Our Symmetries

Arithmetic, Probability, Quantum

John Skilling (john@skilling.co.uk)

MaxEnt2025 Auckland

Start with arithmetic

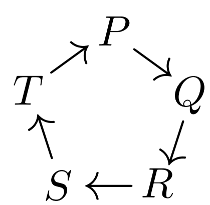
Mathematicians say "*Peano axioms*" !

1: $\exists\, 0$ (✓)

2: $\exists$ successor $S(\cdot)$: $\forall n,\ \exists\, S(n)$. Think $0 \to 1 \to 2 \to 3\ \cdots$ (✓)

3: But what about $\to \bullet \to m \to \bullet \to \bullet \to$?
$\to \bullet \to n \nearrow$ Need to say S is invertable, $\exists$ unique $\leftarrow$. (A fixup)

4: And what about $\to \bullet \to 0 \to 1 \to 2 \to\ \cdots$? Need to say $\nexists\, \bullet \to 0$. (A fixup)

5: What about  along with $0 \to 1 \to 2 \to 3 \to 4 \to\ \cdots$?

Need axiom of induction to exclude disjoint cycles. (A fixup)

Fixups are a disgrace.

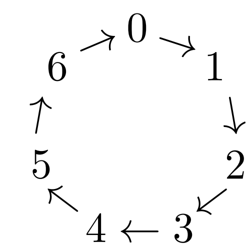
Mathematicians then say "*Zermelo-Fraenkel — welcome to* ∞ *and the Axiom of Choice*" !

Not for me.

Arithmetic from symmetries

1: *We are finite.* Modelling encodes objects from a finite library (size N) of symbols.

2: Demand *lossless communication* (permutations of library).



Fundamental permutation is cyclic with prime length (no subcycles).

Arbitrarily assign labels $\{\underbrace{0, 1, 2, \cdots, N{-}1}_{\text{Library}}\}$ with N prime.

We have Peano #1: $\exists\, 0$ $\quad(\checkmark)$

#2: $\exists$ successor $S(n)$, $n = \underbrace{S(S(\cdots S(}_{n} 0 \underbrace{)\cdots))}_{n}$ $\quad(\checkmark)$

#3: S invertable $\quad(\checkmark)$

#4: **False:** $S(N{-}1) = 0$

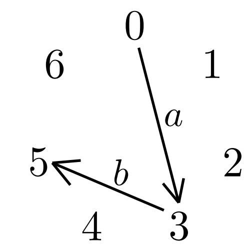
#5: Induction $\quad(\checkmark)$

Begin wraparound arithmetic.

3: *Associativity*

We want to assemble composite objects A $\oplus$ B , P $\oplus$ Q $\oplus$ R , etc, ignoring irrelevant differences.
Demand that representation is associative: $a \oplus (b \oplus c) = (a \oplus b) \oplus c$

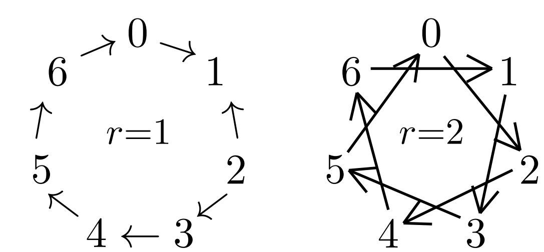
Lossless associativity	$\Longleftrightarrow$	Additive representation $a \oplus b = a + b \pmod N$



Commutativity is emergent.
Subtraction is the inverse. $2 - 5 = 1000000 \pmod{1000003}$.

4: *Distributivity*

We want to be able to communicate additivity by transformations.
Demand that transformations are left-distributive, $T(a+b) = T(a) + T(b)$.
This gives multiplication, $T(x) = rx \pmod N$, with $r \neq 0$.



Left-distributivity over addition	$\Longleftrightarrow$	Linear multiplication $a \otimes b = ab \pmod N$

Right-distributivity and associativity are emergent.
Division is the inverse. $1 \div 3 = 666669 \pmod{1000003}$.

5: *No overflow*

Demand **size of application $< N$** and avoid detailing N.

