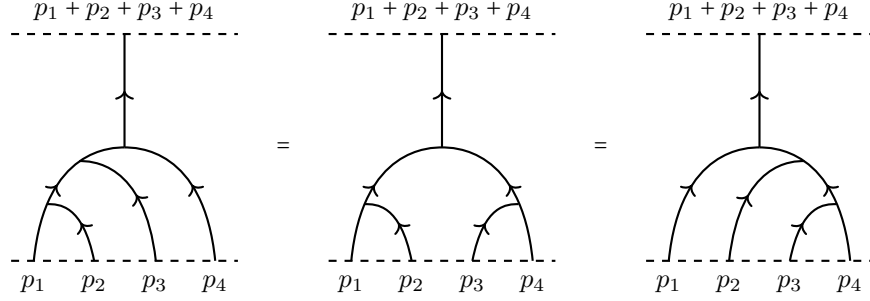


FIGURE 5.0.5. All possible ways of network of additive lines merging into one additive line.

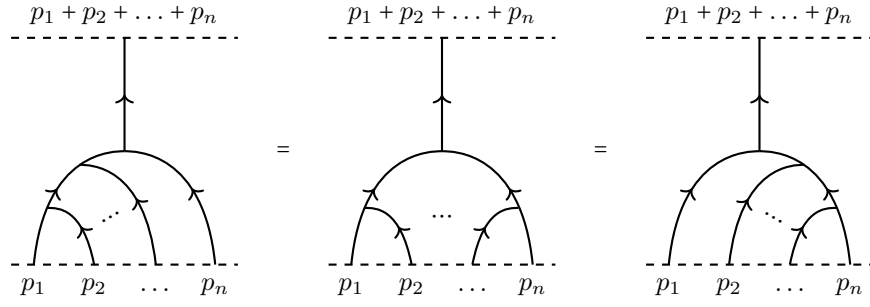


FIGURE 5.0.6. Various diagrams depicting entropy.

$$\begin{aligned}
 &= -p_1 \log |p_1| - p_2 \log |p_2| + \cancel{(p_1 + p_2) \log |p_1 + p_2|} - \cancel{(p_1 + p_2) \log |p_1 + p_2|} - p_3 \log |p_3| \\
 &\quad + \cancel{(p_1 + p_2 + p_3) \log |p_1 + p_2 + p_3|} - \cancel{(p_1 + p_2 + p_3) \log |p_1 + p_2 + p_3|} \\
 &\quad - p_4 \log |p_4| + \underbrace{(p_1 + p_2 + p_3 + p_4) \log |p_1 + p_2 + p_3 + p_4|}_1 \\
 &= -p_1 \log |p_1| - p_2 \log |p_2| - p_3 \log |p_3| - p_4 \log |p_4| = H(p_X).
 \end{aligned}$$

This concludes the proof. \square

PROPOSITION 5.3. *For each diagram in Figure 5.0.6, the additive vertices sum to Shannon entropy.*

The proof of Proposition 5.3 is analogous to Propositions 5.1 and 5.2.

Now, recall equation (3.6) and Cathelineau's symmetry relation (1). They give us the following correspondence:

$$(5.1) \quad \langle a, b \rangle = \langle b, a \rangle \Leftrightarrow (a + b)H\left(\frac{a}{a + b}\right) = (a + b)H\left(\frac{b}{a + b}\right).$$

In particular, in (5.1), take $a = p$ and $b = 1 - p$ to obtain

$$\langle p, 1 - p \rangle = \langle 1 - p, p \rangle \Leftrightarrow H(p) = H(1 - p).$$

We may also introduce the 0-line. See Figures 5.0.7, 5.0.8, and 5.0.9.

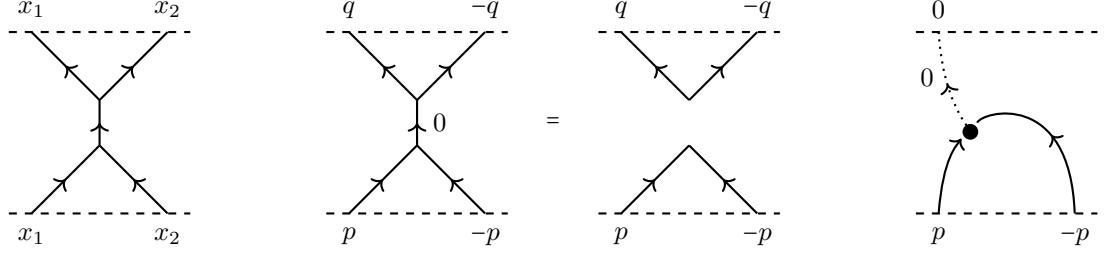


FIGURE 5.0.7. Left: the evaluation of this cobordism is $\langle x_1, x_2 \rangle - \langle x_1, x_2 \rangle = 0$. So the entropy of this system is 0. Middle: 0-lines can be erased. Right: the 0-line can be extended from a $\langle p, -p \rangle$ -line.

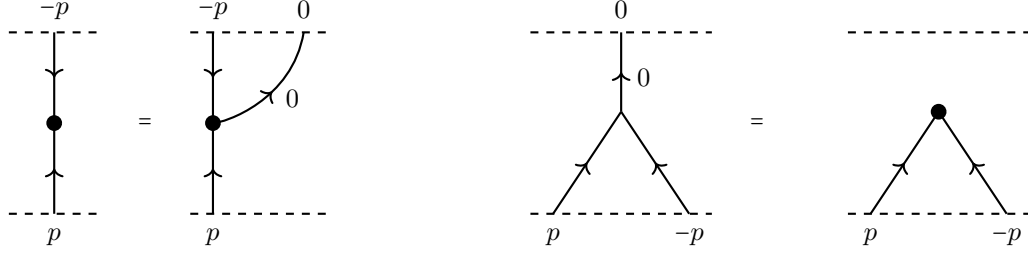


FIGURE 5.0.8. We introduce a dot to represent reversal of orientation. Left: we can erase the 0-line. The 0-line can go in or out, to the top or bottom boundary. Right: we have $\langle p, -p \rangle = 0$.

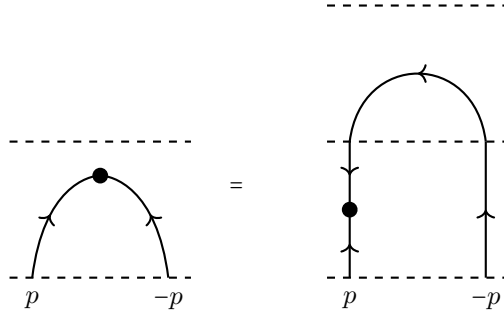


FIGURE 5.0.9. We introduce a dot to represent reversal of orientation. Left: the 0-line has been erased. Right: we have $\langle p, -p \rangle = 0$.

We will now introduce black defects (dots) on additive networks, where whenever we insert a black dot on an additive network, we need to insert another defect on the additive network and reverse orientations on the portion of the additive network connecting these defects. The contribution from Figure 5.0.10 left is $\langle p - q, 1 - p \rangle = (1 - q)H\left(\frac{p-q}{1-q}\right)$. The contribution in Figure 5.0.10 center is $\langle q - 1, p - q \rangle + \langle p - 1, 1 - p \rangle - \langle q - 1, 1 - q \rangle = (p - 1)H\left(\frac{q-1}{p-1}\right)$. Finally, the contribution in Figure 5.0.10 right is $\langle p - q, q - 1 \rangle + \langle 1 - p, p - 1 \rangle - \langle 1 - q, q - 1 \rangle = (p - 1)H\left(\frac{p-q}{p-1}\right)$.

