

An introduction to statistical modelling semantics with higher-order measure theory

Ohad Kammar
University of Edinburgh

Scottish Programming Languages and Verification (SPLV'22)
Summer School
11–16 July, 2022
Heriot-Watt University



THE UNIVERSITY of EDINBURGH
informatics IfCS

Laboratory for Foundations
of Computer Science



supported by:



THE ROYAL
SOCIETY

The
Alan Turing
Institute

Facebook Research NCSC

Recap: Semantic Foundation for Statistics & Probability

finite discrete spaces

Points/
states $\{H, T\}$

Events $\{H\}, \{T\}$

Probability $\frac{3}{4}$ $\frac{1}{4}$

but also:
 $\emptyset \mapsto 0$
 $\{H, T\} \mapsto 1$

Recap: Semantic Foundation for Statistics & Probability

finite discrete spaces

Points/
states

$$\{H, T\}$$

more H
than T

$$\{H, T\}^3$$

Events

$$\{H\}, \{T\}$$

$$\{HHT, HTH, THH, HHH\}$$

$$= \bigcup_{i=1}^3 \{H \dots HTH \dots H\} \cup \{H \dots H\}$$

Probability

$$\frac{3}{4}$$

$$\frac{1}{4}$$

$$\sum_{i=1}^3 \frac{3^2 \times 1}{4^3}$$

$$+ \quad \frac{3^3}{4^3}$$

but also:
 $\emptyset \mapsto 0$
 $\{H, T\} \mapsto 1$

Countable discrete spaces

Points/
states

List $\{H, T\}$

more H
than T

Events

$\bigcup_{n=0}^{\infty} \bigcup_{i=\frac{r_n}{2}}^n \bigcup \{F_1, \dots, F_n\}$ where
 $\rho: \text{Fin } i \rightarrow \text{Fin } n$

$\begin{cases} i = p_j : F_i = H \\ \text{o.w.} : F_i = T \end{cases}$

list
length

Countable discrete spaces

Points/
states

List $\{H, T\}$

moke H
then T

Events

$$\bigcup_{n=0}^{\infty} \bigcup_{i=\frac{r_n}{2}}^n U \{ F_1, \dots, F_n \text{ where } \rho: Fin_i \rightarrow Fin_n$$

$$\begin{cases} i = p_j : F_i = H \\ \text{o.w.} : F_i = T \end{cases}$$

list
length

probability

$$\sum_{n=0}^{\infty} \sum_{i=\frac{r_n}{2}}^n$$

$$\sum_{\rho: Fin_i \rightarrow Fin_n} \frac{999}{1000^n} \cdot \frac{3^i \times 1^{n-i}}{4^n}$$

Continuous spaces

Points/
states

\mathbb{R}

Events Borel subsets $B_{\mathbb{R}}$

$$\text{Probability } E \mapsto \int_E \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}} + \sum_{\substack{n=1 \\ z_n \in E}}^{\infty} \frac{1}{z^n}$$

density
w.r.t.
measure

ω : Points $\rightarrow [0, \infty]$

$$\lambda + \sum_{n=1}^{\infty} \delta_{z_n}$$

Lebesgue ↳ Dirac

Continuous spaces

points/
states

\mathbb{R}

Events Borel subsets $B_{\mathbb{R}}$

$$\text{Probability } E \mapsto \int_E \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}} + \sum_{\substack{n=1 \\ z_n \in E}}^{\infty} \frac{1}{z^n}$$

density
w.r.t.
measure

w : Points $\rightarrow [0, \infty]$

$$\lambda + \sum_{n=1}^{\infty} \delta_{z_n}$$

Lebesgue ↳ Dirac

$$\{H, T\}^N$$

σ -algebra generated
by "open cylinders"
e.g. HTHT $\in \{H, T\}^N$

$$\frac{3^2 \times 1^2}{4}$$

Haar
measure

Discrete finite



Discrete Countable



Continuous

Quasi-Borel Spaces



~~Measurable~~

Agenda

I {
• Borel sets
• Qbs:

def., constructions,

Partiality, type structure

II {
• Measures & integration

see {
• Random variable spaces

CIRM
videos } • Conditional expectation

Slogan:

Measurable by Type

NB:

• Exercise sheets



• #qbs on SPLS

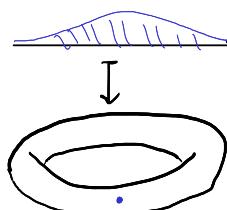
Zulip

Def: Quasi-Borel space $X = (X_1, \mathcal{R}_X)$

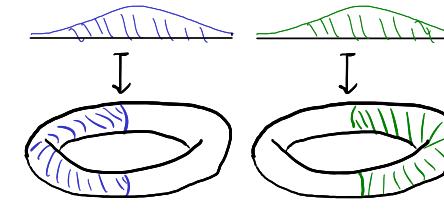
$$\mathcal{R}_X \subseteq L(X_1) \quad \text{Closed under:}$$

Set \curvearrowleft Set of
"carrier"
functions $\alpha: \mathbb{R} \rightarrow X_1$
"random elements"

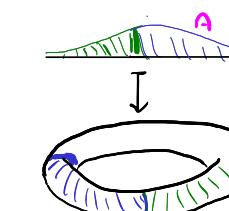
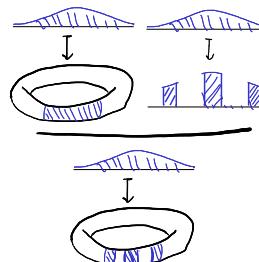
- Constants:



- recombination



- precomposition:



Examples

recombination of
constants

$$- \mathbb{R} = (\mathbb{R}, \text{Meas}(\mathbb{R}, \mathbb{R}))$$

qbs underlying \mathbb{R}

$$- X \in \text{set}, \quad \Gamma_X^{\text{Qbs}} := (X, \sigma\text{-simple}(\mathbb{R}, X))$$

$\lambda r.$ {
 rEA_n: x_n
 :
 :}

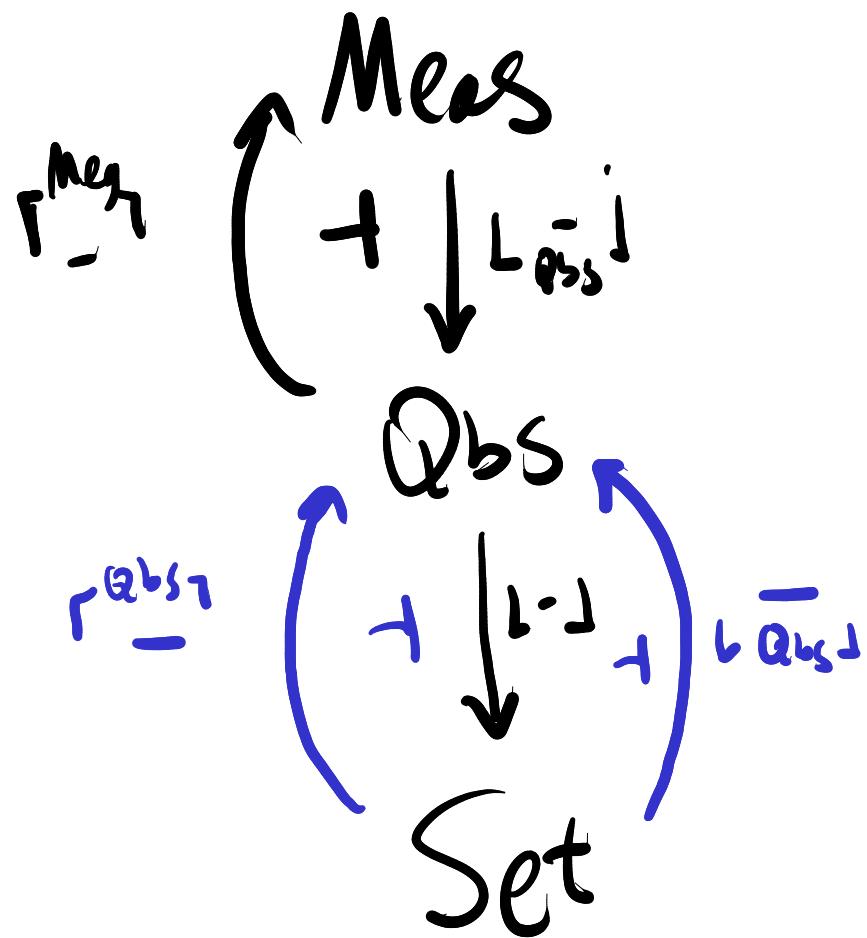
discrete qbs on X

$$- " \quad \Gamma_X^{\text{Qbs}} := (X, X^{|\mathbb{R}|})$$

all functions

Indiscrete qbs on X

Useful adjunctions:



$$\begin{aligned} \text{L-} \underset{\text{Qbs}}{\text{V}} &:= (\text{L-} \text{V}, \text{Meas}(\text{R}, \text{V})) \\ &\quad (\text{V} \in \text{meas}) \\ \Gamma^{\text{res}} X &:= \left\{ A \subseteq \text{L-} X \mid \forall \alpha \in R_X, \alpha^{-1}[A] \in \mathcal{B}_R \right\} \end{aligned}$$

