Derivatives of proofs in linear logic

Daniel Murfet

based on joint work with James Clift

Precis

BHK interpretation: intuitionistic proofs of $A \to B$ give rise to functions $\mathrm{Proofs}(A) \to \mathrm{Proofs}(B)$

- Can these functions be differentiated?
- What would such derivatives be good for?

1. Efficient (re)computation
2. Differentiable reasoning
3. Investigating logic vs physics

Precis

BHK interpretation: intuitionistic proofs of $A \to B$ give rise to functions $\mathrm{Proofs}(A) \to \mathrm{Proofs}(B)$

- Can these functions be differentiated?
- What would such derivatives be good for?

1. Efficient (re)computation
2. Differentiable reasoning
3. Investigating logic vs physics

Curry-Howard correspondence

logic	**programming**	**math**
formula	type	space
proof	program	function
cut-elimination	execution	—
contraction	copying	coalgebra
?	**?**	calculus

Outline

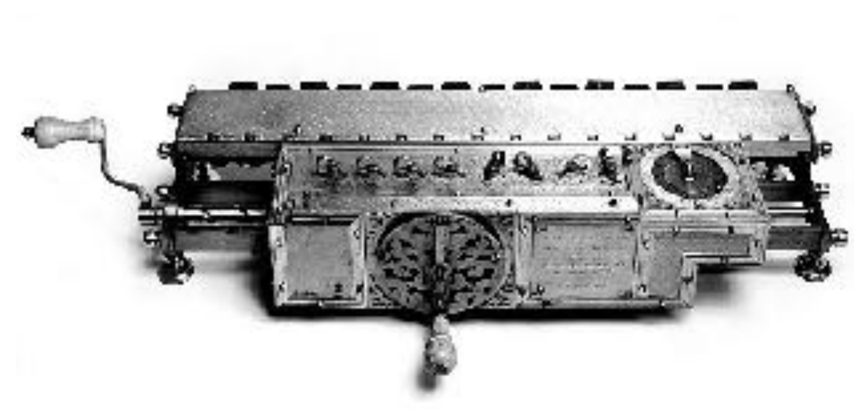
1. History of derivatives in logic
2. Derivatives in the syntax
3. Relation to calculus via coalgebras

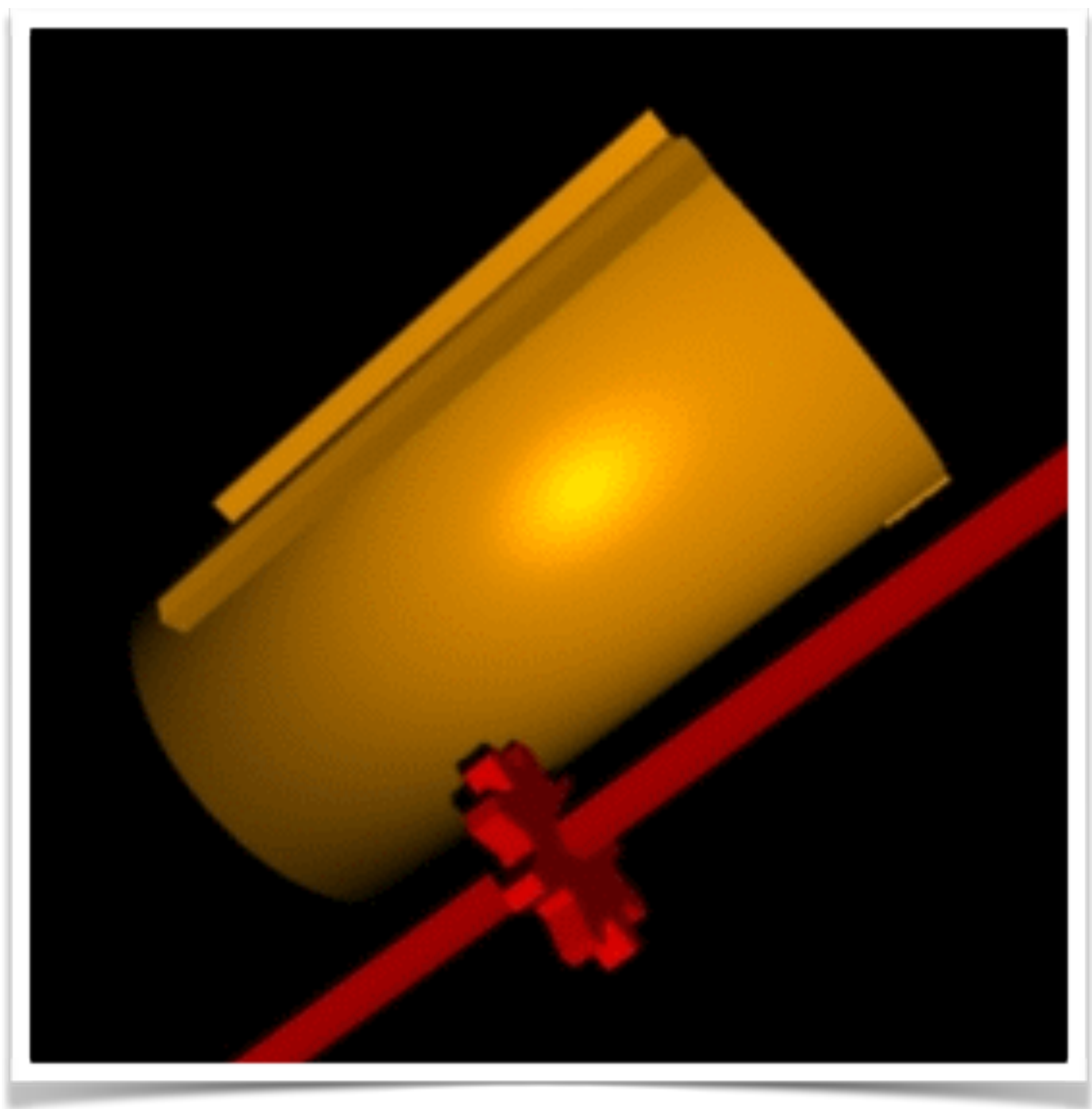
(based on arXiv:1701.01285)

