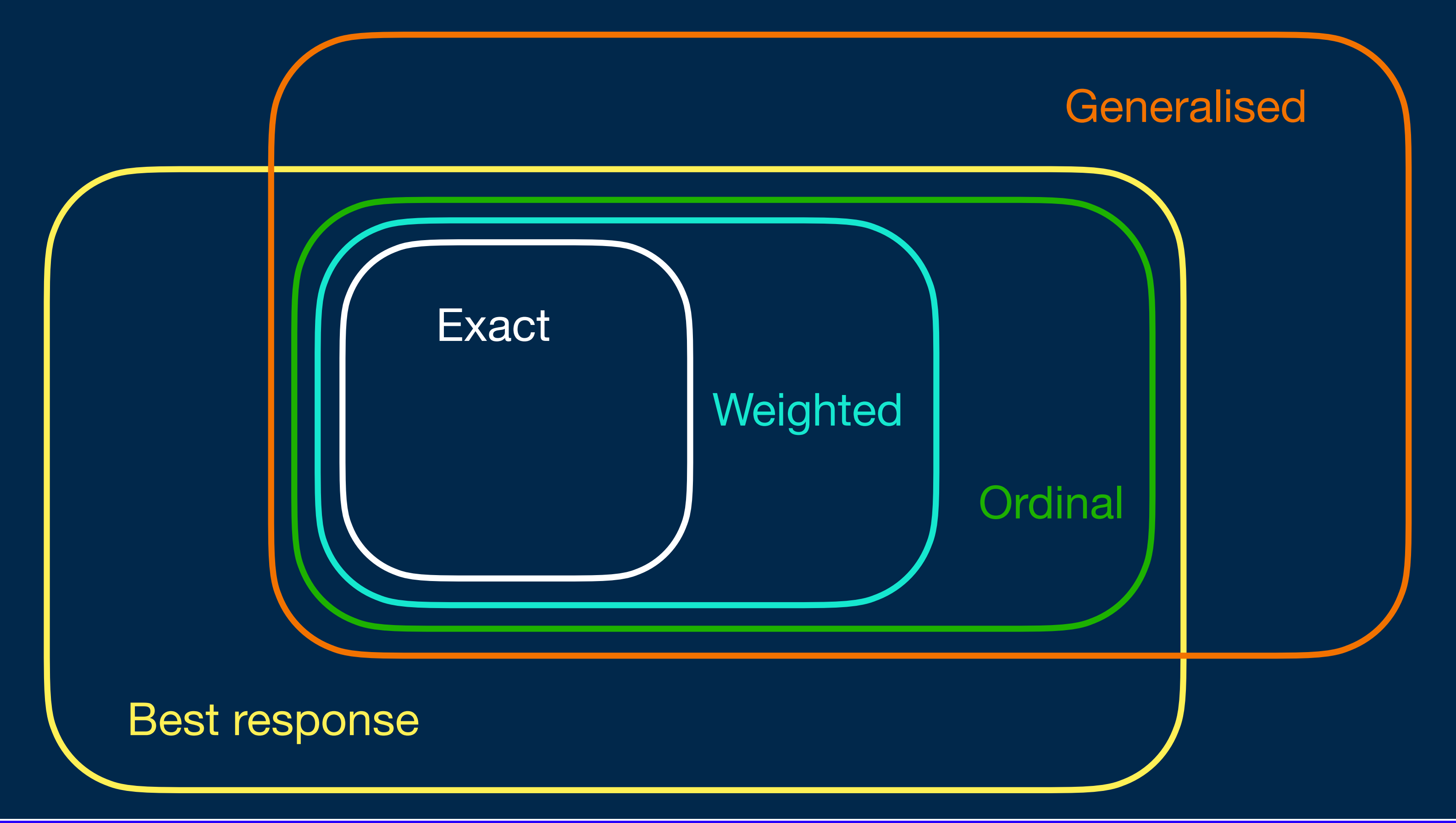

Chapter 2 of "Potential Game Theory"

Theories and applications

Author: Robert P. Gilles

Institute: Department of Economics

The Queen's University of Belfast

Date: April 17, 2026

Version: 3.00

Email: `r.gilles@qub.ac.uk`

To the many shoulders of thinkers on which this work rests

Contents

Meaning of the word "potential".

Meaning 1: Existing in possibility, not in act.
Meaning 2: Having the effect without the external actual property.
Meaning 3: Efficacious; powerful. (Noted as "not in use".)
Meaning 4: Grammatical usage – a mood denoting possibility.