- limits (products, subspaces)
and colimits (co-products, quotients)
- as in Set
- Slogan: every measurable space is carried by a qbs

Example

Product $(X \times Y, \pi_1, \pi_2)$:
necessarily!

$$- L[X \times Y] = L[X_1 \times_1 Y_1]$$

$$- R_{X \times Y} = \{ \lambda r. (\alpha r, \beta r) \mid \alpha \in R_X, \beta \in R_Y \}$$

correlated
random
elements

rest of structure as in Set.

Function Spaces

Straightforward!

- $\lfloor Y^X \rfloor := \text{Qbs}(X, Y)$

- $R_{Y^X} := \text{Uncurry}[\text{Qbs}(R^{XX}, Y)]$

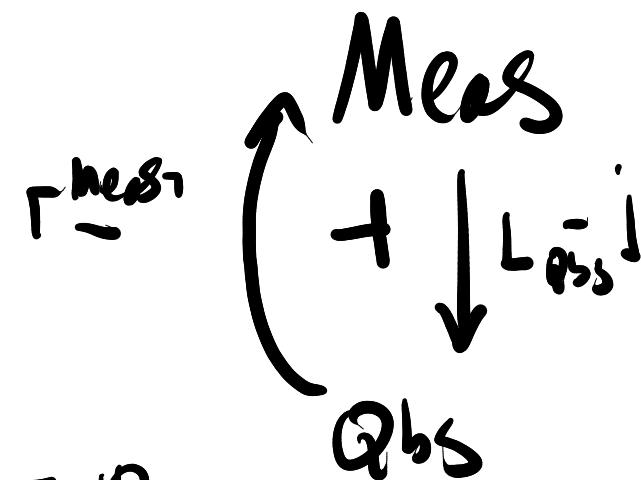
$$= \{ \alpha : R \rightarrow \lfloor Y^X \rfloor \mid \lambda(r, x). \alpha r x : R \times X \rightarrow Y \}$$

- $\text{eval} : Y^X \times X \rightarrow Y$
 $\text{eval}(f, x) := fx$

Meas vs Qbs

By generalities:

$$\begin{array}{c}
 \sigma\text{-algebra} \\
 \text{on } \text{Meas}(\mathbb{R}, \mathbb{R}) \\
 \text{Meas} \\
 \downarrow \\
 \mathbb{R} \times \mathbb{R} \rightarrow \text{Meas}(\mathbb{R}) \\
 \text{Meas} \\
 \downarrow \\
 \mathbb{R} \times \mathbb{R} \rightarrow \text{Meas}(\mathbb{R}) = \mathbb{R} \\
 \text{Meas} \\
 \downarrow \\
 \mathbb{R}^{\mathbb{R}} \times \mathbb{R} \neq \mathbb{R}^{\mathbb{R}} \times \mathbb{R}
 \end{array}$$



No factorisation
by Aumann's
Theorem.

Random Element Space

$R_X := X^R$ since $\lfloor X^R \rfloor = R_X$ as sets.

Why?

(\subseteq) $\alpha \in \lfloor X \rfloor^R \Rightarrow \alpha: \mathbb{R} \rightarrow X$ in Gbs.

$i\delta_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}$ measurable $\Rightarrow i\delta \in R_{\mathbb{R}}$

$\Rightarrow a = \alpha \circ i\delta \in R_X$

(\supseteq) $\alpha \in R_X \Rightarrow \forall \varphi \in R_{\mathbb{R}} = \text{Meas}(\mathbb{R}, \mathbb{R})$. $\alpha \circ \varphi \in R_X \Rightarrow \alpha: \mathbb{R} \rightarrow X$
 $\Rightarrow \alpha \in \lfloor X \rfloor^R$

Subspaces

For $X \in \text{Obs}$, $A \subseteq X$, set:

$$R_A := \{ \alpha: \mathbb{R} \rightarrow A \mid \alpha \in R_X \}$$

Then $A = (A, R_A)$ is the **Subspace qbs**

We write $A \hookrightarrow X$

Borel subspaces ensemble

The σ -algebra $B_X := \{A \subseteq X \mid \forall \alpha \in R_X . \alpha^*[A] \in B_R\}$

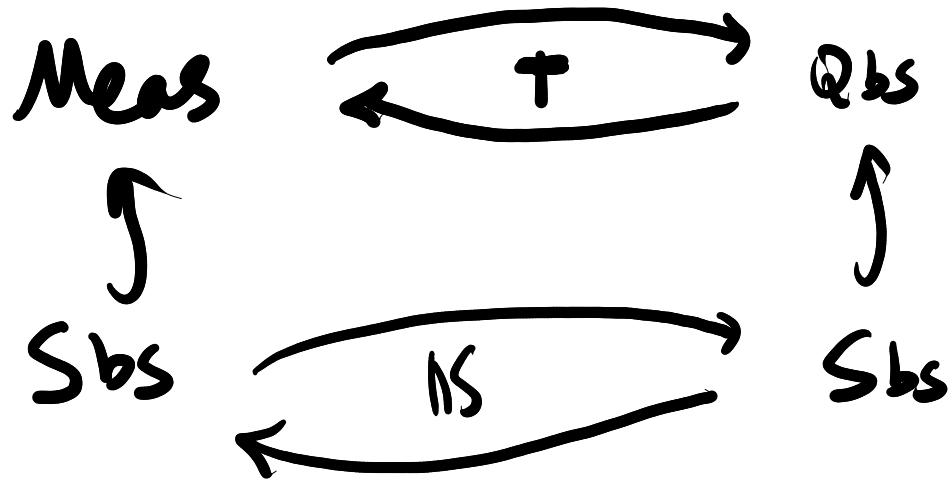
internalises as $B_X = 2^X$, the qbs of
Borel subsets.

$L^{(B_R)}$ are the Borel-on-Borel sets from descriptive set theory.
(cf. [Sabour et al. '21])

Standard Borel spaces

Def: A qbs S is **standard Borel** when

$$S \cong A \text{ for some } A \in \mathcal{B}_{\mathbb{R}}$$



Slogan: Qbs **Conservative extension** of Sbs

Example $C_0 := \{f: \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ continuous}\} \hookrightarrow \mathbb{R}^{\mathbb{R}}$

C_0 is sbs. (Well-known!)

Proof:

$$C_0 \in B_{\mathbb{R}^{\mathbb{Q}}}$$

↑ sbs!

$$C'_0 := \left\{ g \in \mathbb{R}^{\mathbb{Q}} \mid \forall a, b \in \mathbb{Q}, \varepsilon \in \mathbb{Q}^+ \exists \delta \in \mathbb{Q}^+ \forall p, q \in \mathbb{Q} \text{ s.t. } p, q \in [a, b] \text{ and } |p - q| < \delta \Rightarrow |g(p) - g(q)| < \varepsilon \right\}$$

then $C_0 \cong C'_0 \in B_{\mathbb{R}^{\mathbb{Q}}}$:

$$C_0 \rightarrow C'_0$$

$$\psi \mapsto \psi|_{\mathbb{Q}}$$

$$C'_0 \rightarrow C_0$$

$$\psi \mapsto \lambda r. \lim_{n \rightarrow \infty} g(\text{approx}_n^{(\frac{1}{m})_{m \in \mathbb{N}}})_n$$

on closed intervals
(= compact intervals)

continuity

uniform continuity

Borel measurable

by type checks

Example (ctd)

C_0 is sbs, and $\text{eval}: C_0 \times \mathbb{R} \rightarrow \mathbb{R}$
is a measurable.