History of derivatives in logic

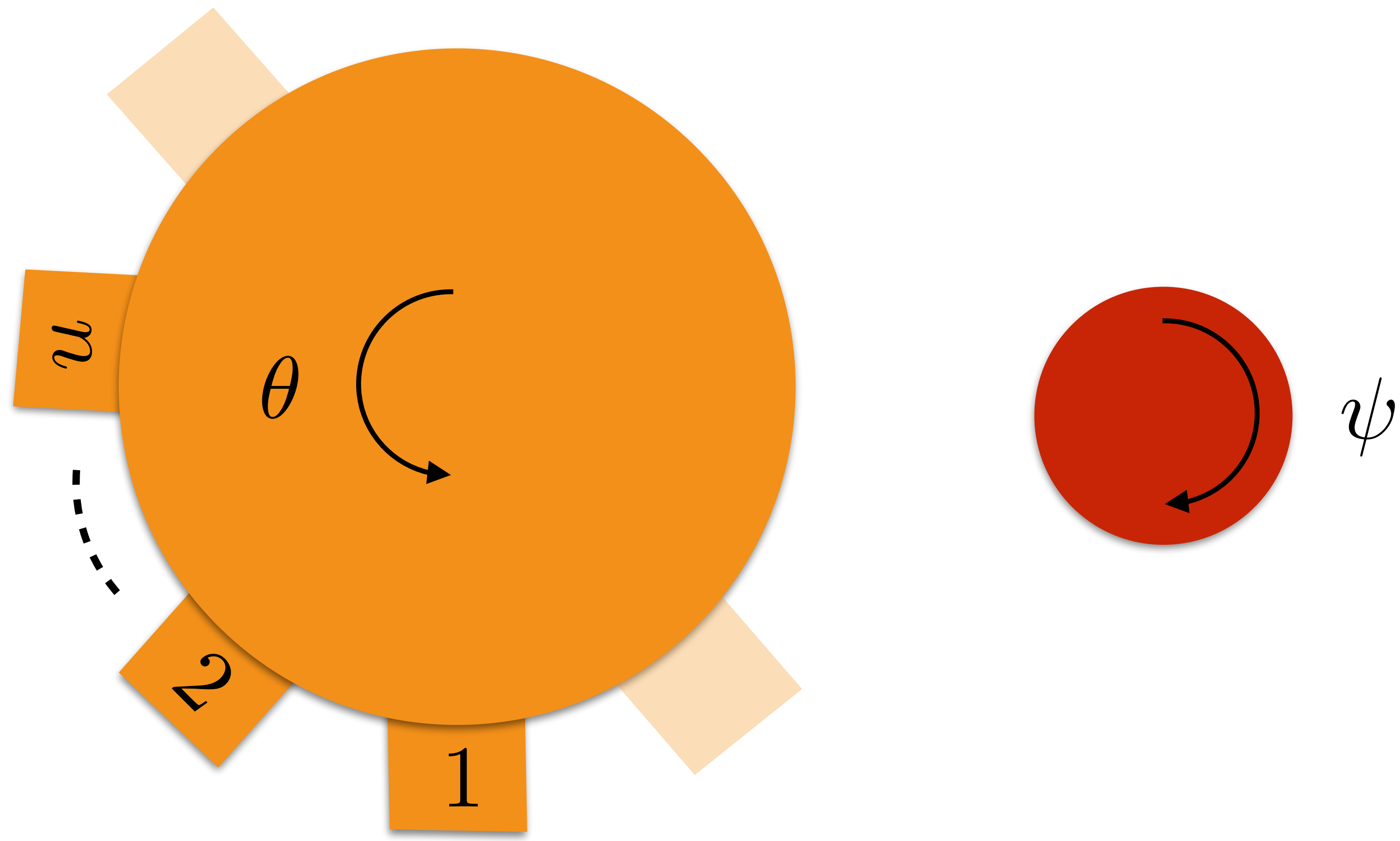
- Leibniz’s stepped reckoner (1670s)
- Babbage’s difference engine (1830s)
- Circuits and 2nd order differential equations
- Automatic differentiation of real-valued programs
- Ehrhard-Regnier’s differential lambda calculus (2003)
- Differential linear logic

History: Leibniz's stepped reckoner





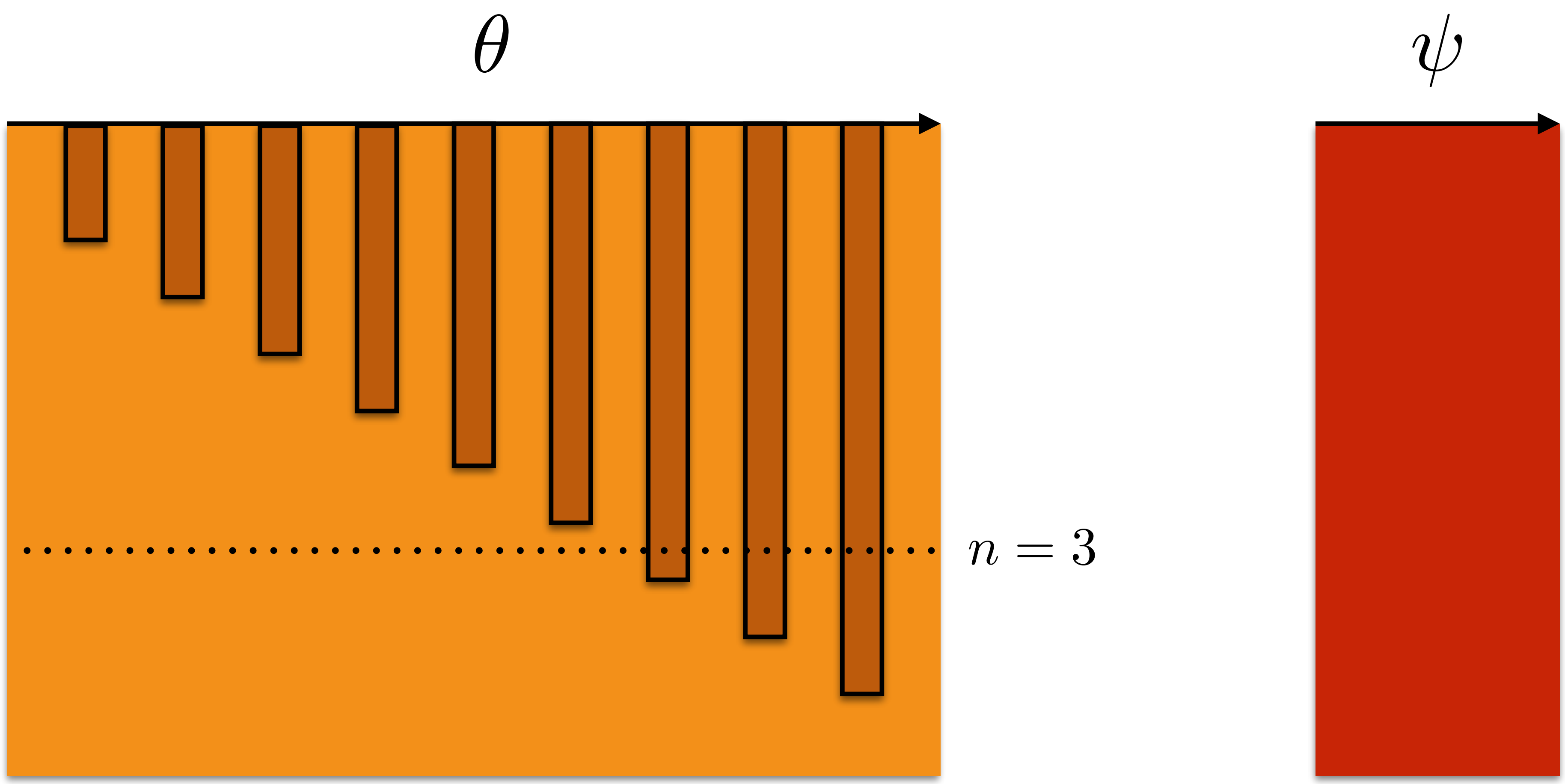
History: Leibniz's stepped reckoner



After a full rotation of the **drum**, the **shaft** rotates by nk

$$\Delta\psi = nk\frac{\Delta\theta}{2\pi} \qquad \theta = 0, 2\pi, 4\pi, \ldots$$

