A polarity theory for sets of desirable gambles

Alessio Benavoli / Alessandro Facchini / Marco Zaffalon

(IDSIA, Lugano - Switzerland)

José Vicente-Perez

(U. of Alicante, Spain)

ISIPTA 2017

Lugano, 10-14 July 2017

Message of this talk

- Sets of lexicographic probability systems stand to sets of desirable gambles just like sets of probabilities stand to sets of almost desirable gambles
- The key ingredient of the correspondence is polarity (duality) theory for convex sets

The case of almost desirability

- Sets of lexicographic probability systems stand to sets of desirable gambles just like sets of probabilities stand to sets of almost desirable gambles
- Closed convex sets of probabilities (credal sets) and coherent sets of almost desirable gambles constitute equivalent theories

The case of almost desirability

- Sets of lexicographic probability systems stand to sets of desirable gambles just like sets of probabilities stand to sets of almost desirable gambles

- Closed convex sets of probabilities (credal sets) and coherent sets of almost desirable gambles constitute equivalent theories/models

The case of almost desirability

- Sets of lexicographic probability systems stand to sets of desirable gambles just like <u>sets of probabilities stand to sets of almost desirable gambles</u>

- Closed convex sets of probabilities (credal sets) and coherent sets of almost desirable gambles constitute equivalent theories/models

- We assume that the set of outcomes Ω of an experiment (e.g., coin tossing) is finite, say $\Omega:=\{1,.., n\}$.

- A gamble on the experiment is thence a vector $g:=(g(1), ..., g(n))$ in the real vector space $\mathbb{R}^n$, where $g(i)$ represents the reward the gambler would obtain if i is the actual outcome of the experiment.

The case of almost desirability

Definition: A set $K \subseteq \mathbb{R}^n$ is called a **coherent set of almost desirable gambles** if it satisfies

[*linearity*] if $f,g \in K$, then $\mu f+\nu g \in K$, for $\mu,\nu>0$

[*accepting partial gain*] if $g > 0_n$ then $g \in K$

[*avoiding sure loss*] $-1_n \notin K$

[*closure*] if $g+f \in K$, for all $f > 0_n$, then $g \in K$

- a coherent set of almost desirable gambles is a *closed convex cone containing the origin and all positive vectors.*

Definition: A set $K \subseteq \mathbb{R}^n$ is called a coherent set of almost desirable gambles if it satisfies

[*linearity*] if $f,g \in K$, then $\mu f+\nu g \in K$, for $\mu,\nu>0$

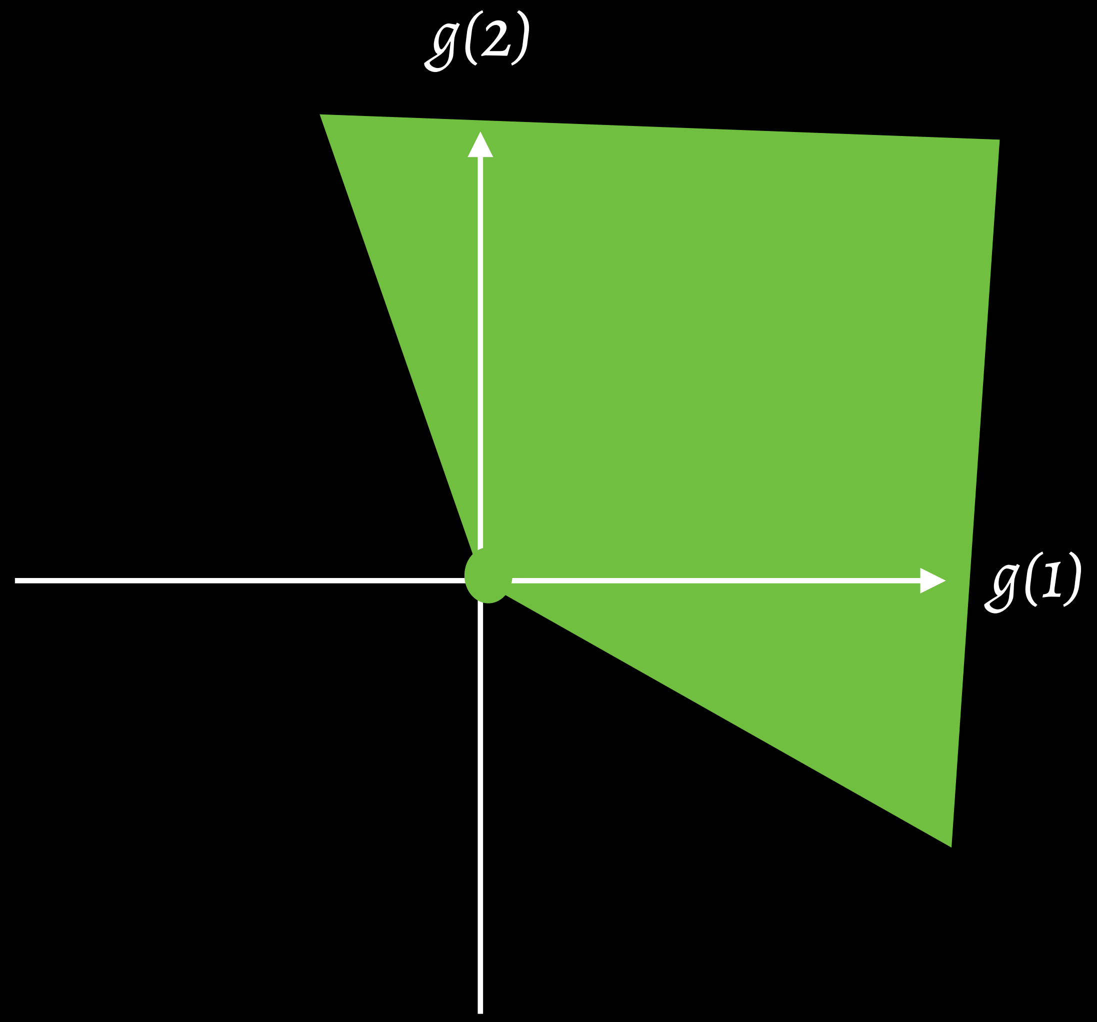
[*accepting partial gain*] if $g > 0_n$ then $g \in K$

[*avoiding sure loss*] $-1_n \notin K$

[*closure*] if $g+f \in K$, for all $f > 0_n$, then $g \in K$

- a coherent set of almost desirable gambles is a *closed convex cone containing the origin and all positive vectors.*

The case of almost desirability



- Closed convex sets of probabilities (credal sets) and coherent sets of almost desirable gambles constitute equivalent theories/models

*What does it mean for two **theories/**models to be **equivalent**?*

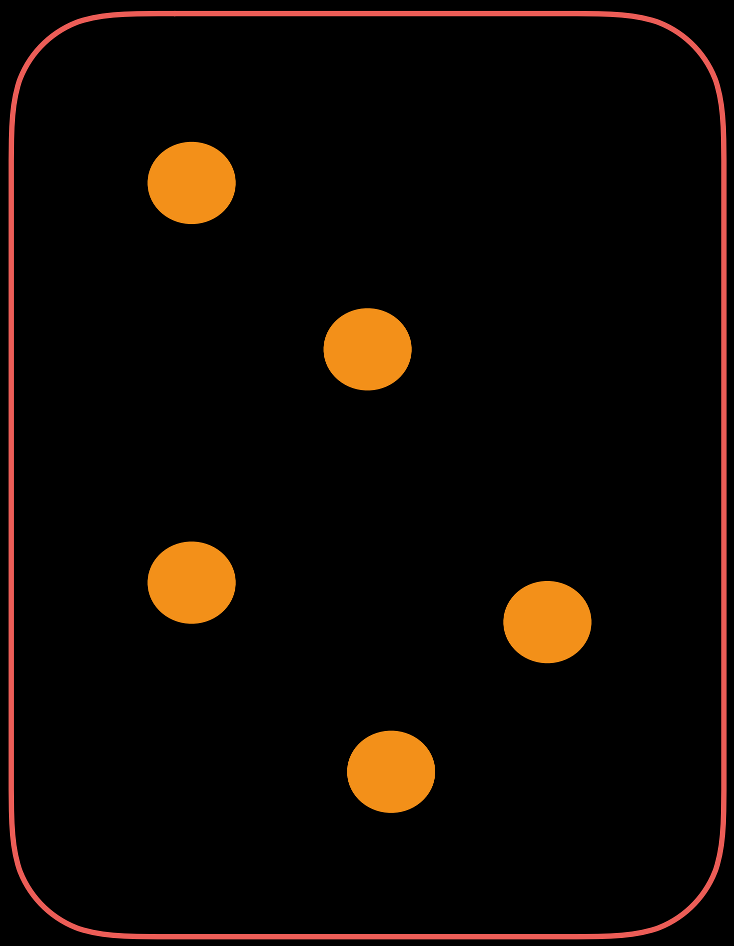
*It means that the two theories, seen as **structures of the same type**, are **isomorphic***

- Closed convex sets of probabilities (credal sets) and coherent sets of almost desirable gambles constitute equivalent theories/models

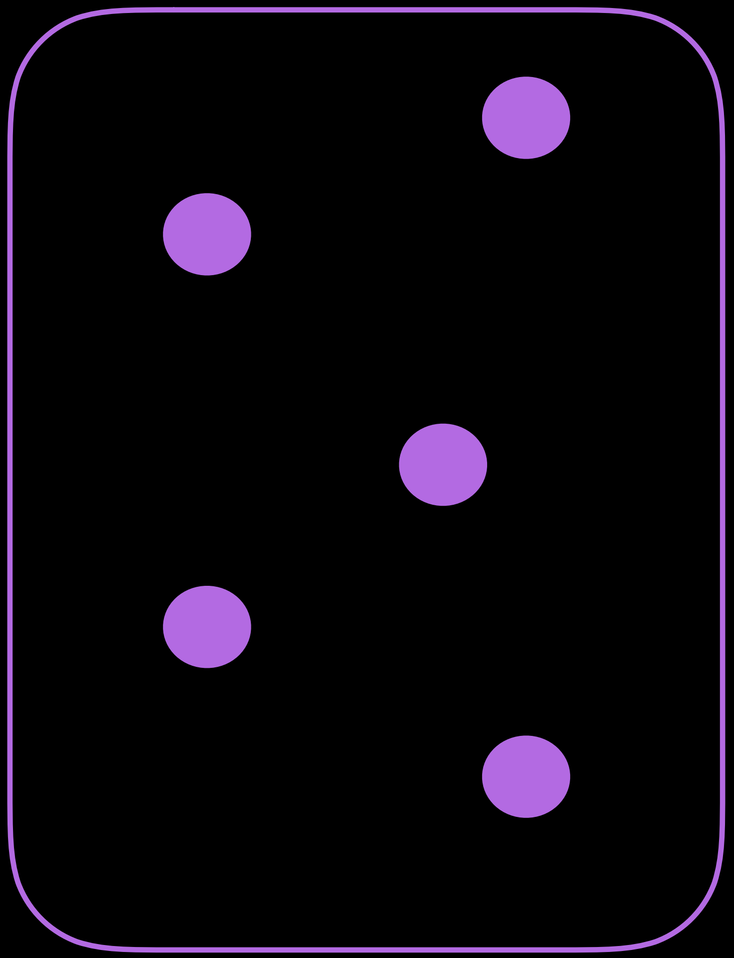
What does it mean for two ***theories/****models to be* ***equivalent****?*

It means that the two theories, seen as ***structures of the same type****, are* ***isomorphic***