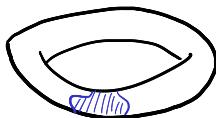
Avoids:

- Constructing complete separable metrics
- Proving that evaluation is measurable w.r.t. metric σ -algebra.

Agenda

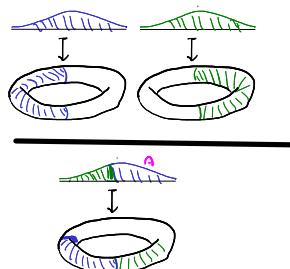
Slogan: Measurable by Type

- Borel sets



- Obs:

def., constructions,
Partiality, type structure



- Measures & integration
- Random variable spaces
- Conditional expectation

Partiality cf. [Väkär et al. '19]

A Borel embedding $e: X \hookrightarrow Y$

- injective function $e: [X] \rightarrow [Y]$
- its image is Borel: $e[X] \in \mathcal{B}_Y$
- e is Strong: $\alpha \in R_X \iff e \circ \alpha \in R_Y$

Examples

- $1 \hookrightarrow 2$
- S is sbs $\iff \exists S \subseteq \mathbb{R}$

Non-examples \sim [Sabok et al.'21]

$$-\left\{ A \in \mathcal{B}_{\mathbb{R}} \mid A \neq \emptyset \right\} \hookrightarrow \mathcal{B}_{\mathbb{R}}$$

$$-\left\{ (A_1, A_2) \in \mathcal{B}_{\mathbb{R}}^2 \mid A \subseteq B \right\} \hookrightarrow \mathcal{B}_{\mathbb{R}}^2$$

$$-\left\{ A \in \mathcal{B}_{\mathbb{R}} \mid A \text{ open} \right\} \hookrightarrow \mathcal{B}_{\mathbb{R}}$$

Def: A Partial map $f: X \rightarrow Y$ is a morphism

$$f: X \rightarrow Y \amalg \{\perp\}$$

Its domain of definition $\text{Dom } f := \{x \mid f_x \neq \perp\}$

$$\begin{array}{c} \uparrow \\ P \\ \downarrow \\ X \end{array}$$

Partial non-sets are ordered:

for $f, g: X \rightarrow Y$ $f \leq g$ When $\forall x. f_x \neq \perp \Rightarrow g_x = f_x.$

[Cockett-Lack '06]

A model of restriction categories / axiomatic domain theory
[Fiore-Plotkin '94] Base embeddings
are the admissible monads

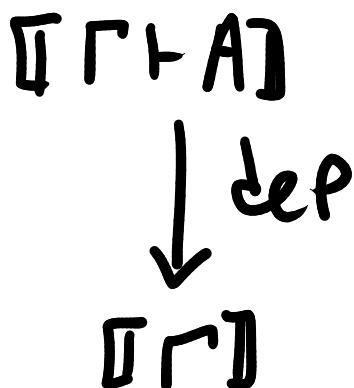
Type Structure

Simple types denote spaces

E.g.: $A \times B$ B^A B_A

Dependent types denote spaces-in-context

$\Gamma \vdash A$



Dependent types denote spaces-in-Content

$\Gamma \vdash \text{Content}$

$\Gamma \vdash A$

type in content

E.g.:

A

↓

1

simple types

$[\Gamma \vdash A]$

dep

$[\Gamma]$

assigns
environment

$[(U : B_A \vdash U)]$

$\{ (U, a) \in B_A^{X_A} \mid a \in U \}$

↓
 π_1

B_A

sbs decoder

Content extension

$$\frac{\Gamma \vdash A}{\Gamma, a:A \vdash}$$

$$\frac{\llbracket \Gamma \vdash A \rrbracket}{\llbracket \Gamma \rrbracket \quad \llbracket \Gamma, a:A \rrbracket := \llbracket \Gamma \vdash A \rrbracket}$$

$\downarrow \text{dep}$

Substitution

E.g. Weakening

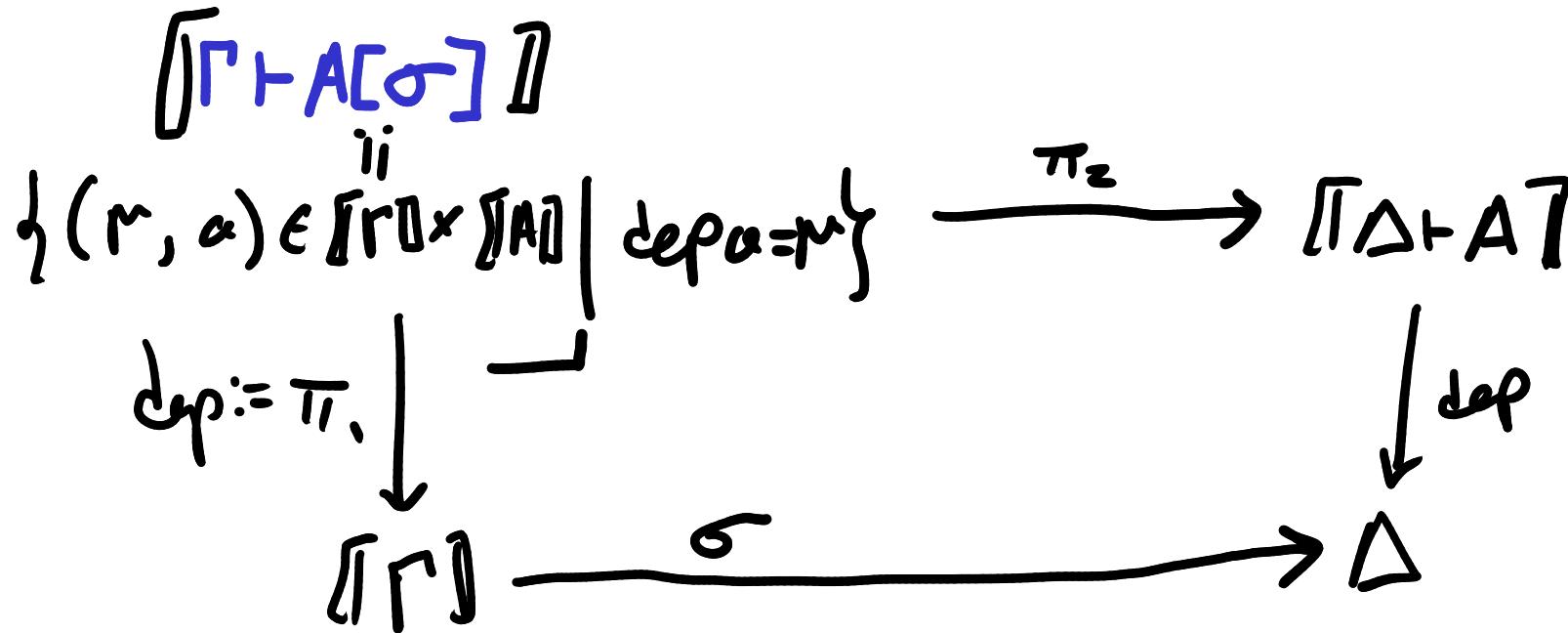
$$\Gamma \vdash \sigma : \Delta$$

$$\llbracket \sigma \rrbracket : \llbracket \Gamma \rrbracket \rightarrow \llbracket \Delta \rrbracket$$

$$\Gamma, a:A \vdash \text{wkn} : \Gamma$$

$$\llbracket \Gamma, a:A \rrbracket := \llbracket \Gamma \vdash A \rrbracket \xrightarrow[\text{dep}]{\text{wkn}} \llbracket \Gamma \rrbracket$$

Action of Substitution on types



E.g.