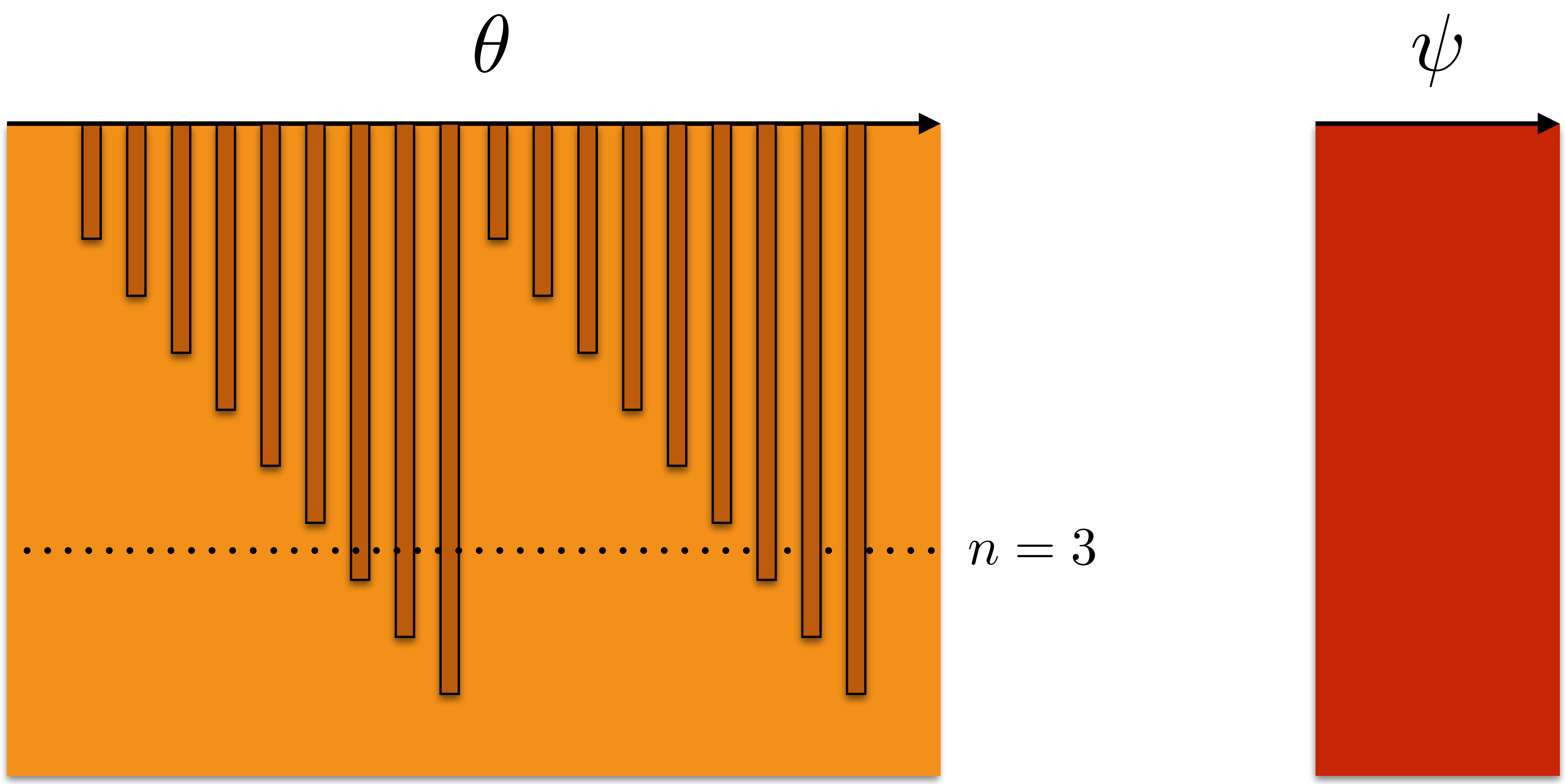
History: Leibniz's stepped reckoner



After a full rotation of the **drum**, the **shaft** rotates by nk

$$\Delta\psi = nk\frac{\Delta\theta}{2\pi} \qquad \Delta\theta = 0, 2\pi, 4\pi, \ldots$$

History: Leibniz's stepped reckoner

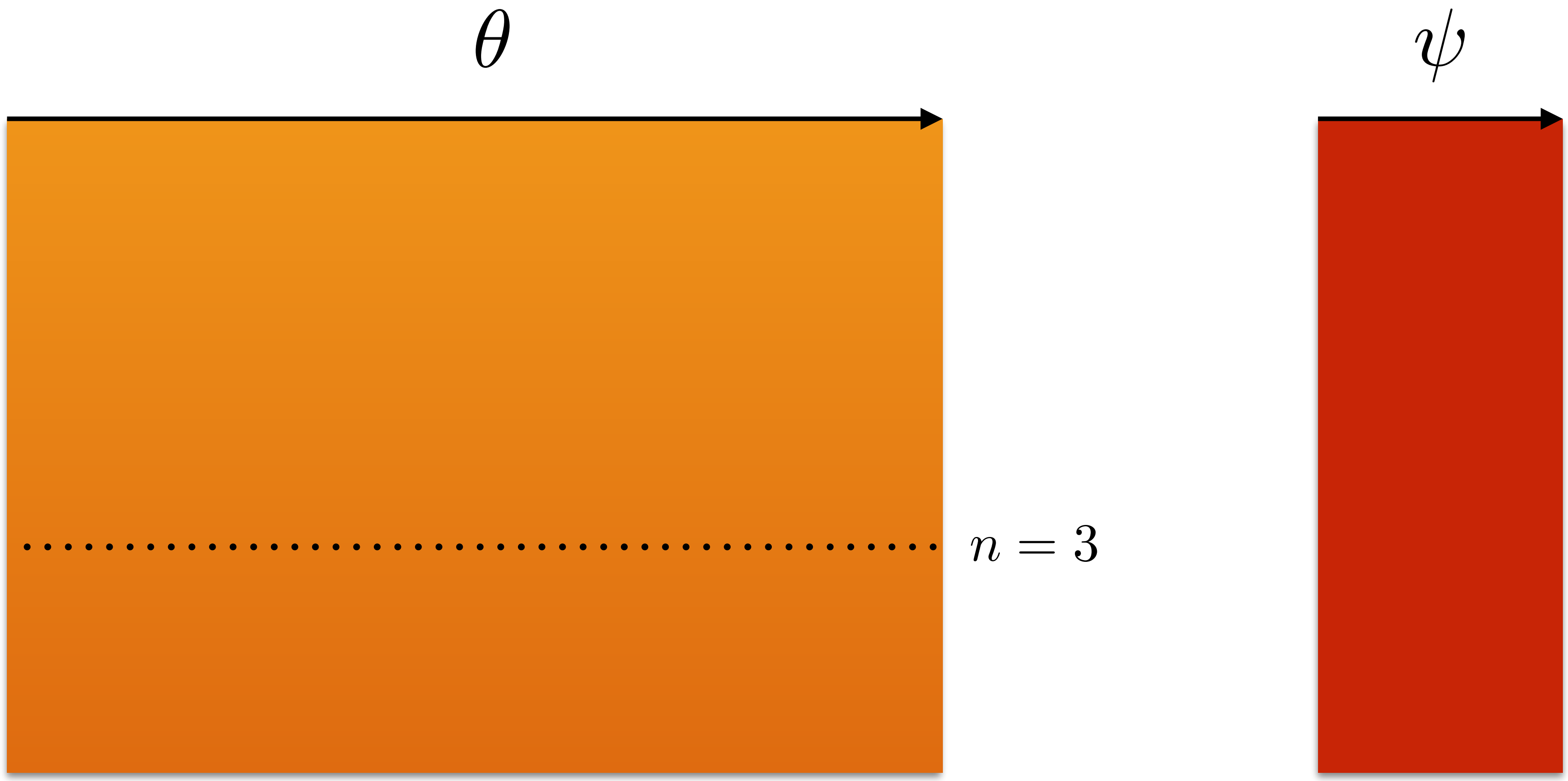


After a full rotation of the **drum**, the **shaft** rotates by nk

(if we halve the rotation caused by each tooth, while doubling the number)

$$\Delta\psi = nk\frac{\Delta\theta}{2\pi} \qquad \Delta\theta = 0, \pi, 2\pi, \ldots$$