Theories as structures, equivalence as isomorphism

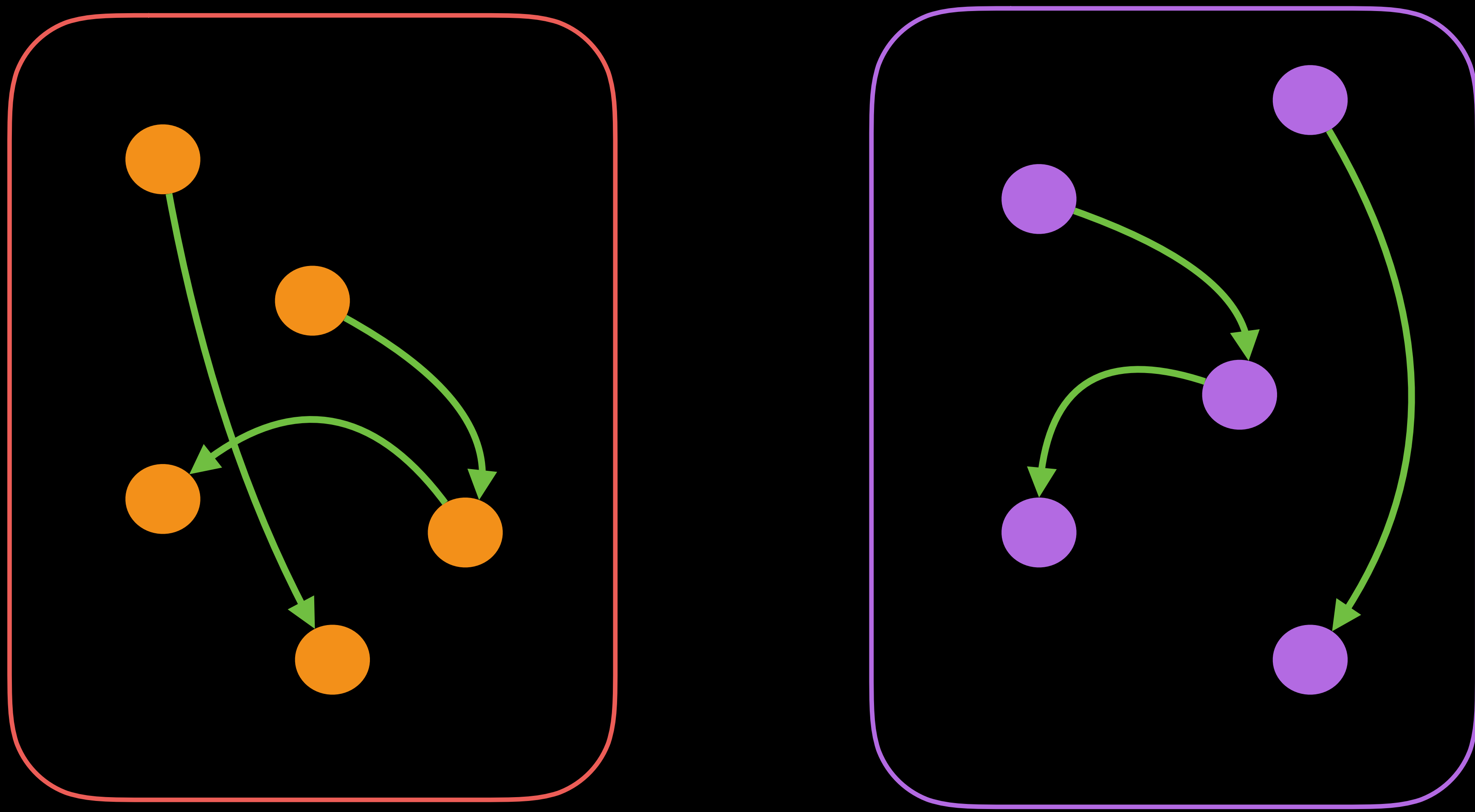


coherent sets of almost desirable gambles



credal sets

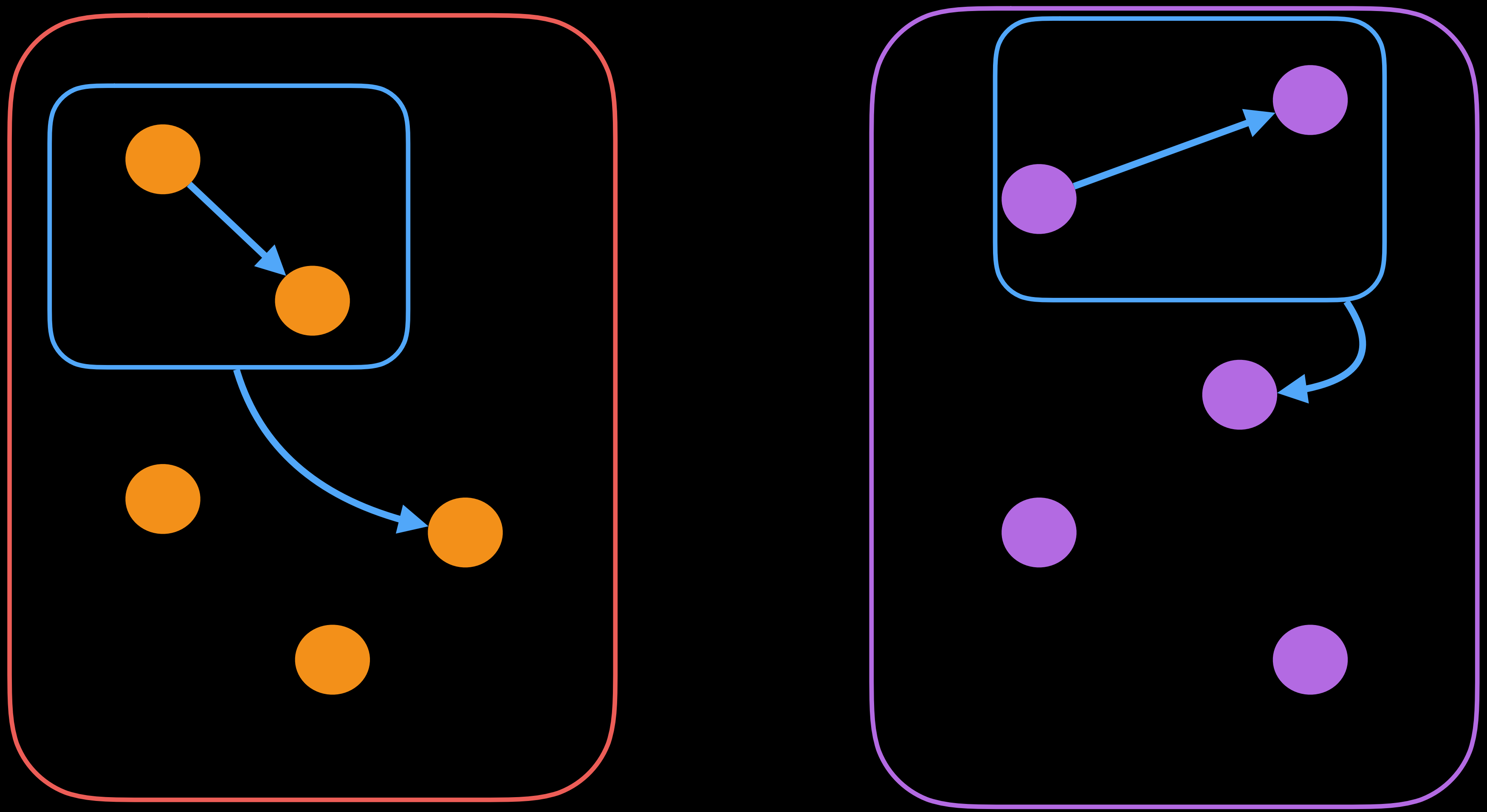
Theories as structures, equivalence as isomorphism



coherent sets of almost desirable gambles

credal sets

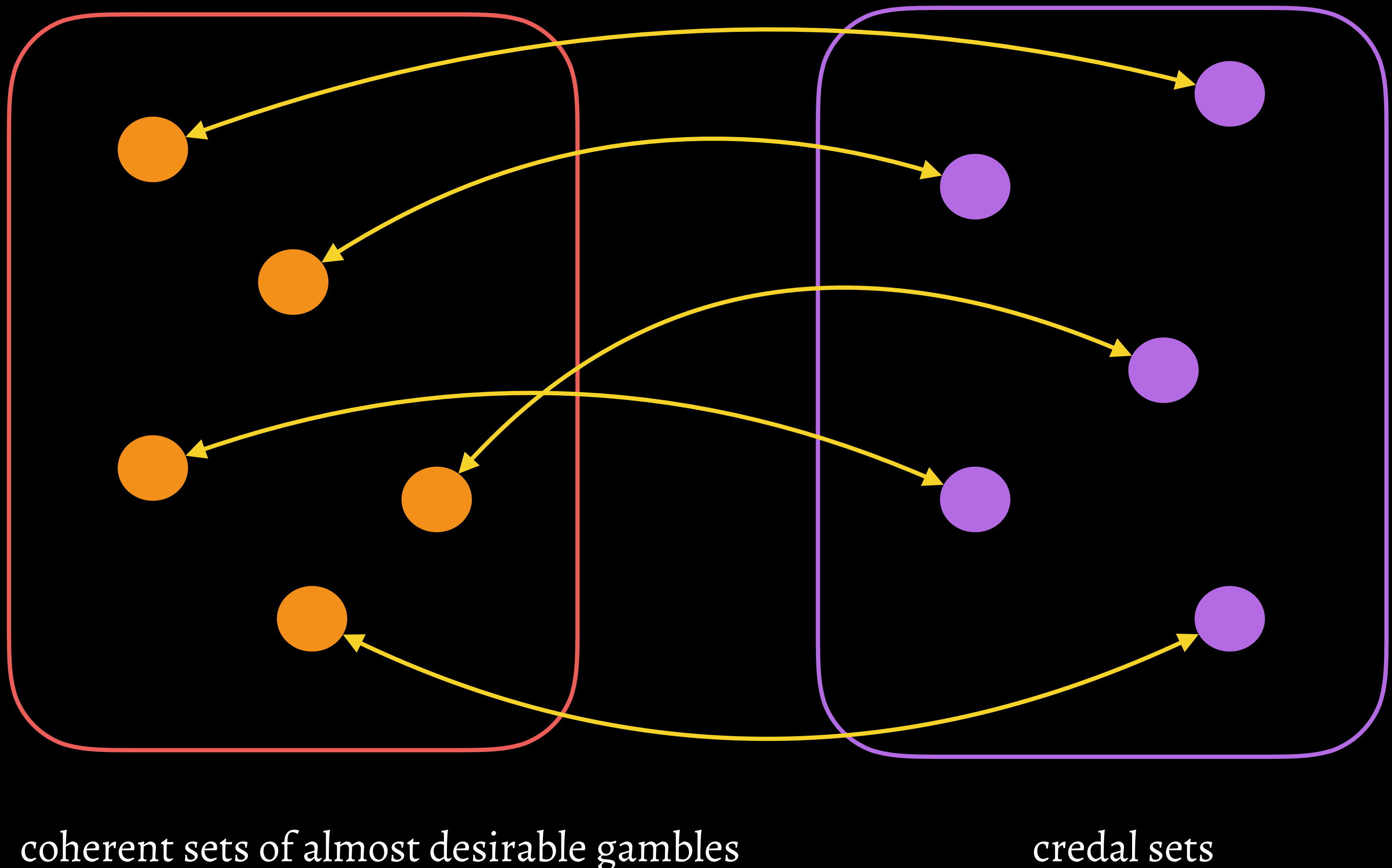
Theories as structures, equivalence as isomorphism



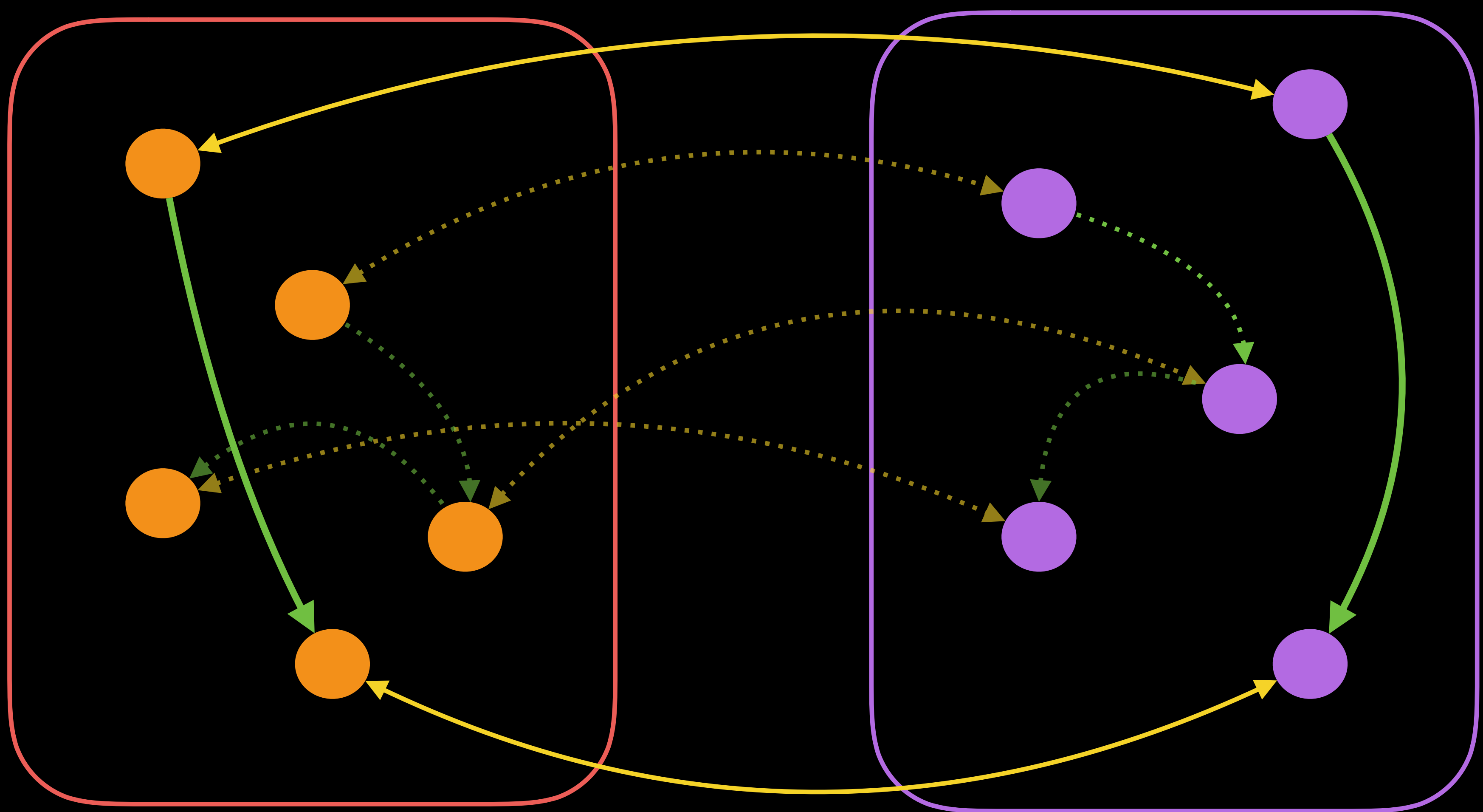
coherent sets of almost desirable gambles

credal sets

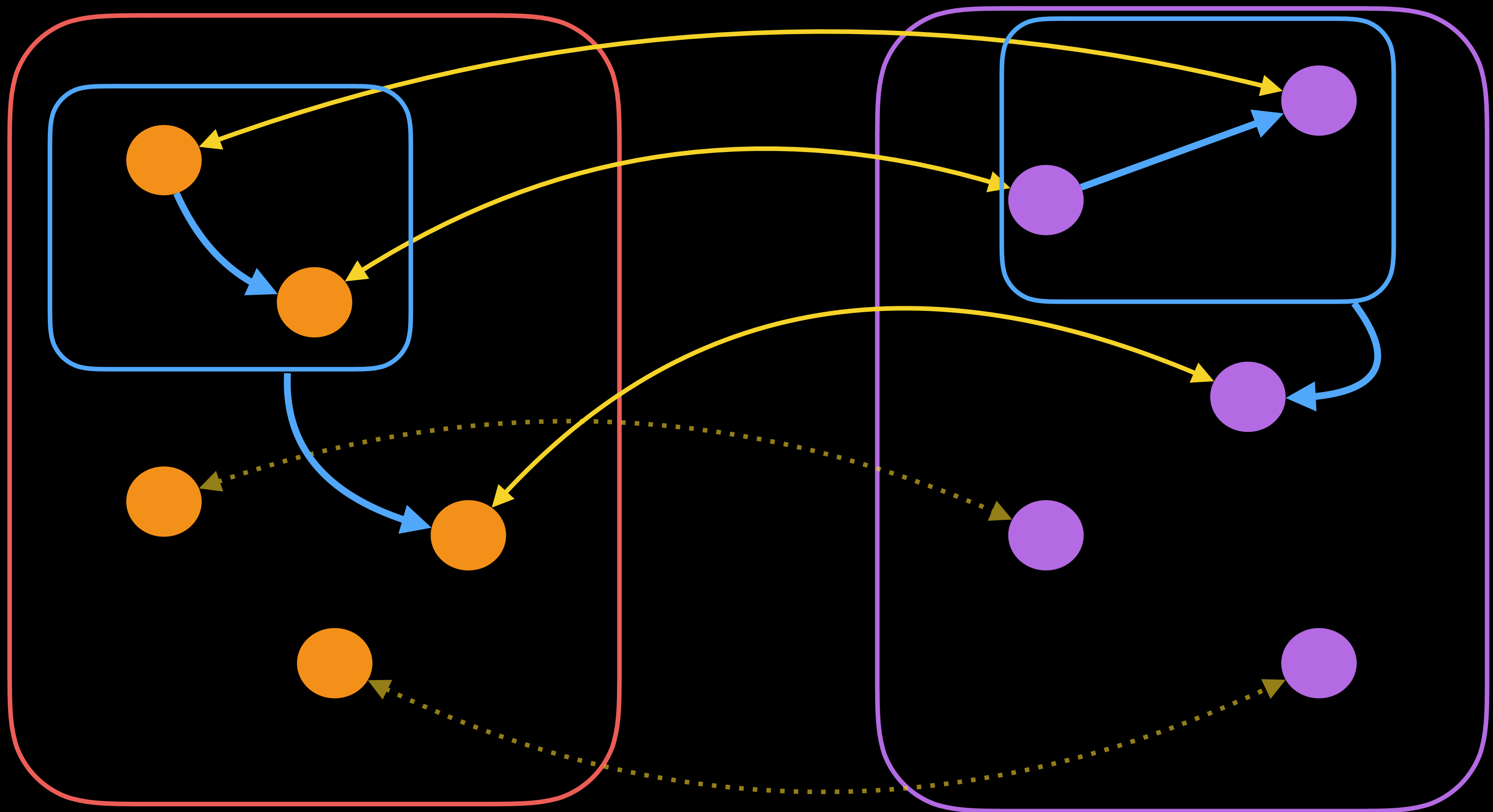
Theories as structures, equivalence as isomorphism



Theories as structures, equivalence as isomorphism



Theories as structures, equivalence as isomorphism



coherent sets of almost desirable gambles

credal sets

The case of almost desirability revisited

Let $\Pi=\{1,.., m\} \subset \Omega$, $g \in \mathbb{R}^m$. We define $g|^{\Pi c} \in \mathbb{R}^n$ as

$$g|^{\Pi c}(\omega):= \begin{cases} g(\omega), \text{ if } \omega \in \Pi \\ 0, \text{ otherwise.} \end{cases}$$