To implement subtraction fully, invent negative numbers: $2 - 5 = -3$.
To implement division fully, invent rational numbers: $1 \div 3 = \frac{1}{3}$.
Continuity and order $(<, =, >)$ are emergent.

Now have real line: proceed to standard mathematics, π, exp, log, cos, sin, etc.

Summary

Set the scene.

1: We are finite $\Longrightarrow$ Library $N < \infty$

2: Lossless communication $\Longrightarrow$ Cyclic permutations

Basic symmetries.

3: Lossless associativity $\Longleftrightarrow$ Additive representation $a \oplus b = a + b \pmod{N}$

4: Left-distributivity over addition $\Longleftrightarrow$ Linear multiplication $a \otimes b = ab \pmod{N}$

Get useful language.

5: Size of application $< N$ $\Longrightarrow$ standard mathematics

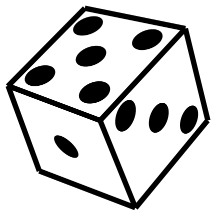
No fixups.

Application — Probability

Set the scene.

Inference is about focussing on posterior subsets $X \in Z$ of prior possibilities Z.

$\{1,3,5\} \in \{1,2,3,4,5,6\}$



Quantify by $\Pr(X \mid Z)$ called *probability*.

Basic symmetries.

Pr is additive over X because disjoint subsets combine associatively.

Pr scales multiplicatively over Z because additivity is preserved over expansion (distributivity).

$$\therefore \qquad \Pr(X \mid Z) = \underset{\nearrow\ \text{measure}}{m(X)}\ \underset{\nwarrow\ \text{function}}{f(Z)}$$

Get useful language.

Consistency during expansion of context $X \in Y \in Z$ requires $f = 1/m$.

$$\therefore \qquad \boxed{\Pr(X \mid Z) = \frac{m(X)}{m(Z)}} \qquad \text{(simple proportion)}$$

Hence sum and product rules of Bayesian probability.

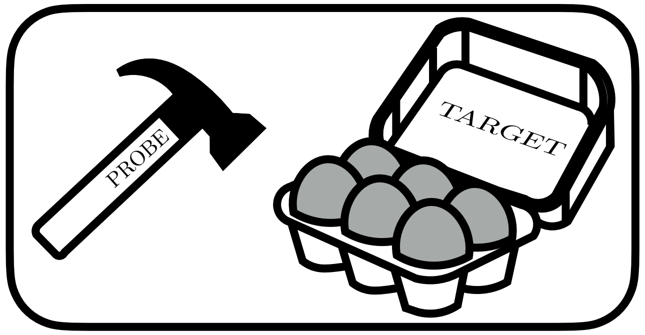
No tedious philosophy (propensity, frequency, belief, plausibility, . . .).

If you have the basic symmetries of arithmetic, then you *have* arithmetic.

Which assumption could a skeptic deny?

Application — Physics

Set the scene.



Physics is about *interactions*, probe∼∼target.

At smallest scale, cannot have full knowledge.

Modelling needs quantity and uncertainty.

Representation of object is based on number pairs. $x = (x_1, x_2), \quad y = (y_1, y_2), \ldots$

Basic symmetries.

Demand lossless associativity $x \oplus (y \oplus z) = (x \oplus y) \oplus z$ of assembly.

∴ Representations add linearly, $(x \oplus y)_i = x_i + y_i$.

Demand that probing is left-distributive, $x \otimes (y + z) = x \otimes y + x \otimes z$, to preserve additivity of targets.

"Probe" and "target" are interchangeable labels, so demand right-distributivity too.

∴ Interaction is bilinear multiplication, $(x \otimes y)_i = \sum_{jk} \varphi_{ijk}\, x_j\, y_k$ with 8 coefficients φ to be defined.

So we have **lossless associativity (linear addition)**
and **left and right distributivity (bilinear multiplication).**

Also demand that **operations chain associatively.** $x \otimes (y \otimes z) = (x \otimes y) \otimes z$

Get useful language.

