# Sophisticated Learning

Duarte Gonçalves

University College London

Topics in Economic Theory

# Overview

# Overview

# Learning in Games

**How do people get to play equilibrium?**

Main question of interest in ‘learning in games’ ($\neq$ games with learning)

**Goals**

Provide foundations for existing equilibrium concepts.

Capture lab behaviour.

Predict adjustment dynamics transitioning to new equilibrium.
(akin to ‘impulse response’ in macro; uncommon but definitely worth investigating)

Select equilibria.

Algorithm to solve for equilibria.

Explain persistence of heuristics/nonequilibrium behaviour.

# Overview

# Overview

Papers:
Kalai and Lehrer (1993, Ecta) "Rational Learning Leads to Nash Equilibria"
Kalai and Lehrer (1993 Ecma) "Subjective Equilibrium in Repeated Games"

# Model and Notation

**Stage game**: finite players $i \in I = \{1, \dots, n\}$; actions $A_i$; $A = \times_i A_i$; payoffs $u_i : A \to \mathbb{R}$.

**Infinite horizon repeated game**: discount $\delta_i \in (0, 1)$; perfect monitoring.

**Histories**: $h^t = (a^0, \dots, a^{t-1}) \in A^t$, with $H^t = A^t$, $H = \cup_{t \geq 0} H^t$; empty history $\emptyset$.

**Behavioural strategies**: $\sigma_i = (\sigma_{i,t})_{t \geq 0}$, $\sigma_{i,t} : H^t \to \Delta(A_i)$; strategy profile $\sigma = (\sigma_i)_i$.

**Outcome measure**: for a fixed $\sigma$, let $\mu^\sigma$ be the induced probability on infinite play paths $\Omega = A^{\mathbb{N}}$.

**Subjective beliefs**: each player $i$ has a prior belief $\nu_i$ over opponents' strategies $\sigma_{-i}$; induces a belief over play paths $\Pi_i$.

**Absolute continuity (truth-compatibility)**: $\mu^\sigma \ll \Pi_i$ for all $i$ (no path of positive $\mu^\sigma$-probability is assigned zero by $\Pi_i$).

**Objective**: players maximise expected discounted payoffs given their posteriors and choose best responses period by period.

# Bayesian Updating and Merging

**Posterior on play paths**: after $h^t$, player $i$ updates $\Pi_i(\cdot \mid h^t)$ by Bayes' rule (well-defined by absolute continuity).

**Merging (KL notion)**: posteriors become *close* on all tail events: for every $\varepsilon > 0$, $\exists T$ s.t. $\forall s \geq T$,
$|\Pi_i(A \mid h^s) - \mu^\sigma(A \mid h^s)| \leq \varepsilon$ for all $A$ in a large (probability-1) class of events.

**Relation to Blackwell–Dubins**: KL's closeness $\iff$ BD merging; KL give an elementary proof and equivalence of topologies.

# Subjective Equilibrium

**Definition (Subjective equilibrium)**

A history-dependent strategy profile $\sigma$ is a **subjective equilibrium** if, along $\mu^{\sigma}$-almost all paths, players' posteriors about future play coincide with the truth (merging), and each $\sigma_i$ is a best response to the posterior over $\sigma_{-i}$.

**Interpretation**: learning exhausted; disagreements (if any) are off path and never observed.

**Consequence**: from some finite time $T$, actions follow best responses to (approximately) correct forecasts of future play.

# Main Result: Rational Learning $\implies$ Nash Equilibrium

**Theorem 1 (Kalai and Lehrer 1993)**

Suppose $\mu^\sigma \ll \Pi_i$ for all $i$ (absolute continuity). Under Bayesian updating and optimal control of expected discounted utility:

Posteriors *merge* with the truth along the realised path.

From some finite time, play is $\varepsilon$-optimal against correct forecasts.

Limit behaviour constitutes a Nash equilibrium of the repeated game.