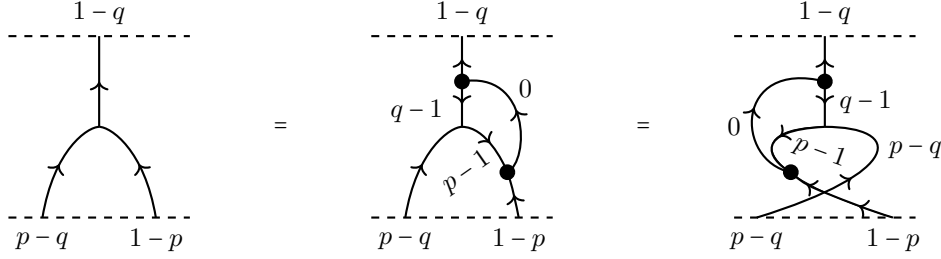


FIGURE 5.0.10. A local symmetry transformation around a trivalent vertex, i.e., change cyclic order at vertex and 2 orientation reversals at a vertex and insert the 0-line in the middle and right figures.

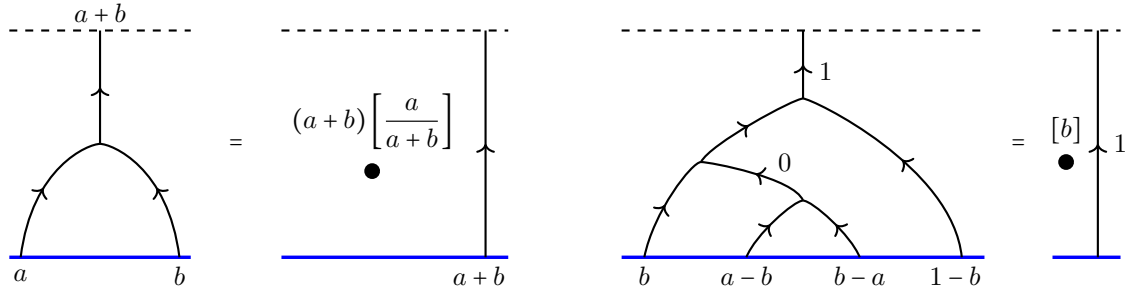


FIGURE 6.0.1. Left: the blue boundary wall absorbs the additive defect on the a -line and b -line, emitting a floating point labeled $(a+b)\left[\frac{a}{a+b}\right]$. Right: The three additive vertices are absorbed, emitting one vertex with the label $[b]$.

They show that entropy also satisfies

$$(1-q)H\left(\frac{p-q}{1-q}\right) = (p-1)H\left(\frac{q-1}{p-1}\right) = (p-1)H\left(\frac{p-q}{p-1}\right).$$

6. Boundary wall

In this section, we introduce a boundary wall, depicted using a blue line at the base of a cobordism. Also see [12, Section 2]. The blue horizontal line in a cobordism is called a boundary wall. It absorbs additive vertices and inserts a labeled floating dot to the left of the network. That is, if two lines weighted a and b merge at an additive vertex, the vertex gets absorbed by the boundary wall and the boundary wall emits a floating defect with the label $(a+b)\left[\frac{a}{a+b}\right]$. See Figure 6.0.1 left. We can also absorb an additive vertex that is connected to the 0-edge (this will give a contribution of 0 so we can draw a vertex with weight 0 or we do not need to draw the 0-vertex). See Figure 6.0.1 right.

In Figure 6.0.2, although the network initially appears complicated, after all additive vertices have been absorbed by the boundary wall, we are left with one line weighted 1 and one vertex with appropriate weight.

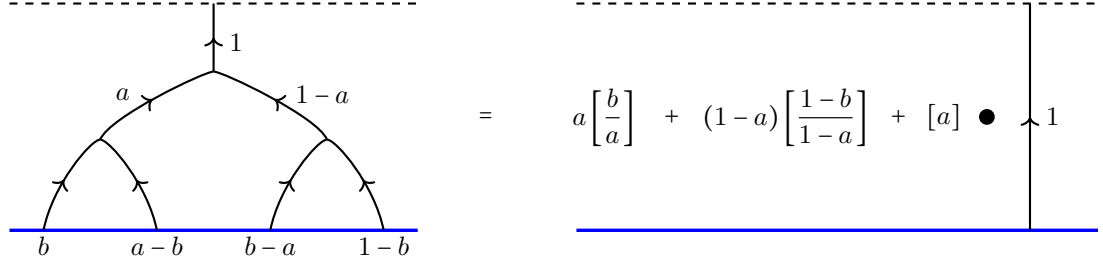


FIGURE 6.0.2. The three additive vertices are absorbed by the blue boundary wall, emitting three floating vertices with the weights $a \left[\frac{b}{a} \right]$, $(1-a) \left[\frac{1-b}{1-a} \right]$, and $[a]$. Since the floating vertices are additive, we can replace three floating vertices with one floating vertex with the label as the sum $a \left[\frac{b}{a} \right] + (1-a) \left[\frac{1-b}{1-a} \right] + [a]$.

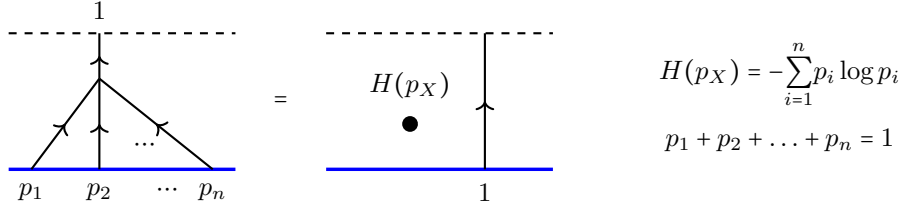


FIGURE 6.0.3. If $p_1 + p_2 + \dots + p_n = 1$, the blue boundary wall absorbs all additive vertices, giving us entropy.

In Figure 6.0.3, we can separate the one additive vertex by using appropriate isotopies, resulting in one floating defect with the weight given by the right-hand side of (3.8), which is Shannon entropy $H(p_X)$.

In Figure 6.0.4, we have a more global picture of our cobordisms. That is, inside the purple dashed ellipse, we have classical physics, involving entropy and information theory. In the larger blue dotted ellipse, we have the theory of probability. When networks turn around, pointing downwards, and when we can rescale the additive lines, we have a larger depiction in the theory of information theory. We are yet to interpret the additive and multiplicative networks outside the blue ellipse, but we believe these observables may have an important interpretation in (quantum) physics.

7. Connections from diagrammatics to $J(\mathbf{k})$

Now, we may want to ask ourselves how do we go from the diagrammatics in Section 5 to elements in the vector space $J(\mathbf{k})$. Given a cobordism (diagram) γ , define $j : \text{Diagrams} \rightarrow J(\mathbf{k})$, where

$$(7.1) \quad j(\gamma) = \sum_{p \in \text{add}(\gamma)} s(p) \omega(p, \gamma) \langle a_p, b_p \rangle \in J(\mathbf{k}),$$

where we sum over all the additive vertices in γ . The coefficient $s(p)$ equals 1 if the vertex p is a merge, and -1 if the additive vertex p is a split. We denote a_p and