Description taken from Samuel Johnson's *A Dictionary of the English Language* (1755).

Preface and acknowledgements

Preface yet to be written.

Chapter 2
Games with Exact Potentials

Chapter summary

- *Definition of exact potential games*
- *Decomposition of exact potential games*
- *Characterisation for differentiable potential games*
- *Ui's characterisation for finite games*
- *General characterisations*
- *Computing exact potentials*
- *The decomposition of arbitrary games into potential and harmonic games*
- *The potentialness of arbitrary games*

In this chapter I discuss the primary class of normal form games associated with a potential function. These games are known as *exact* potential games. We explore the characterisation of this class of games through several approaches, including decomposition methods as well as a constructive technique for the potential function itself.

2.1 Exact potential games

Monderer and Shapley (1996) formally defined the concept of an exact potential game. Recall that a *potential* for a pre-game $(N, \mathbf{A})$ is a function $\Psi\colon \mathbf{A} \to \mathbb{R}$, on which we have not imposed any restrictions. We denoted $\Gamma_\Psi = \langle N, \mathbf{A}, \Psi\rangle$ as a *potential game* and defined its equilibria as the saddle points of Ψ.

Exact potential games are games in which the payoff differentials coincide exactly with the differentials of a potential function. The payoff structure of an exact potential game is therefore entirely summarised in its corresponding potential function.

Definition 2.1

*A game $\Gamma = \langle N, \mathbf{A}, \pi\rangle$ is an **exact potential game** if there exists some potential function $\Psi\colon \mathbf{A} \to \mathbb{R}$ such that for every player $i \in N$, every $\bar{a}_{-i} \in \mathbf{A}_{-i}$ and all $a_i, b_i \in A_i$:*

$$\pi_i(a_i, \bar{a}_{-i}) - \pi_i(b_i, \bar{a}_{-i}) = \Psi(a_i, \bar{a}_{-i}) - \Psi(b_i, \bar{a}_{-i}) \tag{2.1}$$ ♣

The main feature captured by this definition is that in an exact potential game all players share a single incentive: each player improves her payoff only by increasing the value of the potential function. All players therefore follow the same payoff-increasing path, dictated by a common potential function.

The next example explores this definition further by explicitly constructing an exact potential function for the Cournot duopoly with linear market demand.

Example 2.1 Consider again the Cournot duopoly with two firms in a market with a linear demand. In particular, the game is described by $N = \{1, 2\}$, $A_1 = A_2 = \mathbb{R}_+$ and payoff functions given by

$$\pi_1(q_1, q_2) = P(q_1 + q_2)q_1 - c_1(q_1)$$
$$\pi_2(q_1, q_2) = P(q_1 + q_2)q_2 - c_2(q_2)$$

where $P(Q) = \alpha - \beta Q$ is a linear market demand function and $c_1, c_2\colon \mathbb{R}_+ \to \mathbb{R}_+$ are monotone increasing cost functions for both firms in the market.

We now rewrite the two payoff functions as a function of q_i only and a function that describes the interaction of both firms. This results in

$$\pi_1(q_1, q_2) = (\alpha - \beta q_1)q_1 - c_1(q_1) - \beta q_1 q_2 = \Pi_1(q_1) - \beta q_1 q_2$$
$$\pi_2(q_1, q_2) = (\alpha - \beta q_2)q_2 - c_2(q_2) - \beta q_1 q_2 = \Pi_2(q_2) - \beta q_1 q_2$$

We now argue that the function given by

$$\Psi(q_1, q_2) = \Pi_1(q_1) + \Pi_2(q_2) - \beta q_1 q_2 \tag{2.2}$$

is an exact potential for this game. Indeed, we can check that for all $q_1, q_1', q_2 \geqslant 0$ it holds that

$$\begin{aligned}\Psi(q_1, q_2) - \Psi(q_1', q_2) &= [\Pi_1(q_1) + \Pi_2(q_2) - \beta q_1 q_2] - [\Pi_1(q_1') + \Pi_2(q_2) - \beta q_1' q_2] \\ &= [\Pi_1(q_1) - \beta q_1 q_2] - [\Pi_1(q_1') - \beta q_1' q_2] = \pi_1(q_1, q_2) - \pi_1(q_1', q_2)\end{aligned}$$

as required for Ψ to be an exact potential of this Cournot duopoly. ♦

The next insight follows directly from the definition and makes the role of the potential explicit: in an exact potential game, every player's payoff function decomposes into the potential function itself and a so-called *dummy payoff function* that is unresponsive to strategic changes by that player. The strategically relevant payoffs are therefore entirely determined by the potential, while the dummy component merely records the accidental consequences of the opponents' choices.