**Proof Sketch**

**Merging**: apply KL's merging theorem to obtain posterior convergence on tail events.

**Optimality**: best responses w.r.t. posteriors $\implies$ $\varepsilon$-optimality w.r.t. truth for large $t$.

**Equilibrium**: mutual best responses along the limit set $\implies$ Nash equilibrium of the repeated game.

# Corollaries and Immediate Implications

**Corollary 2 (Kalai and Lehrer 1993; BD 1962)**

For $\mu^\sigma$-a.e. path, the posterior probabilities $\Pi_i(\cdot \mid h^t)$ converge uniformly on the large class of events to $\mu^\sigma(\cdot \mid h^t)$. (Version of Blackwell–Dubins' merging.)

# Corollaries and Immediate Implications

**Corollary 2 (Kalai and Lehrer 1993; BD 1962)**

For $\mu^\sigma$-a.e. path, the posterior probabilities $\Pi_i(\cdot \mid h^t)$ converge uniformly on the large class of events to $\mu^\sigma(\cdot \mid h^t)$. (Version of Blackwell–Dubins' merging.)

**Corollary (Incomplete information; Bayesian Nash equilibrium)**

In a discounted repeated game with a finite/countable type space for payoffs, if play starts at a **Bayesian Nash equilibrium** of the incomplete-information repeated game, then eventually players play a Nash equilibrium of the realised complete-information repeated game.

Intuition: at BNE of the incomplete-information repeated game, priors imply absolute continuity on play paths; merging $\implies$ players act *as if* types were known.

# Meaning and Interpretation

**What converges?** Posteriors about future play; best responses to (nearly) correct forecasts $\implies$ Nash play in the repeated game.

**Role of absolute continuity**: rules out dogmatic priors that assign zero to realised events; ensures Bayes can learn from data.

**Why repeated games?** Stationarity of opponents' *strategies* (not actions) makes learning feasible despite strategic feedback.

**Experimentation**: endogenous via dynamic optimisation of discounted utility; no ad hoc trembles needed.

# Additional Details (Notation as in §3.1)

**Spaces**: $\Omega = A^{\mathbb{N}}$ with product $\sigma$-algebra; cylinders generated by finite histories.

**Outcome law**: $\mu^{\sigma}(\cdot) = \mathbb{P}_{\sigma}(\cdot)$ over $\Omega$; filtration $\mathcal{F}_t$ from $H^t$.

**Beliefs on strategies $\rightarrow$ beliefs on paths**: priors $\nu_i$ over $\sigma_{-i}$ induce $\Pi_i$ over $\Omega$ (via mapping $\sigma_{-i} \mapsto \mu^{(\sigma_i, \sigma_{-i})}$).

**Absolute continuity on $\Omega$**: $\mu^{\sigma} \ll \Pi_i$; equivalently, every cylinder $C(h^t)$ with $\mu^{\sigma}(C(h^t)) > 0$ has $\Pi_i(C(h^t)) > 0$.

**Payoffs**: $U_i(\sigma) = \mathbb{E}_{\mu^{\sigma}}\left[\sum_{t\geq 0} \delta_i^t \, u_i(a^t)\right]$.

# Critiques (Fudenberg and Levine 1998)

**Endogeneity of absolute continuity**: AC must hold for the *realised* play path $\Longrightarrow$ fixed-point flavour; as hard as equilibrium selection.

**Grain of truth**: desirable to ensure AC *regardless* of opponents' play is impossible on uncountable history spaces; weaker classes of priors may work only in truncated/favourable settings.

**Example (Chicken)**: plausible sets "insist $n$ periods then yield"; symmetric beliefs $\Longrightarrow$ optimal stopping leads to paths of measure $0$ under priors $\Rightarrow$ AC fails.

**Interpretation**: best seen as a descriptive result on eventual consensus, not as a *learning path* to equilibrium with exogenously specified priors.