Definition [*De Cooman and Quaeghebeur (2012)*]: Let $K \subset \mathbb{R}^n$. The conditioned set of K with respect to $\Pi \subset \Omega$ is the set

$$K|_{\Pi}:=\{ g \in \mathbb{R}^m : g|^{\Pi c} \in K \}$$

The case of almost desirability revisited

Let $\Pi = \{1, .., m\} \subset \Omega$, $g \in \mathbb{R}^m$. We define $g|^{\Pi c} \in \mathbb{R}^n$ as

$$g|^{\Pi c}(\omega) := \begin{cases} g(\omega), \text{ if } \omega \in \Pi \\ 0, \text{ otherwise.} \end{cases}$$

Definition: Let $M \subset \mathbb{P}_n$ be a credal set. The conditioning of M on $\Pi \subset \Omega$ is the projection on Π of all $q(.|\Pi) \in \mathbb{P}_n$ with $q \in M$, that is

$$M|_{\Pi} := \{ p \in \mathbb{P}_m : \exists q \in M, q(.|\Pi) = p|^{\Pi c} \}$$

Proving the correspondence

Central tools:

Classical separation theorem for closed convex sets, and the (positive) polarity operator $(\,)^{\bullet}$

If $K \subset \mathbb{R}^n$ is a nonempty **closed convex set**, then for every $g \notin K$ there exist $v \in \mathbb{R}^n$ (non-null) and $b \in \mathbb{R}$ such that for all $f \in K$

$$v \cdot f \geq b > v \cdot g$$

Proving the correspondence

Central tools:

Classical separation theorem for closed convex sets, and the (positive) polarity operator $(\ )^{\bullet}$

Given a subset $A \subseteq \mathbb{R}^n$, its *polar* is the closed convex cone

$$A^{\bullet} = \{ g \in \mathbb{R}^n : g \cdot f \geq 0, \text{ for every } f \in A \}$$

Reformulating the separation theorem:

If $K \subset \mathbb{R}^n$ is a nonempty **closed convex cone**, then for every $g \notin K$ there exists $v \in \mathbb{R}^n$ (non-null) such that $K \subset \{v\}^\bullet$ but $g \notin \{v\}^\bullet$

Hence:

K is a *closed convex cone* iff it is the polar of some set $A \subset \mathbb{R}^n$ (i.e. $K = A^\bullet$)

Reformulating the separation theorem:

If $K \subset \mathbb{R}^n$ is a nonempty **closed convex cone**, then for every $g \notin K$ there exists $v \in \mathbb{R}^n$ (non-null) such that $K \subset \{v\}^\bullet$ but $g \notin \{v\}^\bullet$

Hence:

K is a *closed convex cone* iff it is the polar of some set $A \subset \mathbb{R}^n$ (i.e. $K = A^\bullet$)

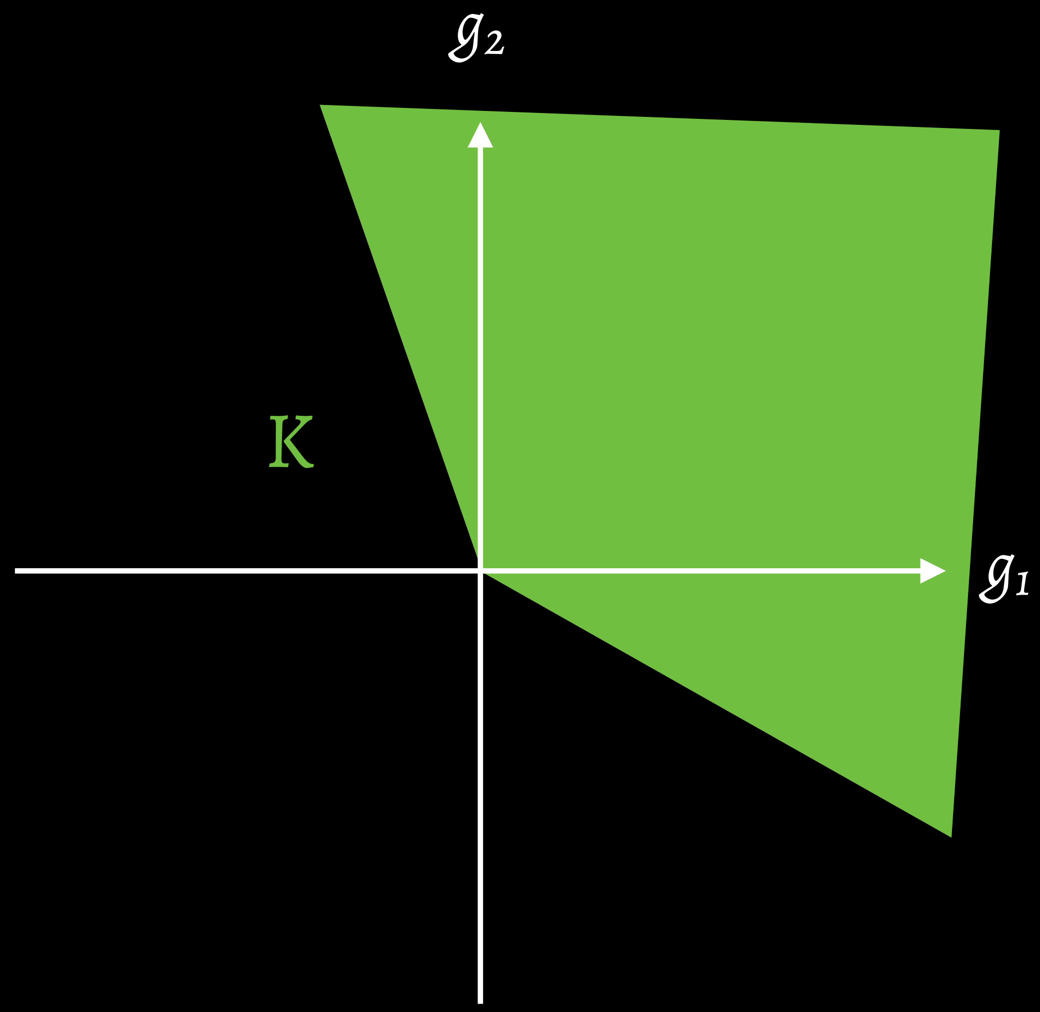
Reformulating the separation theorem:

If $K \subset \mathbb{R}^n$ is a nonempty **closed convex cone**, then for every $g \notin K$ there exists $v \in \mathbb{R}^n$ (non-null) such that $K \subset \{v\}^{\bullet}$ but $g \notin \{v\}^{\bullet}$

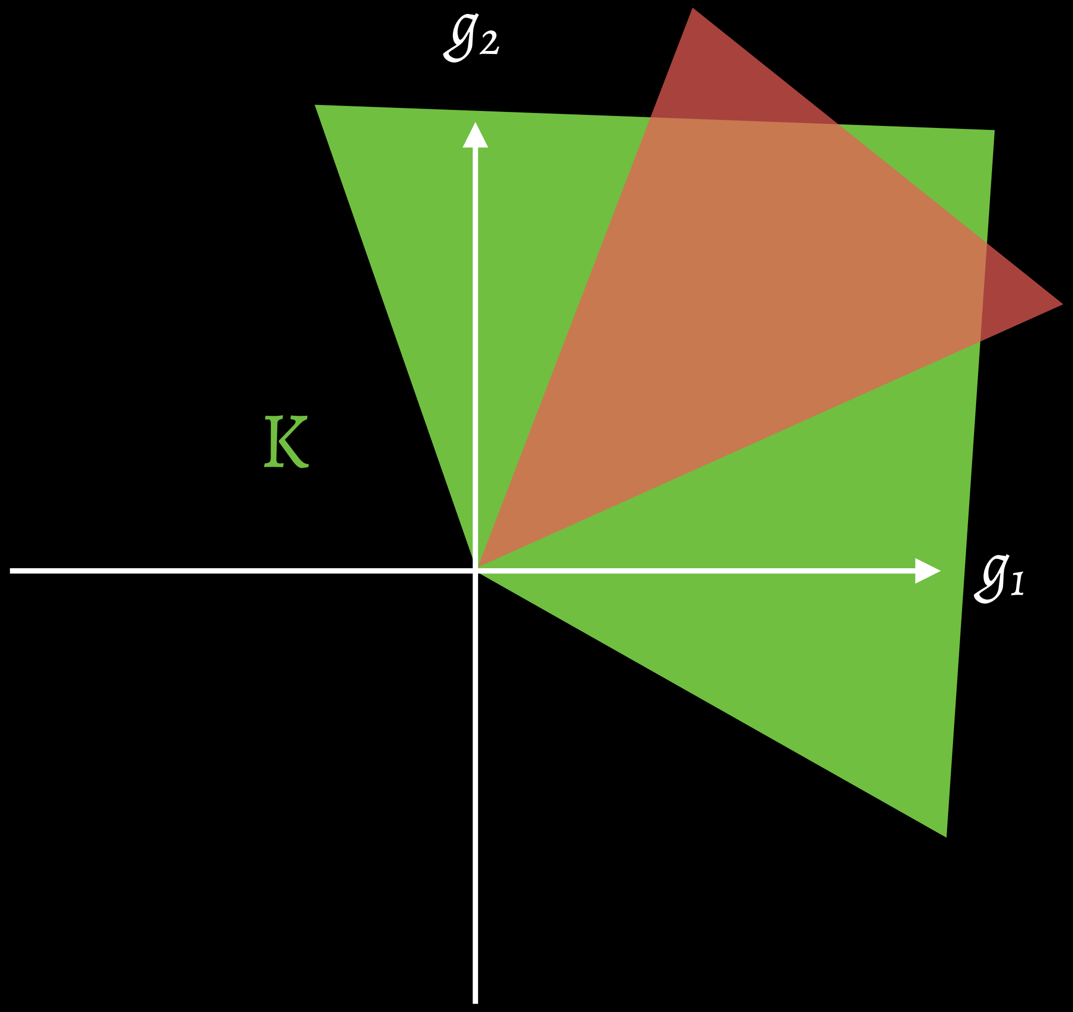
Hence:

M is a *credal set* iff it is the intersection of the set $\mathbb{P}_n$ of all pmf over Ω with the polar of some set.