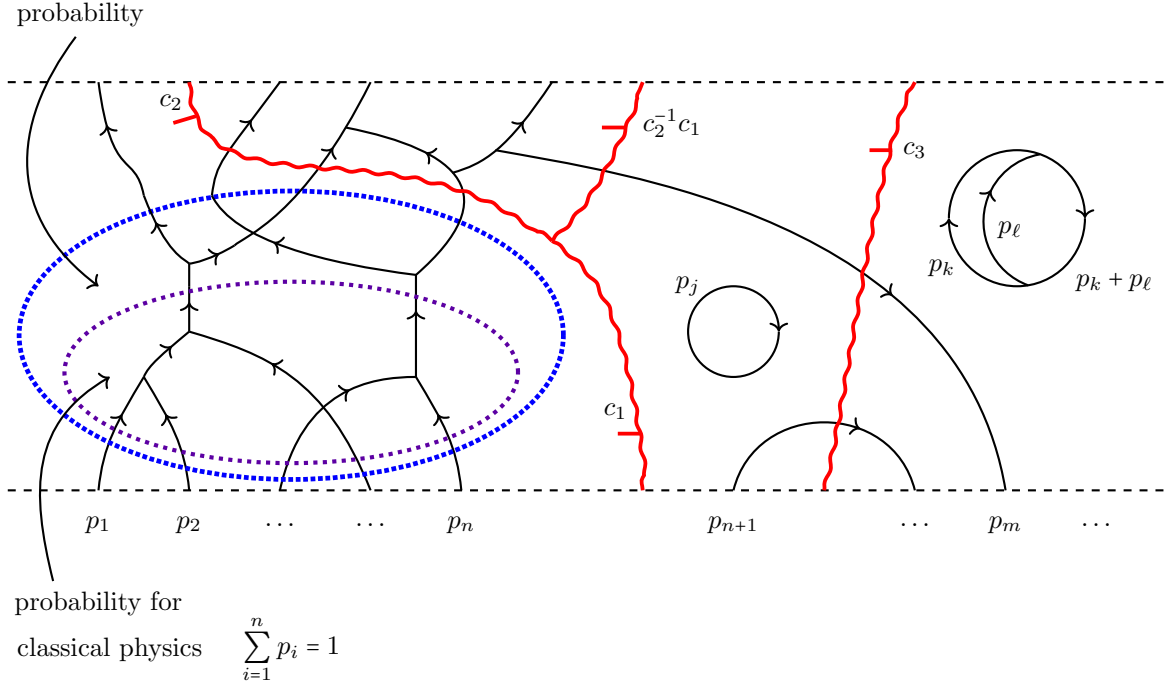


FIGURE 6.0.4. The diagram shows a morphism (cobordism). Small purple ellipse encloses a part of the morphism that admits an interpretation in earlier categorical approaches to entropy. Larger blue ellipse encloses a part of the diagram that can be interpreted in classical probability. Parts of the diagram that include wavy (red) multiplicative lines, U-turns of additive lines and additive lines pointing down or closing upon themselves make sense in a larger category, but not in the familiar classical probability, since rescaling is involved and one could remove the renormalization to 1. Multiplicative lines scale labels of additive lines.

b_p the two lines weighted a and b , respectively, at point p . The coefficient $\omega(p, \gamma)$ is the product of all the multiplicative red lines as we virtually move the additive vertex p far to the left as possible, crossing over any multiplicative lines and taking into account the conormal direction of these lines.

The value $j(\gamma)$ encodes contribution from additive vertices. It computes 2-cocycles and it is an invariant under isotopies. So to a diagrammatic morphism γ , assign element $j(\gamma)$ of $J(\mathbf{k})$. Generating objects of this monoidal category are labelled by $a \in \mathbf{k}$ (endpoints of additive networks) and $c \in \mathbf{k}^*$ (endpoints of multiplicative networks).

From [12, Section 5], we have the following:

PROPOSITION 7.1 (Im–Khovanov). *The invariant $j(\gamma)$ depends only on the source and target objects of the morphism γ .*

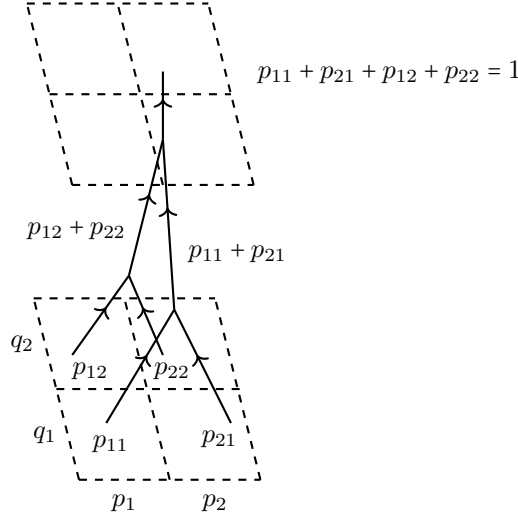


FIGURE 8.0.1. Joint entropy for two random variables X and Y with probabilities p_1 and p_2 and q_1 and q_2 , respectively.

Because $j(\gamma)$ only depends on the source and the target objects of morphism γ , and not on any of the interior network, j is an invariant. We leave it as future work to explore other invariants of these cobordisms.

8. Diagrammatics of conditional entropy

Let $H(Y|X)$ be the conditional entropy of a discrete random variable Y given the random variable X . In terms of probability, conditional entropy is defined to be

$$(8.1) \quad H(X|Y) = - \sum_{i=1}^n \sum_{j=1}^m P(X_i, Y_j) \log P(X_i|Y_j),$$

where

$$(8.2) \quad H(X, Y) = H(Y|X) + H(X).$$

This is reminiscent of an additive version of the ordinary chain rule.

Diagram associated to Proposition 4.2 is Figure 8.0.1. In order to understand conditional entropy in terms of joint entropy (8.2) using diagrammatics, we provide the diagrammatics for

$$(8.3) \quad H(Y|X) = H(X, Y) - H(X).$$

In Figure 8.0.2, we give an example of how conditional entropy of two random variables can be interpreted. In Figure 8.0.3, we provide two different ways to obtain entropy of two random variables: $H(X, Y) = H(X|Y) + H(Y)$ and $H(X, Y) = H(Y|X) + H(X)$. In Figure 8.0.4, the morphism simplifies to the identity cobordism.

We see that one can draw diagrammatics for the information content, i.e., self-information, which is $I(X) = \log \frac{1}{p_X}$. In terms of entropy, it can be reinterpreted as $H(p_X) = E[I(X)]$. We leave it as an exercise to draw the diagrams for mutual

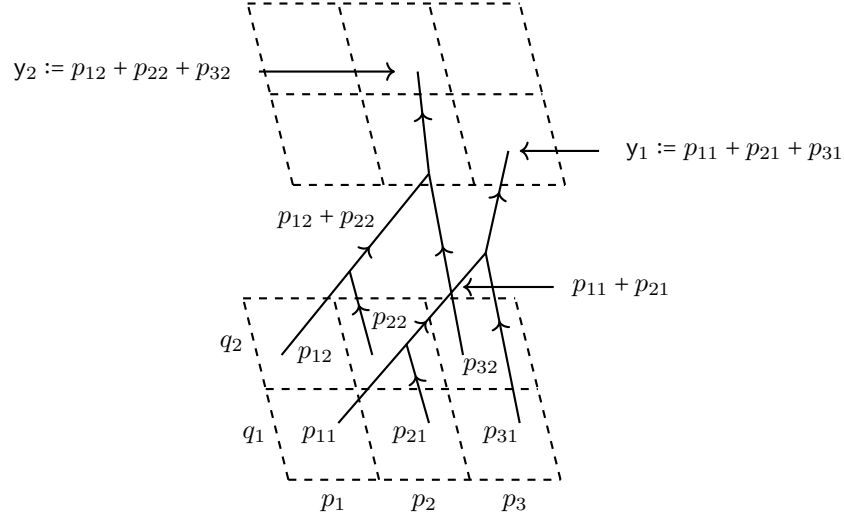


FIGURE 8.0.2. A diagrammatic for the conditional entropy $H(X|Y) = H(X, Y) - H(Y)$. That is, we group along Y 's, but we do not group the Y 's themselves.