# Takeaways

Under absolute continuity, Bayesian learning merges beliefs with the truth on play paths.

Optimal control with merged beliefs $\implies$ eventual play of a Nash equilibrium of the repeated game.

Application to *Bayesian Nash equilibria* with incomplete information: eventually play NE of the realised complete-information game.

AC is strong/endogenous; caution interpreting KL as a general path-to-equilibrium theory.

# Model and Notation

**Stage game**: finite players $i \in I = \{1, \dots, n\}$; actions $A_i$; $A = \times_i A_i$; payoffs $u_i : A \to \mathbb{R}$.

**Infinite horizon repeated game**: discount $\delta_i \in (0, 1)$; perfect monitoring.

**Histories**: $h^t = (a^0, \dots, a^{t-1}) \in A^t$, with $H^t = A^t$, $H = \cup_{t \geq 0} H^t$; empty history $\emptyset$.

**Behavioural strategies**: $\sigma_i = (\sigma_{i,t})_{t \geq 0}$, $\sigma_{i,t} : H^t \to \Delta(A_i)$; strategy profile $\sigma = (\sigma_i)_i$.

**Outcome measure**: for a fixed $\sigma$, let $\mu^\sigma$ be the induced probability on infinite play paths $\Omega = A^{\mathbb{N}}$.

**Subjective beliefs**: each player $i$ has a prior belief $\nu_i$ over opponents' strategies $\sigma_{-i}$; induces a belief over play paths $\Pi_i$.

**Absolute continuity (truth-compatibility)**: $\mu^\sigma \ll \Pi_i$ for all $i$ (no path of positive $\mu^\sigma$-probability is assigned zero by $\Pi_i$).

**Objective**: players maximise expected discounted payoffs given their posteriors and choose best responses period by period.

# Bayesian Updating and Merging

**Posterior on play paths**: after $h^t$, player $i$ updates $\Pi_i(\cdot \mid h^t)$ by Bayes' rule (well-defined by absolute continuity).

**Merging (KL notion)**: posteriors become *close* on all tail events: for every $\varepsilon > 0$, $\exists T$ s.t. $\forall s \geq T$,
$|\Pi_i(A \mid h^s) - \mu^\sigma(A \mid h^s)| \leq \varepsilon$ for all $A$ in a large (probability-1) class of events.

**Relation to Blackwell–Dubins**: KL's closeness $\iff$ BD merging; KL give an elementary proof and equivalence of topologies.

# Subjective Equilibrium

**Definition (Subjective equilibrium)**

A history-dependent strategy profile $\sigma$ is a **subjective equilibrium** if, along $\mu^{\sigma}$-almost all paths, players' posteriors about future play coincide with the truth (merging), and each $\sigma_i$ is a best response to the posterior over $\sigma_{-i}$.

**Interpretation**: learning exhausted; disagreements (if any) are off path and never observed.

**Consequence**: from some finite time $T$, actions follow best responses to (approximately) correct forecasts of future play.

# Main Result: Rational Learning $\implies$ Nash Equilibrium

**Theorem 1 (Kalai and Lehrer 1993)**

Suppose $\mu^\sigma \ll \Pi_i$ for all $i$ (absolute continuity). Under Bayesian updating and optimal control of expected discounted utility:

- Posteriors *merge* with the truth along the realised path.
- From some finite time, play is $\varepsilon$-optimal against correct forecasts.
- Limit behaviour constitutes a Nash equilibrium of the repeated game.

**Proof Sketch**

- **Merging**: apply KL's merging theorem to obtain posterior convergence on tail events.
- **Optimality**: best responses w.r.t. posteriors $\implies$ $\varepsilon$-optimality w.r.t. truth for large $t$.
- **Equilibrium**: mutual best responses along the limit set $\implies$ Nash equilibrium of the repeated game.