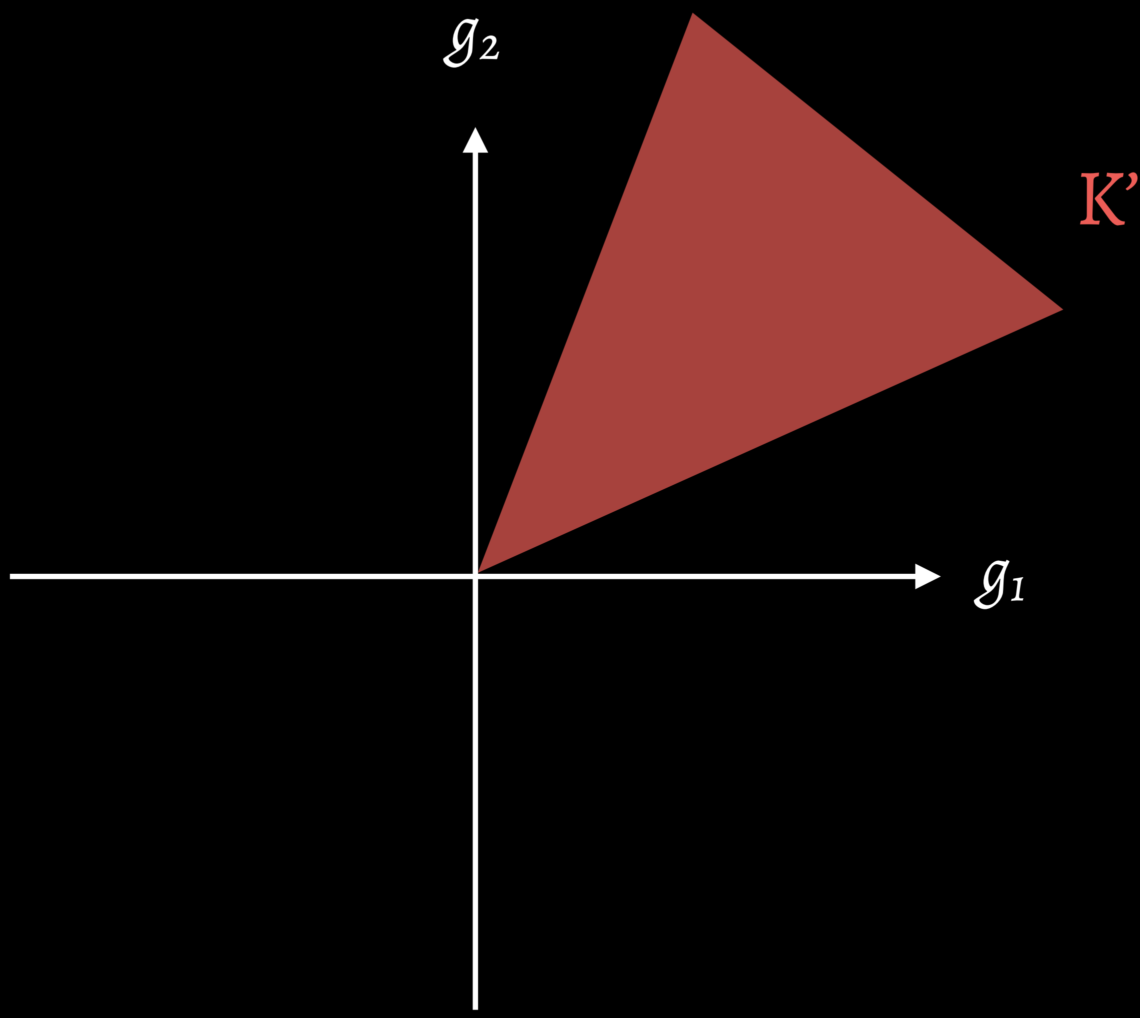
Proving the correspondence



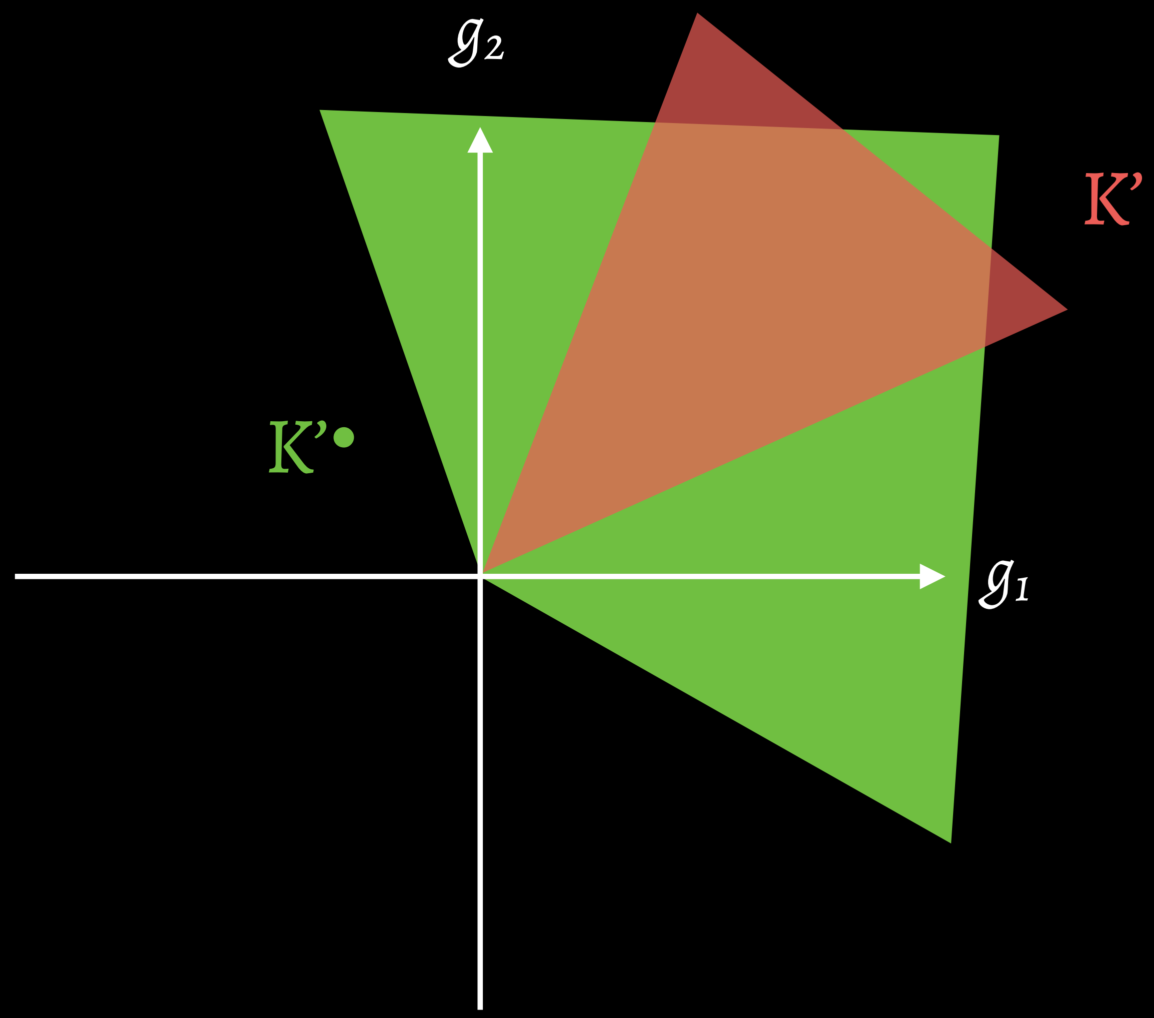
Proving the correspondence



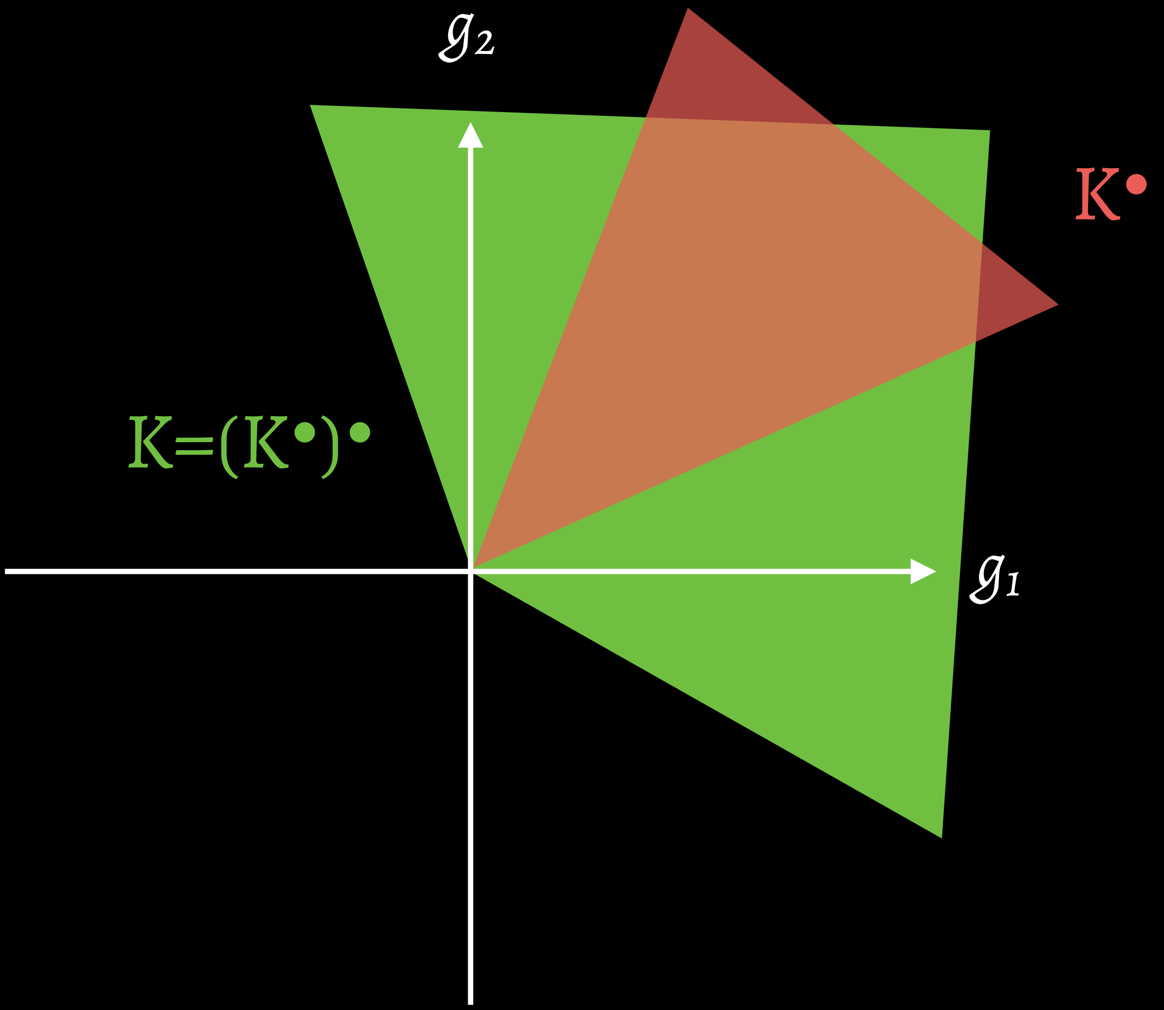
Proving the correspondence



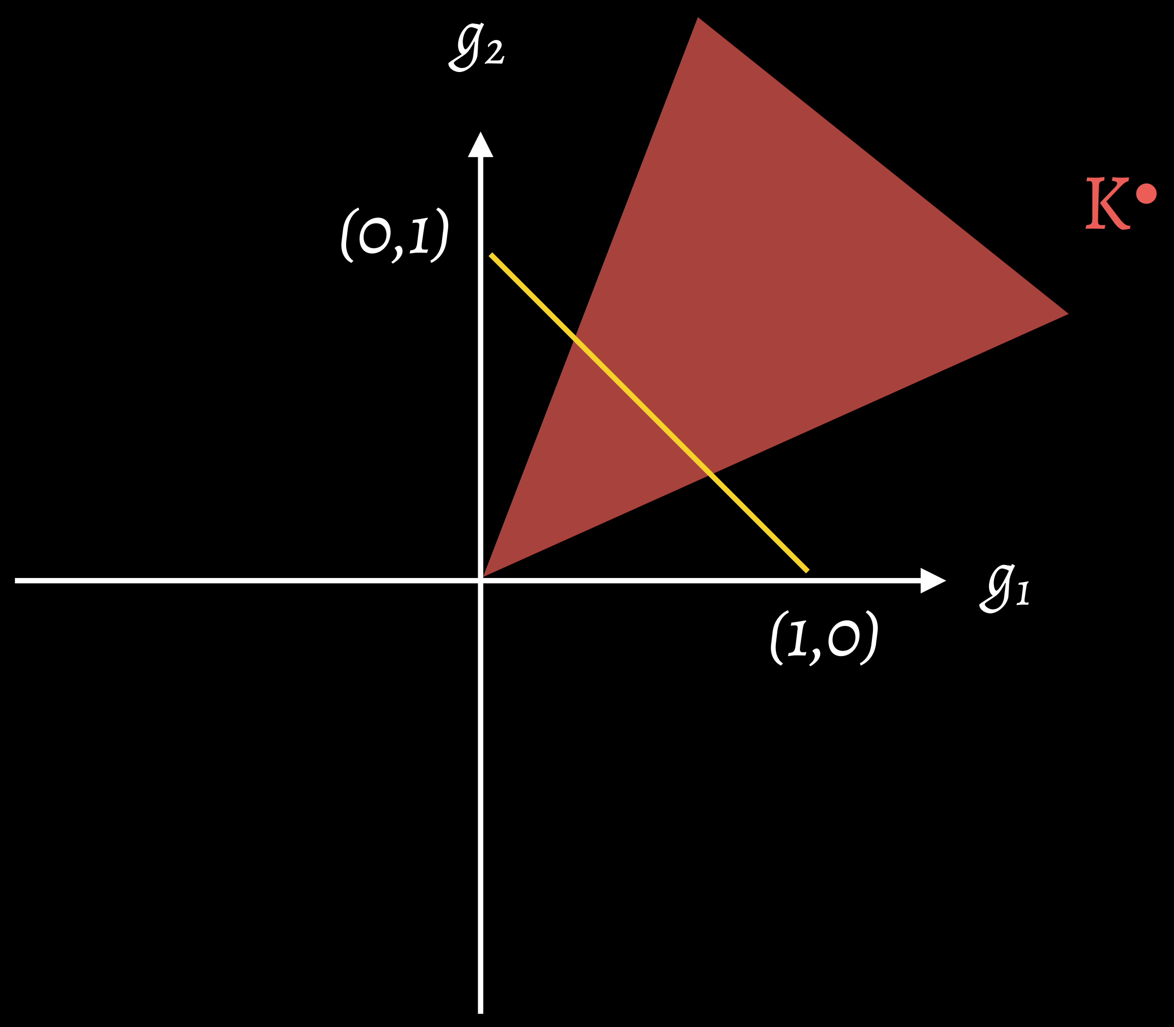
Proving the correspondence



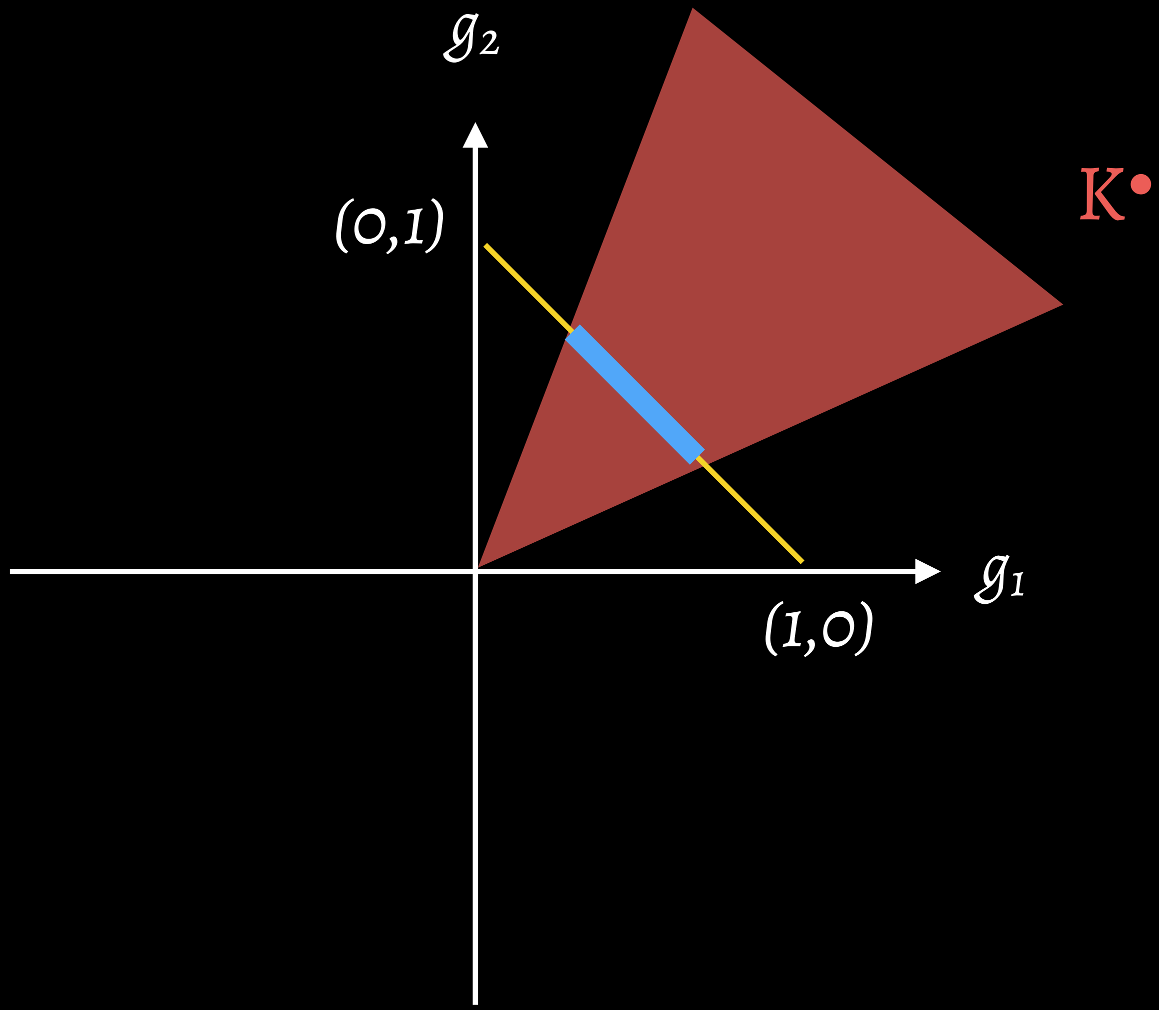
Proving the correspondence



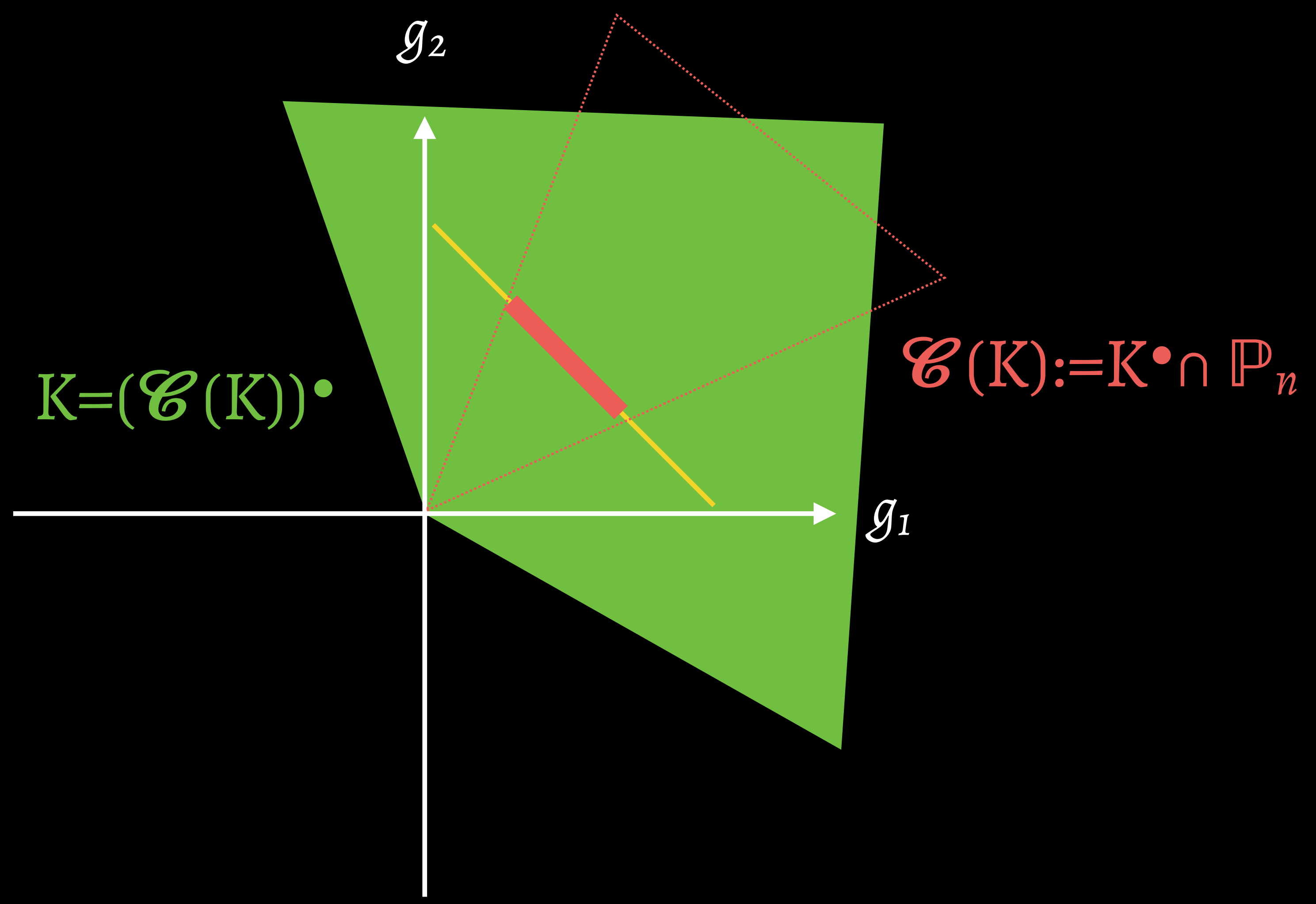
Proving the correspondence



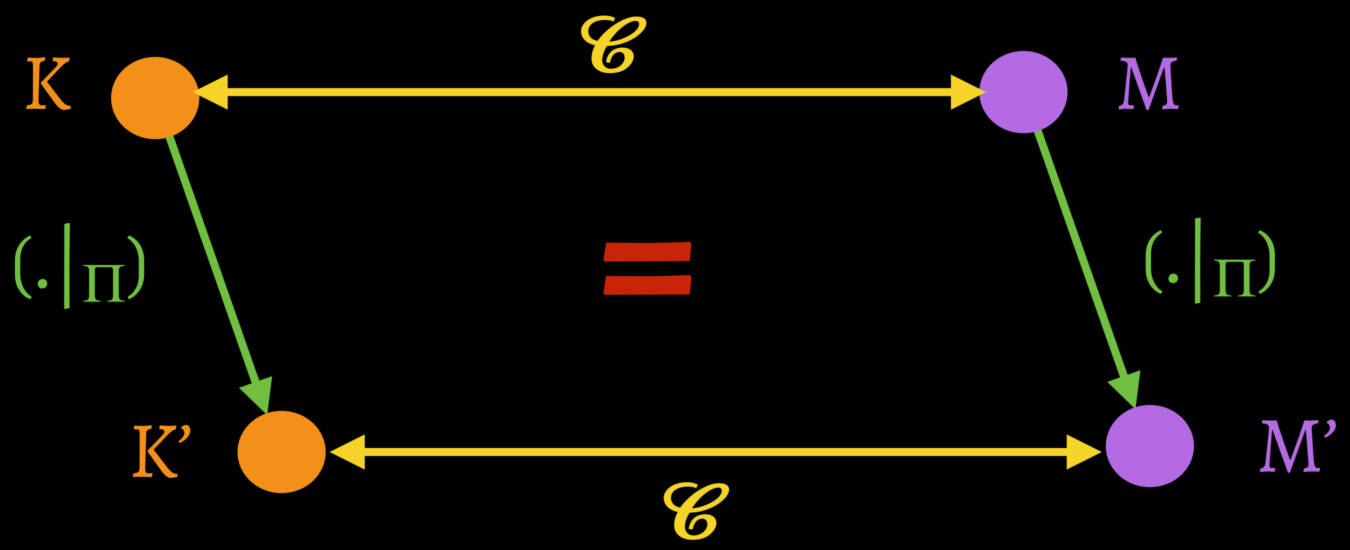
Proving the correspondence



Proving the correspondence



Theorem: The map $\mathscr{C}: \mathrm{K} \longmapsto \mathrm{K}^{\bullet} \cap \mathbb{P}_n$ is an isomorphism between the collection of coherent sets of almost desirable gambles equipped with the conditioning operation and the collection of credal sets equipped with the conditioning operation.



The case of desirability

Definition: A set $K \subseteq \mathbb{R}^n$ is called a coherent set of almost desirable gambles if it satisfies

[*linearity*] if $f,g \in K$, then $\mu f+\nu g \in K$, for $\mu,\nu>0$

[*accepting partial gain*] if $g > 0_n$ then $g \in K$