Proposition 2.1 (Standard decomposition of exact potential games)

A game $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ is an exact potential game with corresponding potential function $\Psi \colon \mathbf{A} \to \mathbb{R}$ if and only if for every player $i \in N$ there exists a game $\delta_i \in \mathbb{G}^N$ such that for every $\bar{a}_{-i} \in \mathbf{A}_{-i}$ and all $a_i, b_i \in A_i$ it holds that $\delta_i(a_i, \bar{a}_{-i}) = \delta_i(b_i, \bar{a}_{-i})$ and $\pi_i = \Psi + \delta_i$. ♠

Proof Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be some game.

First, assume that there exist a functions $\Psi \colon \mathbf{A} \to \mathbb{R}$ and $\delta_i \colon \mathbf{A} \to \mathbb{R}^N$ such that for every $\bar{a}_{-i} \in \mathbf{A}_{-i}$ and all $a_i, b_i \in A_i$ it holds that $\delta_i(a_i, \bar{a}_{-i}) = \delta_i(b_i, \bar{a}_{-i})$ and $\pi_i = \Psi + \delta_i$. We will show that necessarily Γ is an exact potential game with potential function Ψ. Indeed,

$$\begin{aligned}\pi_i(a_i, \bar{a}_{-i}) - \pi_i(b_i, \bar{a}_{-i}) &= (\Psi(a_i, \bar{a}_{-i}) + \delta_i(a_i, \bar{a}_{-i})) - (\Psi(b_i, \bar{a}_{-i}) + \delta_i(b_i, \bar{a}_{-i})) \\ &= \Psi(a_i, \bar{a}_{-i}) - \Psi(b_i, \bar{a}_{-i}) + \delta_i(a_i, \bar{a}_{-i}) - \delta_i(b_i, \bar{a}_{-i}) \\ &= \Psi(a_i, \bar{a}_{-i}) - \Psi(b_i, \bar{a}_{-i})\end{aligned}$$

showing that Γ is an exact potential game with potential Ψ.

Next, Assume that Γ is an exact potential game with potential Ψ. Then for every player $i \in N$ we can write $\pi_i = \Psi + (\pi_i - \Psi)$. Define $\delta_i = \pi_i - \Psi$. We show that δ_i is a dummy game. Indeed, for every $\bar{a}_{-i} \in \mathbf{A}_{-i}$ and all $a_i, b_i \in A_i$ it follows that

$$\delta_i(a_i, \bar{a}_{-i}) = \pi_i(a_i, \bar{a}_{-i}) - \Psi(a_i, \bar{a}_{-i})$$
$$= \pi_i(b_i, \bar{a}_{-i}) - \Psi(b_i, \bar{a}_{-i}) = \delta_i(b_i, \bar{a}_{-i})$$

by the defining property of an exact potential game (2.1). ¶

The decomposition of exact potential games into a coordination game based on the potential itself and individual dummy games allows us to identify potentials for a certain class of games. The next example collects the insights from Balder (1997).

Example 2.2 A game $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ is called a *game with additively coupled payoffs* if for every $i \in N$:

$$\pi_i(a) = \sum_{j \in N} v_{ij}(a_j) \qquad \text{with } v_{ij} \colon A_j \to \mathbb{R} \text{ for every } j \in N.$$

It is clear that for player $i \in N$ the composites v_{ij} form dummy games for all $j \neq i$. This implies with Proposition 2.1 that an exact potential for this game is now given by $\Psi(a) = \sum_{i \in N} v_{ii}(a_i)$. This allows the decomposition of such a game as

$$\pi_i(a) = \Psi(a) + \left[\sum_{j \neq i} v_{ij}(a_j) - \sum_{j \neq i} v_{jj}(a_j) \right]$$

where the bracketed expression is clearly a dummy game payoff for player i. ♦

Furthermore, Proposition 2.1 provides a straightforward test for whether a matrix game has an exact potential. The next example explores this in more detail.