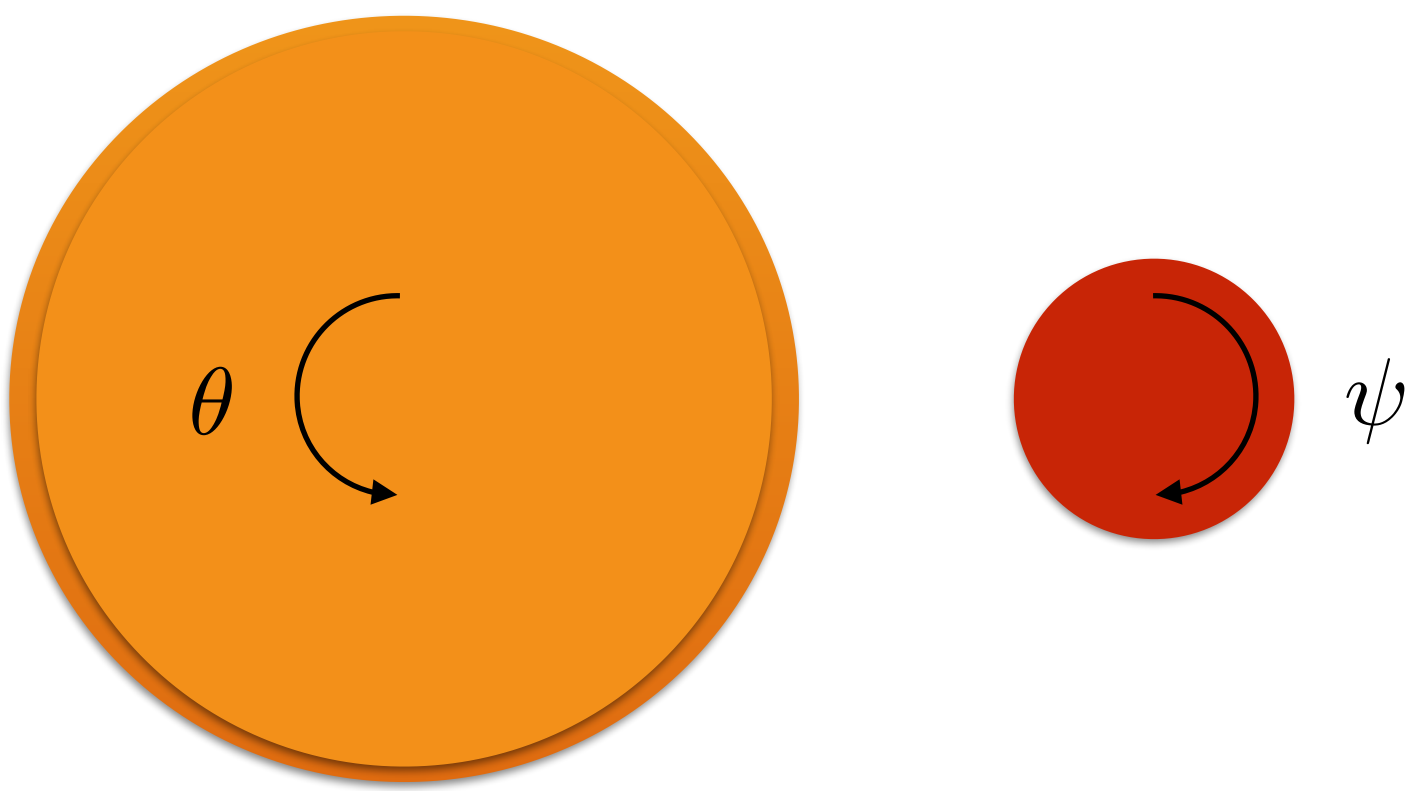
History: Leibniz's stepped reckoner



In the limit of infinitely many repetitions of this group of nine teeth

$$d\psi = nk\frac{d\theta}{2\pi} = nk' d\theta \qquad k' = \frac{k}{2\pi} \qquad \boxed{\frac{d\psi}{d\theta} = nk'}$$

History: Leibniz's stepped reckoner



$$\frac{d\psi}{d\theta} = n$$

$$\psi = n\theta$$

$$e^{i\psi} = e^{in\theta} = (e^{i\theta})^n$$

$$\begin{array}{ccc}
U(1) & \xrightarrow{(-)^n} & U(1) \\
\downarrow & & \downarrow \\
SO(2) & \xrightarrow{(-)^n} & SO(2) \\
\downarrow & & \downarrow \\
M_2(\mathbb{R}) & \xrightarrow{(-)^n = [\![n]\!]} & M_2(\mathbb{R})
\end{array}$$

Upshot: The stepped reckoner gives a "physical semantics" of the Church numerals matching the denotational semantics in vector spaces

Derivatives in the syntax

- Differential linear logic adds a new deduction rule, which produces the derivative of a proof in a direction specified by a new (linear) hypothesis.

$$\dfrac{\begin{array}{c}\pi \\ \vdots \\ !A \vdash B\end{array}}{!A, A \vdash B}\ \mathrm{diff} \qquad (a_1, a_2) \longrightarrow \lim_{h \to 0} \frac{\pi(a_1 + h a_2) - \pi(a_1)}{h}$$

(this is meaningless)

- In the best formulation diff is derived from *dereliction, cocontraction* and *coweakening*.

Deduction rules for (intuitionistic, first-order) linear logic

(Dereliction): $$\frac{\Gamma, A, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\ \text{der}$$

(Contraction): $$\frac{\Gamma, !A, !A, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\ \text{ctr}$$

(Weakening): $$\frac{\Gamma, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\ \text{weak}$$

(Axiom): $$\frac{}{A \vdash A}$$

(Cut): $$\frac{\Gamma \vdash A \quad \Delta', A, \Delta \vdash B}{\Delta', \Gamma, \Delta \vdash B}\ \text{cut}$$

(Promotion): $$\frac{!\Gamma \vdash A}{!\Gamma \vdash !A}\ \text{prom}$$

(Left $\multimap$): $$\frac{\Gamma \vdash A \quad \Delta', B, \Delta \vdash C}{\Delta', \Gamma, A \multimap B, \Delta \vdash C}\ \multimap L$$

(Left $\otimes$): $$\frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, A \otimes B, \Delta \vdash C}\ \otimes\text{-}L$$

(Right $\multimap$): $$\frac{A, \Gamma \vdash B}{\Gamma \vdash A \multimap B}\ \multimap R$$

(Right $\otimes$): $$\frac{\Gamma \vdash A \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \otimes B}\ \otimes\text{-}R$$

Deduction rules for (intuitionistic, first-order) linear logic

(Dereliction): $\dfrac{\Gamma, A, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\,\text{der}$

(Contraction): $\dfrac{\Gamma, !A, !A, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\,\text{ctr}$

(Weakening): $\dfrac{\Gamma, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\,\text{weak}$

$\underline{001} : \mathbf{bint}_A = !(A \multimap A) \multimap (!(A \multimap A) \multimap (A \multimap A))$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overline{A \vdash A} \quad \dfrac{\overline{A \vdash A} \quad \dfrac{\overline{A \vdash A} \quad \overline{A \vdash A}}{A, A \multimap A \vdash A}\multimap L}{A, A \multimap A, A \multimap A \vdash A}\multimap L}{A, A \multimap A, A \multimap A, A \multimap A \vdash A}\multimap L}{A \multimap A, A \multimap A, A \multimap A \vdash A \multimap A}\multimap R}{!(A \multimap A), !(A \multimap A), !(A \multimap A) \vdash A \multimap A}\,3\times\ \text{der}}{!(A \multimap A), !(A \multimap A) \vdash A \multimap A}\,\text{ctr}}{\vdash \mathbf{bint}_A}\,2\times \multimap R$$

(Axiom): $\dfrac{}{A \vdash A}$

(Cut): $\dfrac{\Gamma \vdash A \qquad \Delta', A, \Delta \vdash B}{\Delta', \Gamma, \Delta \vdash B}\,\text{cut}$

(Promotion): $\dfrac{!\Gamma \vdash A}{!\Gamma \vdash !A}\,\text{prom}$

(Left $\multimap$): $\dfrac{\Gamma \vdash A \qquad \Delta', B, \Delta \vdash C}{\Delta', \Gamma, A \multimap B, \Delta \vdash C}\multimap L$

(Left $\otimes$): $\dfrac{\Gamma, A, B, \Delta \vdash C}{\Gamma, A \otimes B, \Delta \vdash C}\otimes\text{-}L$