[*avoiding sure loss*] $-1_n \notin K$

[*closure*] if $g+f \in K$, for all $f > 0_n$, then $g \in K$

The case of desirability

Definition: A set $K \subseteq \mathbb{R}^n$ is called a coherent set of desirable gambles if it satisfies

[*linearity*] if $f,g \in K$, then $\mu f+\nu g \in K$, for $\mu,\nu>0$

[*accepting partial gain*] if $g > 0_n$ then $g \in K$

[*avoiding status quo*] $0_n \notin K$

Mathematically, it means that

- a coherent set of desirable gambles is a *convex cone omitting its apex and containing all positive vectors.*

The case of desirability

Definition: A set $K \subseteq \mathbb{R}^n$ is called a coherent set of desirable gambles if it satisfies

[*linearity*] if $f, g \in K$, then $\mu f + \nu g \in K$, for $\mu, \nu > 0$

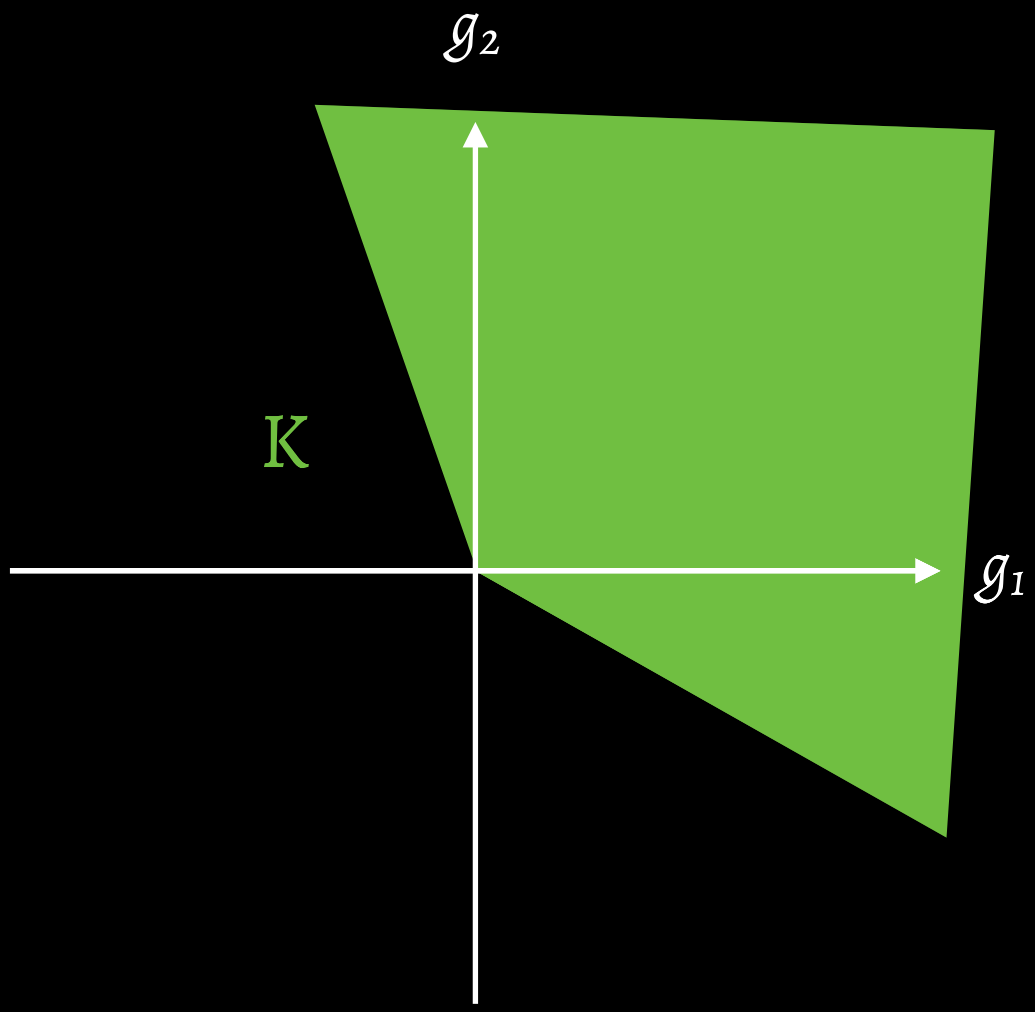
[*accepting partial gain*] if $g > 0_n$ then $g \in K$

[*avoiding status quo*] $0_n \notin K$

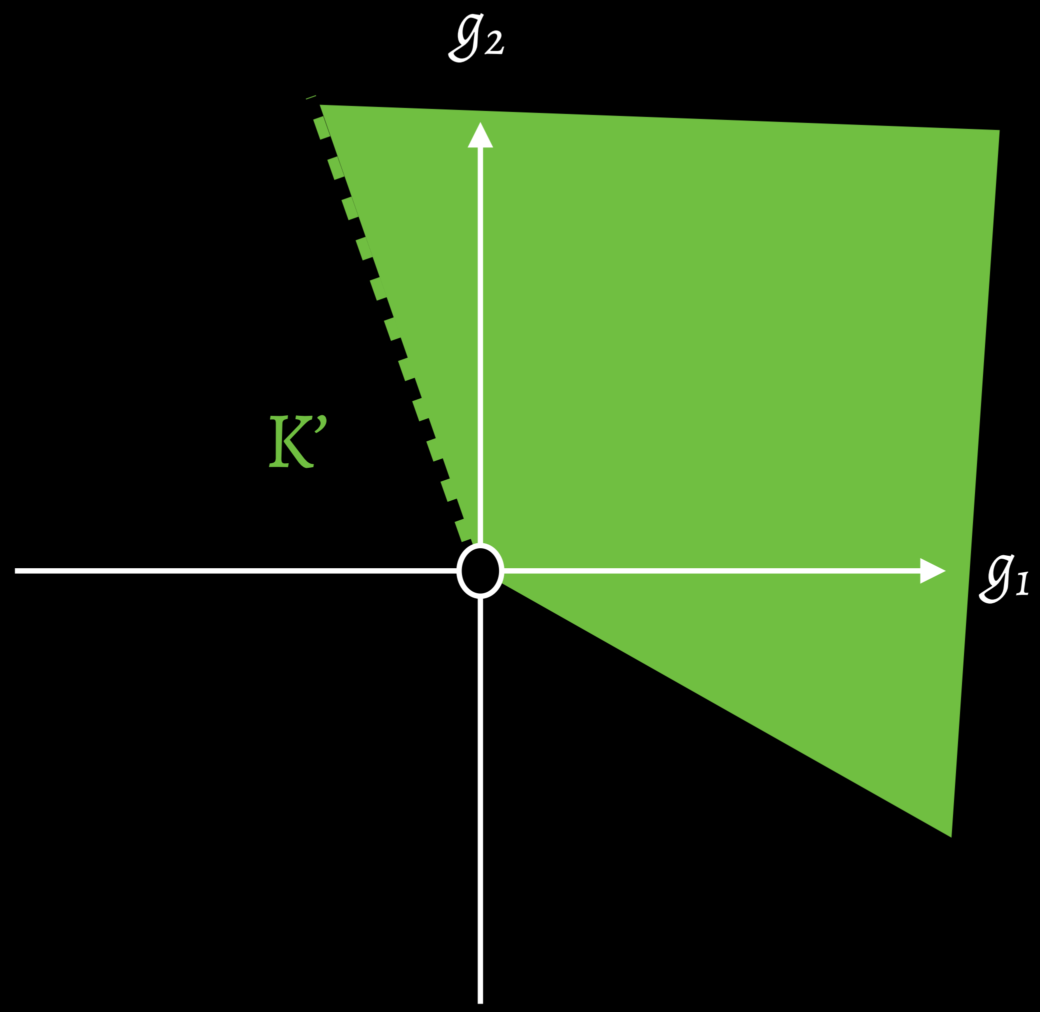
Mathematically, it means that

- a coherent set of desirable gambles is a *convex cone omitting its apex and containing all positive vectors.*