Example 2.3 Consider a simple 2×2 matrix game with abstract payoffs. Then this matrix game is an exact potential game according to the decomposition proposed in Proposition 2.1 if and only if these payoffs can be written as

	L	**R**
U	$\Psi_{11} + \alpha\,,\ \Psi_{11} + \delta$	$\Psi_{12} + \beta\,,\ \Psi_{12} + \delta$
D	$\Psi_{21} + \alpha\,,\ \Psi_{21} + \varepsilon$	$\Psi_{22} + \beta\,,\ \Psi_{22} + \varepsilon$

More explicitly, the original matrix game can be decomposed in a sum of a potential game and a dummy game:

	L	**R**
U	Ψ_{11}, Ψ_{11}	Ψ_{12}, Ψ_{12}
D	Ψ_{21}, Ψ_{21}	Ψ_{22}, Ψ_{22}

$+$

	L	**R**
U	α, δ	β, δ
D	α, ε	β, ε

We remark that this puts clear restrictions on which payoff structures are supporting an exact potential and which do not. For example, the standard Prisoners' Dilemma admits an exact potential, as shown below:

	L	**R**
U	$1, 1$	$3, 0$
D	$0, 3$	$2, 2$

$=$

	L	**R**
U	$2, 2$	$1, 1$
D	$1, 1$	$0, 0$

$+$

	L	**R**
U	$-1, -1$	$2, -1$
D	$-1, 2$	$2, 2$

In this expression, the first matrix—the standard Prisoners' Dilemma—is decomposed in a potential game (second matrix) and a dummy game (third matrix). This decomposition also makes evident that the potential maximiser (U, L) coincides with the unique strict Nash equilibrium in this game.
The next example shows that there are simple games in which the potential maximiser is a Nash equilibrium, but not all Nash equilibria are potential maximisers:

	L	**R**
U	$1, 1$	$1, -1$
D	$-1, 1$	$2, 2$

$=$

	L	**R**
U	$3, 3$	$1, 1$
D	$1, 1$	$2, 2$

$+$

	L	**R**
U	$-2, -2$	$0, -2$
D	$-2, 0$	$0, 0$

The potential has a unique maximiser (U, L), which is also a Nash equilibrium. However, there is a secondary Nash equilibrium in this game (D, R) that is not a potential maximiser and it is in fact socially or Pareto optimal in this game. ♦

The simple decomposition of Proposition 2.1 also determines the dimension of the linear subspace of exact potential games built on a given pre-game. The next corollary records this without proof.[1]

Corollary 2.1 (Facchini et al. 1997)

Let $\langle N, \mathbf{A}\rangle$ be a finite pre-game. Then the set of exact potential games for forms a strict linear subspace of $\mathbb{G}^N_E(\mathbf{A}) \subset \mathbb{G}^N$ with dimension

$$\dim\left(\mathbb{G}^N_E(\mathbf{A})\right) = \prod_{i\in N} |A_i| + \sum_{i\in N}\prod_{j\neq i} |A_j| - 1 \tag{2.3}$$

♡

2.2 Characterisation of exact potential games

Proposition 2.1 provided a simple yet effective characterisation of an exact potential game. Its decomposition forms the basis of more fundamental tests and characterisations, which this section surveys.

2.2.1 The Monderer-Shapley characterisation for differentiable games

Monderer and Shapley (1996) is the seminal paper on exact potential games and their direct generalisations. It investigates in detail the conditions under which exact potential games with differentiable and smooth payoff functions can be characterised.

Formally, we refer to a game $\Gamma = \langle N, \mathbf{A}, \pi\rangle$ as a *differentiable game* if Γ is a one-dimensional Euclidean game with convex action sets $A_i \subseteq \mathbb{R}$, $i \in N$, and continuously differentiable payoff functions π on the convex continuum $\mathbf{A} \subseteq \mathbb{R}^N$. Furthermore, we refer to a differentiable game Γ as a *smooth game* if the following properties hold:

- All action sets are closed intervals on the real line in the sense that for every $i \in N\colon A_i = [\underline{a}_i, \bar{a}_i] \subset \mathbb{R}$;
- All payoff functions are twice continuously differentiable, meaning that for every $i \in N\colon \pi_i \in \mathcal{C}^2(\mathbf{A})$ are of class $\mathcal{C}^2$.

[1] For a proof we refer to the original paper Facchini et al. (1997) or the clear exposition in Voorneveld (1999, page 11).

From the definition of a derivative we immediately deduce from the definition of an exact potential game that at every action tuple in $\mathbf{A}$, the payoff functions have a derivative that is equal to the derivative of the potential function. This leads to the following characterisation of exact potential games that are differentiable.