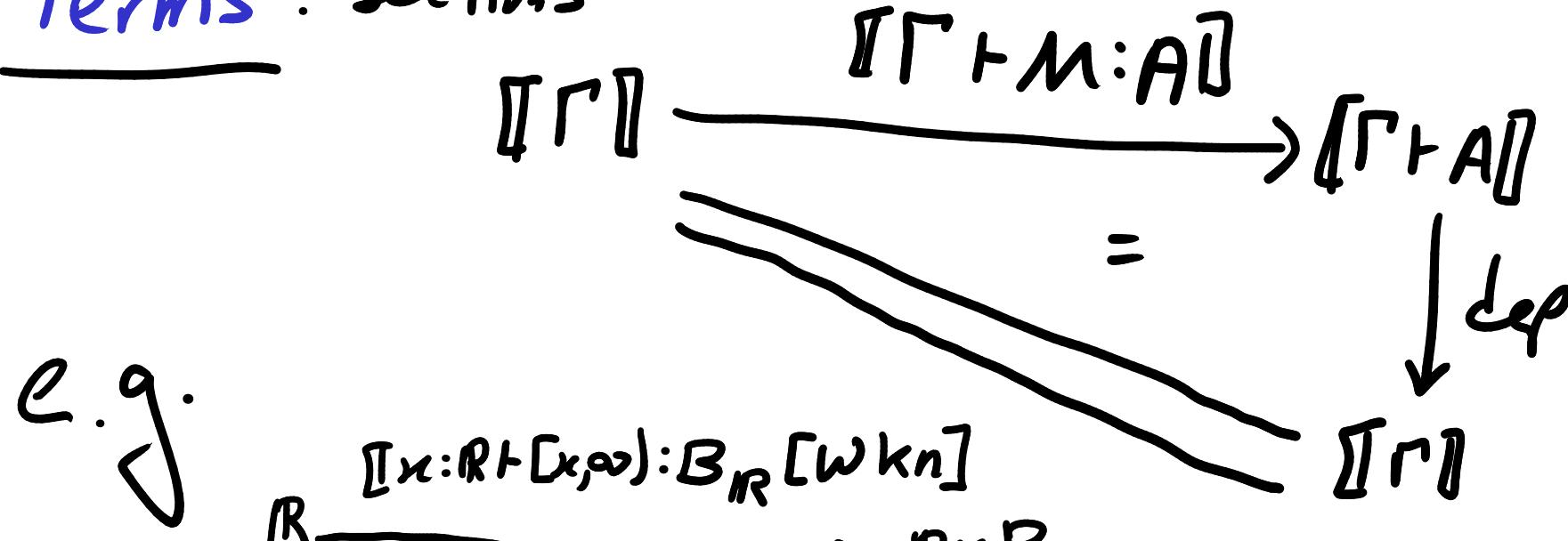
$$\boxed{\Gamma \vdash \lambda x:A. B[xm]} := A \times B \xrightarrow{\pi_2} B$$

$\downarrow \text{dep} := \pi_1$

$x:A \xrightarrow{\sim} 1$

Simple type

Terms : sections



e.g.

$$R \xrightarrow{[\lambda x : R \vdash [x, \alpha] : B_R \text{ [wkh]}]} R \times B_R$$

$$= \downarrow \pi_1$$

E.g. Variables: $\boxed{\Gamma, \alpha : A \vdash \alpha : A}$

$$\boxed{\Gamma, \alpha : A} \xrightarrow{< \text{id}, \text{dep}_{\Gamma \vdash A} >} \boxed{\Gamma, \alpha : A \vdash A \text{ [wkh]}}$$

$$= \downarrow \text{dep}$$

Exercise:

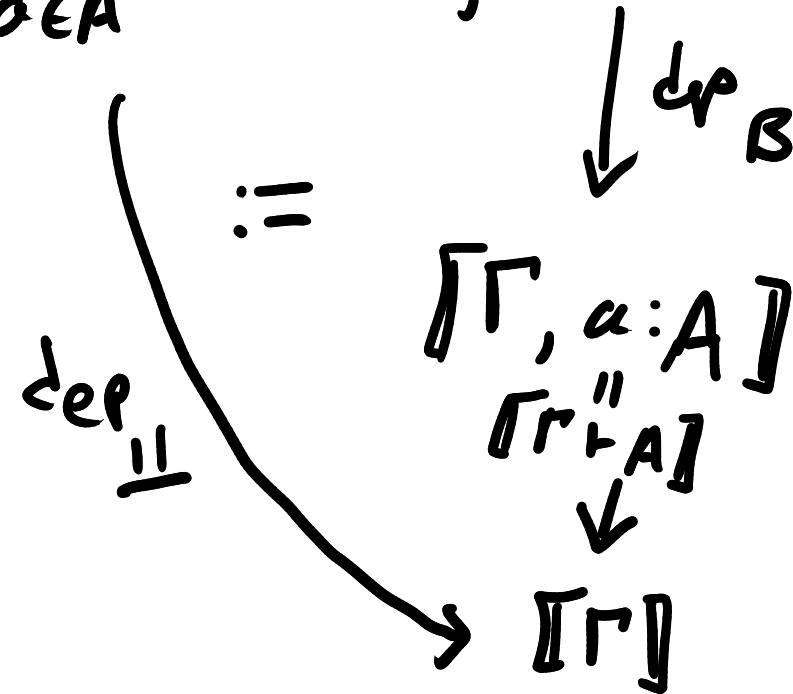
action of substitution

$M[\sigma]$

Dependent Pairs

$$\frac{\Gamma, a:A \vdash B}{\Gamma \vdash \prod_{a:A} B}$$

$$[\Gamma \prod_{a:A} A] := [\Gamma, a:A \vdash B]$$



Dependent Products

$$\frac{\Gamma, a:A \vdash B}{\Gamma \vdash \prod_{a:A} B}$$

$$\prod_{a:A} B$$

$$\lfloor \delta \Gamma \vdash \prod_{a:A} B \rfloor :=$$

$$\left\{ (r_0, f : \{ a \in \llbracket A \rrbracket \mid \text{dep } a = r_0 \} \rightarrow \llbracket \Gamma, a : A \vdash B \rrbracket) \middle| \right. \\ \left. \forall a \in \llbracket \Gamma, a : A \rrbracket . \text{dep } a = r_0 \Rightarrow \text{dep } (f a) = a \right\}$$

Exercise: find the random elements.

aha: $(a:A) \rightarrow B$

Example

($\Omega \in \text{Obs}$)

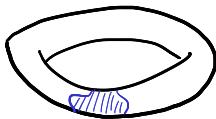
Converging $\hookrightarrow (\mathbb{R}^{[-\infty, \infty]^{\mathbb{N}}})^{\mathbb{N}}$

Converging := $\prod_{w \in \Omega} \{ \vec{f} \mid \exists \lim_{n \rightarrow \infty} f_n(w) \in \{$

Agenda

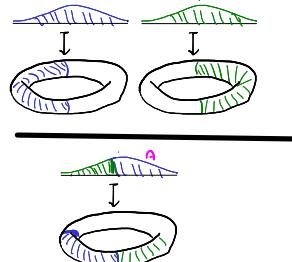
Slogan: Measurable by Type

- Borel sets



- Obs:

def., constructions,



Partiality, refinement $\rightarrow, \Leftarrow, \mathbb{H}, \mathbb{T}$

- Measures & integration
- Random variable spaces
- Conditional expectation

Def: A measure μ over \mathbb{R} is a function

$$\mu : \mathcal{B}_{\mathbb{R}} \rightarrow \mathbb{W} := [0, \infty]$$

s.t. - $\mu \emptyset = 0$

- $A \in \mathcal{B}_{\mathbb{R}}^{\mathbb{N}}$ $A_n \cap A_m = \emptyset$
 $(n \neq m)$

$$\mu \left(\bigcup_{n=0}^{\infty} A_n \right) = \sum_{n=0}^{\infty} \mu A_n$$

For measurable spaces, replace \mathbb{R} with V

We write $\mathcal{L}(V)$ for the set of measures on V

For abs X , take $\mathcal{L}(G^{\text{meas}} X)$

The unrestricted Giry space

Equip $\llcorner GV \lrcorner$ with

$$R_{GV} := \left\{ \alpha: R \rightarrow GV \mid \forall A \in \mathcal{B}_V. \lambda r. \alpha(r, A): R \rightarrow W \right\}$$

↗ α is a kernel.