The three product rules

We have bilinear multiplication $(x \otimes y)_i = \sum_{jk} \varphi_{ijk}\, x_j\, y_k$ with φ to be defined,

with associativity $x \otimes (y \otimes z) = (x \otimes y) \otimes z$

Associativity imposes 16 quadratic constraints on the 8 φ's. $\sum_{t=1}^{2} \varphi_{ixt}\varphi_{tyz} = \sum_{t=1}^{2} \varphi_{itz}\varphi_{txy} \quad \forall\ i, x, y, z \in \{1, 2\}$

They allow three product rules $\boxed{\begin{matrix} x_1 \\ x_2 \end{matrix}} \otimes \boxed{\begin{matrix} y_1 \\ y_2 \end{matrix}} = \left(\underbrace{\boxed{\begin{matrix} x_1 y_1 - x_2 y_2 \\ x_1 y_2 + x_2 y_1 \end{matrix}}}_{\text{A}} \text{ or } \underbrace{\boxed{\begin{matrix} x_1 y_1 + x_2 y_2 \\ x_1 y_2 + x_2 y_1 \end{matrix}}}_{\text{B}} \text{ or } \underbrace{\boxed{\begin{matrix} x_1 y_1 \\ x_1 y_2 + x_2 y_1 \end{matrix}}}_{\text{C}} \right)$ [algebra!]

Extract operator x: $x = \left(\underbrace{\boxed{\begin{matrix} x_1 & -x_2 \\ x_2 & x_1 \end{matrix}}}_{\text{A}} \text{ or } \underbrace{\boxed{\begin{matrix} x_1 & x_2 \\ x_2 & x_1 \end{matrix}}}_{\text{B}} \text{ or } \underbrace{\boxed{\begin{matrix} x_1 & 0 \\ x_2 & x_1 \end{matrix}}}_{\text{C}} \right)$

Use polar coordinates. $x = r \left(\underbrace{\boxed{\begin{matrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{matrix}}}_{\substack{\text{A} \\ \text{complex}}} \text{ or } \underbrace{\boxed{\begin{matrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{matrix}}}_{\substack{\text{B} \\ \text{split-complex}}} \text{ or } \underbrace{\boxed{\begin{matrix} 1 & 0 \\ \theta & 1 \end{matrix}}}_{\substack{\text{C} \\ \text{(see later)}}} \right)$

Complex numbers from ignorance

For each product rule, phase $\theta = \arg(x)$ is additive, $\arg(x \otimes y) = \arg(x) + \arg(y)$.

Hence representation of phase interval $\Delta\theta = \theta_2 - \theta_1$ is invariant to offsets.

Hence prior probability that we (initially ignorant) assign to a phase interval is invariant to offsets.

$$\Pr(\theta) = \text{constant}$$

Try rule A (complex numbers): range is cyclic from 0 to 2π. $\Pr(\theta) = \dfrac{1}{2\pi}$, uniform from 0 to 2π.

Try rule B or rule C: range unlimited $\theta \in (-\infty, \infty)$. No proper prior.

Rule A alone allows identification of uncertainty, as phase θ of a pair.

Representation of object is based on complex numbers. !

Quantity $\sim r$, uncertainty $\sim \theta$

Want A and B and C instead of A **or** B **or** C.

Rules A and B give us generators of the form $X = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and $Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $YX = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = Z$

These define a four-element group spanned by $\{\mathbf{1}, X, Y, Z\}$ with multiplication table

$\downarrow\cdot\rightarrow$	$\cdot\mathbf{1}$	$\cdot X$	$\cdot Y$	$\cdot Z$
$\mathbf{1}\cdot$	$\mathbf{1}$	X	Y	Z
$X\cdot$	X	$-\mathbf{1}$	$-Z$	Y
$Y\cdot$	Y	Z	$\mathbf{1}$	X
$Z\cdot$	Z	$-Y$	$-X$	$\mathbf{1}$

This demands a **4-parameter representation.**

Rules A and B

All this still works even if (as will be the case) parameters are complex instead of real.

The four-element group $\{\mathbf{1},\ X,\ Y,\ Z\}$ is upgraded to $\{\mathbf{1},\ X,\ Y,\ Z;\ i,\ iX,\ iY,\ iZ\}$ where $i^2 = -1$.

The multiplication table is upgraded to 8×8.