(Right $\multimap$): $\dfrac{A, \Gamma \vdash B}{\Gamma \vdash A \multimap B}\multimap R$

(Right $\otimes$): $\dfrac{\Gamma \vdash A \qquad \Delta \vdash B}{\Gamma, \Delta \vdash A \otimes B}\otimes\text{-}R$

Deduction rules for differential linear logic

(Dereliction): $$\frac{\Gamma, A, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\ \text{der}$$

(Codereliction): $$\frac{\Gamma, !A, \Delta \vdash B}{\Gamma, A, \Delta \vdash B}\ \text{coder}$$

(Contraction): $$\frac{\Gamma, !A, !A, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\ \text{ctr}$$

(Cocontraction): $$\frac{\Gamma, !A, \Delta \vdash B}{\Gamma, !A, !A, \Delta \vdash B}\ \text{coctr}$$

(Weakening): $$\frac{\Gamma, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\ \text{weak}$$

(Coweakening): $$\frac{\Gamma, !A, \Delta \vdash B}{\Gamma, \Delta \vdash B}\ \text{coweak}$$

(Axiom): $$\frac{}{A \vdash A}$$

(Cut): $$\frac{\Gamma \vdash A \qquad \Delta', A, \Delta \vdash B}{\Delta', \Gamma, \Delta \vdash B}\ \text{cut}$$

(Promotion): $$\frac{!\Gamma \vdash A}{!\Gamma \vdash !A}\ \text{prom}$$

(Left $\multimap$): $$\frac{\Gamma \vdash A \qquad \Delta', B, \Delta \vdash C}{\Delta', \Gamma, A \multimap B, \Delta \vdash C}\ \multimap L$$

(Left $\otimes$): $$\frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, A \otimes B, \Delta \vdash C}\ \otimes\text{-}L$$

(Right $\multimap$): $$\frac{A, \Gamma \vdash B}{\Gamma \vdash A \multimap B}\ \multimap R$$

(Right $\otimes$): $$\frac{\Gamma \vdash A \qquad \Delta \vdash B}{\Gamma, \Delta \vdash A \otimes B}\ \otimes\text{-}R$$

Deduction rules for differential linear logic

(Dereliction): $$\frac{\Gamma, A, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\ \mathrm{der}$$

(Codereliction): $$\frac{\Gamma, !A, \Delta \vdash B}{\Gamma, A, \Delta \vdash B}\ \mathrm{coder}$$

(Contraction): $$\frac{\Gamma, !A, !A, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\ \mathrm{ctr}$$

(Cocontraction): $$\frac{\Gamma, !A, \Delta \vdash B}{\Gamma, !A, !A, \Delta \vdash B}\ \mathrm{coctr}$$

(Weakening): $$\frac{\Gamma, \Delta \vdash B}{\Gamma, !A, \Delta \vdash B}\ \mathrm{weak}$$

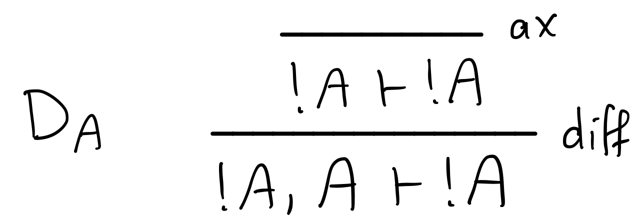
(Coweakening): $$\frac{\Gamma, !A, \Delta \vdash B}{\Gamma, \Delta \vdash B}\ \mathrm{coweak}$$

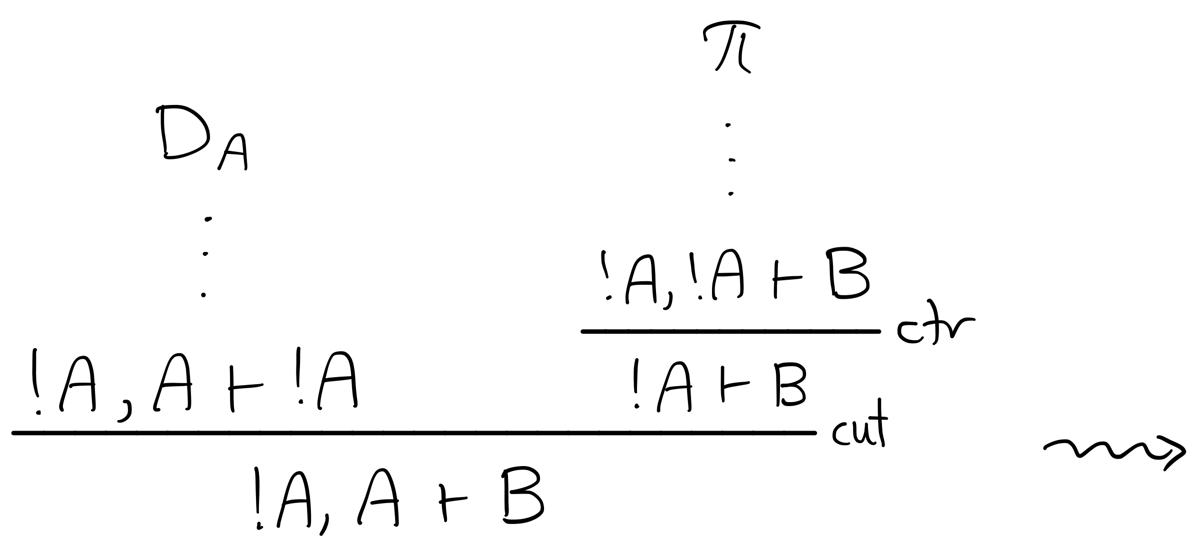
$$\dfrac{\begin{array}{c}\pi \\ \vdots \\ !A \vdash B\end{array}}{!A, A \vdash B}\ \mathrm{diff}$$

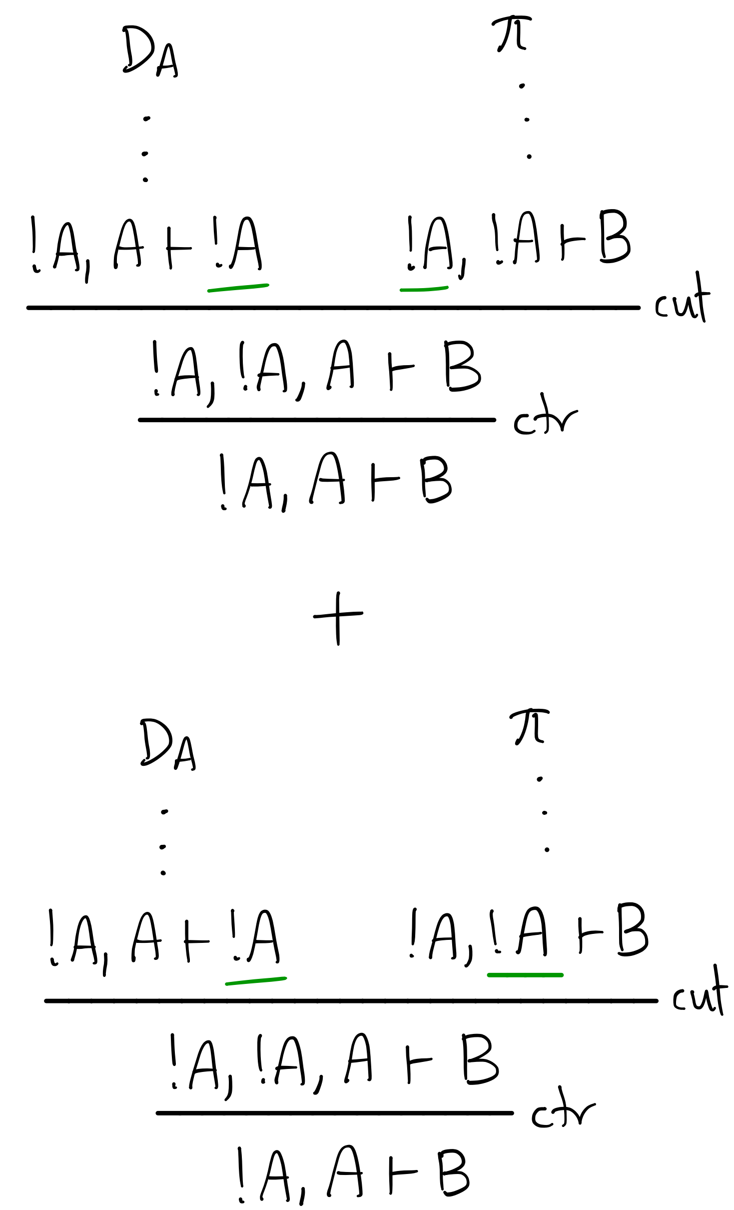
is defined to be

$$\dfrac{\dfrac{\begin{array}{c}\pi \\ \vdots \\ !A \vdash B\end{array}}{!A, !A \vdash B}\ \mathrm{coctr}}{!A, A \vdash B}\ \mathrm{coder}$$