information:

$$\begin{aligned}
 I(X, Y) &= H(X) - H(X|Y) = H(Y) - H(Y|X) \\
 (8.4) \quad &= H(X) + H(Y) - H(X, Y) \\
 &= H(X, Y) - H(X|Y) - H(Y|X).
 \end{aligned}$$

9. Additional properties of $J(\mathbf{k})$

From Im–Khovanov in [12], we saw that the two symbols $\langle \cdot, \cdot \rangle$ in Cathelineau's \mathbf{k} -vector space is related to $H(p_X)$ via the equation

$$\langle a, b \rangle = (a + b)H\left(\frac{a}{a + b}\right)$$

if $a + b \neq 0$; otherwise, $\langle a, b \rangle = 0$. This shows that rescaling the two symbols by $a + b$ gives us a rescaled entropy:

$$(9.1) \quad \frac{1}{a + b} \langle a, b \rangle = H\left(\frac{a}{a + b}\right) = H\left(\frac{b}{a + b}\right).$$

In particular, when $a = b$, then

$$(9.2) \quad \langle a, a \rangle = 2aH\left(\frac{1}{2}\right),$$

or equivalently,

$$\frac{1}{2a} \langle a, a \rangle = H\left(\frac{1}{2}\right) \quad \text{or} \quad \frac{1}{2} \langle 1, 1 \rangle = H\left(\frac{1}{2}\right).$$

Although this is an interesting observation, length or norm does not make sense since the formal pair $\langle \cdot, \cdot \rangle$ of symbols is not linear. Perhaps there may be other interesting properties about Cathelineau's vector space that could easily be observed from the diagrammatics.

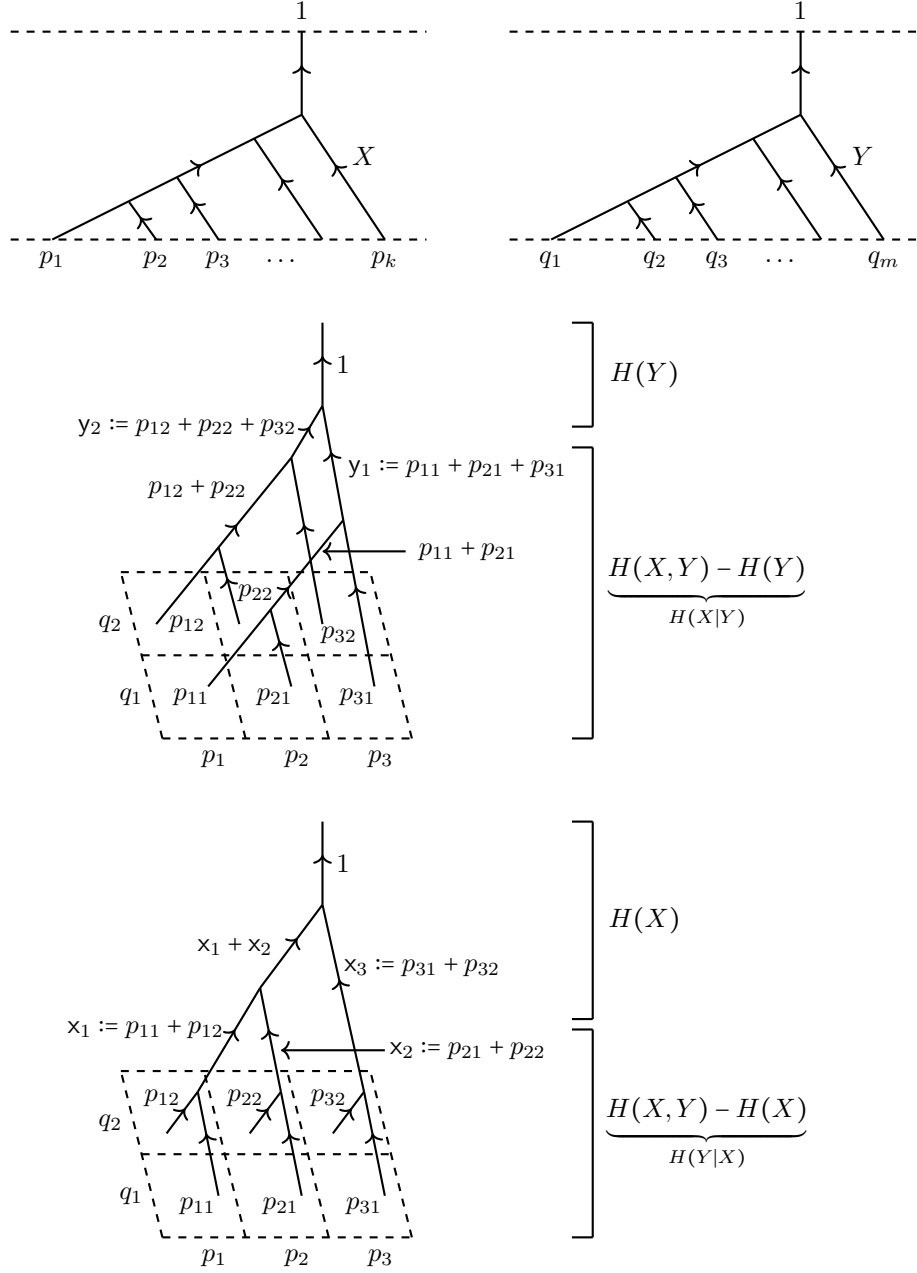


FIGURE 8.0.3. Let $\{p_1, p_2, p_3\}$ be probabilities for the random variable X and $\{q_1, q_2, q_3\}$ be probabilities for the random variable Y . Let $p_{ij} := (p_i, q_j)$. Top left: random variable X with probabilities $\{p_1, p_2, \dots, p_k\}$. Top right: random variable Y with probabilities $\{q_1, q_2, \dots, q_m\}$. Middle: absorbing all the additive vertices gives us $H(X, Y)$. Absorbing all the additive vertices except the top one gives $H(X, Y) - H(Y)$. Bottom: absorbing all the additive vertices except the top one gives $H(X, Y) - H(X)$.

