Chapter 4

Nash Equilibrium

Dominance arguments are compelling but often silent: in the Battle of the Sexes or the Stag Hunt, nothing is dominated, and rationalizability permits everything. The central solution concept of game theory adds one more requirement: beliefs must be *correct*.

4.1 Definition and interpretation

Definition 4.1: Best response correspondence

Player i's *best response correspondence* maps each profile of opponents' actions to the set of i's optimal actions:

$$\mathrm{BR}_i(a_{-i}) \;=\; \arg\max_{a_i \in A_i} u_i(a_i, a_{-i}).$$

Definition 4.2: Nash equilibrium

An action profile $a^* = (a_1^*, \ldots, a_n^*)$ is a *Nash equilibrium* if for every player i,

$$u_i(a_i^*, a_{-i}^*) \;\geq\; u_i(a_i, a_{-i}^*) \qquad \text{for all } a_i \in A_i,$$

equivalently $a_i^* \in \mathrm{BR}_i(a_{-i}^*)$ for every i: *every player's action is a best response to the others' equilibrium actions.*

Key Idea: No profitable unilateral deviation

A Nash equilibrium is a profile of mutually consistent plans: if everyone expects a^* to be played, no individual can gain by deviating *alone*. It is a *stability* or *self-enforcement* condition—not a claim that the outcome is good (the Prisoner's Dilemma equilibrium is bad for both players) nor that players somehow negotiate it.

Why should we expect equilibrium play? Several complementary interpretations:

- **Self-enforcing agreement**: if players discuss beforehand and reach a non-binding agreement, only Nash profiles are viable—anything else, someone would break.
- **Stable social convention**: in a population repeatedly playing the game (drivers at intersections, market participants), a convention persists only if no one benefits from unilaterally deviating—i.e., only if it is Nash.
- **Rational expectations**: if a theory of the game's play were known to all players, rational

players would follow it only if it prescribed a Nash equilibrium; any other prediction refutes itself.

- **Limit of learning/evolution**: many adaptive dynamics, when they converge, converge to Nash equilibria (Chapters 16, 17).

4.2 Finding pure-strategy Nash equilibria

For matrix games, the mechanical method is *best-response marking*: for each column, underline the row player's best payoff(s); for each row, underline the column player's best payoff(s); cells where both payoffs are underlined are Nash equilibria.

Example 4.1: Marking best responses

	L	C	R
T	$\underline{7},\underline{6}$	$5,5$	$0,4$
M	$5,1$	$\underline{6},\underline{8}$	$1,2$
B	$3,3$	$4,4$	$\underline{2},\underline{5}$

Check: against L, row's best is T (7); against C, M (6); against R, B (2). Against T, column's best is L (6); against M, C (8); against B, R (5). All three diagonal cells are Nash equilibria—a reminder that equilibria need not be unique and can be Pareto-ranked.

Revisiting the classics of Chapter 3:

- **Prisoner's Dilemma**: unique equilibrium (D, D).
- **Stag Hunt**: two pure equilibria, (S, S) and (H, H).
- **Battle of the Sexes**: two pure equilibria, (B, B) and (F, F).
- **Matching Pennies**: *no* pure equilibrium—the motivation for mixed strategies (Chapter 5).

4.3 Games with continuous action spaces

When actions are continuous, we find equilibria by solving the players' first-order conditions simultaneously.

Example 4.2: Cournot duopoly

Firms 1 and 2 simultaneously choose outputs $q_1, q_2 \geq 0$. The market price is $P = a - q_1 - q_2$ (when positive) and both firms have unit cost $c < a$, so

$$u_i(q_i, q_j) = q_i\,(a - q_i - q_j - c).$$

Best response: maximizing over q_i gives $\mathrm{BR}_i(q_j) = \frac{a-c-q_j}{2}$ (when nonnegative). Solving $q_1 = \mathrm{BR}_1(q_2)$ and $q_2 = \mathrm{BR}_2(q_1)$:

$$q_1^* = q_2^* = \frac{a-c}{3}, \qquad P^* = \frac{a+2c}{3}, \qquad u_i^* = \frac{(a-c)^2}{9}.$$

Each firm produces more than half the monopoly quantity $\frac{a-c}{2}$ in total; joint profit is below the monopoly level—a continuous cousin of the Prisoner's Dilemma (each firm's output exerts a negative externality on the other). Full treatment in Chapter 7.