Product rule as cut-elimination rule







Proofs in differential linear logic are formal linear sums of proof trees

$$\mathbf{bint}_A = !(A \multimap A) \multimap (!(A \multimap A) \multimap (A \multimap A)).$$

$$E = A \multimap A$$

$$\underline{\mathrm{repeat}} : !\mathbf{bint}_A \multimap \mathbf{bint}_A$$

$$\dfrac{\dfrac{}{!E \vdash !E} \qquad \dfrac{\underline{\mathrm{comp}}^2_A}{\vdots \atop E, E \vdash E}}{!E, !E \multimap E, E, E \vdash E} \multimap L$$

$$\dfrac{\dfrac{}{!E \vdash !E} \qquad !E, !E \multimap E, E, E \vdash E}{!E, !E, \mathbf{bint}_A, E \vdash E} \multimap L$$

$$\dfrac{\dfrac{}{!E \vdash !E} \qquad !E, !E, \mathbf{bint}_A, E \vdash E}{!E, !E, !E, \mathbf{bint}_A, !E \multimap E \vdash E} \multimap L$$

$$\dfrac{\dfrac{}{!E \vdash !E} \qquad !E, !E, !E, \mathbf{bint}_A, !E \multimap E \vdash E}{!E, !E, !E, !E, \mathbf{bint}_A, \mathbf{bint}_A \vdash E} \multimap L$$

$$\dfrac{!E, !E, !E, !E, \mathbf{bint}_A, \mathbf{bint}_A \vdash E}{!E, !E, !E, \mathbf{bint}_A, \mathbf{bint}_A \vdash E} \mathrm{ctr}$$

$$\dfrac{!E, !E, !E, \mathbf{bint}_A, \mathbf{bint}_A \vdash E}{!E, !E, \mathbf{bint}_A, \mathbf{bint}_A \vdash E} \mathrm{ctr}$$

$$\dfrac{!E, !E, \mathbf{bint}_A, \mathbf{bint}_A \vdash E}{\mathbf{bint}_A, \mathbf{bint}_A \vdash \mathbf{bint}_A} \, 2\times \multimap R$$

$$\dfrac{\mathbf{bint}_A, \mathbf{bint}_A \vdash \mathbf{bint}_A}{!\mathbf{bint}_A, !\mathbf{bint}_A \vdash \mathbf{bint}_A} \, 2\times \ \mathrm{der}$$

$$\dfrac{!\mathbf{bint}_A, !\mathbf{bint}_A \vdash \mathbf{bint}_A}{!\mathbf{bint}_A \vdash \mathbf{bint}_A} \mathrm{ctr}$$

$$\underline{\text{repeat}}$$

$$\vdots$$

$$\dfrac{!\mathbf{bint}_A \vdash \mathbf{bint}_A}{!\mathbf{bint}_A, \mathbf{bint}_A \vdash \mathbf{bint}_A}\ \text{diff}$$

$$(\underline{S}, \underline{T}) \longmapsto \underline{ST} + \underline{TS}$$

$$(S + \varepsilon T)(S + \varepsilon T) = SS + \varepsilon(ST + TS) + \varepsilon^2 TT$$

Relation to calculus via coalgebras

- Following Ehrhard-Regnier we have defined derivatives in the syntax, via new deduction rules and cut-elimination rules.
- Do these syntactic derivatives capture the logical content lying behind the *semantic* derivatives?
- In particular, are they consistent with the role of Church numerals in Leibniz's stepped reckoner?
- **Yes**: because coalgebras

Algebras over a field k

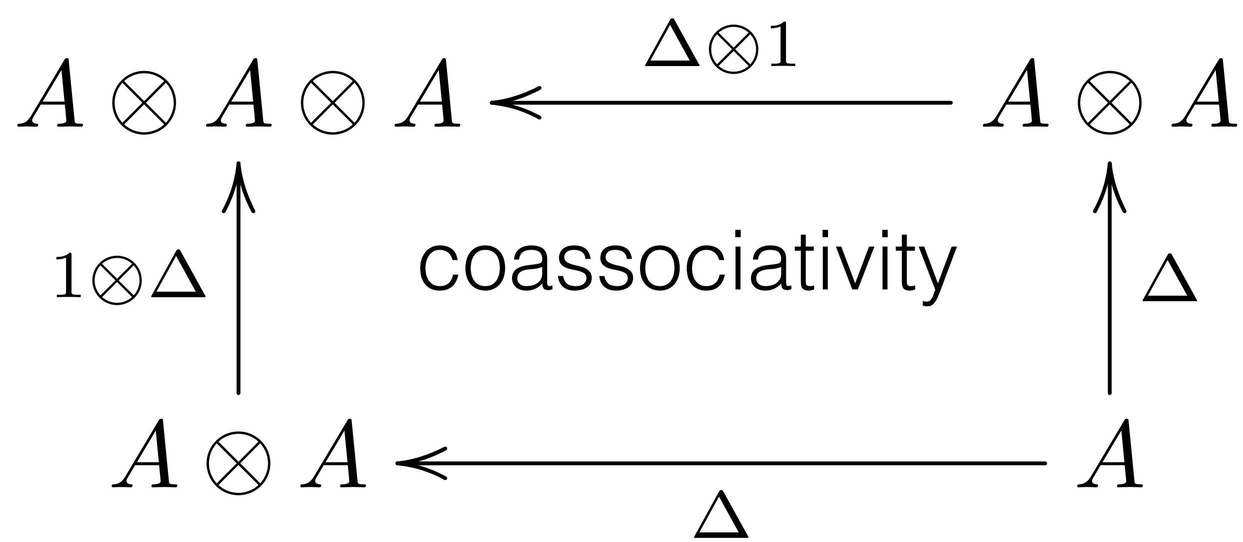
multiplication $m : A \otimes A \longrightarrow A$ $\quad$ $u : k \longrightarrow A$ unit