Theorem 2.1 (Monderer and Shapley 1996, Lemma 4.4)

Let Γ be a differentiable game and let $\Psi \colon \mathbf{A} \to \mathbb{R}$. Then Ψ is an exact potential for Γ if and only if Ψ is continuously differentiable and

$$\frac{\partial \pi_i}{\partial a_i} = \frac{\partial \Psi}{\partial a_i} \tag{2.4}$$

for all players $i \in N$. ♡

Monderer and Shapley also characterised exact potential games that are smooth. Differentiating the condition stated in Theorem 2.1 once more, we obtain an equality of mixed second derivatives for games whose payoff functions are twice continuously differentiable. This leads to the reformulation of Theorem 2.1 as Theorem 2.2 below, applicable to smooth games. A potential function for such a game can then be constructed by integrating (2.4); Monderer and Shapley used a line integral for this purpose.

Theorem 2.2 (Monderer and Shapley 1996, Theorem 4.5)

Let Γ be a smooth game. The smooth game Γ is an exact potential game if and only if for all players $i, j \in N$:

$$\frac{\partial^2 \pi_i}{\partial a_i \, \partial a_j} = \frac{\partial^2 \pi_j}{\partial a_i \, \partial a_j} \tag{2.5}$$

♡

Theorem 2.2 shows that the interaction between the actions chosen by players is expressed in the mixed second derivative. A common term such as a bi-quadratic expression is supported through an exact potential, but deviations or different weights on these components would violate (2.5). A simple example illustrates this.

Example 2.4 Consider a two-player smooth game with $A_1 = A_2 = [0, 1]$. Denoting actions by the two players as x and y, respectively, payoff functions are given by

$$\pi_1(x, y) = x^2y^2 + 2x - 6y^2 \qquad \text{and} \qquad \pi_2 = x^2y^2 + 6x - y^2.$$

Note that the potential determining interaction component is the common x^2y^2. This has to be symmetric, by (2.5), meaning that one player cannot have a different weight on the common interaction component than the other player. It has quantitatively be identical.

We now compute that

$$\frac{\partial \pi_1}{\partial x} = 2xy^2 + 2 \qquad \text{and} \qquad \frac{\partial \pi_2}{\partial y} = 2x^2y - 2y$$

From this it is easily checked that (2.5) holds:

$$\frac{\partial^2 \pi_1}{\partial x \, \partial y} = \frac{\partial^2 \pi_2}{\partial y \, \partial x} = 4xy$$

Therefore, we conclude that this game has to be an exact potential game. We can use the formula stated in Theorem 2.2 to compute the exact potential function. For every action tuple (x, y), let $\bar{a} = (0, 0)$ and let $\gamma(t) = (tx, ty)$ such that $\gamma(0) = \bar{a} = (0, 0)$ and $\gamma(1) = (x, y)$. Clearly, $\gamma'(t) = (x, y)$. Thus, we derive that

$$\begin{aligned}\Psi(x,y) &= \int_0^1 \left[\frac{\partial \pi_1}{\partial x}(tx,ty)\cdot x + \frac{\partial \pi_2}{\partial y}(tx,ty)\cdot y\right] dt \\ &= \int_0^1 \left[2(tx)(ty)^2+2\right] x\,dt + \int_0^1 \left[2(tx)^2(ty) - 2(ty)\right] y\,dt \\ &= x\int_0^1 \left(2xy^2t^3+2\right)\,dt + y^2\int_0^1 \left(2x^2-2t\right)\,dt \\ &= x\left(\tfrac{1}{2}xy^2+2\right) + y^2\left(\tfrac{1}{2}x^2-1\right) = x^2y^2+2x-y^2\end{aligned}$$

It is easy to check that this is indeed an exact potential for the smooth game Γ as given above. ♦

An interesting application of the characterisation method of Monderer and Shapley (1996) is the class of games describing the contest for a prize, known as *rent seeking* (Buchanan 1980). Rent seeking refers to the economic practice of investing resources in the pursuit of a rent or prize; with only a single winner, the activity is highly inefficient, yet many areas of economic life—from government contracting to school placement—are organised as rent-seeking contests.