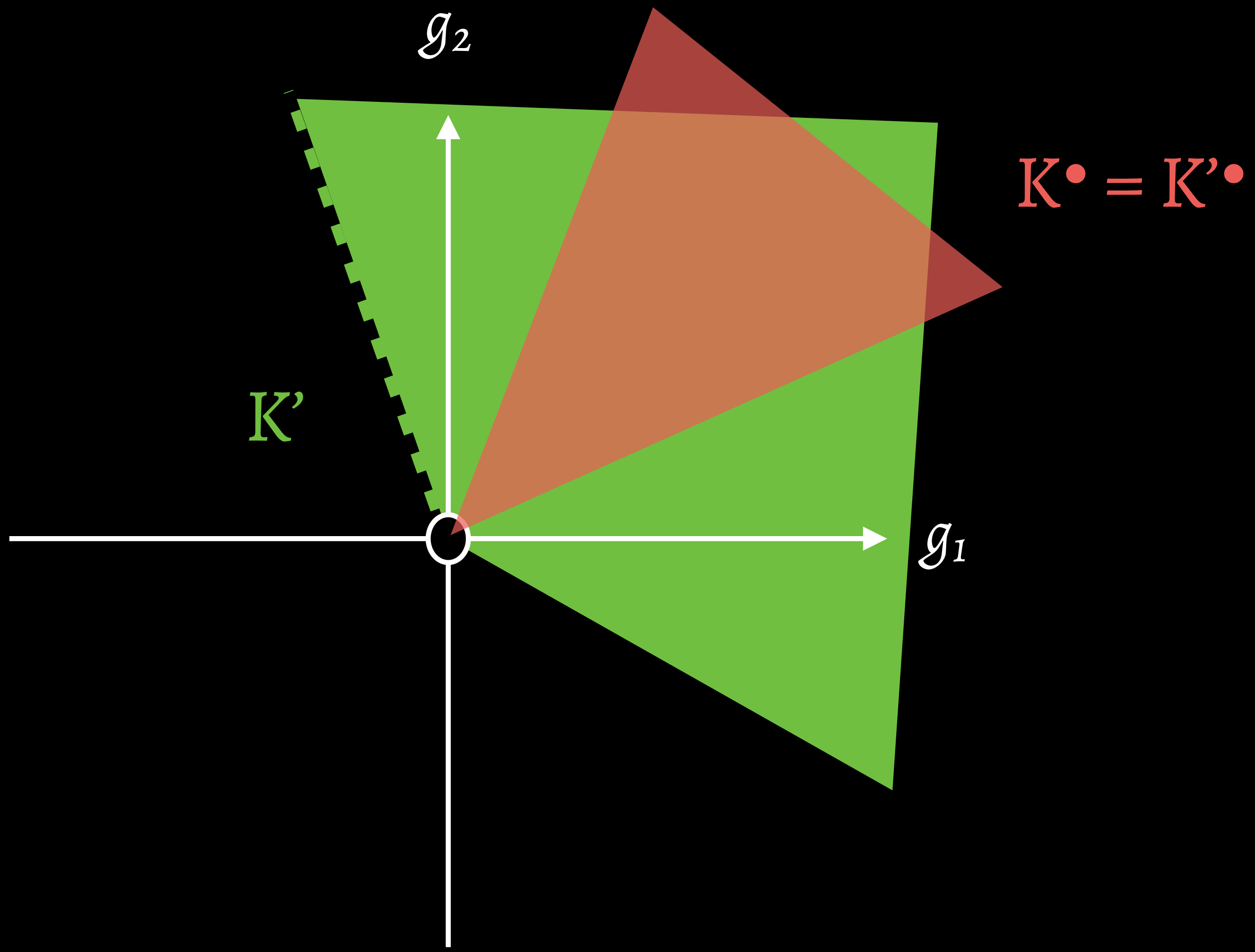
The case of desirability



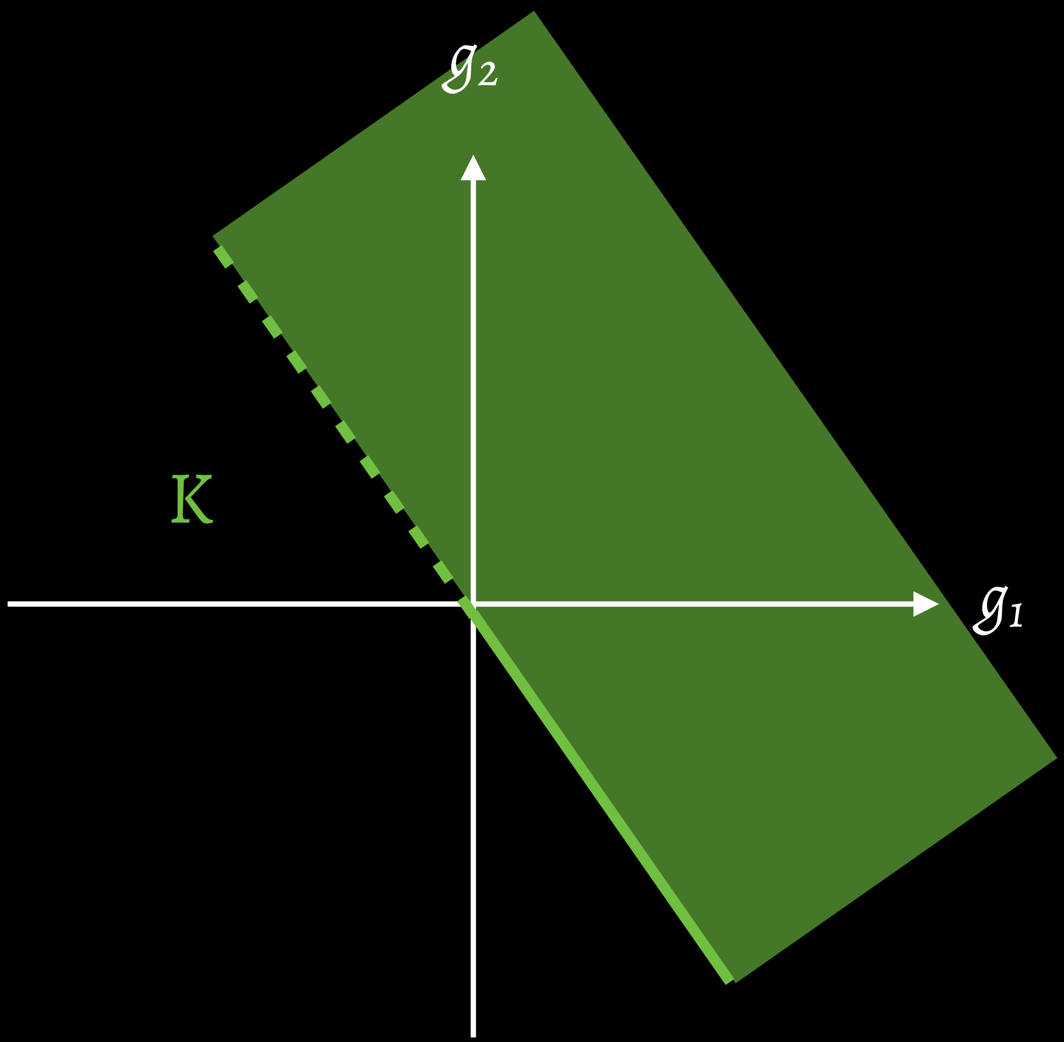
The case of desirability



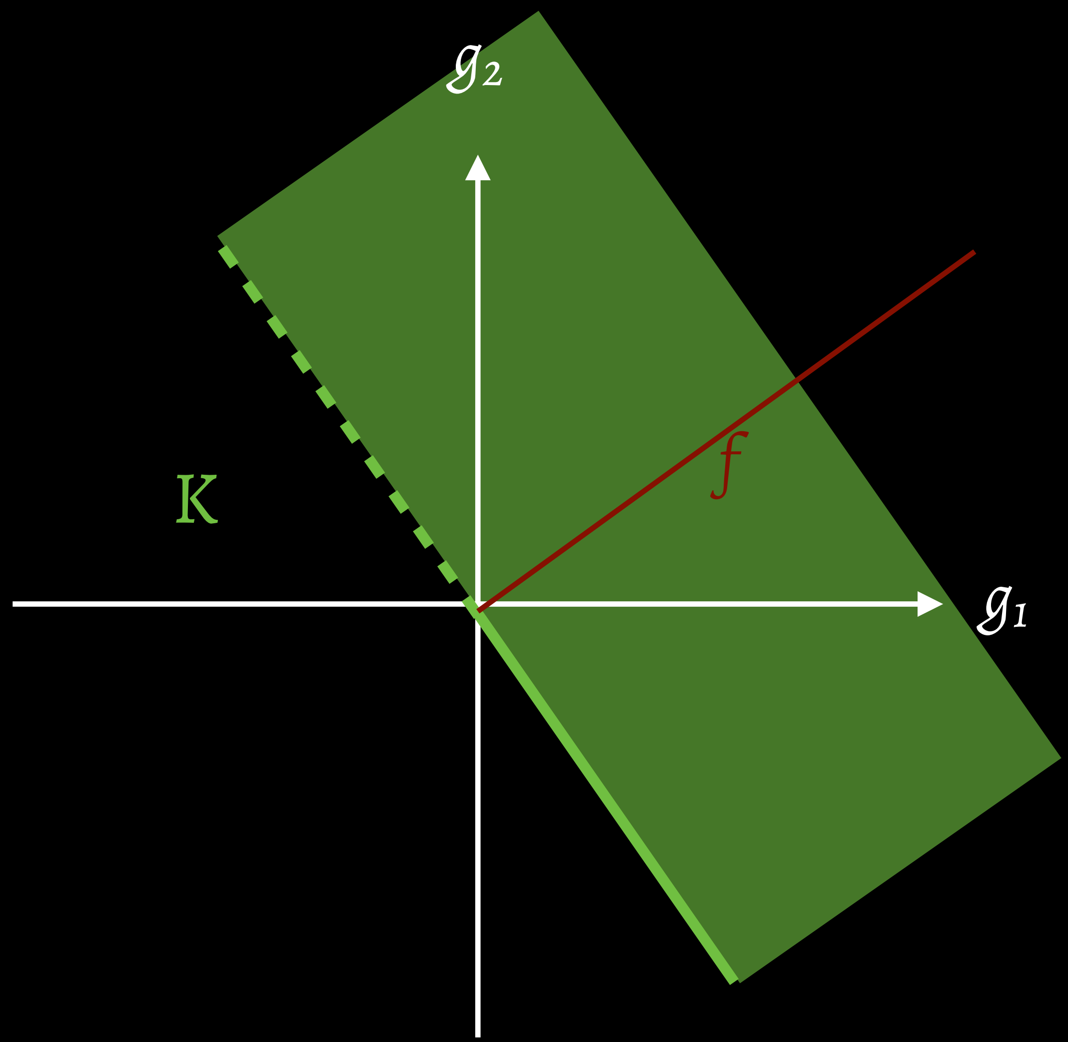
The case of desirability



Lexicographic polarity

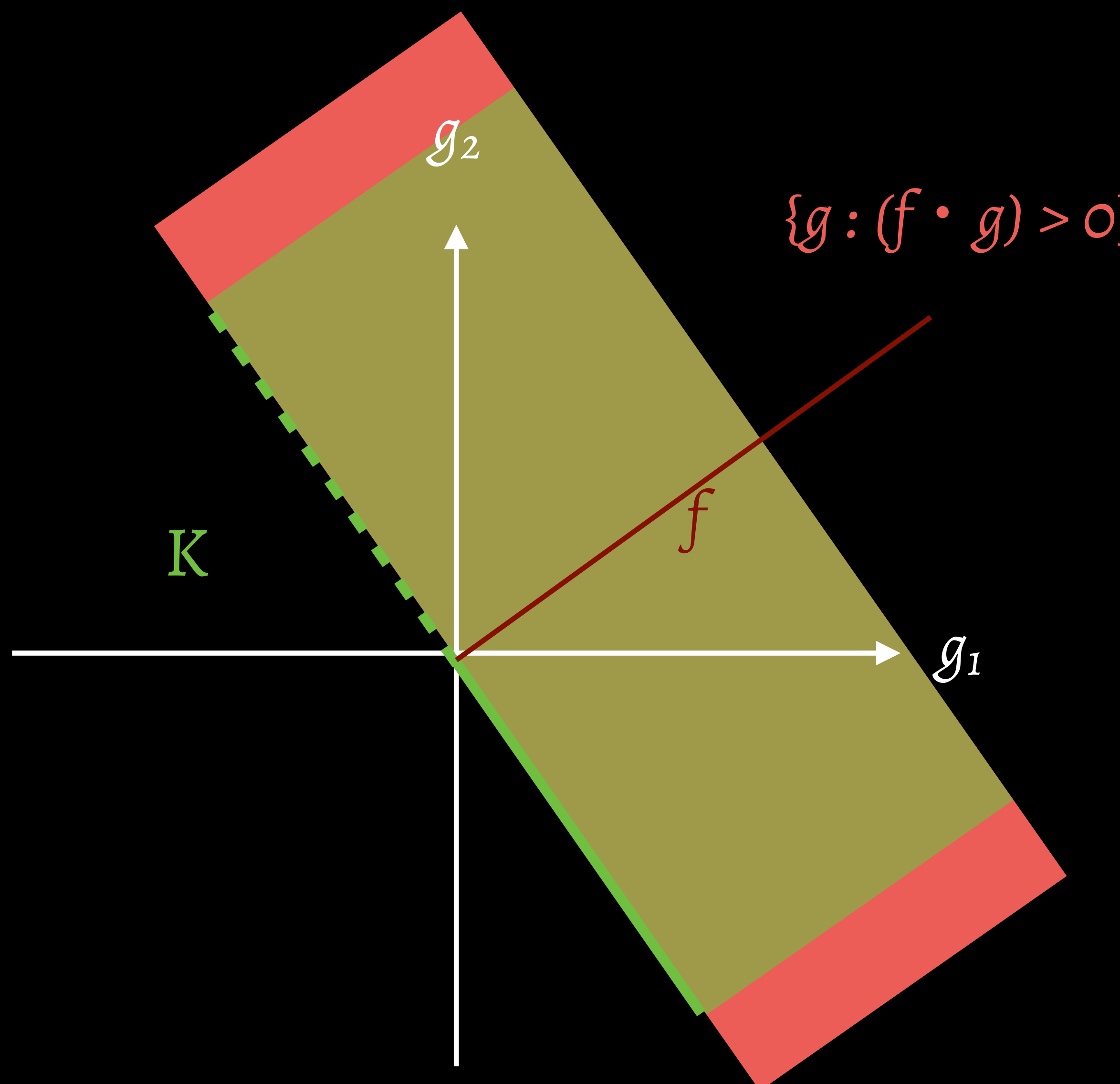


Lexicographic polarity

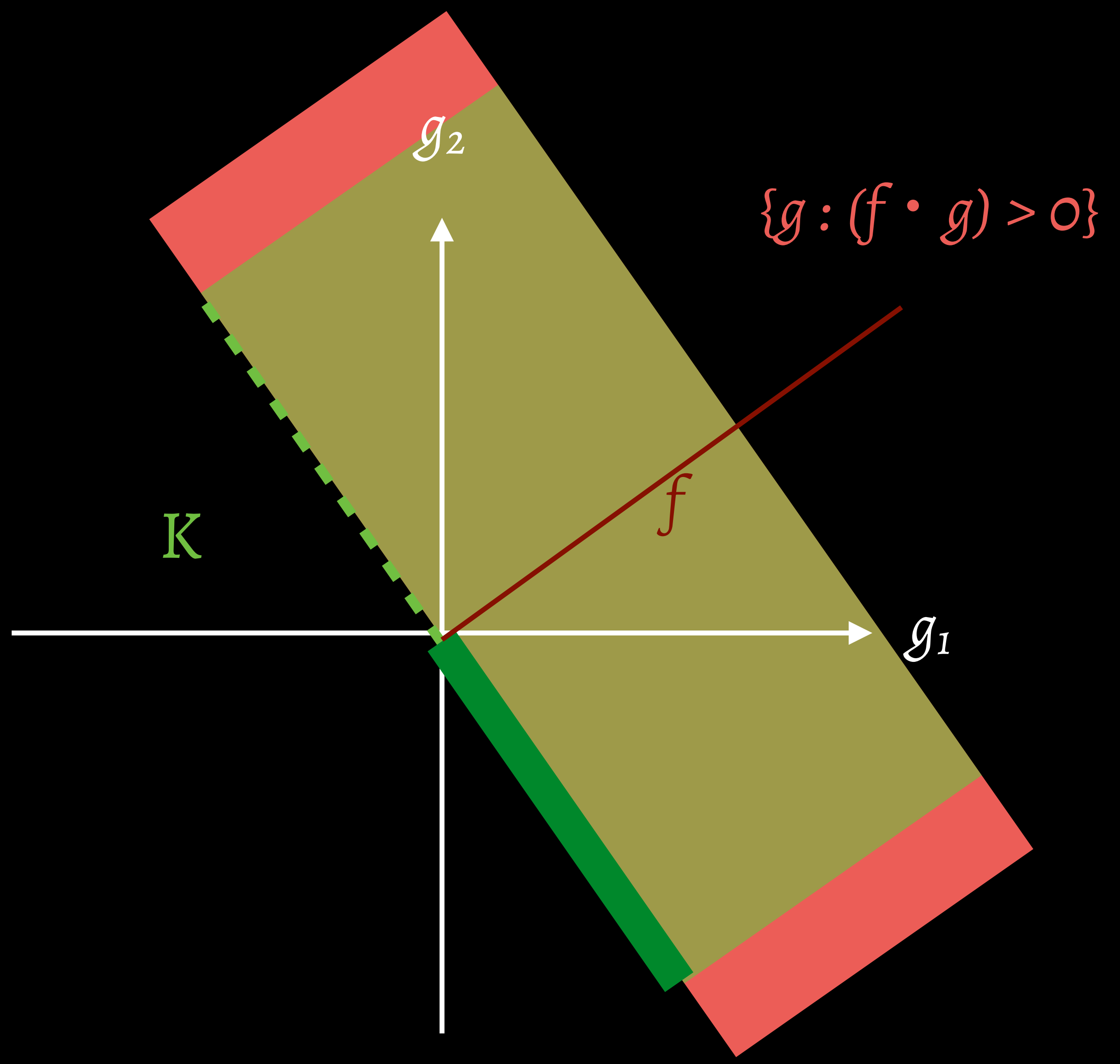


Alessandro Facchini, IDSIA

Lexicographic polarity