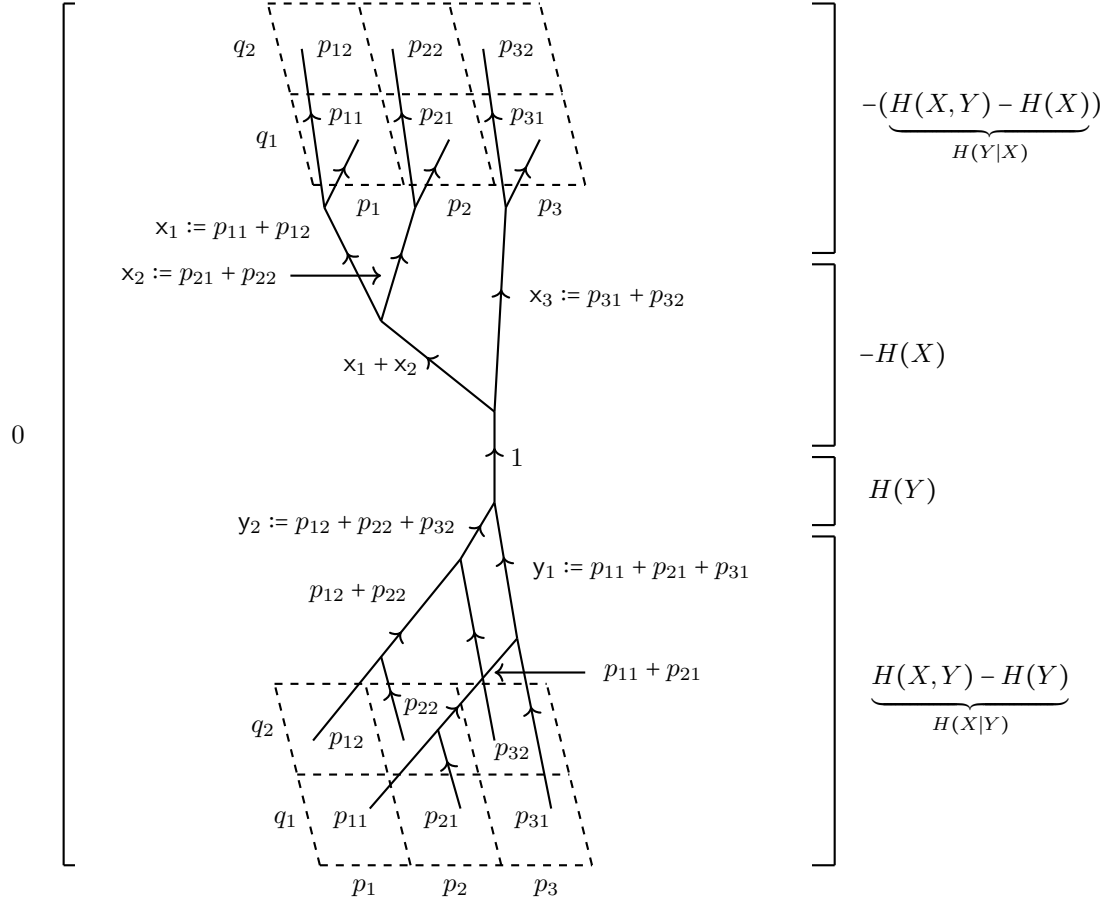


FIGURE 8.0.4. Let $\{p_1, p_2, p_3\}$ be probabilities for the random variable X and $\{q_1, q_2, q_3\}$ be probabilities for the random variable Y . Let $p_{ij} := (p_i, q_j)$. Above is a diagrammatic of 0 since the top and the bottom collections of boundary points in it are identical.

10. Deforming 5-term dilogarithm into 4-term infinitesimal dilogarithm

We refer to [2, Section 1.3] where Cathelineau derives the dilogarithm to infinitesimal dilogarithm. In this section, we give a completely written out derivation of this deformation.

Let A be a unital commutative ring. Let P_A be an abelian group generated by $\{z\}$, where z and $1 - z$ are invertible in A , with relations

- (1) $\{z\} + \left\{\frac{1}{z}\right\} = 0$,
- (2) $\left\{\frac{1}{z}\right\} - \{1 - z\} = 0$,
- (3) $\{z_1\} - \{z_2\} + \left\{\frac{z_2}{z_1}\right\} - \left\{\frac{1-z_2}{1-z_1}\right\} + \left\{\frac{(1-z_2)z_1}{(1-z_1)z_2}\right\} = 0$,

where the last relation (3) is called the 5-term dilogarithm equation. A homomorphism of unital rings $A \rightarrow A'$ induces a homomorphism of abelian groups $P_A \rightarrow P_{A'}$.

We now specialize to the case when $A = \mathbf{k}[\varepsilon]$, where $\varepsilon^2 = 0$, and $A' = \mathbf{k}$. Let $\tilde{TP}(\mathbf{k}) := \ker(P_{\mathbf{k}[\varepsilon]} \rightarrow P_{\mathbf{k}})$, where the homomorphism $\mathbf{k}[\varepsilon] \rightarrow \mathbf{k}$ is identity on \mathbf{k} and $\varepsilon \mapsto 0$. This induces the short exact sequence

$$0 \longrightarrow \tilde{TP}(\mathbf{k}) \xrightarrow{\quad \sigma \quad} P_{\mathbf{k}[\varepsilon]} \twoheadrightarrow P_{\mathbf{k}} \longrightarrow 0$$

with the section σ given by

$$(10.1) \quad \sigma(\{a + b\varepsilon\}) = \{a + b\varepsilon\} - \{a\}, \quad a \neq 0, 1.$$

The kernel $\tilde{TP}(\mathbf{k})$ is also generated by the images of $\{a + b\varepsilon\}$ under σ as in (10.1) over all $a \in \mathbf{k}^* \setminus \{1\}$ and $b \in \mathbf{k}$. We also have \mathbf{k}^* acting on $\tilde{TP}(\mathbf{k})$ via the translation on the coefficient of ε :

$$c \sigma(\{a + b\varepsilon\}) = c(\{a + b\varepsilon\} - \{a\}) = \{a + cb\varepsilon\} - \{a\} = \sigma(\{a + cb\varepsilon\}).$$

Let $\mu : \tilde{TP}(\mathbf{k}) \rightarrow \mathbf{k}^+ \wedge_{\mathbb{Z}} \mathbf{k}^+$ be a morphism of abelian groups, where $\mu(\{a + b\varepsilon\} - \{a\}) = \frac{b}{a} \wedge \frac{b}{1-a}$, or alternatively, $\mu(\{a + 2b\varepsilon\} + \{a\} - 2\{a + b\varepsilon\}) = 2\left(\frac{b}{a} \wedge \frac{b}{1-a}\right) \neq 0$. Let $N_{\mathbf{k}}$ be the subgroup of $\tilde{TP}(\mathbf{k})$ generated by expressions of the form

$$(10.2) \quad \{a + (b + b')\varepsilon\} + \{a\} - \{a + b\varepsilon\} - \{a + b'\varepsilon\}.$$

In $\tilde{TP}(\mathbf{k})$, we can view $\{a + b\varepsilon\} - \{a\}$ as equal to its deformation. That is, let $TP(\mathbf{k}) := \tilde{TP}(\mathbf{k})/N_{\mathbf{k}}$. Then in $TP(\mathbf{k})$, we have:

$$\{a + (b + b')\varepsilon\} - \{a + b'\varepsilon\} = \{a + b\varepsilon\} - \{a\}.$$

PROPOSITION 10.1. *Let \mathbf{k} be a field of characteristic 0. Then there exists an isomorphism of vector spaces*

$$(10.3) \quad \varphi : \beta_{\mathbf{k}} \xrightarrow{\cong} TP(\mathbf{k}), \quad \text{where } \varphi([a]) = \{a + a(1-a)\varepsilon\} - \{a\}.$$

This leads us to a commutative diagram with two short exact sequences:

$$\begin{array}{ccccccc} & & 0 & & & & \\ & & \downarrow & & & & \\ & & N_{\mathbf{k}} & & & & \\ & & \downarrow & & & & \\ 0 & \longrightarrow & \tilde{TP}(\mathbf{k}) & \xrightarrow{\quad \sigma \quad} & P_{\mathbf{k}[\varepsilon]} & \twoheadrightarrow & P_{\mathbf{k}} \longrightarrow 0 \\ & & \downarrow \text{quotient by } N_{\mathbf{k}} & & & & \\ \beta_{\mathbf{k}} & \xrightarrow{\cong \varphi} & TP(\mathbf{k}) & & & & \\ & & \downarrow & & & & \\ & & 0. & & & & \end{array}$$

We will now prove Proposition 10.1.