Example 2.5 The basic game theoretic formulation of rent seeking as a contest has been set out by Tullock (1980). There are two players $N = \{1,2\}$ who pursue to win a prize $V > 0$. This pursuit is executed through the exertion of a certain level of effort $a_i \in A_i = \mathbb{R}_+$ for both players $i = 1,2$. Effort results in two outcomes. The first is the prime outcome, namely the factor that affects the probability of winning the prize V. This is for each player $i \in N$ described by the same continuously differentiable function $f\colon \mathbb{R}_{++} \to \mathbb{R}_+$ with continuous derivative f'.[2]
Second, the exertion of effort a_i results in costs, described by a continuously differentiable cost function $c_i\colon \mathbb{R}_+ \to \mathbb{R}_+$ with $c_i(0) = 0$ and continuous derivative c_i', for each player $i \in N$.
Payoffs are now determined as the expected payoff from winning the prize, minus the cost related to the exertion of effort in this rent seeking enterprise:

$$\pi_i(a) = \begin{cases} \frac{f(a_i)}{f(a_1)+f(a_2)}\,V - c_i(a_i) & \text{for } a \neq (0,0) \\ \frac{1}{2}\,V & \text{if } a = (0,0)\end{cases} \tag{2.6}$$

Tullock (1980) originally introduced $f(x) = x^\rho$ for $x > 0$, where $\rho > 0$ is an efficiency parameter. The Nash equilibrium for the standard cost function $c(x) = x$ can be determined as $a^* = \frac{\rho V}{4}$. We refer to the detailed analysis in Baye, Kovenock, and Vries (1994) for further conclusions about the existence of mixed strategy Nash equilibria for the standard game.
We use Theorem 2.2 to check whether this rent seeking game has an exact potential. Note that we have to check the second derivative of the two payoff functions and check whether they are equal. First, we investigate the derivative of π_i for $i = 1,2$:

[2] Here we use the notation of $\mathbb{R}_{++} = (0,\infty)$ as the open positive orthant of the real line.

$$\frac{\partial \pi_1}{\partial a_1} = \frac{f'(a_1)(f(a_1)+f(a_2)) - f'(a_1)f(a_1)}{(f(a_1)+f(a_2))^2} V - c_1'(a_1)$$

$$= \frac{f'(a_1)f(a_2)}{(f(a_1)+f(a_2))^2} V - c_1'(a_1) \quad \text{and, similarly,}$$

$$\frac{\partial \pi_2}{\partial a_2} = \frac{f(a_1)f'(a_2)}{(f(a_1)+f(a_2))^2} V - c_2'(a_2)$$

This allows us to determine the required second derivatives for this rent seeking game:

$$\frac{\partial^2 \pi_1}{\partial a_1 \partial a_2} = \frac{f'(a_1)f'(a_2)(f(a_1)+f(a_2))^2 - 2f'(a_1)f(a_2)f'(a_2)(f(a_1)+f(a_2))}{(f(a_1)+f(a_2))^4} V$$

$$= \frac{f'(a_1)f'(a_2)(f(a_1)-f(a_2))}{(f(a_1)+f(a_2))^3} V$$

$$\frac{\partial^2 \pi_2}{\partial a_2 \partial a_1} = \frac{f'(a_1)f'(a_2)(f(a_2)-f(a_1))}{(f(a_1)+f(a_2))^3} V = -\frac{\partial^2 \pi_1}{\partial a_1 \partial a_2}$$

We therefore conclude that rent-seeking games are *not* exact potential games. The Monderer-Shapley criterion of Theorem 2.2 has made this quite straightforward to establish; absent the criterion it would be considerably harder. ♦

Generalisation of the Monderer-Shapley characterisations

Differentiable games, as defined above, are one-dimensional: all action sets are convex subsets of the real line. Can the Monderer-Shapley characterisations (Theorems 2.1 and 2.2) be extended to multi-dimensional Euclidean spaces? This is the subject of Arefizadeh, Nedić, and Dasarathy (2024), which informs the discussion in this subsection.

For these results we turn to general Euclidean games with convex action sets. Recall that in a Euclidean game each player $i \in N$ has an action set $A_i \subseteq \mathbb{R}^{k_i}$, where $k_i \in \mathbb{N}$ is the dimension of the actions available. Now denote $\bar{k} = \sum_{i \in N} k_i$, then $\mathbf{A} = \prod_{i \in N} A_i \subseteq \mathbb{R}^{\bar{k}}$. For any continuously differentiable function $f \colon \mathbb{R}^{\bar{k}} \to \mathbb{R}$ we let

$$D_i f(a) = \left(\frac{\partial f}{\partial a_{i,1}}, \frac{\partial f}{\partial a_{i,2}}, \ldots, \frac{\partial f}{\partial a_{i,k_i}} \right)$$

where we write $a_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,k_i}) \in \mathbb{R}^{k_i}$. The operator D_i is the partial differential with regard to a_i.