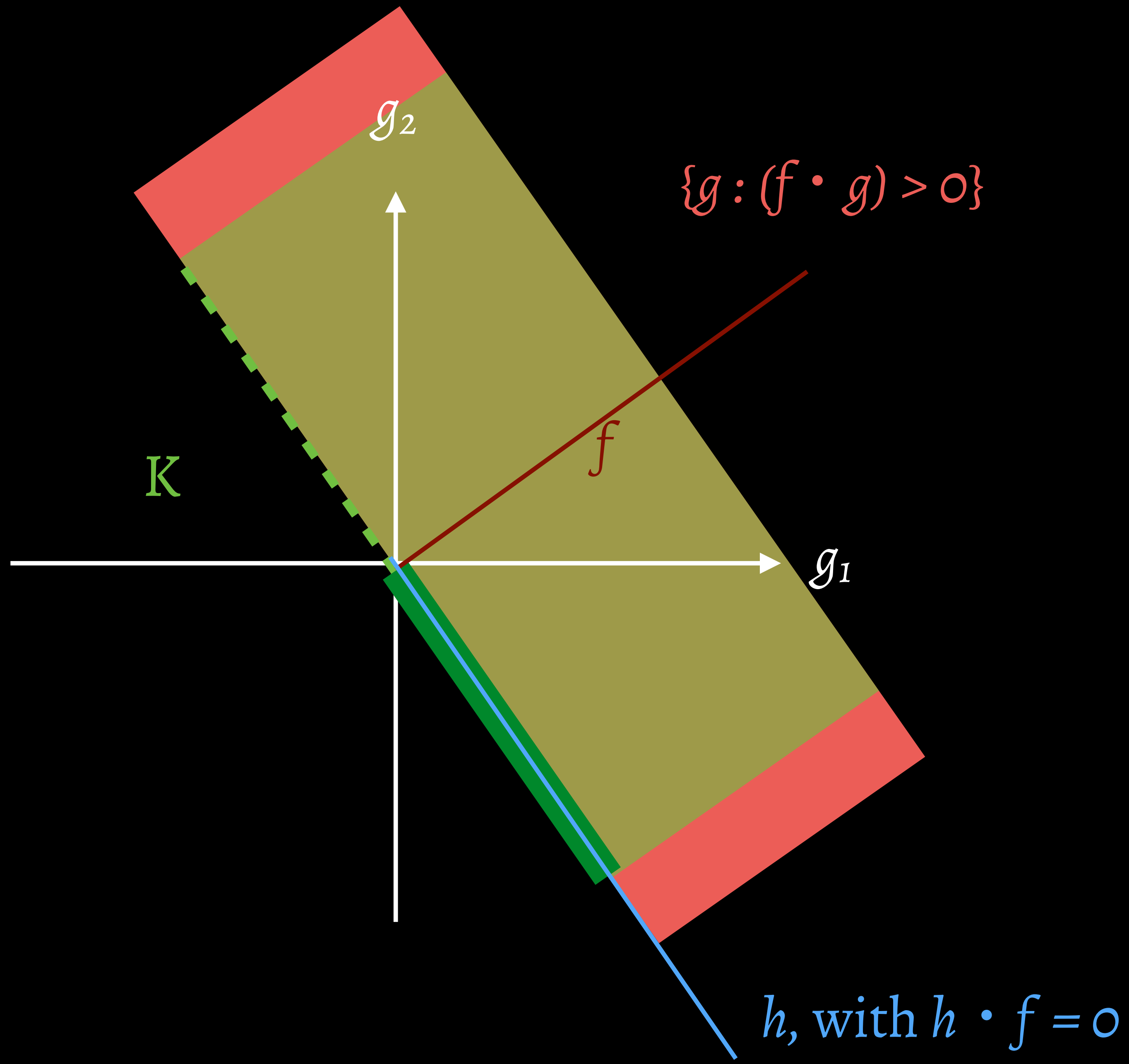


Lexicographic polarity

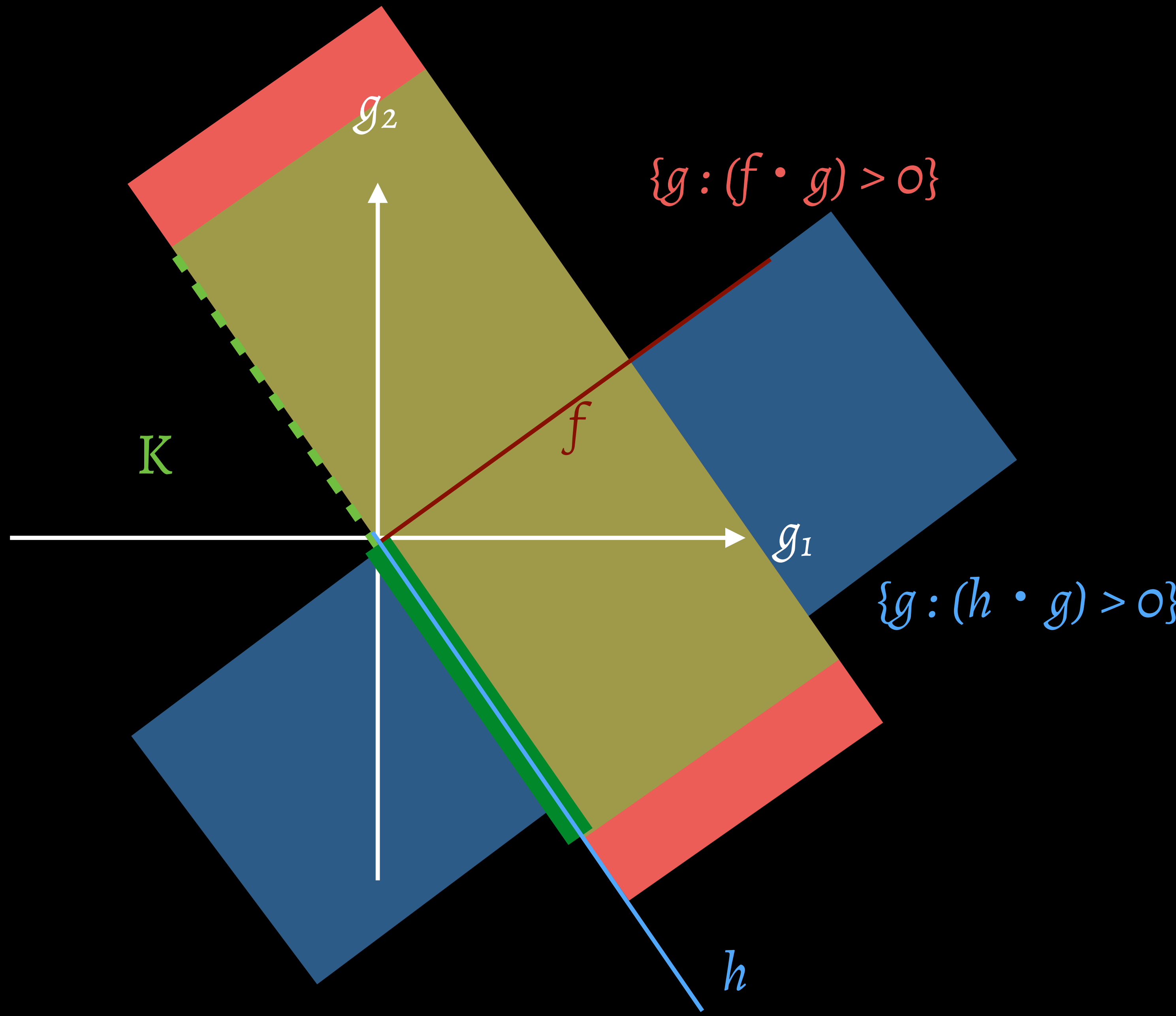


Alessandro Facchini, IDSIA

Lexicographic polarity

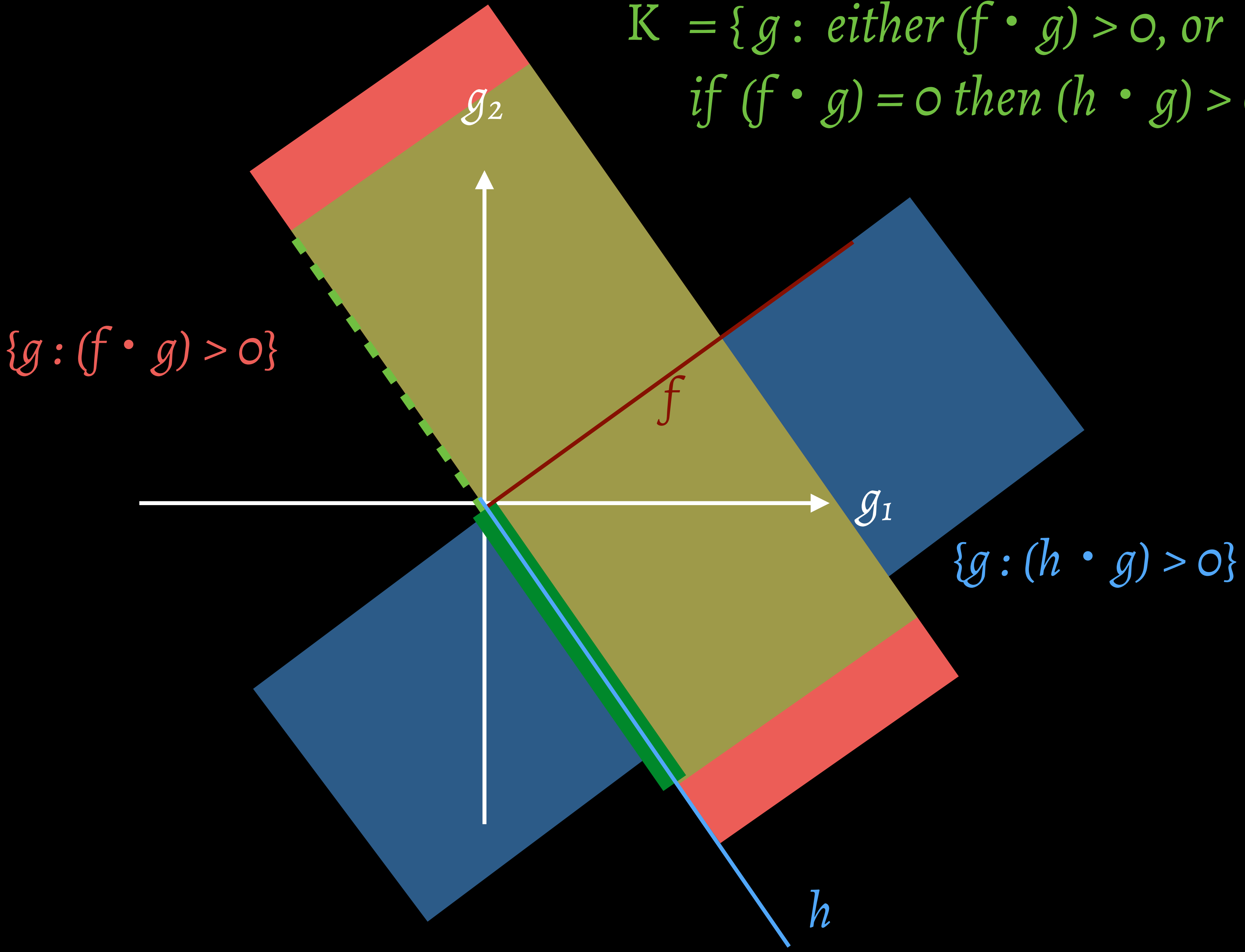


Lexicographic polarity



Lexicographic polarity

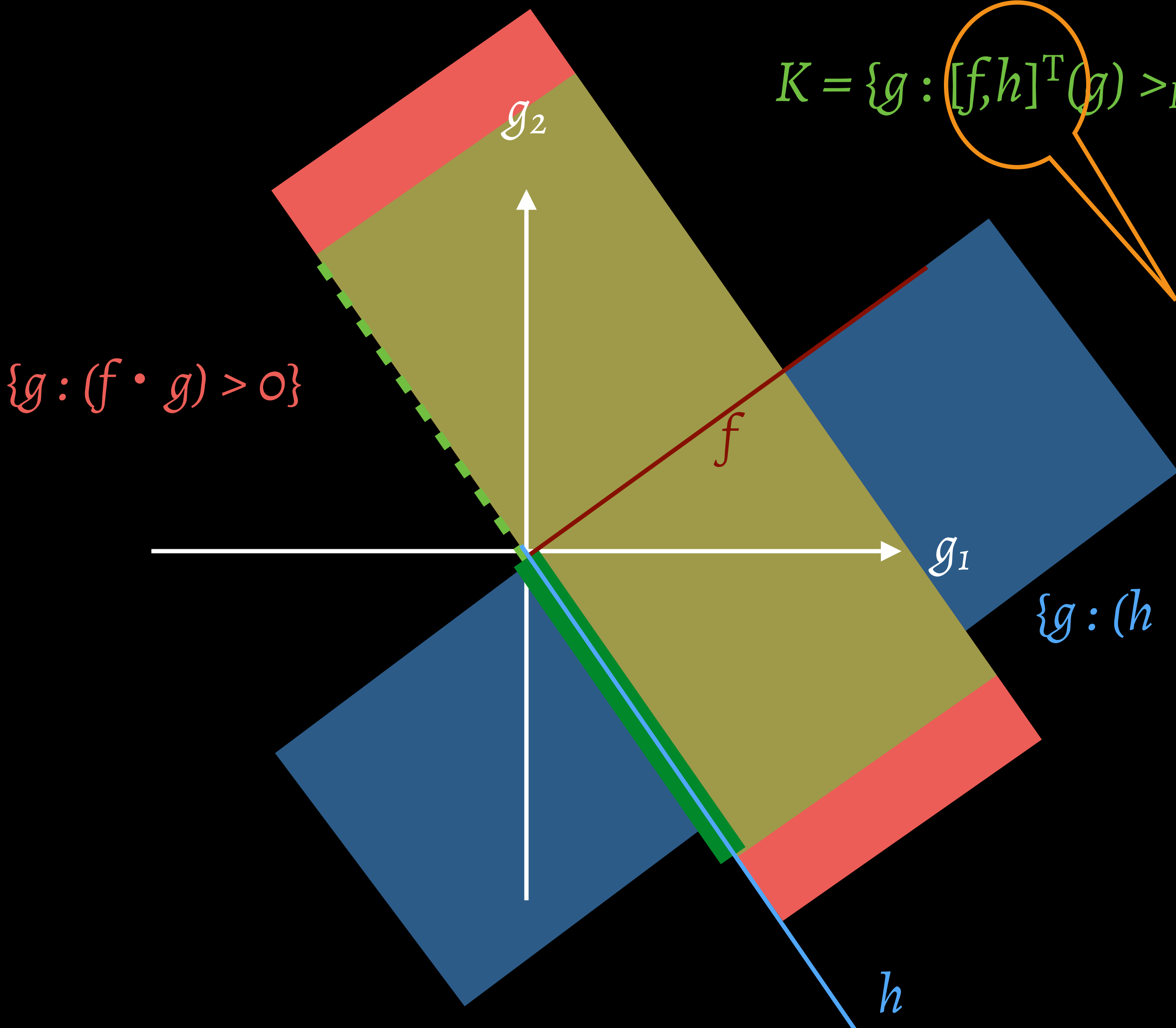
$$K = \{ g : \text{ either } (f \bullet g) > 0, \text{ or}$$
$$\text{if } (f \bullet g) = 0 \text{ then } (h \bullet g) > 0 \}$$



