# Categorical Magnitude and Entropy

**Juan Pablo Vigneaux**

Department of Mathematics
California Institute of Technology
Pasadena, CA, USA
vigneaux@caltech.edu

GSI 2023 - Statistics, Information and Topology

Saint-Malo, August 30th, 2023

# Outline

# Magnitude

Let $\mathbf{A}$ be a finite category.

## Definition

The **zeta function** associated with $\mathbf{A}$ is $\zeta : \mathrm{Ob}\,\mathbf{A} \times \mathrm{Ob}\,\mathbf{A} \to \mathbb{Z}$, $(a, b) \mapsto |\mathrm{Hom}(a, b)|$.

## Definition

A **weighting** $k^\bullet : \mathrm{Ob}\,\mathbf{A} \to \mathbb{Q}$ satisfies $\sum_{b \in \mathrm{Ob}\,\mathbf{A}} \zeta(a, b) k^b = 1$.
Similarly, a **coweighting** $k_\bullet : \mathrm{Ob}\,\mathbf{A} \to \mathbb{Q}$ satisfies $\sum_{a \in \mathrm{Ob}\,\mathbf{A}} k_a \zeta(a, b) = 1$.

Leinster [Lei08]: When $\mathbf{A}$ has a weighting $k^\bullet$ and a coweighting $k_\bullet$, it holds that $\sum_{a \in \mathrm{Ob}\,\mathbf{A}} k^a = \sum_{a \in \mathrm{Ob}\,\mathbf{A}} k_a$ and its common value is the **magnitude** $\chi(\mathbf{A})$ of $\mathbf{A}$.
We say in this case that $\mathbf{A}$ has magnitude.

# Origins

The magnitude of posets was already introduced by Rota in the 60s, under the name "Euler characteristic".

In this setting, the function $\zeta$ has an inverse $\mu$ (as a matrix), called the **Möbius function**.

## Theorem (P. Hall)

$$\mu(a,b) = \sum_{k\geq 0} (-1)^k \#\{\textit{nondegenerate paths of length } k \textit{ between } a \textit{ and } b\}.$$

It follows that $\chi(\mathbf{A}) = \sum_{a,b\in \mathrm{Ob}\,\mathbf{A}} \mu(a,b)$ is the topological Euler characteristic of $B\mathbf{A}$, the geometric realization of $\mathbf{A}$'s nerve.

# A more general formula

For a general category, the matrix $\zeta$ is not invertible. However, it has a unique Moore-Penrose pseudoinverse $\zeta^+$, that satisfies the equations

$$\zeta\zeta^+\zeta = \zeta, \quad \zeta^+\zeta\zeta^+ = \zeta^+, \quad (\zeta\zeta^+)^* = \zeta\zeta^+, \quad (\zeta^+\zeta)^* = \zeta^+\zeta.$$

**Theorem (Chen & V. '23; Akkaya & Ünlü '23)**

*Whenever* **A** *has magnitude,*

$$\chi(\mathbf{A}) = \sum_{a,b \in \mathrm{Ob}\,\mathbf{A}} \zeta^+(a,b). \tag{1}$$

In fact, (1) extends the definition of magnitude to all finite categories. Akkaya & Ünlü [AÜ23] proved that it is invariant under equivalence of categories.

Ongoing work: a generalization of Hall's theorem.

# Properties of magnitude

Magnitude—defined via $\zeta^+$—has the following properties (see Chen & V. [CV23]):

1. $\chi(\mathbf{A} \times \mathbf{B}) = \chi(\mathbf{A})\chi(\mathbf{B})$.
   Because $\zeta_{\mathbf{A}\times\mathbf{B}} = \zeta_{\mathbf{A}} \otimes \zeta_{\mathbf{B}}$ (Kronecker product),
   hence $\zeta^+_{\mathbf{A}\times\mathbf{B}} = \zeta^+_{\mathbf{A}} \otimes \zeta^+_{\mathbf{B}}$.

# Properties of magnitude

Magnitude—defined via $\zeta^+$—has the following properties (see Chen & V. [CV23]):

1. $\chi(\mathbf{A} \times \mathbf{B}) = \chi(\mathbf{A})\chi(\mathbf{B})$.
   Because $\zeta_{\mathbf{A}\times\mathbf{B}} = \zeta_{\mathbf{A}} \otimes \zeta_{\mathbf{B}}$ (Kronecker product),
   hence $\zeta^+_{\mathbf{A}\times\mathbf{B}} = \zeta^+_{\mathbf{A}} \otimes \zeta^+_{\mathbf{B}}$.
2. $\chi(\mathbf{A} + \mathbf{B}) = \chi(\mathbf{A}) + \chi(\mathbf{B})$.
   Because $\zeta_{\mathbf{A}+\mathbf{B}} = \zeta_{\mathbf{A}} \oplus \zeta_{\mathbf{B}}$ (block diagonal matrix),
   hence $\zeta^+_{\mathbf{A}+\mathbf{B}} = \zeta^+_{\mathbf{A}} \oplus \zeta^+_{\mathbf{B}}$.