PROOF. We will prove that φ is well-defined by checking compatibility with the relation (2) in Section 2.2. Let

$$u = a + a(1-a)\varepsilon, \quad v = b + b(1-b)\varepsilon.$$

Then

$$\begin{aligned} \frac{u(1-v)}{v(1-u)} &= \frac{(a + a(1-a)\varepsilon)(1 - (b + b(1-b)\varepsilon))}{(b + b(1-b)\varepsilon)(1 - (a + a(1-a)\varepsilon))} \\ &= \frac{a - ab + a(1-b)(1-a-b)\varepsilon}{b - ab + b(1-a)(1-a-b)\varepsilon} \\ &= \frac{a(1-b)}{b(1-a)} \\ &= \frac{a(1-b)}{b(1-a)}. \end{aligned}$$

Now,

$$\begin{aligned} &\varphi\left([a] - [b] + (1-a)\left[\frac{1-b}{1-a}\right] + a\left[\frac{b}{a}\right]\right) \\ &= \varphi([a]) - \varphi([b]) + (1-a)\varphi\left(\left[\frac{1-b}{1-a}\right]\right) + a\varphi\left(\left[\frac{b}{a}\right]\right) \\ &= \{a + a(1-a)\varepsilon\} - \{b\} - (\{b + b(1-b)\varepsilon\} - \{b\}) \\ &\quad + (1-a)\left(\left\{\frac{1-b}{1-a} + \frac{1-b}{1-a}\left(1 - \frac{1-b}{1-a}\right)\varepsilon\right\} - \left\{\frac{1-b}{1-a}\right\}\right) + a\left(\left\{\frac{b}{a} + \frac{b}{a}\left(1 - \frac{b}{a}\right)\varepsilon\right\} - \left\{\frac{b}{a}\right\}\right) \\ &= \sigma(\{a + a(1-a)\varepsilon\}) - \sigma(\{b + b(1-b)\varepsilon\}) + (1-a)\sigma\left(\left\{\frac{1-b}{1-a} + \frac{1-b}{1-a}\left(1 - \frac{1-b}{1-a}\right)\varepsilon\right\}\right) \\ &\quad + a\sigma\left(\left\{\frac{b}{a} + \frac{b}{a}\left(1 - \frac{b}{a}\right)\varepsilon\right\}\right) \\ &= \sigma(\{u\}) - \sigma(\{v\}) + \sigma\left(\left\{\frac{1-b}{1-a} + (1-b)\left(1 - \frac{1-b}{1-a}\right)\varepsilon\right\}\right) + \sigma\left(\left\{\frac{b}{a} + b\left(1 - \frac{b}{a}\right)\varepsilon\right\}\right) \\ &= \sigma(\{u\}) - \sigma(\{v\}) + \sigma\left(\left\{\frac{1-b}{1-a} - \frac{1-b}{1-a}(a-b)\varepsilon\right\}\right) + \sigma\left(\left\{\frac{b}{a} + \frac{b}{a}(a-b)\varepsilon\right\}\right) \\ &\stackrel{+}{=} \sigma(\{u\}) - \sigma(\{v\}) + \sigma\left(\left\{\frac{b}{a} + \frac{b}{a}(a-b)\varepsilon\right\}\right) - \sigma\left(\left\{\frac{1-b}{1-a} + \frac{1-b}{1-a}(a-b)\varepsilon\right\}\right) \\ &\quad + \left\{\frac{a(1-b)}{b(1-a)}\right\} - \left\{\frac{a(1-b)}{b(1-a)}\right\} \\ &= \sigma(\{u\}) - \sigma(\{v\}) + \sigma\left(\left\{\frac{b}{a} + \frac{b}{a}(a-b)\varepsilon\right\}\right) - \sigma\left(\left\{\frac{1-b}{1-a} + \frac{1-b}{1-a}(a-b)\varepsilon\right\}\right) \\ &\quad + \sigma\left(\left\{\frac{a(1-b)}{b(1-a)}\right\}\right) \\ &= \sigma\left(\{u\} - \{v\} + \left\{\frac{b}{a}(1 + (a-b)\varepsilon)\right\} - \left\{\frac{(1-b) + (1-b)(a-b)\varepsilon}{1-a}\right\} + \left\{\frac{a(1-b)}{b(1-a)}\right\}\right) \\ &= \sigma\left(\{u\} - \{v\} + \left\{\frac{v}{u}\right\} - \left\{\frac{1-v}{1-u}\right\} + \left\{\frac{u(1-v)}{v(1-u)}\right\}\right) \end{aligned}$$

since

$$\begin{aligned}
\frac{v}{u} &= \frac{b + b(1-b)\varepsilon}{a + a(1-a)\varepsilon} = \frac{b + b(1-b)\varepsilon}{a + a(1-a)\varepsilon} \frac{a - a(1-a)\varepsilon}{a - a(1-a)\varepsilon} \\
&= \frac{ab + ab(a-b)\varepsilon}{a^2} = \frac{b}{a} + \frac{b}{a}(a-b)\varepsilon, \\
\frac{1-v}{1-u} &= \frac{1 - (b + b(1-b)\varepsilon)}{1 - (a + a(1-a)\varepsilon)} = \frac{1-b-b(1-b)\varepsilon}{1-a-a(1-a)\varepsilon} \\
&= \frac{1-b-b(1-b)\varepsilon}{1-a-a(1-a)\varepsilon} \frac{1-a+a(1-a)\varepsilon}{1-a+a(1-a)\varepsilon} \\
&= \frac{(1-a)(1-b) + (1-a)(1-b)(a-b)\varepsilon}{(1-a)^2} \\
&= \frac{(1-b) + (1-b)(a-b)\varepsilon}{1-a} \\
&= \frac{1-b}{1-a} + \frac{1-b}{1-a}(a-b)\varepsilon, \\
\frac{u(1-v)}{v(1-u)} &= \frac{(a + a(1-a)\varepsilon)(1 - (b + b(1-b)\varepsilon))}{(b + b(1-b)\varepsilon)(1 - (a + a(1-a)\varepsilon))} \\
&= \frac{a(1-b)(1 + (1 - (a+b))\varepsilon)}{b(1-a)(1 + (1 - (a+b))\varepsilon)} \\
&= \frac{a(1-b)}{b(1-a)},
\end{aligned}$$

and for $b' = -b$ in (10.2), we have

$$\{a + (b-b)\varepsilon\} + \{a\} - \{a + b\varepsilon\} - \{a - b\varepsilon\} = \{a\} + \{a\} - \{a + b\varepsilon\} - \{a - b\varepsilon\}.$$

So in $TP(\mathbf{k})$,

$$\begin{aligned}
0 &= \sigma(\{a\} + \{a\} - \{a + b\varepsilon\} - \{a - b\varepsilon\}) \\
&= 2\sigma(\{a\}) - \sigma(\{a + b\varepsilon\}) - \sigma(\{a - b\varepsilon\}) \\
&= 2(\{a\} - \{a\}) - \sigma(\{a + b\varepsilon\}) - \sigma(\{a - b\varepsilon\}) \\
&= 0 - \sigma(\{a + b\varepsilon\}) - \sigma(\{a - b\varepsilon\}).
\end{aligned}$$

The equality \dagger holds in $TP(\mathbf{k})$ since

$$-\sigma(\{a + b\varepsilon\}) = \sigma(\{a - b\varepsilon\}).$$

Now to prove that φ is injective, let

$$(10.4) \quad \rho: TP(\mathbf{k}) \longrightarrow \mathbf{k}^* \otimes \mathbf{k}^+, \quad \rho(\{a + b\varepsilon\} - \{a\}) = a \otimes \frac{b}{1-a} + (1-a) \otimes \frac{b}{a},$$

be a map of vector spaces over \mathbf{k} . Then we have a commutative diagram

$$\begin{array}{ccc}
\beta_{\mathbf{k}} & \xrightarrow{D} & \mathbf{k}^* \otimes \mathbf{k}^+, \\
& \searrow \varphi & \nearrow \rho \\
& TP(\mathbf{k}) &
\end{array}$$

where D and ρ are injective. So φ is also injective. \square

REMARK 10.2. Note that in general, ρ may not be surjective, but D and ρ are injective imply that $TP(\mathbf{k})$ is not too small.