Farewell Meas

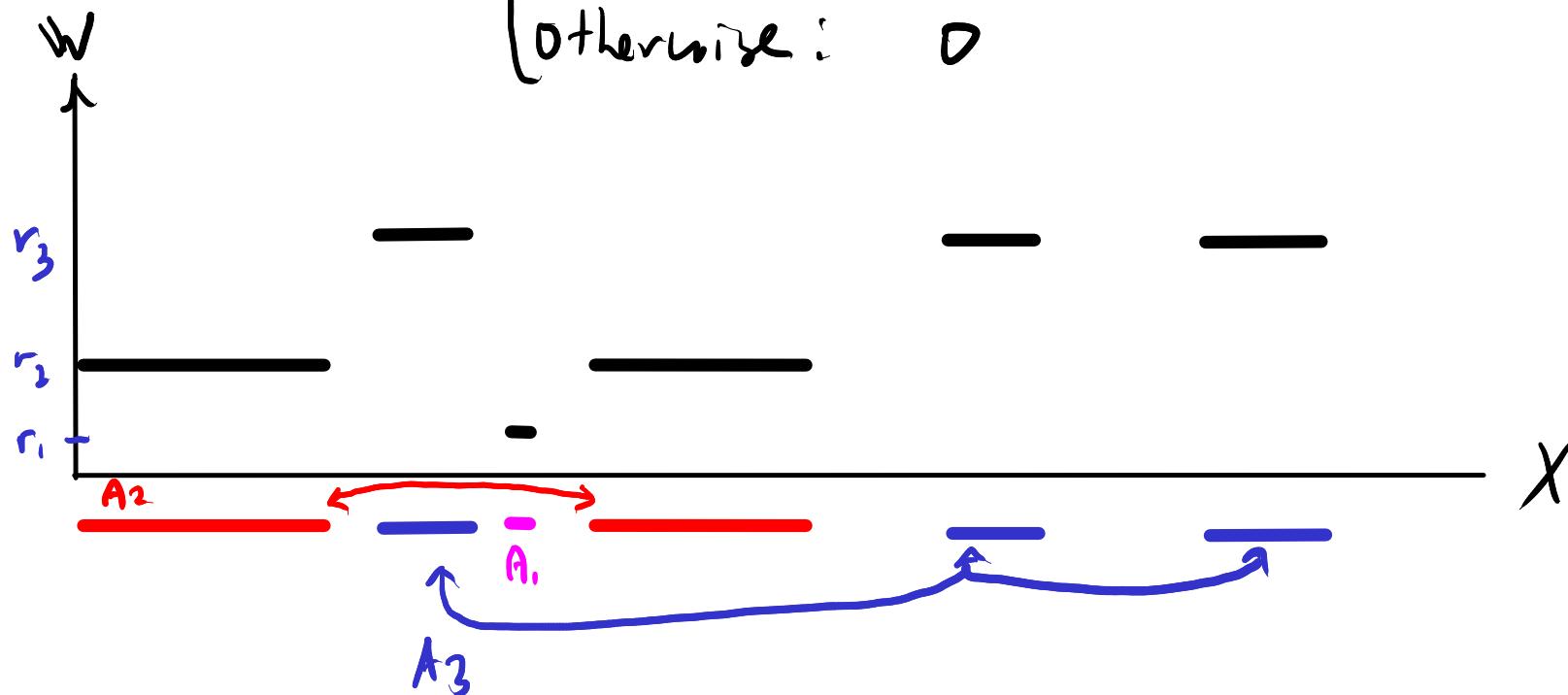
Now on:

1. All spaces are quasi-Borel
2. "measurable function" meas qbs morphism!

Def: Simple function $\varphi: X \rightarrow W$ when

$\exists n \in \mathbb{N}, \vec{A} \in \mathcal{B}_X^n, A_i \cap A_j = \emptyset, r_i \in W$ s.t.
 $(i \neq j)$

$$\varphi(x) = \begin{cases} \vdots & \vdots \\ x \in A_i & r_i \\ \vdots & \vdots \\ \text{otherwise: } & 0 \end{cases}$$



Encoder into a space:

$$\text{SimpleCode} := \bigsqcup_{n \in \mathbb{N}} \mathcal{B}_X^n \times \mathcal{W}^n$$

$$\text{Simple} := \{ f \in \mathcal{W}^X \mid f \text{ simple} \} \hookrightarrow \mathcal{W}^X$$

and define an interpretation:

$$[\![\cdot]\!]: \text{SimpleCode} \longrightarrow \text{Simple}$$

$$[\![(\mathbf{n}, \vec{\mathbf{A}}, \vec{r})]\!] := \sum_{i=1}^n r_i \cdot [\![\cdot \in A_i]\!]$$

↳ characteristic function
for A_i

Lemma: $f: X \rightarrow W$ is measurable → remember!
continuous
morphisms!

iff $f = \lim_{n \rightarrow \infty} f_n$ for some monotone sequence

$\stackrel{\rightarrow}{f} \in \text{Simple}$.

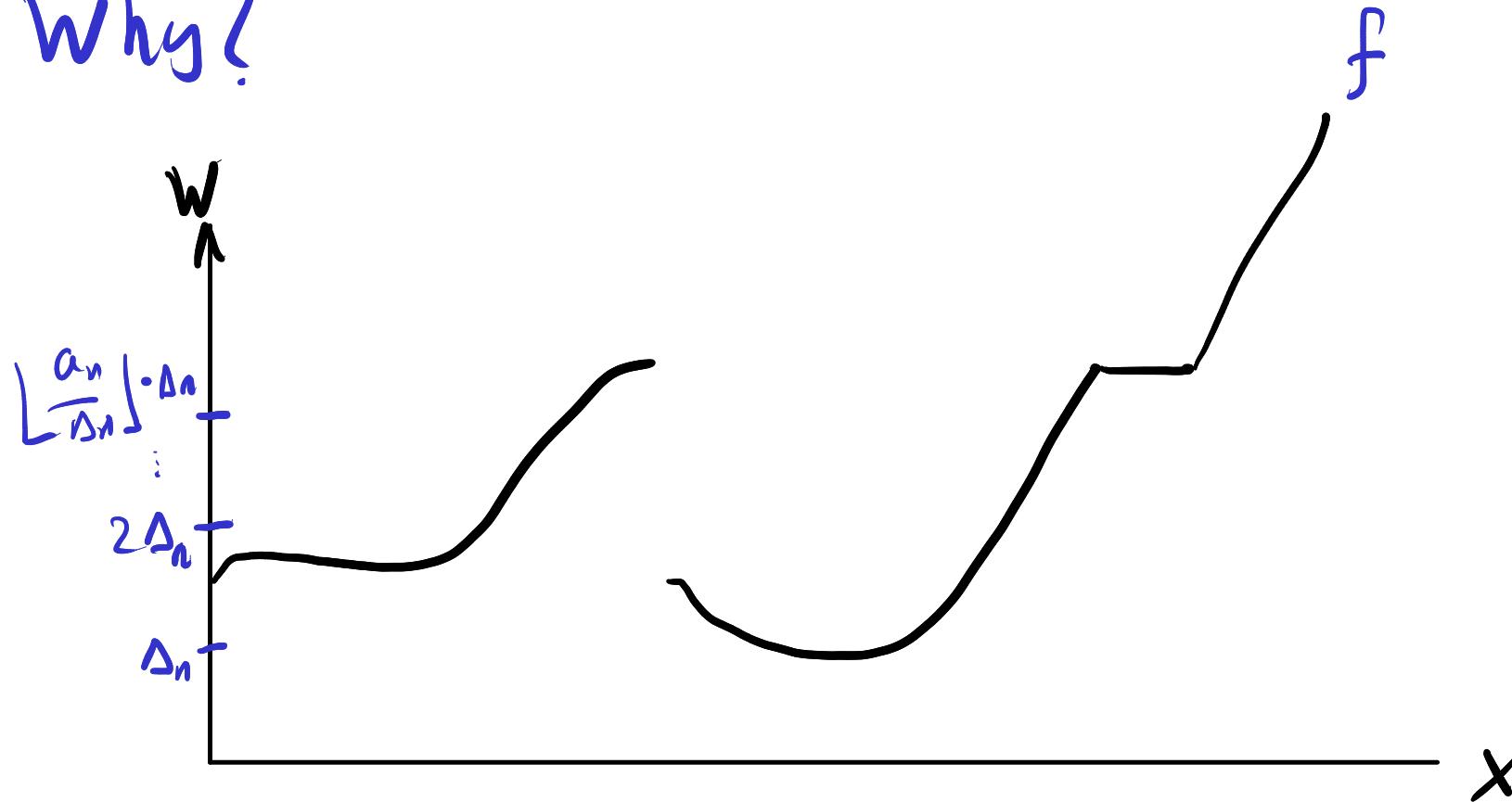
Moreover, we have measurable such choice:

Simple Approx:

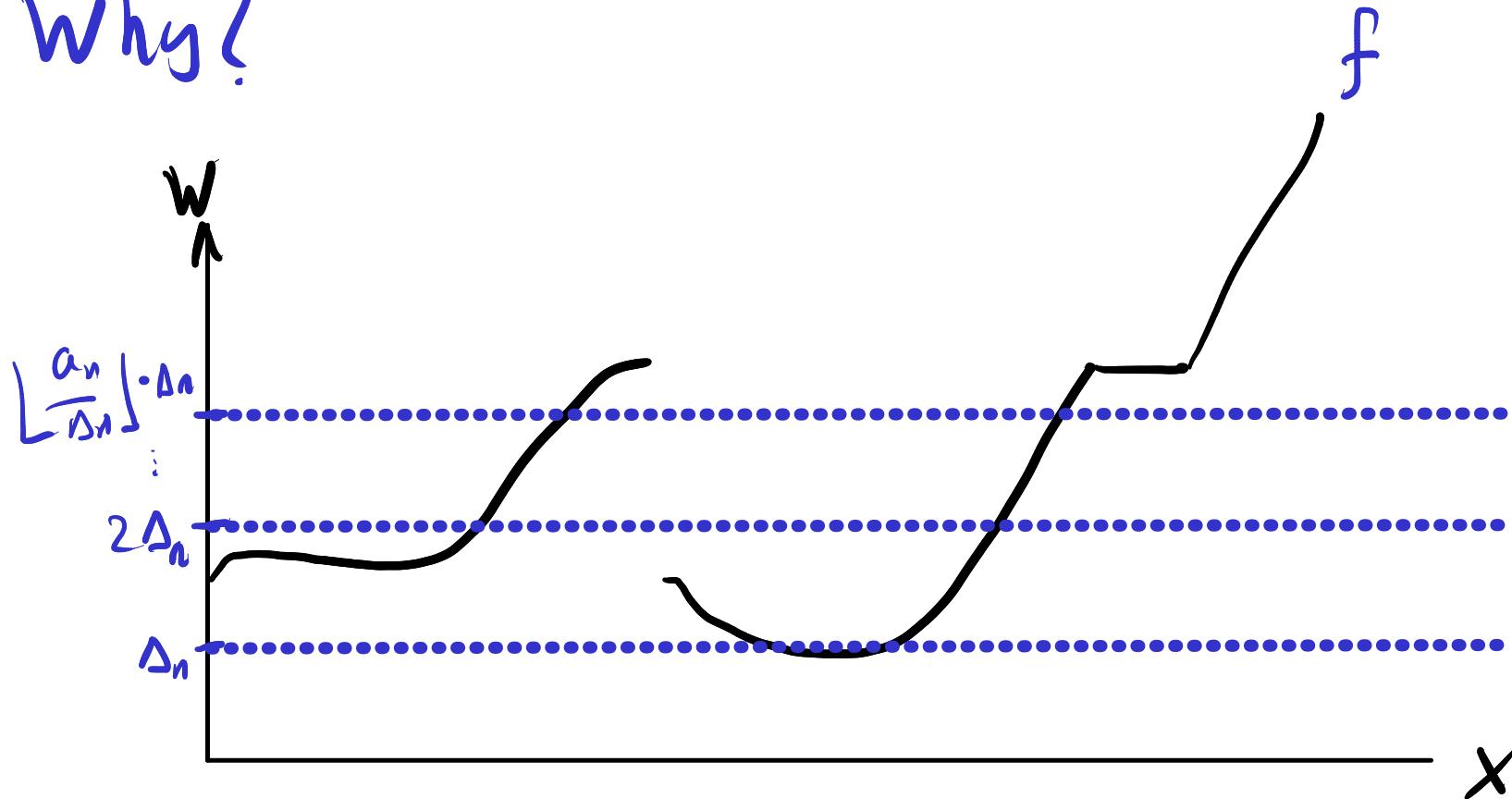
$$\left\{ \vec{\Delta} \in \mathbb{R}^+ \mid \Delta \rightarrow 0 \right\} \times \left\{ \vec{a} \in W^{\mathbb{N}} \mid \begin{array}{l} \vec{a} \text{ monotone} \\ a_n \rightarrow \infty \end{array} \right\} \times W^X \rightarrow \text{SimpleCode}$$

\uparrow
rate of convergence
 \uparrow
range of approximation

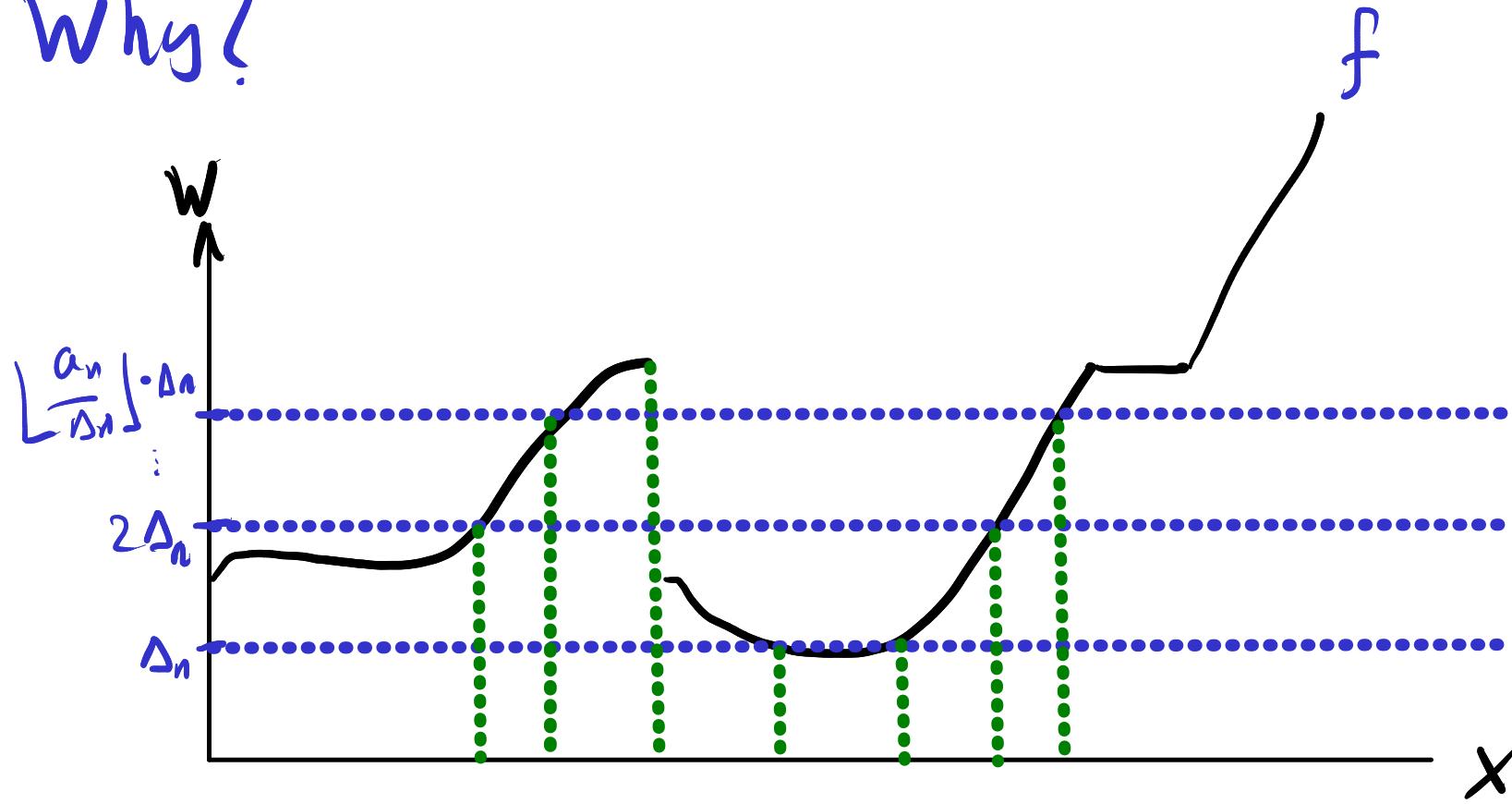
Why?