Magnitude—defined via $\zeta^+$—has the following properties (see Chen & V. [CV23]):

1. $\chi(\mathbf{A} \times \mathbf{B}) = \chi(\mathbf{A})\chi(\mathbf{B})$.
   Because $\zeta_{\mathbf{A}\times\mathbf{B}} = \zeta_{\mathbf{A}} \otimes \zeta_{\mathbf{B}}$ (Kronecker product),
   hence $\zeta^+_{\mathbf{A}\times\mathbf{B}} = \zeta^+_{\mathbf{A}} \otimes \zeta^+_{\mathbf{B}}$.
2. $\chi(\mathbf{A} + \mathbf{B}) = \chi(\mathbf{A}) + \chi(\mathbf{B})$.
   Because $\zeta_{\mathbf{A}+\mathbf{B}} = \zeta_{\mathbf{A}} \oplus \zeta_{\mathbf{B}}$ (block diagonal matrix),
   hence $\zeta^+_{\mathbf{A}+\mathbf{B}} = \zeta^+_{\mathbf{A}} \oplus \zeta^+_{\mathbf{B}}$.
3. $\chi(\mathbf{A}) = |\operatorname{Ob}\mathbf{A}|$, the cardinality, when **A** is a discrete category.

So magnitude can be seen as a categorical generalization of cardinality.

# Magnitude of enriched categories and metric spaces

In this framework, $\mathrm{Ob}\,\mathbf{A}$ is still finite, but one supposes that for any objects $a$ and $b$, $\mathrm{Hom}(a, b)$ is an object of a symmetric monoidal category $(\mathbf{V}, \otimes, 1)$.

Given a monoid morphism $|\cdot| : (\mathbf{V}/\cong, \otimes, 1) \to (\mathbb{K}, \cdot, 1)$, where $\mathbb{K}$ is a field, one introduces again the $\zeta$ function via $\zeta(a, b) = |\mathrm{Hom}(a, b)|$. One may then define the magnitude via weightings or the pseudoinverse.

**Example**

An ordinary category is an enriched category over sets equipped with the cartesian product. The map $|\cdot|$ is the cardinality.

**Example**

A finite metric space $(X, d)$ can be seen as a category $\mathbf{X}$ enriched over the poset $([0, \infty], \geq)$ equipped with the monoidal structure given by the sum. Then $\mathrm{Ob}\,\mathbf{X} = X$ and $\mathrm{Hom}(x, y) = d(x, y) \in [0, \infty)$. An $\mathbb{R}$-valued map $|\cdot|$ is necessarily of the form $x \mapsto e^{kx}$ for some $k \in \mathbb{R}$.

# Outline

# Entropy

Let $X$ be a finite set, equipped with a probability mass function $p : X \to [0,1]$. Their **entropy** is:

$$H(X,p) = -\sum_{x \in X} p(x) \log p(x). \tag{2}$$

When $u(x) = 1/|X|$ for all $x$ (uniform law), then $H(X,u) = \log |X|$.

Moreover,

$$H(X \times Y, p \otimes q) = H(X,p) + H(Y,q),$$

which gives in particular

$$\log |X \times Y| = \log |X| + \log |Y|,$$

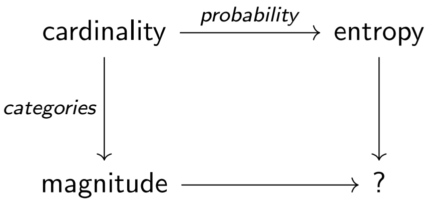
of course equivalent to $|X \times Y| = |X||Y|$.

So it makes sense to regard $H$ as a *probabilistic* generalization of (the log of) cardinality.

# Problem

To find a *categorical* generalization of entropy that might also be regarded as a *probabilistic* generalization of magnitude. In particular, such that:

- It coincides with Shannon entropy for discrete categories.
- Gives $\log \chi(\mathbf{A})$ for a particular choice of its arguments, analogous to the uniform probability.



# Outline

## Definition (Categorical probabilistic triples)

A **categorical probabilistic triple** $(\mathbf{A}, p, \phi)$ consists of

1. a finite category $\mathbf{A}$,
2. a p.m.f. $p$ on $\mathrm{Ob}(\mathbf{A})$, and
3. a function $\phi : \mathrm{Ob}(\mathbf{A}) \times \mathrm{Ob}(\mathbf{A}) \to [0, \infty)$ such that $\phi(a, a) > 0$ for all objects $a$ of $\mathbf{A}$, and $\phi(b, b') = 0$ whenever there is no arrow from $b$ to $b'$ in $\mathbf{A}$.