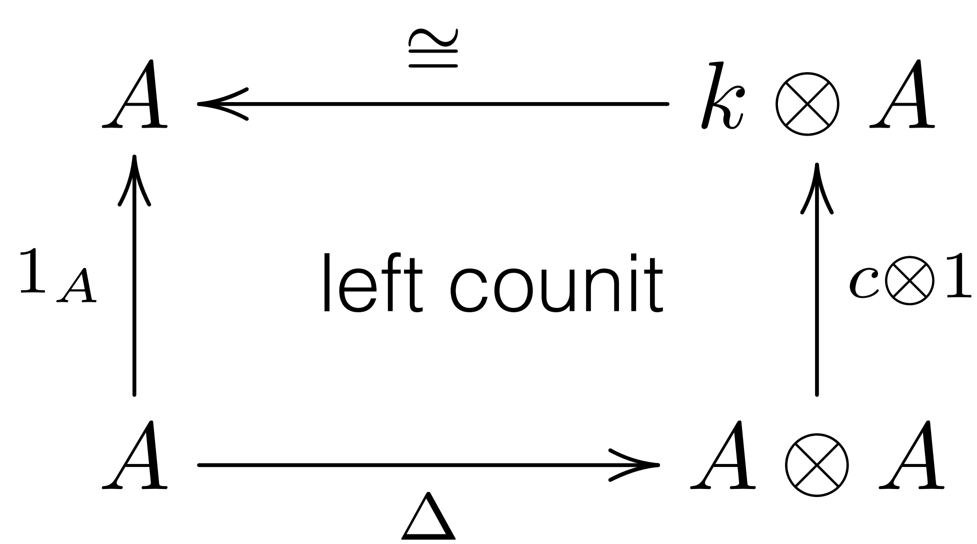
$$
\begin{array}{ccc}
A \otimes A \otimes A & \xrightarrow{m \otimes 1} & A \otimes A \\
{\scriptstyle 1 \otimes m}\big\downarrow & \text{associativity} & \big\downarrow{\scriptstyle m} \\
A \otimes A & \xrightarrow[m]{} & A
\end{array}
$$

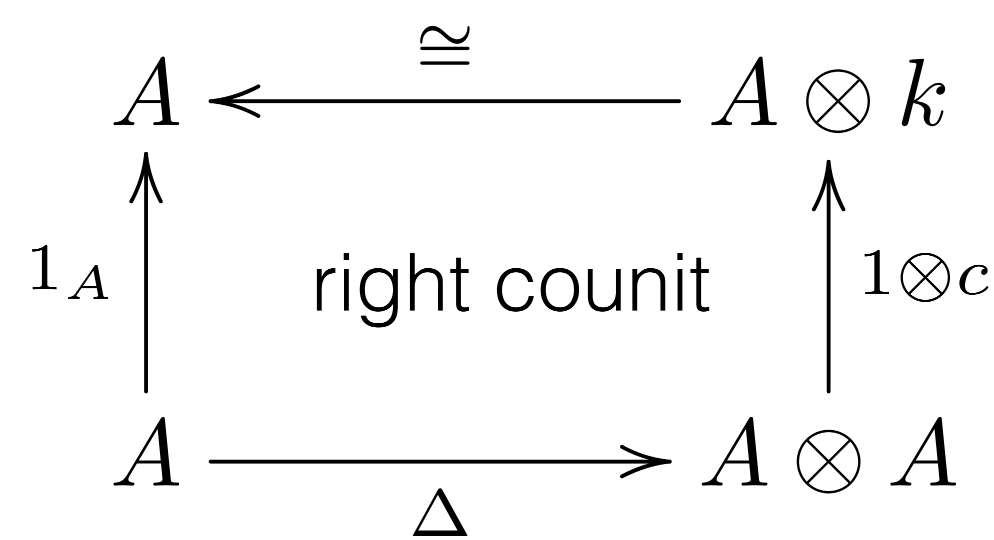
$$
\begin{array}{ccc}
A & \xrightarrow{\cong} & k \otimes A \\
{\scriptstyle 1_A}\big\downarrow & \text{left unit} & \big\downarrow{\scriptstyle u \otimes 1} \\
A & \xleftarrow[m]{} & A \otimes A
\end{array}
\qquad
\begin{array}{ccc}
A & \xrightarrow{\cong} & A \otimes k \\
{\scriptstyle 1_A}\big\downarrow & \text{right unit} & \big\downarrow{\scriptstyle 1 \otimes u} \\
A & \xleftarrow[m]{} & A \otimes A
\end{array}
$$

Coalgebras over a field k

comultiplication $\Delta : A \longrightarrow A \otimes A$ $\qquad$ $c : A \longrightarrow k$ counit







Examples

polynomial algebra

$$k[x_1, \ldots, x_n]$$

ring of dual numbers

$$k[\varepsilon]/(\varepsilon^2) = k \cdot 1 \oplus k \cdot \varepsilon$$

$$\varepsilon^2 = 0$$

polynomial coalgebra

$$k[x_1, \ldots, x_n]$$

$$\Delta(x^n) = \sum_{i=0}^{n} x^i \otimes x^{n-i}$$

dual of the ring of dual numbers

$$(k[\varepsilon]/(\varepsilon^2))^* = k \cdot 1^* \oplus k \cdot \varepsilon^*$$

$$\Delta(1) = 1 \otimes 1$$

$$\Delta(\varepsilon^*) = 1 \otimes \varepsilon^* + \varepsilon^* \otimes 1$$

Consider a morphism of k-algebras

$$k[x_1, \ldots, x_n] \xrightarrow{\varphi} k[\varepsilon]/(\varepsilon^2) = k \cdot 1 \oplus k \cdot \varepsilon$$

$$\varphi(x_i) = \lambda_i + \mu_i \varepsilon$$

it is straightforward to see that, for any polynomial f,

$$\varphi(f) = f(\lambda_1, \ldots, \lambda_n) + \sum_i \mu_i \frac{\partial f}{\partial x_i}\Big|_{x=\vec{\lambda}} \cdot \varepsilon$$

this gives rise to a bijection of k-algebra morphisms with pairs

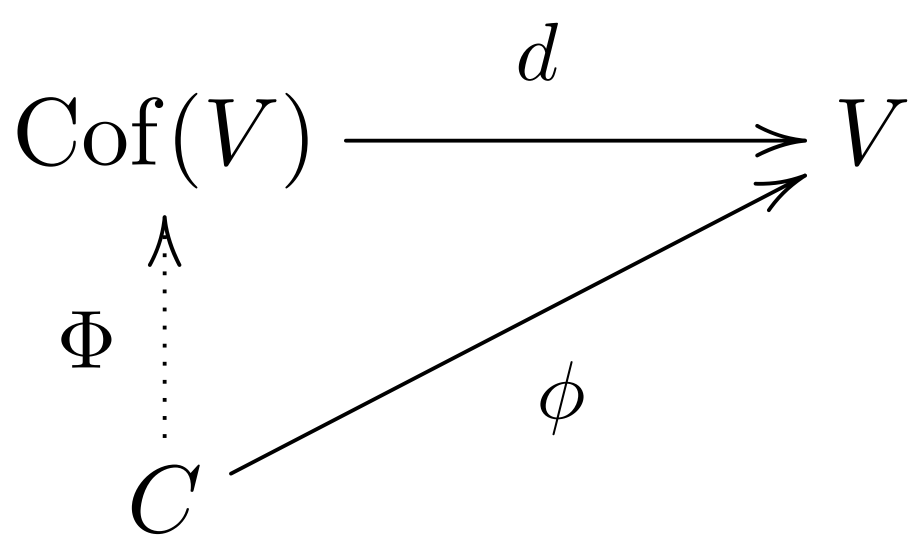
$$\mathrm{Hom}_{k-\mathrm{Alg}}(k[x_1, \ldots, x_n], k[\varepsilon]/(\varepsilon^2)) \xleftrightarrow{1:1} k^n \times k^n$$

$\varphi \longleftrightarrow (\vec{\lambda}, \vec{\mu})$ **(point, tangent vector)**

Universal coalgebra

The *cofree coalgebra* $\mathrm{Cof}(V)$ over a vector space V is a coalgebra together with a linear map $d : \mathrm{Cof}(V) \to V$ which is universal, in the sense that for any coalgebra C and linear $\phi : C \to V$ there unique morphism of coalgebras Φ such that

$$d \circ \Phi = \phi$$



Theorem: $\mathrm{Cof}(V)$ is the space of distributions with finite support on V, i.e. all derivatives of Dirac distributions