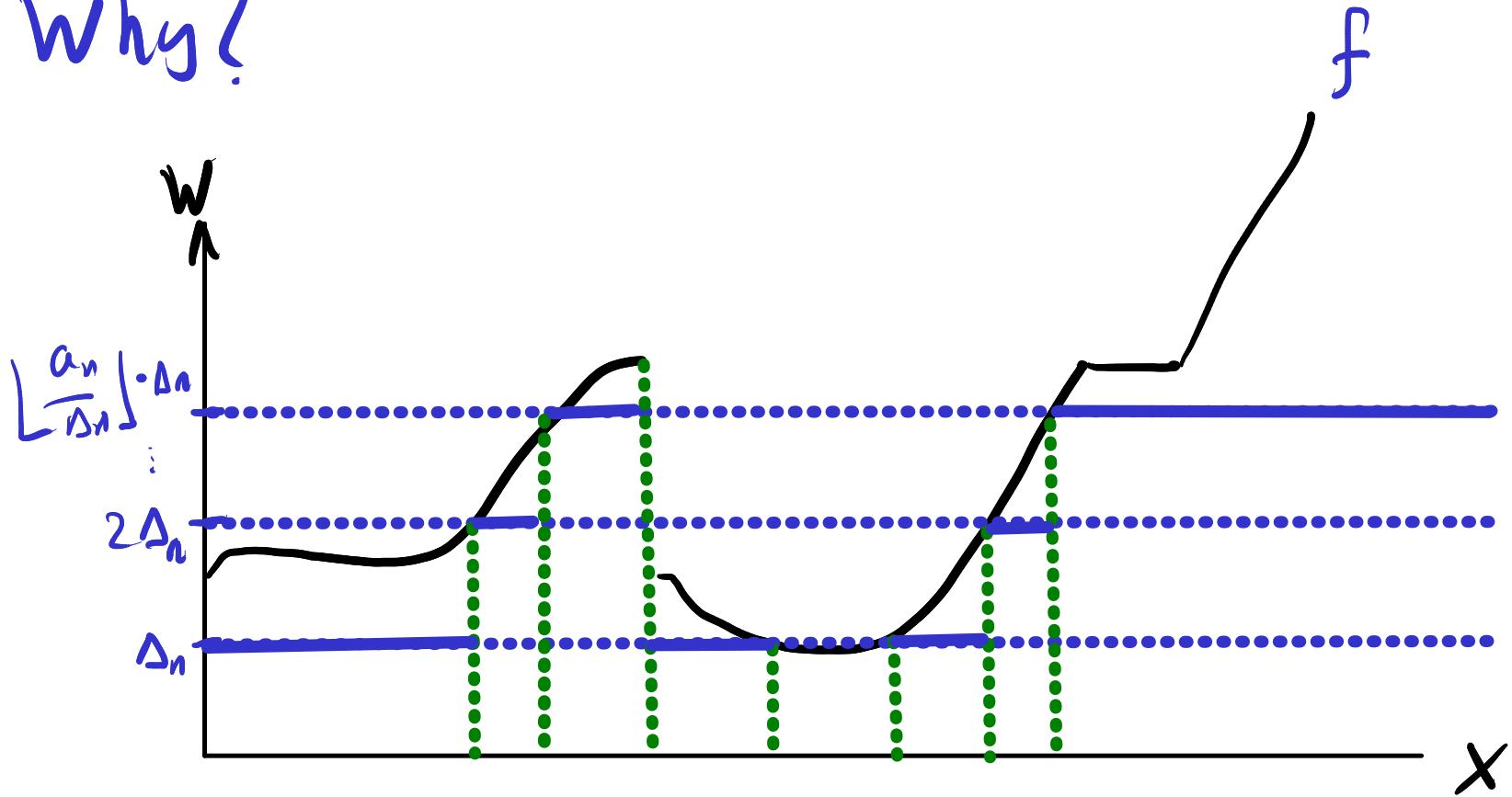
Why?



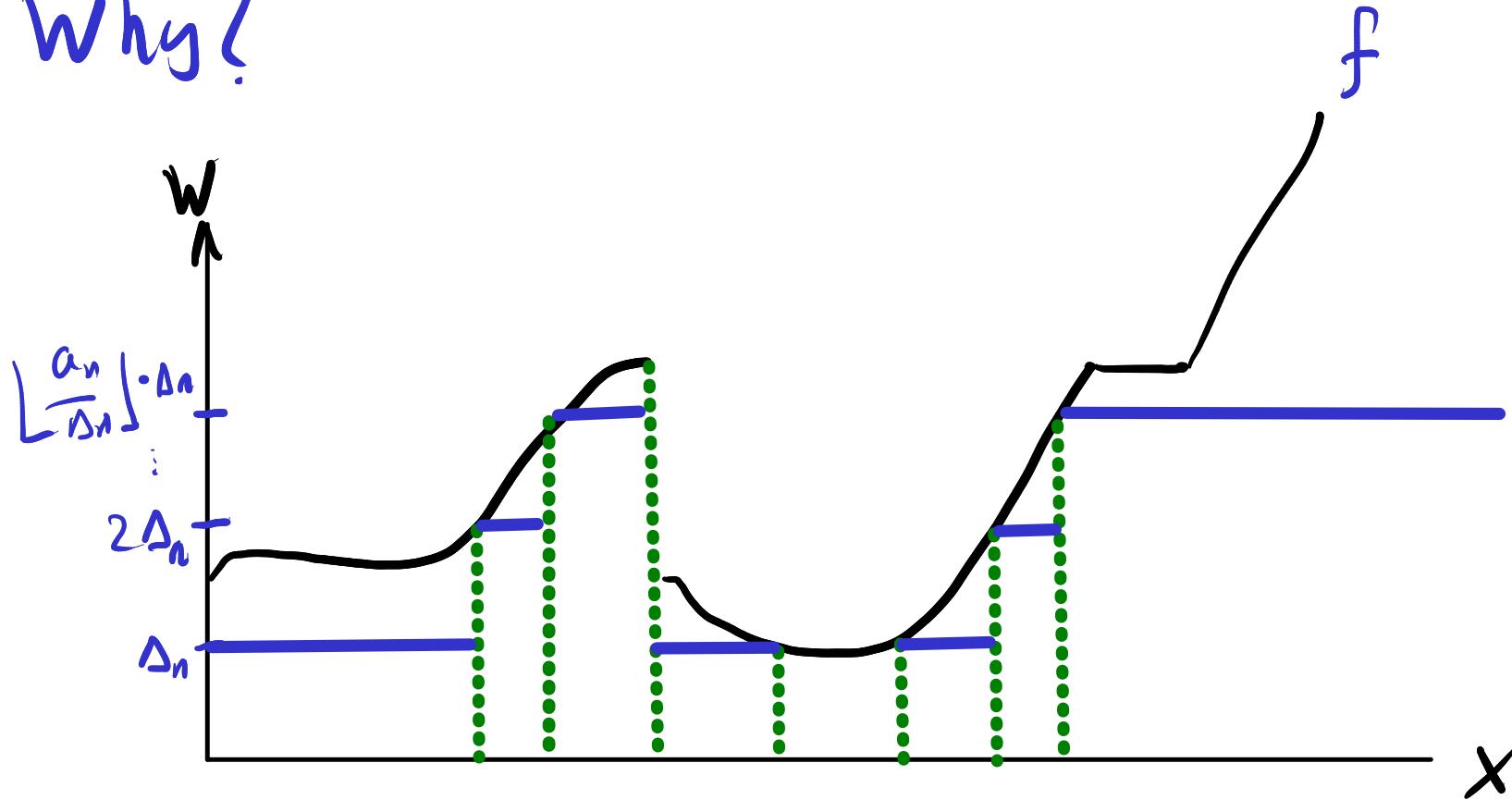
Why?