The next result generalises Theorem 2.1 to Euclidean games with arbitrary action sets in multi-dimensional Euclidean spaces.

Theorem 2.3 (Arefizadeh, Nedić, and Dasarathy 2024, Theorem 2(a))

Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be a Euclidean game with $A_i \subseteq \mathbb{R}^{k_i}$ for each $i \in N$ and $\mathbf{A} \subseteq \mathbb{R}^{\bar{k}}$ with $\bar{k} = \sum_{i \in N} k_i$. Suppose that for every $i \in N$: π_i is a continuously differentiable function on an open set containing $\mathbf{A}$. Then Γ is an exact potential game if and only if there exists some function $\Psi \colon \mathbb{R}^{\bar{k}} \to \mathbb{R}$ such that for every player $i \in N$ and every $a \in \mathbf{A}$:

$$D_i \pi_i(a) = D_i \Psi(a) \tag{2.7}$$

♡

Theorem 2.2 can also be generalised in a similar fashion. For that we need to define the partial Hessian of a twice continuously differentiable function $f\colon \mathbb{R}^{\bar{k}} \to \mathbb{R}$. We denote for all players $i, j \in N$:

$$D^2_{i,j}\, f(a) = \begin{bmatrix} \frac{\partial^2 f}{\partial a_{i,1}\,\partial a_{j,1}} & \frac{\partial^2 f}{\partial a_{i,1}\,\partial a_{j,2}} & \cdots & \frac{\partial^2 f}{\partial a_{i,1}\,\partial a_{j,k_j}} \\ \frac{\partial^2 f}{\partial a_{i,2}\,\partial a_{j,1}} & \frac{\partial^2 f}{\partial a_{i,2}\,\partial a_{j,2}} & \cdots & \frac{\partial^2 f}{\partial a_{i,2}\,\partial a_{j,k_j}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial a_{i,k_i}\,\partial a_{j,1}} & \frac{\partial^2 f}{\partial a_{i,k_i}\,\partial a_{j,2}} & \cdots & \frac{\partial^2 f}{\partial a_{i,k_i}\,\partial a_{j,k_j}} \end{bmatrix}$$

as the partial Hessian $k_i \times k_j$ matrix of second order derivatives of f. This leads to the following generalisation of Theorem 2.2.

Theorem 2.4 (Arefizadeh, Nedić, and Dasarathy 2024, Theorem 2(b))

Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be a Euclidean game with convex action sets $A_i \subseteq \mathbb{R}^{k_i}$ for every player $i \in N$ and $\mathbf{A} \subseteq \mathbb{R}^{\bar{k}}$ with $\bar{k} = \sum_{i \in N} k_i$. Suppose that for every $i \in N$: π_i is a twice continuously differentiable function on an open set containing $\mathbf{A}$.
Then Γ is an exact potential game if and only if for all players $i, j \in N$ and all $a \in \mathbf{A}$:

$$D^2_{i,j}\, \pi_i(a) = D^2_{j,i}\, \pi_j(a) \tag{2.8}$$

♡

We illustrate Theorem 2.4 with a simple mathematical example as well as a modified Cournot duopoly.

Example 2.6 Consider a two-player Euclidean game $\Gamma = \langle \{1, 2\}, \mathbf{A}, \pi \rangle$ that is characterised by $A_1 = [0, 1]^2$, $A_2 = [0, 1]$ and payoff functions given by

$$\pi_1(x_1, x_2; y) = y(x_1 + x_2)^2 - x_1 x_2 + x_2 y^2$$
$$\pi_2(x_1, x_2; y) = y(x_1 + x_2)^2 + x_2 y^2 + y$$

Note here that player 1's action set is two-dimensional and that player 2's action set is one-dimensional. This discrepancy is expressed by making player 1's action two dimensional, represented by (x_1, x_2), while player 2's strategy is just the single number y.