## Example

1. $\theta = \zeta$, the zeta function.
2. $\theta = \delta$, the Kronecker delta. (Zeta function of a discrete category.)
3. Starting with a metric space $(M, d)$, set $\mathrm{Ob}\,\mathbf{A} = M$ and $\theta(m, m') = e^{-d(m,m')}$.

   This is the zeta function of the metric space when seen as a category enriched over the monoidal category $([0, \infty), +)$.
4. $\theta(a, b) = \mathbb{P}(a|b)$, a probability of transition from $b$ to those $a$ such that $a \to b$ (i.e. under the arrows of $\mathbf{A}^{op}$).

# Operations

Given triples $(\mathbf{A}, p, \phi), (\mathbf{B}, q, \theta)$, one can define:

1. **Probability preserving product**
$(\mathbf{A}, p, \phi) \otimes (\mathbf{B}, q, \theta) := (A \times B, p \otimes q, \phi \otimes \theta)$, where

$$(p \otimes q)(\langle a, b \rangle) = p(a)q(b)$$
$$(\theta \otimes \phi)(\langle a, b \rangle, \langle a', b' \rangle) = \theta(a, a')\phi(b, b').$$

2. **Weighted sum** $(\mathbf{A}, p, \phi) \oplus_\lambda (\mathbf{B}, q, \theta) := (\mathbf{A} \coprod \mathbf{B}, p \oplus_\lambda q, \phi \oplus \theta)$, for any $\lambda \in [0, 1]$, where

$$(p \oplus_\lambda q)(x) = \begin{cases} \lambda p(x) & x \in \mathrm{Ob}(\mathbf{A}) \\ (1 - \lambda)q(x) & x \in \mathrm{Ob}(\mathbf{B}) \end{cases}$$

$$(\phi \oplus \theta)(x, y) = \begin{cases} \phi(x, y) & x, y \in \mathrm{Ob}(\mathbf{A}) \\ \theta(x, y) & x, y \in \mathrm{Ob}(\mathbf{B}) \\ 0 & \text{otherwise} \end{cases}.$$

# Convergence

A sequence of triples $(A_k, p_k, \phi_k)$ converges to $(A, p, \phi)$ if $A_k = A$ for all $k$ that are large enough, and $p_k \to p$ and $\phi_k \to \phi$ pointwise.

# Categorical entropy

Defined as:

$$\mathcal{H}(\mathbf{A}, p, \phi) = -\sum_{a \in \mathrm{Ob}\,\mathbf{A}} p(a) \ln \left( \sum_{b \in \mathrm{Ob}\,\mathbf{A}} \phi(a,b) p(b) \right). \tag{3}$$

We assume $p(a)\ln(\ldots) = 0$ whenever $p(a) = 0$.

This quantity appears as a diversity measure (on a finite set $\mathrm{Ob}\,\mathbf{A}$ of species, $\phi$ similarity matrix, $p$ relative abundance).

# Main results

## Proposition

*$\mathcal{H}$ has the following properties:*

1. $\mathcal{H}(\mathbf{A}, p, \delta) = -\sum_{a \in \mathrm{Ob}\,\mathbf{A}} p(a) \ln p(a)$, *Shannon entropy.*
2. $\mathcal{H}((\mathbf{A}, p, \phi) \otimes (\mathbf{B}, q, \theta)) = \mathcal{H}(\mathbf{A}, p, \phi) + \mathcal{H}(\mathbf{B}, q, \theta)$.
3. $\mathcal{H}((\mathbf{A}, p, \phi) \oplus_\lambda (\mathbf{B}, q, \theta)) = \lambda \mathcal{H}(\mathbf{A}, p, \phi) + (1-\lambda)\mathcal{H}(\mathbf{B}, q, \theta) + \mathcal{H}(\mathbf{2}, \Lambda, \delta)$, *where* $\mathbf{2}$ *discrete category with two objects, with probabilities* $\lambda$ *and* $1-\lambda$.
4. *It is (sequentially) continuous.*

## Proposition

*Let* $(\mathbf{A}, p, \phi)$ *be a probabilistic category. Suppose* $\phi$ *has a weighting and a coweighting, and magnitude* $\|\phi\|$.
*If* $\phi$ *has a nonnegative weighting* $w$, *then* $u = w/\|\phi\|$ *is a probability distribution and* $\mathcal{H}(\mathbf{A}, u, \phi) = \ln \|\phi\|$.

Remark that in general $u$ is a signed probability.