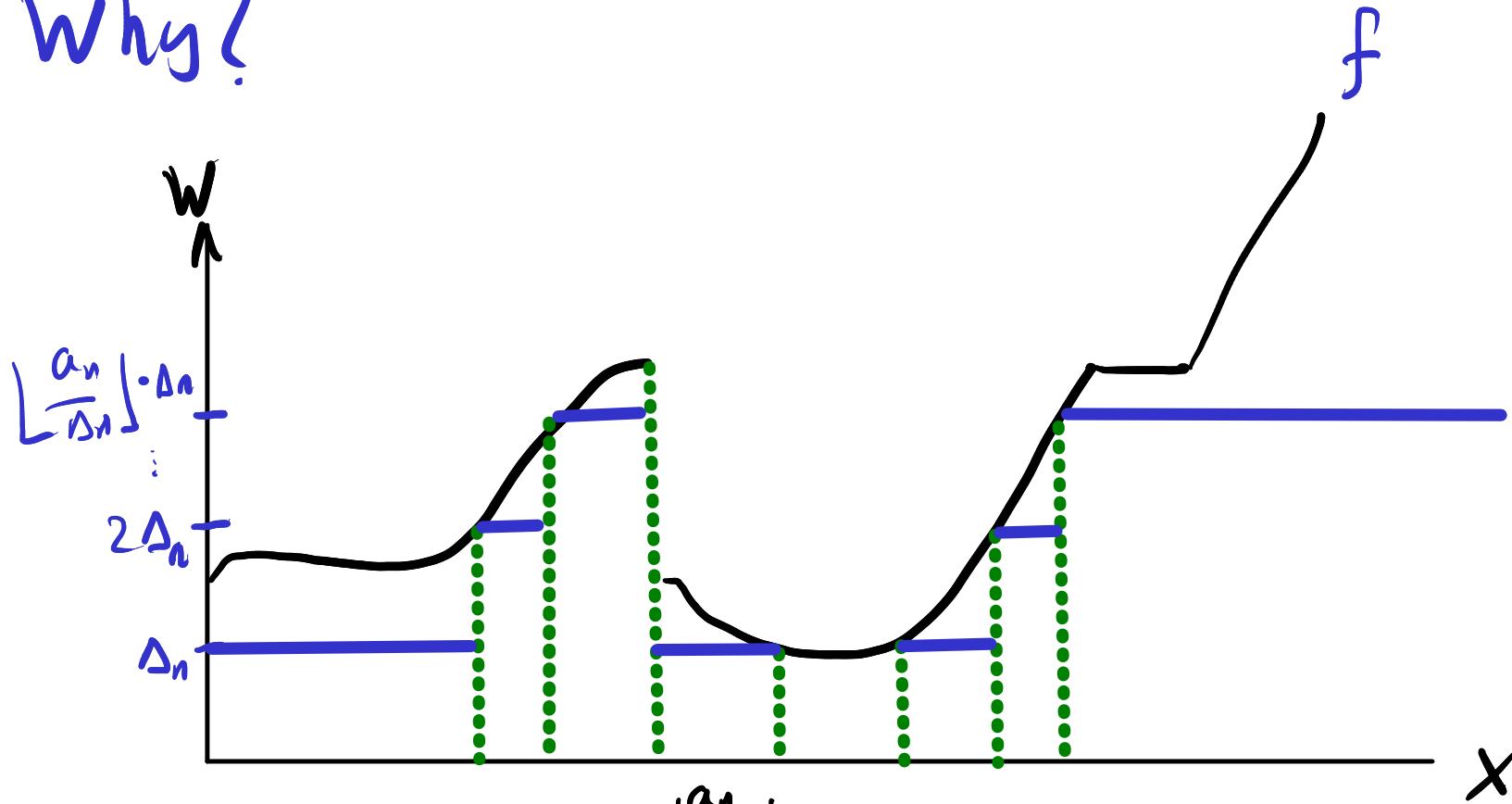
Why?



Why?

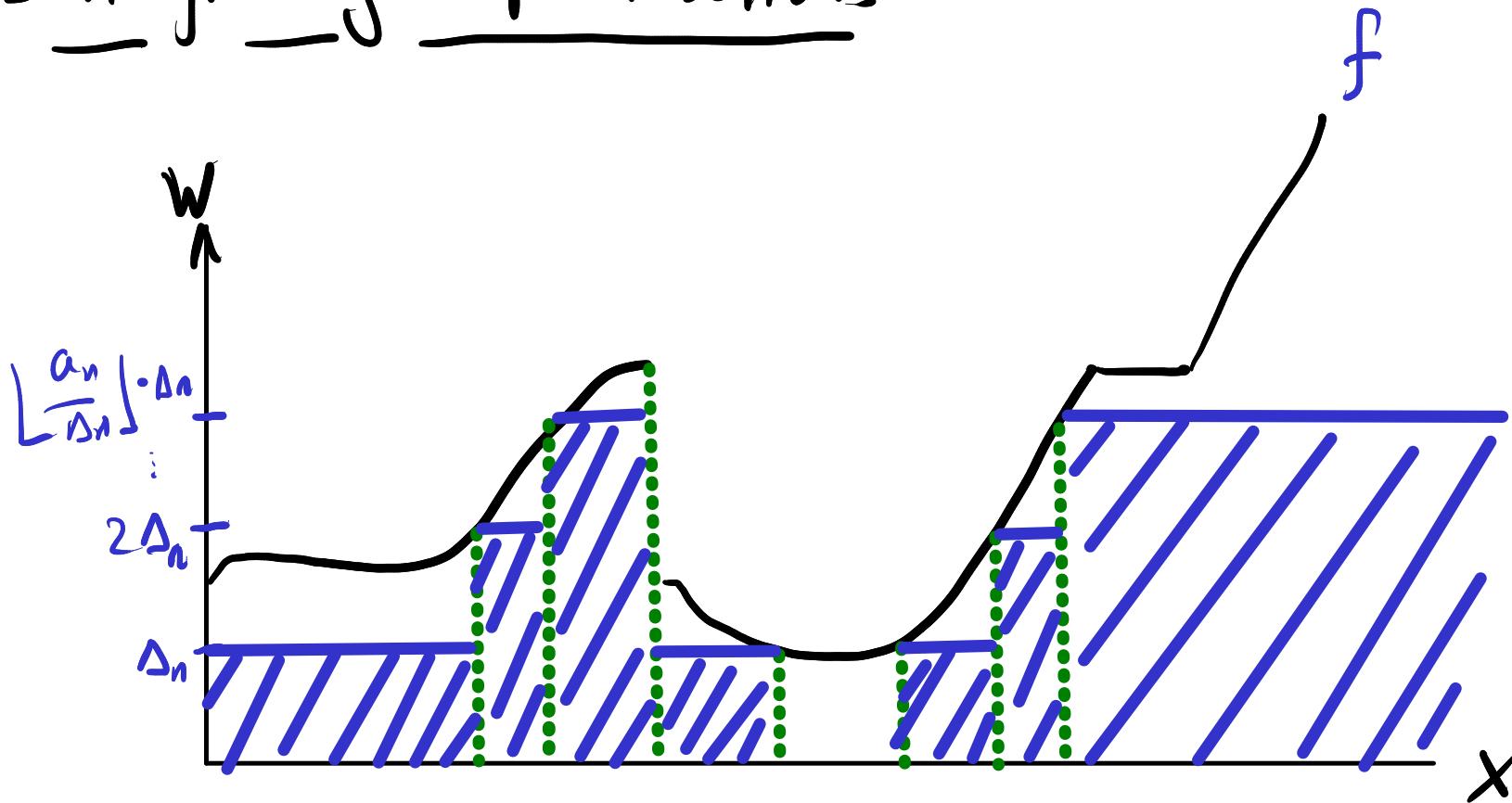


Why?



$$\left[\text{Simple Approx} \xrightarrow{\Delta, \alpha} f \right] := \sum_{i=1}^{\lfloor \frac{a_n}{\Delta_n} \rfloor} i \cdot \Delta_n [i \cdot \Delta_n \leq f < (i+1) \Delta_n] + \left[\frac{a_n}{\Delta_n} \Delta_n \cdot [f \geq \lfloor \frac{a_n}{\Delta_n} \rfloor \cdot \Delta_n] \right] \in \text{Simple}$$

Integrating Simple Functions



$\int : G X \times \text{Simple Code} \rightarrow W$

$$\int \mu(n, \vec{A}, \vec{r}) := \sum_{I \subseteq \{1, \dots, n\}} \left(\sum_{i \in I} r_i \right) \cdot \mu \left(\bigcap_{i \in I} A_i \setminus \bigcup_{i \notin I} A_i \right)$$

Integration



Property higher-order operation

$$\int : Gx \times W^X \rightarrow W$$

$$\int \mu f := \sup \left\{ \int \mu q \mid q \in \text{Simple}, \quad q \leq f \right\}$$

measurable by type

$$= \lim_{n \rightarrow \infty} \int \mu (\text{Simple Approx}_{\vec{\Delta}, \vec{a}} f)_n$$

we also write

$$\int \mu(dx) t$$

for $\int \mu(\lambda x, t)$

$$\text{for } \frac{a_n}{\Delta_n} \rightarrow 0, \text{ e.g. } \Delta_n = \frac{1}{2^n}, \quad a_n = n.$$

resolution