Sweedler semantics $[\![-]\!] : \mathrm{LL} \longrightarrow \mathrm{Vect}$

$$[\![A \multimap B]\!] = \mathrm{Hom}_k([\![A]\!], [\![B]\!])$$

$$[\![A \otimes B]\!] = [\![A]\!] \otimes [\![B]\!]$$

$$[\![!A]\!] = \mathrm{Cof}([\![A]\!])$$

dereliction = universal linear map $[\![!A]\!] \longrightarrow [\![A]\!]$

contraction = comultiplication $[\![!A]\!] \longrightarrow [\![!A]\!] \otimes [\![!A]\!]$

weakening = counit $[\![!A]\!] \longrightarrow k$

promotion = lifting of $[\![!A]\!] \longrightarrow [\![B]\!]$ to $[\![!A]\!] \longrightarrow [\![!B]\!]$

Sweedler semantics $[\![-]\!] : \mathrm{LL} \longrightarrow \mathrm{Vect}$

$$[\![A \multimap B]\!] = \mathrm{Hom}_k([\![A]\!], [\![B]\!])$$

$$[\![A \otimes B]\!] = [\![A]\!] \otimes [\![B]\!]$$

$$[\![!A]\!] = \mathrm{Cof}([\![A]\!])$$

The Sweedler semantics is also a semantics of differential linear logic, as follows:

$$\dfrac{\begin{array}{c}\pi \\ \vdots \\ !A \vdash B\end{array}}{!A, A \vdash B}\ \text{diff}$$

$$\begin{array}{ccccc}
[\![!A]\!] \otimes [\![A]\!] & \longrightarrow & [\![!A]\!] & \xrightarrow{[\![\pi]\!]} & [\![B]\!] \\
\| & & \| & & \\
\mathrm{Cof}([\![A]\!]) \otimes [\![A]\!] & \longrightarrow & \mathrm{Cof}([\![A]\!]) & &
\end{array}$$

$$D \otimes \nu \longmapsto \partial_\nu D$$

(point, tangent vector) $V \times V$ (λ, μ) $V = k^n$

$$\updownarrow \cong$$

$$\mathrm{Hom}_{k-\mathrm{Alg}}(k[x_1, \ldots, x_n], k[\varepsilon]/(\varepsilon^2))$$

$$\updownarrow \cong$$

$$\mathrm{Hom}_k(\mathrm{Sym}(V^*), k[\varepsilon]/(\varepsilon^2))$$

$$\updownarrow \cong$$

$$\mathrm{Hom}_k(V^*, k[\varepsilon]/(\varepsilon^2))$$

$$\updownarrow \cong$$

$$\mathrm{Hom}_k((k[\varepsilon]/(\varepsilon^2))^*, V)$$

$$\updownarrow \cong$$

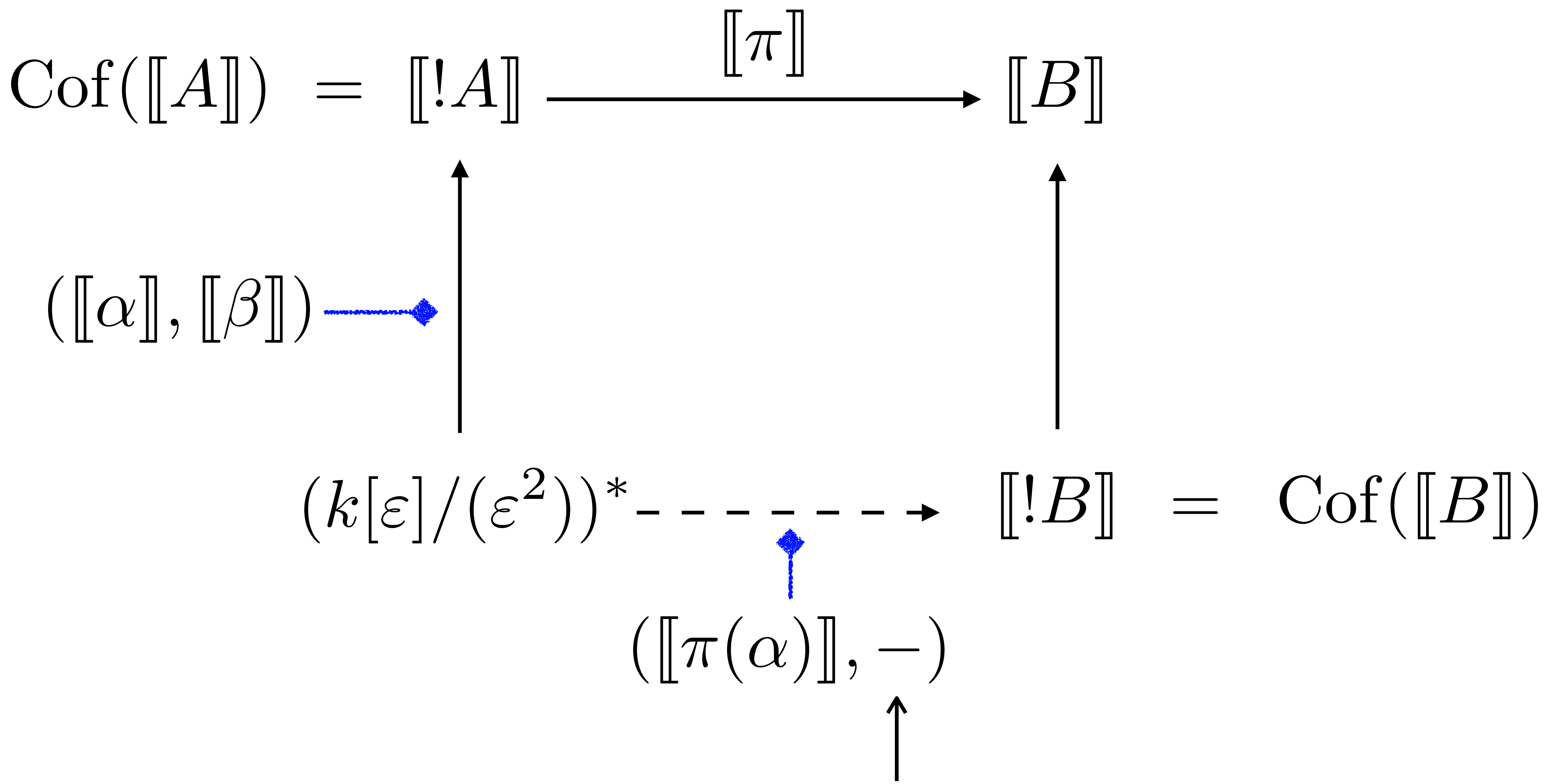
$$\mathrm{Hom}_{k-\mathrm{Coalg}}((k[\varepsilon]/(\varepsilon^2))^*, \mathrm{Cof}(V))$$

$1^* \mapsto \mathrm{Dirac}_\lambda$

$\varepsilon^* \mapsto \partial_\mu \, \mathrm{Dirac}_\lambda$

How to differentiate a proof denotation

Given $\pi : !A \multimap B$, $\alpha, \beta : A$ so that $[\![\alpha]\!], [\![\beta]\!] \in [\![A]\!]$



The directional derivative of π at α in the direction of β

Conciliation: syntax vs semantics

- The semantics of (intuitionistic, first-order) linear logic in vector spaces uses cofree coalgebras to model contraction, weakening and dereliction.
- Since the cofree coalgebra is made up of Dirac distributions and their derivatives, this semantics is naturally a model of *differential* linear logic.
- Linear logic secretly wants to be differentiated!

Conclusion/Questions

- Derivatives are natural in (linear) logic.
- Examples like the stepped reckoner suggest the use of calculus in logic is justified. Are there more convincing mechanical examples of this kind?
- The Sweedler semantics is a step in the direction of more interesting algebra and geometry. What is the logical content of distributions with more general support?
- Differential linear logic forms the basis for one approach to integrating symbolic reasoning with neural networks (work in progress with H. Hu).