$\downarrow\cdot\rightarrow$	$\cdot\mathbf{1}$	$\cdot X$	$\cdot Y$	$\cdot Z$	$\cdot i$	$\cdot iX$	$\cdot iY$	$\cdot iZ$
$\mathbf{1}\cdot$	$\mathbf{1}$	X	Y	Z	i	iX	iY	iZ
$X\cdot$	X	$-\mathbf{1}$	$-Z$	Y	iX	$-i$	$-iZ$	iY
$Y\cdot$	Y	Z	$\mathbf{1}$	X	iY	iZ	i	iX
$Z\cdot$	Z	$-Y$	$-X$	$\mathbf{1}$	iZ	$-iY$	$-iX$	i
$i\cdot$	i	iX	iY	iZ	$-\mathbf{1}$	$-X$	$-Y$	$-Z$
$iX\cdot$	iX	$-i$	$-iZ$	iY	$-X$	$\mathbf{1}$	Z	$-Y$
$iY\cdot$	iY	iZ	i	iX	$-Y$	$-Z$	$-\mathbf{1}$	$-X$
$iZ\cdot$	iZ	$-iY$	$-iX$	i	$-Z$	Y	X	$-\mathbf{1}$

This is the Lorentz group !

As in all groups, the identity $\mathbf{1}$ is special. Its coefficient gives *quantity*.
The pseudoscalar i commutes with everything so is also special.
Its coefficient is *rate of change*, with respect to phase. For any complex number(s), $\frac{d}{d\theta}(re^{i\theta}) = i\,re^{i\theta}$.

Rule C

$i = \frac{d}{d\theta}$ implements rule C operating on $\boxed{\begin{matrix} r \\ \theta \end{matrix}}$.

Lorentz factorisation

The group was $\{\mathbf{1}, X, Y, Z; i, iX, iY, iZ\}$.

$\mathbf{1}$ was interpreted as *quantity*.

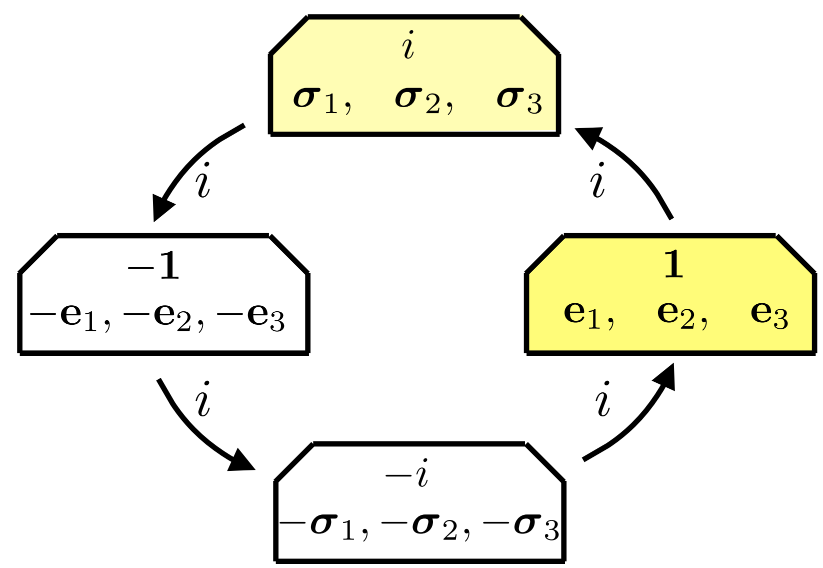
i was interpreted as *evolution*.

Of the other elements, $(X, iY, -iZ)$ square to -1 (4th order); relabel as $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$;

while $(iX, -Y, Z)$ square to $+1$ (2nd order); relabel as $(i\mathbf{e}_1, i\mathbf{e}_2, i\mathbf{e}_3) = \underbrace{(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3)}_{\text{Pauli matrices}}$.