Second proof. We give a second argument that the 5-term dilogarithm deforms to the infinitesimal 4-term dilogarithm. Let \mathbf{k} be a field, and let $\mathbf{k}^b = \mathbf{k}^* \setminus \{1\} = \mathbf{k} \setminus \{0, 1\}$.

Let $\beta_2(\mathbf{k})$ be the quotient of $\mathbb{Q}[\mathbf{k}^b]$ by

$$[a] - [b] + \left[\frac{b}{a} \right] + \left[\frac{1-b^{-1}}{1-a^{-1}} \right] - \left[\frac{1-b}{1-a} \right] = 0,$$

with $(1-a)(1-b)\left(1-\frac{b}{a}\right) \in \mathbf{k}^*$.

Let $\mathbf{k}_2 := \mathbf{k}[t]/(t^2)$ be the ring of dual numbers. If $a \in \mathbf{k}^b$, define

$$\langle a \rangle := a + a(1-a)t = a(1 + (1-a)t) \in \mathbf{k}_2.$$

Note that $(1+ct)^{-1} = 1-ct$ in \mathbf{k}_2 since $t^2 = 0$. We see that this element $\langle a \rangle$ is invertible in \mathbf{k}_2 , with $\langle a \rangle^{-1} = a^{-1}(1 - (1-a)t)$.

Define the action of \mathbf{k}^* on the ring \mathbf{k}_2 of dual numbers as $\lambda * (b_0 + b_1 t) = b_0 + \lambda b_1 t$, where $\lambda \in \mathbf{k}^*$. That is, $\lambda * 1 = 1$ and $\lambda * t = \lambda t$. So λ only translates the coefficient of t .

LEMMA 10.3. *We have $\frac{\langle b \rangle}{\langle a \rangle} = a * \left\langle \frac{b}{a} \right\rangle$.*

PROOF. The left hand side shows

$$\frac{\langle b \rangle}{\langle a \rangle} = \frac{b(1 + (1-b)t)}{a(1 + (1-a)t)} = \frac{b}{a} (1 + ((1-b) - (1-a))t) = \frac{b}{a} (1 + (a-b)t).$$

On the other hand, the right hand side shows

$$a * \left\langle \frac{b}{a} \right\rangle = a * \left(\frac{b}{a} \left(1 + \left(1 - \frac{b}{a} \right) t \right) \right) = \frac{b}{a} \left(1 + \left(1 - \frac{b}{a} \right) at \right) = \frac{b}{a} \left(1 + \frac{a-b}{a} at \right) = \frac{b}{a} (1 + (a-b)t).$$

\square

LEMMA 10.4. *The relation*

$$\frac{1 - \langle b \rangle}{1 - \langle a \rangle} = (a - 1) * \left\langle \frac{1 - b}{1 - a} \right\rangle$$

holds.

PROOF. On the left hand side, we have

$$\begin{aligned} \frac{1 - \langle b \rangle}{1 - \langle a \rangle} &= \frac{1 - b(1 + (1-b)t)}{1 - a(1 + (1-a)t)} \\ &= \frac{1 - b - b(1-b)t}{1 - a - a(1-a)t} \cdot \frac{1 - a + a(1-a)t}{1 - a + a(1-a)t} \\ &= \frac{(1-a)(1-b) + (a(1-a)(1-b) - (1-a)b(1-b))t}{(1-a)^2} \\ &= \frac{\cancel{(1-a)}(1-b) + (a-b)\cancel{(1-a)}(1-b)t}{(1-a)^2} \\ &= \frac{1-b}{1-a} (1 + (a-b)t) \end{aligned}$$

while on the right hand side, we have

$$(a-1) * \left\langle \frac{1-b}{1-a} \right\rangle = (a-1) * \frac{1-b}{1-a} \left(1 + \left(1 - \frac{1-b}{1-a} \right) t \right) = \frac{1-b}{1-a} \left(1 + \left(1 - \frac{1-b}{1-a} \right) (a-1)t \right),$$

which are indeed equal. \square

LEMMA 10.5. *We have $1 - \langle a \rangle^{-1} = (1 - a^{-1})(1 - t)$.*

PROOF. Since

$$\begin{aligned} 1 - \langle a \rangle^{-1} &= 1 - a^{-1}(1 + (1-a)t)^{-1} = 1 - a^{-1}(1 - (1-a)t) \\ &= 1 - a^{-1} + a^{-1}(1-a)t = 1 - a^{-1} + (a^{-1} - 1)t \\ &= (1 - a^{-1})(1 - t), \end{aligned}$$

the lemma holds. \square

LEMMA 10.6. *The equation*

$$\frac{1 - \langle b \rangle^{-1}}{1 - \langle a \rangle^{-1}} = \frac{1 - b^{-1}}{1 - a^{-1}}$$

holds.

PROOF. The lemma is immediate since

$$\frac{1 - \langle b \rangle^{-1}}{1 - \langle a \rangle^{-1}} = \frac{(1 - b^{-1})(\cancel{1-t})}{(1 - a^{-1})(\cancel{1-t})} = \frac{1 - b^{-1}}{1 - a^{-1}}.$$

\square

LEMMA 10.7. *We have $[b * \langle c \rangle] = b[\langle c \rangle]$ where $b \in \mathbf{k}$ and $c \in \mathbf{k}_2$ and $(-1)[\langle 1-a \rangle] = -[\langle a \rangle]$ and $a[\langle a^{-1} \rangle] = -[\langle a \rangle]$.*

PROOF. This is clear using the construction of $\beta(\mathbf{k})$. \square

LEMMA 10.8. *The 5-term dilogarithm*

$$(10.5) \quad [\langle a \rangle] - [\langle b \rangle] + \left[\frac{\langle b \rangle}{\langle a \rangle} \right] + \left[\frac{1 - \langle b \rangle^{-1}}{1 - \langle a \rangle^{-1}} \right] - \left[\frac{1 - \langle b \rangle}{1 - \langle a \rangle} \right] = 0$$

holds in the commutative ring $\beta_2(\mathbf{k}_2)$.

PROOF. This is clear by the definition of the ring $\beta_2(\mathbf{k}_2)$. \square

PROPOSITION 10.9. *Taking the limit as $t \rightarrow 0$, the 5-term dilogarithm in (10.5) deforms to the 4-term infinitesimal dilogarithm*

$$(10.6) \quad [a] - [b] + a \left[\frac{b}{a} \right] + (1-a) \left[\frac{1-b}{1-a} \right] = 0, \quad a \in \mathbf{k} \setminus \{0, 1\}, \quad b \in \mathbf{k}^*.$$

PROOF. We have $\left[\frac{\langle b \rangle}{\langle a \rangle} \right] = \left[a * \left\langle \frac{b}{a} \right\rangle \right] = a \left[\left\langle \frac{b}{a} \right\rangle \right]$ by Lemmas 10.3 and 10.7. We also have

$$\left[\frac{1 - \langle b \rangle^{-1}}{1 - \langle a \rangle^{-1}} \right] = \left[\frac{(1 - b^{-1})(\cancel{1-t})}{(1 - a^{-1})(\cancel{1-t})} \right] = \left[\frac{1 - b^{-1}}{1 - a^{-1}} \right]$$

by Lemma 10.6. Finally, we have

$$\left[\frac{1 - \langle b \rangle}{1 - \langle a \rangle} \right] = \left[(a-1) * \left\langle \frac{1-b}{1-a} \right\rangle \right] = (a-1) \left[\left\langle \frac{1-b}{1-a} \right\rangle \right]$$

by Lemmas 10.4 and 10.7. Putting these together, we have

$$[\langle a \rangle] - [\langle b \rangle] + a \left[\left\langle \frac{b}{a} \right\rangle \right] + \left[\frac{1-b^{-1}}{1-a^{-1}} \right] + (1-a) \left[\left\langle \frac{1-b}{1-a} \right\rangle \right] = 0.$$