Example 4.3: Tragedy of the commons

n herders each choose how many goats $g_i \geq 0$ to graze on a common pasture. The value per goat is $v(G) = 100 - G$ where $G = \sum_j g_j$, so $u_i = g_i(100 - G)$. The first-order condition for herder i is $100 - G - g_i = 0$. By symmetry $g_i = g$ and $(n+1)g = 100$:

$$g^* = \frac{100}{n+1}, \qquad G^* = \frac{100n}{n+1} \xrightarrow{n\to\infty} 100.$$

The socially optimal total (maximizing $G(100 - G)$) is $G^{\text{opt}} = 50$; the commons is overgrazed, the more so the more herders there are. Each herder ignores the cost his goats impose on others—an *externality*, the structural cause of inefficient equilibria.

4.4 Existence, multiplicity, and selection

Three basic facts frame how we use Nash equilibrium.

Theorem 4.1: Nash's existence theorem (1950) — statement

Every finite game has at least one Nash equilibrium, possibly in mixed strategies.

The proof, via fixed-point theorems, is given in Chapter 13. Existence makes Nash equilibrium a universally applicable prediction tool.

Multiplicity. Many games have several equilibria (Stag Hunt, Battle of the Sexes). Game theory then needs a theory of *selection*:

- **Focal points** (Schelling 1960): cultural or contextual salience can pick an equilibrium. Asked to meet a stranger in New York with no communication, many people say "noon at Grand Central"—nothing in the mathematics, but a powerful coordinating force.
- **Payoff dominance vs. risk dominance**: in the Stag Hunt, (S, S) gives higher payoffs, but (H, H) is safer against strategic uncertainty. Evolutionary and learning models often select the risk-dominant equilibrium.
- **Communication and cheap talk** can help select equilibria (though not always credibly—Chapter 11).
- **Refinements** sharpen predictions in dynamic games (Chapters 6, 13).

Efficiency. Nash equilibria may be socially inefficient (Prisoner's Dilemma, commons). Quantifying this inefficiency—the *price of anarchy*—is a central theme of algorithmic game theory (Chapter 17).

Historical Note

John Nash's 1950 PhD thesis at Princeton was 27 pages long. Its central concept was initially dismissed by von Neumann as "just a fixed point theorem," yet it became the cornerstone of modern economics. Nash shared the 1994 Nobel Prize with Harsanyi and Selten. His life—including his struggle with schizophrenia and late-life recognition—is

portrayed in *A Beautiful Mind*.

4.5 Correlated equilibrium: a first look

Nash equilibrium assumes independent randomization. Aumann (1974) observed that a *correlation device* can enlarge the set of stable outcomes.

Example 4.4: Traffic light

In the Chicken/Hawk–Dove game

	Yield	Dare
Yield	$4,4$	$2,5$
Dare	$5,2$	$0,0$

the pure equilibria are (Dare, Yield) and (Yield, Dare); the symmetric mixed equilibrium gives expected payoff $10/3$ each. Now install a fair traffic light: with probability $\frac{1}{2}$ it shows (Green, Red), else (Red, Green); each driver sees only their own light, and the convention is Green $\to$ Dare, Red $\to$ Yield. Given a red light you know the other is daring, so yielding is optimal ($2 > 0$); given green, the other yields, so daring is optimal ($5 > 4$). Nobody deviates, and the average payoff is $\frac{5+2}{2} = 3.5 > 10/3$. This is a *correlated equilibrium* that is not a Nash equilibrium. The general theory—including why correlated equilibria are computationally far more tractable than Nash—appears in Chapter 17.

Exercises

Exercise 4.1: Find all pure equilibria

	X	Y	Z
A	$3,2$	$0,0$	$1,1$
B	$0,0$	$2,3$	$1,0$
C	$1,1$	$0,1$	$2,2$

Use best-response marking to find all pure-strategy Nash equilibria.

Exercise 4.2: Partnership game

Two partners each choose effort $e_i \in [0,4]$. Output is $e_1 + e_2 + \frac{1}{2}e_1e_2$, shared equally; effort costs e_i^2. So $u_i = \frac{1}{2}(e_1 + e_2 + \frac{1}{2}e_1e_2) - e_i^2$. Find the Nash equilibrium efforts and compare with the efforts that maximize the sum of payoffs. Explain the direction of the inefficiency.

Exercise 4.3: Hotelling's beach

Two ice-cream vendors choose locations on a beach $[0,1]$; customers are uniformly distributed and walk to the nearest vendor (splitting equally if equidistant). Each vendor maximizes market share. Show that both locating at $\frac{1}{2}$ is the unique Nash equilibrium. Relate this to the observed similarity of political platforms (the median voter theorem).

Exercise 4.4: Equilibrium is not optimality

Construct a 2×2 game in which the unique Nash equilibrium gives both players their *worst* payoff among all four cells, or prove that this is impossible. What is the strongest statement of this kind that *is* possible?