Lorentz group can be relabelled $\underbrace{\underbrace{\{\mathbf{1}, i\}}_{\text{complex}} \times \underbrace{\{\mathbf{1}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}}_{\text{quaternion}}}_{\text{biquaternion}} = \{\underbrace{\mathbf{1}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3}_{\text{real}}; \underbrace{i, \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3}_{\text{imaginary}}\}$

$\downarrow\cdot\rightarrow$	$\cdot\mathbf{1}$	$\cdot\mathbf{e}_1$	$\cdot\mathbf{e}_2$	$\cdot\mathbf{e}_3$	$\cdot i$	$\cdot\boldsymbol{\sigma}_1$	$\cdot\boldsymbol{\sigma}_2$	$\cdot\boldsymbol{\sigma}_3$
$\mathbf{1}\cdot$	$\mathbf{1}$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	i	$\boldsymbol{\sigma}_1$	$\boldsymbol{\sigma}_2$	$\boldsymbol{\sigma}_3$
$\mathbf{e}_1\cdot$	$\mathbf{e}_1$	$-\mathbf{1}$	$\mathbf{e}_3$	$-\mathbf{e}_2$	$\boldsymbol{\sigma}_1$	$-i$	$\boldsymbol{\sigma}_3$	$-\boldsymbol{\sigma}_2$
$\mathbf{e}_2\cdot$	$\mathbf{e}_2$	$-\mathbf{e}_3$	$-\mathbf{1}$	$\mathbf{e}_1$	$\boldsymbol{\sigma}_2$	$-\boldsymbol{\sigma}_3$	$-i$	$\boldsymbol{\sigma}_1$
$\mathbf{e}_3\cdot$	$\mathbf{e}_3$	$\mathbf{e}_2$	$-\mathbf{e}_1$	$-\mathbf{1}$	$\boldsymbol{\sigma}_3$	$\boldsymbol{\sigma}_2$	$-\boldsymbol{\sigma}_1$	$-i$
$i\cdot$	i	$\boldsymbol{\sigma}_1$	$\boldsymbol{\sigma}_2$	$\boldsymbol{\sigma}_3$	$-\mathbf{1}$	$-\mathbf{e}_1$	$-\mathbf{e}_2$	$-\mathbf{e}_3$
$\boldsymbol{\sigma}_1\cdot$	$\boldsymbol{\sigma}_1$	$-i$	$\boldsymbol{\sigma}_3$	$-\boldsymbol{\sigma}_2$	$-\mathbf{e}_1$	$\mathbf{1}$	$-\mathbf{e}_3$	$\mathbf{e}_2$
$\boldsymbol{\sigma}_2\cdot$	$\boldsymbol{\sigma}_2$	$-\boldsymbol{\sigma}_3$	$-i$	$\boldsymbol{\sigma}_1$	$-\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{1}$	$-\mathbf{e}_1$
$\boldsymbol{\sigma}_3\cdot$	$\boldsymbol{\sigma}_3$	$\boldsymbol{\sigma}_2$	$-\boldsymbol{\sigma}_1$	$-i$	$-\mathbf{e}_3$	$-\mathbf{e}_2$	$\mathbf{e}_1$	$\mathbf{1}$



$\{\mathbf{1}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ factors out as the subgroup of *quaternions*.

$$\mathbb{L} = \mathbb{C} \times \mathbb{H} \quad !$$

The witches' brew

$$\underbrace{\underbrace{\{\mathbf{1}\ ,\ i\}}_{\text{uncertainty}} \times \underbrace{\{\mathbf{1}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}}_{\text{mathematics}}}_{\text{the language of physics}}$$

Add logic and stir.

Here as he walked by
on the 16th of October 1843
Sir William Rowan Hamilton
in a flash of genius discovered
the fundamental formula for
quaternion multiplication
$$i^2 = j^2 = k^2 = ijk = -1$$
& cut it on a stone of this bridge

John Skilling and Kevin Knuth at the quaternion plaque in Dublin, 13 April 2024.

Relativistic quantum formalism is just the arithmetic of number pairs !

Sum rule from associative commutativity of content.
Product rules from associative distributivity of operators.
Number pairs, for quantity and uncertainty.

Simple and general.
No other assumptions.

We recognise:

- complex numbers underlying physics
- phase as ignorance accompanying quantity
- quantification by Born rule
- Lorentz group
- 4-spin and 4-momentum
- three-dimensional space
- special relativity with Minkowski metric
- matter and antimatter
- the Dirac equation
- conservation of quantity