The element $\left[\frac{1-b^{-1}}{1-a^{-1}} \right]$ is the zero element in

$$\ker(\beta(\mathbf{k}_2) \longrightarrow \beta(\mathbf{k})) / ([a + (b + b')t] + [a] - [a + bt] - [a + b't]).$$

We thus obtain

$$[\langle a \rangle] - [\langle b \rangle] + a \left[\left\langle \frac{b}{a} \right\rangle \right] + (1-a) \left[\left\langle \frac{1-b}{1-a} \right\rangle \right] = 0,$$

which t -deforms to

$$[a] - [b] + a \left[\frac{b}{a} \right] + (1-a) \left[\frac{1-b}{1-a} \right] = 0 \text{ for } a \neq b \text{ and } a \in \mathbf{k}^b.$$

□

11. Future Work

It would be interesting to extend the work of Im–Khovanov of diagrammatic interpretation in [12] to Rényi entropy, which is a generalization of Shannon entropy, von Neumann entropy, which is a quantized analogue of Gibbs entropy, and relative entropy, which is also known as Kullback–Leibler divergence. There may be a deeper interpretation if one were to extend the diagrammatics of entropy to their higher-dimensional analogues.

References

1. Leon Brillouin, *Science and information theory*, Academic Press, Inc., New York, 1956.
2. Jean-Louis Cathelineau, *Sur l'homologie de SL_2 à coefficients dans l'action adjointe*, Math. Scand. **63** (1988), no. 1, 51–86.
3. ———, *Infinitesimal dilogarithms, extensions and cohomology*, J. Algebra **332** (2011), 87–113.
4. David Deutsch and Patrick Hayden, *Information flow in entangled quantum systems*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **456** (2000), no. 1999, 1759–1774.
5. Dmitry K. Faddeev, *On the concept of entropy of a finite probabilistic scheme*, Uspehi Mat. Nauk (N.S.) **67** (1956), no. 1, 227–231.
6. David Gepner, Mee Seong Im, Mikhail Khovanov, and Nitu Kitchloo, *Foams with flat connections and algebraic K-theory*, arXiv preprint [arXiv:2405.14465](https://arxiv.org/abs/2405.14465) (2024), 1–57.
7. Paul Gustafson, Mee Seong Im, Remy Kaldawy, Mikhail Khovanov, and Zachary Lihn, *Automata and one-dimensional TQFTs with defects*, Lett. Math. Phys. **113** (2023), no. 93, 1–38.
8. Paul Gustafson, Mee Seong Im, and Mikhail Khovanov, *Boolean TQFTs with accumulating defects, sofic systems, and automata for infinite words*, Lett. Math. Phys. **114** (2024), no. 6, Paper No. 135.
9. Aleksandr Y. Hinich, *On the basic theorems of information theory*, Uspehi Mat. Nauk (N.S.) **11** (1956), no. 1(67), 17–75.
10. Mee Seong Im and Mikhail Khovanov, *Foams, iterated wreath products, field extensions and Sylvester sums*, arXiv preprint [arXiv:2107.07845](https://arxiv.org/abs/2107.07845) (2021), 1–78.
11. ———, *Topological theories and automata*, arXiv preprint [arXiv:2202.13398](https://arxiv.org/abs/2202.13398) (2022), 1–70.
12. ———, *Entropy, cocycles, and their diagrammatics*, arXiv preprint [arXiv:2409.08462](https://arxiv.org/abs/2409.08462) (2024), 1–82.
13. ———, *Foam cobordism and the Sah–Arnoux–Fathi invariant*, arXiv preprint [arXiv:2403.06030](https://arxiv.org/abs/2403.06030), to appear in *Arnold Math.* (2024), 1–38.

14. ———, *One-dimensional topological theories with defects: the linear case*, Algebraic and topological aspects of representation theory, Contemp. Math., vol. 791, Amer. Math. Soc., Providence, RI, 2024, pp. 105–146.
15. ———, *From finite state automata to tangle cobordisms: a TQFT journey from one to four dimensions*, From representation theory to mathematical physics and back, Contemp. Math., vol. 817, Amer. Math. Soc., Providence, RI, 2025, pp. 179–220.
16. Mee Seong Im, Mikhail Khovanov, and Victor Ostrik, *Universal construction in monoidal and non-monoidal settings, the Brauer envelope, and pseudocharacters*, Theory Appl. Categ. **44** (2025), no. 2, 15–83.
17. Mee Seong Im and Paul Zimmer, *One-dimensional topological theories with defects and linear generating functions*, Involve **15** (2022), no. 2, 319–331.
18. Aleksandr I. Khinchin, *Mathematical foundations of information theory*, Dover Publications, Inc., New York, 1957, Translated by R. A. Silverman and M. D. Friedman.
19. Mikhail Khovanov, *Universal construction of topological theories in two dimensions*, arXiv preprint [arXiv:2007.03361](https://arxiv.org/abs/2007.03361) (2020), 1–56.
20. Maxim Kontsevich, *The $1\frac{1}{2}$ -logarithm. Appendix to: “On poly(ana)logs. I” [Compositio Math **130** (2002), no. 2, 161–210] by P. Elbaz-Vincent and H. Gangl*, Compositio Math. **130** (2002), no. 2, 211–214.
21. Harvey S. Leff and Andrew F. Rex, *Maxwell’s demon and the culture of entropy*, Phys. Essays **10** (1997), no. 1, 125–149.
22. Kumar Vijay Mishra, M. Ashok Kumar, and Ting-Kam Leonard Wong, *Information Geometry for the Working Information Theorist*, arXiv preprint [arXiv:2310.03884](https://arxiv.org/abs/2310.03884) (2023), 1–12.
23. Frank Nielsen, *Geometric Theory of Information*, Springer International Publishing Switzerland, Signals and Communication Technology, 2014.
24. ———, *An elementary introduction to information geometry*, Entropy **22** (2020), no. 10, Paper No. 1100, 61.
25. Frank Nielsen and Frédéric Barbaresco (eds.), *Geometric science of information*, Lecture Notes in Computer Science, vol. 12829, Springer, Cham, [2021] ©2021.
26. Frank Nielsen and Frédéric Barbaresco (eds.), *Geometric science of information. Part I*, Lecture Notes in Computer Science, vol. 14071, Springer, Cham, [2023] ©2023.
27. Frank Nielsen and Frédéric Barbaresco (eds.), *Geometric science of information. Part II*, Lecture Notes in Computer Science, vol. 14072, Springer, Cham, [2023] ©2023.
28. Inder Jeet Taneja, *On the branching property of entropy*, Ann. Polon. Math. **35** (1977/78), no. 1, 67–75.

DEPARTMENT OF MATHEMATICS, JOHNS HOPKINS UNIVERSITY, BALTIMORE, MD 21218, USA (CURRENT)

DEPARTMENT OF MATHEMATICS, UNITED STATES NAVAL ACADEMY, ANNAPOLIS, MD 21402, USA
Email address: meeseong@jhu.edu

U.S. NAVAL RESEARCH LABORATORY, WASHINGTON, DC 20375, USA
Email address: clement.k.kam.civ@us.navy.mil

U.S. NAVAL RESEARCH LABORATORY, WASHINGTON, DC 20375, USA
Email address: caden.j.pici.civ@us.navy.mil