# Corollaries and Immediate Implications

**Corollary 2 (Kalai and Lehrer 1993; BD 1962)**

For $\mu^{\sigma}$-a.e. path, the posterior probabilities $\Pi_i(\cdot \mid h^t)$ converge uniformly on the large class of events to $\mu^{\sigma}(\cdot \mid h^t)$. (Version of Blackwell–Dubins' merging.)

# Corollaries and Immediate Implications

**Corollary 2 (Kalai and Lehrer 1993; BD 1962)**

For $\mu^\sigma$-a.e. path, the posterior probabilities $\Pi_i(\cdot \mid h^t)$ converge uniformly on the large class of events to $\mu^\sigma(\cdot \mid h^t)$. (Version of Blackwell–Dubins' merging.)

**Corollary (Incomplete information; Bayesian Nash equilibrium)**

In a discounted repeated game with a finite/countable type space for payoffs, if play starts at a **Bayesian Nash equilibrium** of the incomplete-information repeated game, then eventually players play a Nash equilibrium of the realised complete-information repeated game.

Intuition: at BNE of the incomplete-information repeated game, priors imply absolute continuity on play paths; merging $\implies$ players act *as if* types were known.

# Meaning and Interpretation

**What converges?** Posteriors about future play; best responses to (nearly) correct forecasts $\implies$ Nash play in the repeated game.

**Role of absolute continuity**: rules out dogmatic priors that assign zero to realised events; ensures Bayes can learn from data.

**Why repeated games?** Stationarity of opponents' *strategies* (not actions) makes learning feasible despite strategic feedback.

**Experimentation**: endogenous via dynamic optimisation of discounted utility; no ad hoc trembles needed.

# Additional Details (Notation as in §3.1)

**Spaces**: $\Omega = A^{\mathbb{N}}$ with product $\sigma$-algebra; cylinders generated by finite histories.

**Outcome law**: $\mu^{\sigma}(\cdot) = \mathbb{P}_{\sigma}(\cdot)$ over $\Omega$; filtration $\mathcal{F}_t$ from $H^t$.

**Beliefs on strategies → beliefs on paths**: priors $\nu_i$ over $\sigma_{-i}$ induce $\Pi_i$ over $\Omega$ (via mapping $\sigma_{-i} \mapsto \mu^{(\sigma_i,\sigma_{-i})}$).

**Absolute continuity on $\Omega$**: $\mu^{\sigma} \ll \Pi_i$; equivalently, every cylinder $C(h^t)$ with $\mu^{\sigma}(C(h^t)) > 0$ has $\Pi_i(C(h^t)) > 0$.

**Payoffs**: $U_i(\sigma) = \mathbb{E}_{\mu^{\sigma}}\left[\sum_{t\geq 0} \delta_i^t\, u_i(a^t)\right]$.

# Critiques (Fudenberg and Levine 1998)

**Endogeneity of absolute continuity**: AC must hold for the *realised* play path $\Longrightarrow$ fixed-point flavour; as hard as equilibrium selection.

**Grain of truth**: desirable to ensure AC *regardless* of opponents' play is impossible on uncountable history spaces; weaker classes of priors may work only in truncated/favourable settings.

**Example (Chicken)**: plausible sets "insist $n$ periods then yield"; symmetric beliefs $\Longrightarrow$ optimal stopping leads to paths of measure $0$ under priors $\Rightarrow$ AC fails.

**Interpretation**: best seen as a descriptive result on eventual consensus, not as a *learning path* to equilibrium with exogenously specified priors.

# Takeaways

- Under absolute continuity, Bayesian learning merges beliefs with the truth on play paths.
- Optimal control with merged beliefs $\implies$ eventual play of a Nash equilibrium of the repeated game.
- Application to *Bayesian Nash equilibria* with incomplete information: eventually play NE of the realised complete-information game.
- AC is strong/endogenous; caution interpreting KL as a general path-to-equilibrium theory.

# Sophisticated Learning

Duarte Gonçalves

University College London

Topics in Economic Theory