# Outline

# A category of probabilistic triples

**ProbFinCat**: probabilistic finite categories.
Objects: categorical probabilistic triples.
Morphisms: $F : (\mathbf{A}, p, \phi) \to (\mathbf{B}, q, \theta)$ given by functor $F : \mathbf{A} \to \mathbf{B}$ that preserves the probability and the kernel:

- For all $b \in \mathrm{Ob}(\mathbf{B})$,

$$q(b) = F_* p(b) = \sum_{a \in F^{-1}(b)} p(a).$$

- For all $b, b' \in \mathrm{Ob}(\mathbf{B})$,

$$\theta(b, b') = F_* \phi(b, b') = \begin{cases} \sum_{a' \in F^{-1}(b')} \frac{p(a')}{F_* p(b')} \sum_{a \in F^{-1}(b)} \phi(a, a') & F_* p(b') > 0 \\ 1 & b = b', F_* p(b') = 0 \\ 0 & b \neq b', F_* p(b') = 0 \end{cases}.$$

Lemmas 1 and 2: This is well-defined.

# Convergence

## Definition

A sequence of morphisms $f_n : (\mathbf{A}_n, p_n, \phi_n) \to (\mathbf{B}_n, q_n, \theta_n)$ converges to $f : (\mathbf{A}, p, \phi) \to (\mathbf{B}, q, \theta)$ if $\mathbf{A}_n = \mathbf{A}$, $\mathbf{B}_n = \mathbf{B}$ and $f_n = f$ provided $n$ is big enough, and $p_n \to p$, $q_n \to q$, $\phi_n \to \phi$ and $\theta_n \to \theta$ pointwise.

**TransFinCat**: full subcategory of **ProbFinCat** of triples $(\mathbf{A}, p, \phi)$ with $\phi$ probabilistic transition kernel.

**FinProb**: category whose objects are pairs $(X, p)$ of a finite set with a p.m.f., and whose morphisms are probability preserving maps.

One naturally has **FinProb** $\hookrightarrow$ **TransFinCat** $\hookrightarrow$ **ProbFinCat**.

In particular, convergence in **FinProb** as defined by Baez, Fritz and Leinster [BFL11] is recovered.

# A characterization of classical entropy

$\mathbf{R}_+$: Category with one object $*$, and $\mathrm{Hom}(*, *) = \mathbb{R}$ with composition given by the sum.

A functor $F : \mathbf{ProbFinCat} \to \mathbf{R}_+$ is:

- **convex-linear** if for all morphisms $f$, $g$ and all scalars $\lambda \in [0, 1]$,

$$F(\lambda f \oplus (1 - \lambda)q) = \lambda F(f) + (1 - \lambda)F(g).$$

- **continuous** if $F(f_n) \to F(f)$ whenever $f_n \to f$.

Via embedding $\mathbf{FinProb} \hookrightarrow \mathbf{ProbFinCat}$, similar definitions for **FinProb**.

**Proposition (Baez-Fritz-Leinster '08)**

*If a functor $F : \mathbf{FinProb} \to \mathbf{R}_+$ is convex-linear and continuous, then there is $c \in \mathbb{R}$ such that each arrow $f : (A, p) \to (B, q)$ is mapped to $F(f) = c(H(p) - H(q))$.*

## Proposition (Baez-Fritz-Leinster '08)

*If a functor $F$ : $\mathbf{FinProb} \to \mathbf{R}_+$ is convex-linear and continuous, then there is $c \in \mathbb{R}$ such that each arrow $f : (A, p) \to (B, q)$ is mapped to $F(f) = c(H(p) - H(q))$.*

## Proof.

Let $\top$ be the terminal object, set $I(A, p) = F((A, p) \to \top)$. Then convex-linearity implies a functional equation for $I$. □

Remark: **TransFinCat** has a terminal object too.

## Proposition (Chen-V.)

*The functor $G$ : $\mathbf{TransFinCat} \to \mathbf{R}_+$ that maps $f : (\mathbf{A}, p, \phi) \to (\mathbf{B}, q, \theta)$ to $c(\mathcal{H}(\mathbf{A}, p, \phi) - \mathcal{H}(\mathbf{B}, q, \theta))$, for some $c \in \mathbb{R}$, is convex-linear and continuous.*

But is it the only one?

# References I

Mustafa Akkaya and Özgün Ünlü, *The Euler characteristic of finite categories*, arXiv preprint arXiv:2301.08966 (2023).

John C Baez, Tobias Fritz, and Tom Leinster, *A characterization of entropy in terms of information loss*, Entropy **13** (2011), no. 11, 1945–1957.

Stephanie Chen and Juan Pablo Vigneaux, *A formula for the categorical magnitude in terms of the Moore-Penrose pseudoinverse*, Bulletin of the Belgian Mathematical Society - Simon Stevin (2023), To appear in Vol. 30, Issue 3.

Tom Leinster, *The Euler characteristic of a category*, Documenta Mathematica **13** (2008), 21–49.