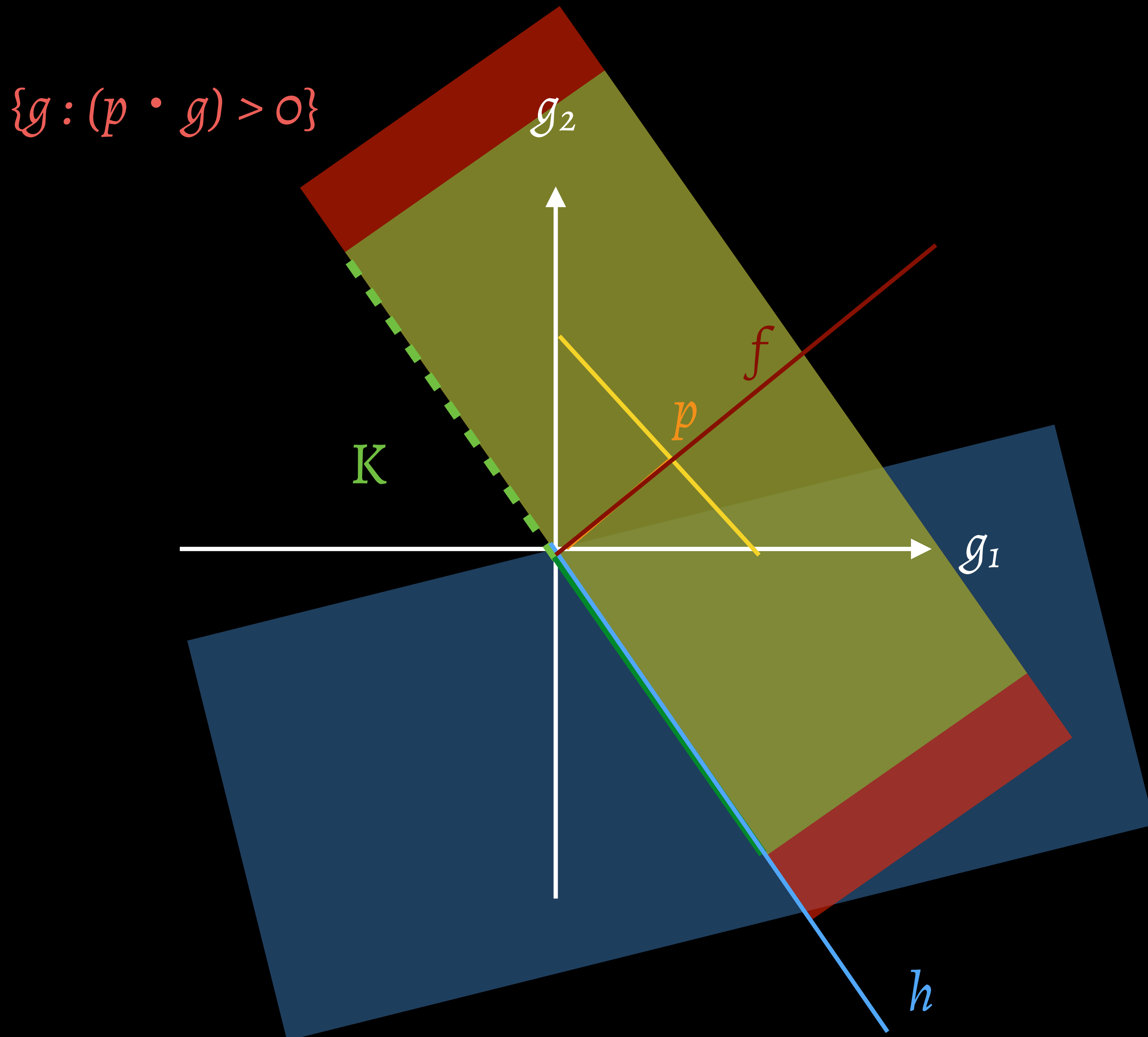
Lexicographic polarity

$K = \{g : [f,h]^T(g) >_L 0\}$

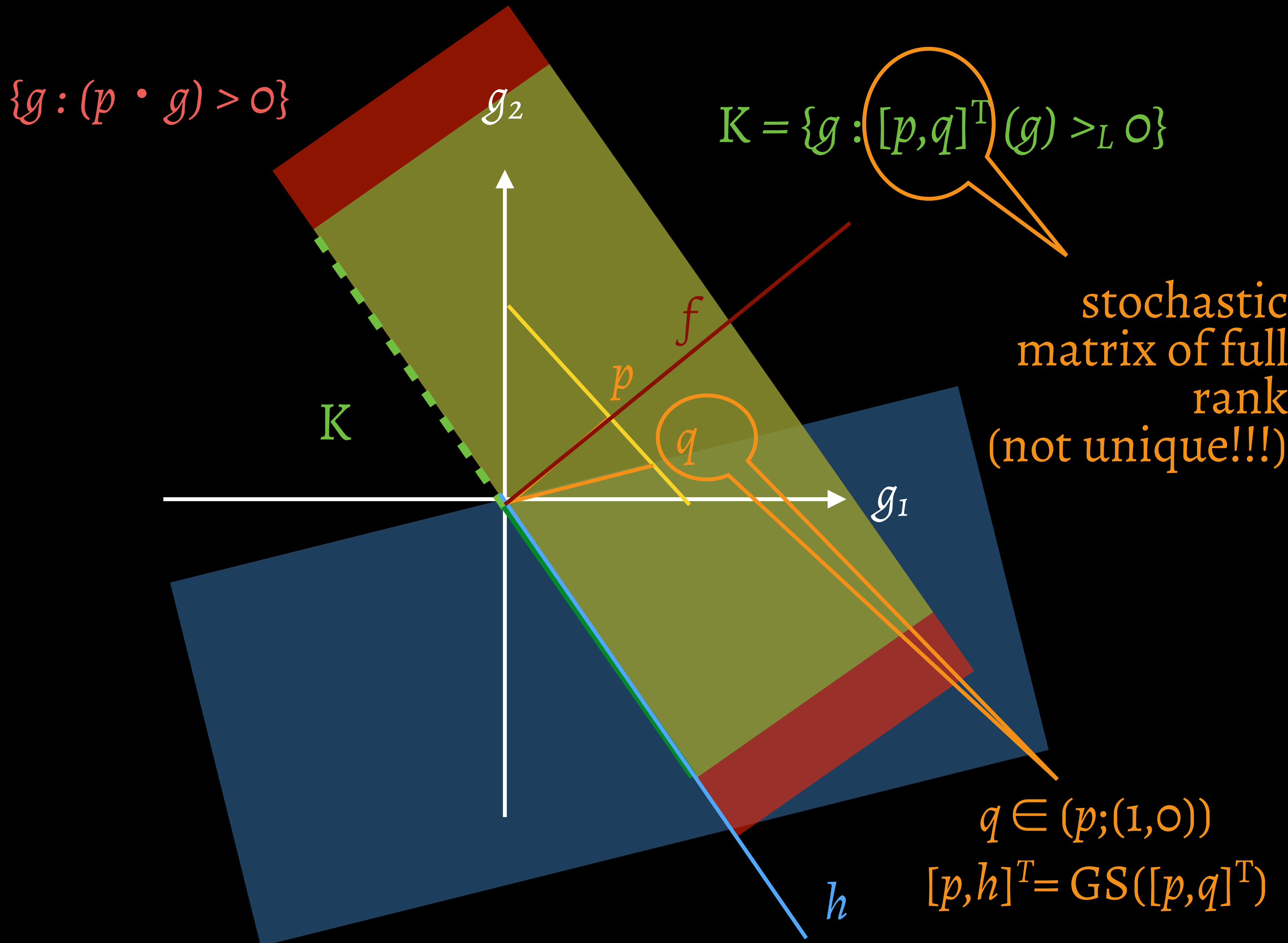
orthogonal matrix (unique when orthonormal)



Lexicographic polarity



Lexicographic polarity



The case of desirability revisited

Central tools:

Lexicographic separation theorem for convex sets, and the (lexicographic) polarity operator $(\,)^{\blacktriangledown}$

Theorem [*Martinez-Legaz (1983)*]: If $K \subset \mathbb{R}^n$ is a nonempty convex set, then for every $g \notin K$ there exists $A \in \mathbb{M}_n$ (even $A \in \mathbb{O}_n$) and $b \in \mathbb{R}^n$ such that for all $f \in K$

$$A(f) >_L b \geq_L A(g).$$

Central tools:

Lexicographic separation theorem for convex sets, and the (lexicographic) polarity operator $(\ )^{\blacktriangledown}$

Given $K \subseteq \mathbb{R}^n$, its *L-polar* is the set

$$K^{\blacktriangledown} = \{ A \in \mathbb{M}_n : A(f) >_L 0, \forall f \in K \}$$

We call a set $M \subseteq \mathbb{M}_n$ a *L-convex cone* if it is the L-polar of some (i.e. $M = K^{\blacktriangledown}$)

Central tools:

Lexicographic separation theorem for convex sets, and the (lexicographic) polarity operator $(\,)^{\blacktriangledown}$

Given $K \subseteq \mathbb{R}^n$, its *L-polar* is the set

$$K^{\blacktriangledown} = \{ A \in \mathbb{M}_n : A(f) >_L 0, \forall f \in K \}$$

We call a set $M \subseteq \mathbb{M}_n$ a *L-convex cone* if it is the L-polar of some set (i.e. $M = K^{\blacktriangledown}$)

Central tools:

Lexicographic separation theorem for convex sets, and the (lexicographic) polarity operator $(\ )^{\blacktriangledown}$

Given $K \subseteq \mathbb{R}^n$, its *L-polar* is the set

$$K^{\blacktriangledown} = \{ A \in \mathbb{M}_n : A(f) >_L 0, \forall f \in K \}$$

We call a set $M \subseteq \mathbb{M}_n$ a *L-credal set* if it is the intersection of the set $\mathbb{T}_n$ of all stochastic matrices of full rank with some L-convex cone.

Central tools:

Lexicographic separation theorem for convex sets,
and the (lexicographic) polarity operator $(\ )^{\blacktriangledown}$

Given $M \subseteq \mathbb{M}_n$ we also define the convex cone omitting its apex

$$M^{\triangle} = \{ g \in \mathbb{R}^n : A(g) >_L 0, \forall A \in M\}$$

Proposition: $K \subseteq \mathbb{R}^n$ is a convex cone omitting its apex if and only if $K = (K^{\blacktriangledown})^{\triangle}$

Reformulating the lexicographic separation theorem:

Proposition: Let $K \in \mathbb{D}_n$ and $g \notin K$, then there exists a matrix $A \in \mathbb{O}_n$ with $A >_L 0$ such that

$$K \subset \{A\}^{\triangle} \text{ but } g \notin \{A\}^{\triangle}.$$

If K is maximal, A is unique.

Where $A >_L 0$ means that each column a of A is such that $a >_L 0$.

Theorem: The map $\mathrm{K} \longmapsto \mathscr{G}(\mathrm{K}) := K^{\blacktriangledown} \cap \mathbb{T}_n$ is a bijection between coherent sets of desirable gambles and L-credal sets whose inverse is $(\ldots)^{\triangle}$.

The definition of conditioning for stochastic matrices is a variation of the definition by Blume et al. (1991) of conditioning for lexicographic systems.

$$P = \begin{array}{ccc} (\omega_1) & (\omega_2) & (\omega_3) \\ 0.5 & 0.25 & 0.25 \\ 0 & 0.5 & 0.5 \\ 0.25 & 0.5 & 0.25 \end{array}$$

we condition with respect to $\Pi=\{\omega_2,\omega_3\}$

The definition of conditioning for stochastic matrices is a variation of the definition by Blume et al. (1991) of conditioning for lexicographic systems.

	(ω_1)	(ω_2)	(ω_3)
$P =$	0.5	0.25	0.25
	0	0.5	0.5
	0.25	0.5	0.25

we condition with respect to $\Pi=\{\omega_2,\omega_3\}$

The definition of conditioning for stochastic matrices is a variation of the definition by Blume et al. (1991) of conditioning for lexicographic systems.

$$P_1 = \begin{array}{ccc} (\omega_1) & (\omega_2) & (\omega_3) \\ 0 & 0.5 & 0.5 \\ 0 & 0 & 0 \\ 0 & 0.67 & 0.33 \end{array}$$

we condition with respect to $\Pi=\{\omega_2,\omega_3\}$

1. for each row *p*, we take $p(.|\Pi)$ if defined, else 0

The definition of conditioning for stochastic matrices is a variation of the definition by Blume et al. (1991) of conditioning for lexicographic systems.

$$P_2 = \begin{array}{ccc} (\omega_1) & (\omega_2) & (\omega_3) \\ 0 & 0.5 & 0.5 \\ 0 & 0.67 & 0.33 \end{array}$$

we condition with respect to $\Pi=\{\omega_2,\omega_3\}$

2. we discard each row which is a linear combination of the rows preceding it

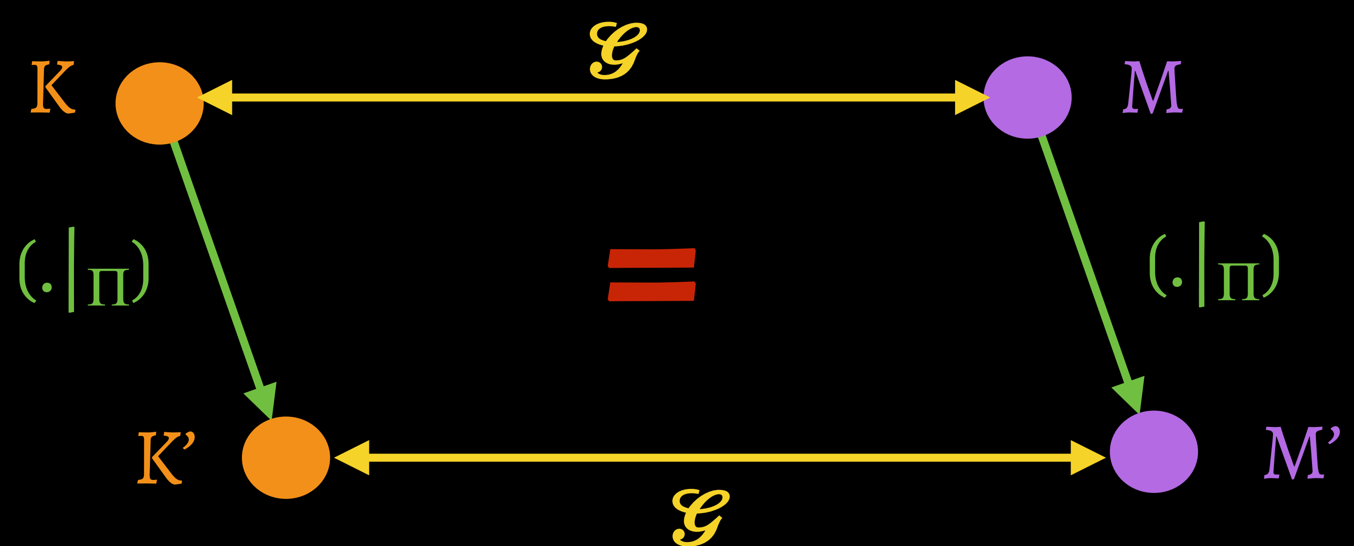
The definition of conditioning for stochastic matrices is a variation of the definition by Blume et al. (1991) of conditioning for lexicographic systems.

	(ω_2)	(ω_3)
$P\|\Pi =$	0.5	0.5
	0.67	0.33

we condition with respect to $\Pi=\{\omega_2,\omega_3\}$

3. we take the projections over Π

Theorem: The map $\mathrm{K} \longmapsto \mathscr{G}(\mathrm{K}) := K^{\blacktriangledown} \cap \mathbb{T}_n$ is an isomorphism between the collection of all coherent sets of desirable gambles equipped with the conditioning operation and the collection of all L-credal sets equipped with the corresponding conditioning operation.



What next

- Complete the analysis by considering e.g. marginalisation and independence
- Geometrical properties of L-credal sets
- Obtain similar correspondences when the sample space is infinite
- Describe correspondence within Category Theory, hence subsuming both **classical** and **quantum** probabilistic cases [joint work w/ F. Zanasi] (*see Alessio Benavoli's invited talk for the latter case and the role of polarity/duality for deriving QM*)

What next

- Complete the analysis by considering e.g. marginalisation and independence
- Geometrical properties of L-credal sets
- Obtain similar correspondences when the sample space is infinite
- Describe correspondence within Category Theory, hence subsuming both **classical** and **quantum** probabilistic cases [joint work w/ F. Zanasi]
  (*see Alessio Benavoli's invited talk for the latter case and the role of polarity/duality for deriving* QM)