We argue that Γ is an exact potential game. We use Theorem 2.4 to verify this. Indeed note that we have the following derivatives for the two payoff functions:

$$D_x\, \pi_1(x_1, x_2; y) = \left[\, 2y(x_1 + x_2) - x_2 \,,\; 2y(x_1 + x_2) - x_1 + y^2 \,\right]$$
$$D_y\, \pi_2(x_1, x_2; y) = (x_1 + x_2)^2 + 2x_2 y + 1$$

This leads to the conclusion that

$$D^2_{yx}\, \pi_1(x_1, x_2; y) = \begin{bmatrix} 2(x_1 + x_2) \\ 2(x_1 + x_2) + 2y \end{bmatrix} \quad \text{and} \quad D^2_{xy}\, \pi_2(x_1, x_2; y) = \begin{bmatrix} 2(x_1 + x_2) \\ 2(x_1 + x_2) + 2y \end{bmatrix}$$

Clearly these second derivatives are identical. Using Theorem 2.4 it follows immediately that Γ has an exact potential. After some contemplation, it is indeed easy to see that $\Psi\colon [0, 1]^3 \to \mathbb{R}$ given by

$$\Psi(x_1, x_2; y) = y(x_1 + x_2)^2 - x_1 x_2 + x_2 y^2 + y$$

is an exact potential function for this game. ♦

2.2.2 Ui's characterisation of finite exact potential games

Ui (2000) is a seminal contribution to the theory of exact potential games. Beyond the characterisation discussed here, Ui introduces a fundamentally different perspective on exact potential games by establishing a profound relationship with *cooperative games with transferable utility* (TU-games).[3] Every exact potential game corresponds to a subclass of TU-games whose Shapley values (Shapley 1953) coincide with the payoff function of the exact potential game. Conversely, the space of all TU-games maps onto the class of exact potential games through the Shapley value, underlining the special role of the Shapley value in bridging cooperative and non-cooperative game theory.[4]

Furthermore, Ui's main theorem provides a foundation for characterising exact potential games through so-called *interaction potentials*. For this we recall that the restricted action space $\mathbf{A}_S$ for a certain coalition of players $S \subseteq N$ is defined as $\mathbf{A}_S = \prod_{j \in S} A_j$. The next theorem states this in precise terms:

Theorem 2.5 (Ui 2000, Theorem 3)

A game $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ is an exact potential game if and only if for every coalition of players $S \subseteq N$ there exists an interaction potential function $\Psi_S \colon \mathbf{A}_S \to \mathbb{R}$ such that for every $a \in \mathbf{A}$ and every player $i \in N$:

$$\pi_i(a) = \sum_{S \subseteq N \colon i \in S} \Psi_S(a_S) \tag{2.9}$$

An exact potential function for the game Γ is for every action tuple $a \in \mathbf{A}$ given by

$$\Psi(a) = \sum_{S \subseteq N} \Psi_S(a_S) \tag{2.10}$$

♡

Ui's characterisation of exact potential games is rather useful to construct potential functions for various games by introspection of the payoff functions. I explore two examples of well-known exact potential games, using Theorem 2.5.

Example 2.7 Cournot Oligopoly

Consider a Cournot oligopoly, which generalises the linear duopoly considered in the previous chapter. For this let $N = \{1, \ldots, n\}$ be a finite set of firms. The cost of producing q_i is given by $c_i(q_i) > 0$, where c_i is any monotone increasing function on $\mathbb{R}_+$. Each firm $i \in N$ strategically selects an output level $q_i \geqslant 0$ with a total market supply of $Q = \sum_{i \in N} q_i$.

Assuming that the inverse market demand function is linear and given by $P(Q) = \alpha - \beta Q$, we arrive at a profit function for each firm $i \in N$ as

$$\pi_i(q_1, \ldots, q_n) = P(Q)q_i - c_i(q_i) = \left(\alpha - \beta \sum_{j \in N} q_j \right) q_i - c_i(q_i)$$

[3] For an introduction to cooperative game theory and its main methods and solution concepts I refer to Gilles (2010).

[4] I do not explore the details of Ui's analysis here, since this would take us into cooperative game theory, which is not the focus of the present text. See Ui (2000) for an exposition.

From this, with Ui's characterisation in mind, we can now decompose each firm's profit function into interaction potential elements:

$$\pi_i(q_1,\ldots,q_n) = \sum_{j\neq i}(-\beta q_j q_i) + \big(\alpha q_i - \beta q_i^2 - c_i(q_i)\big) = \sum_{j\neq i}\Psi_{ji}(q_j,q_i) + \Psi_i(q_i)$$

where $\Psi_{ij}(q_i,q_j) = -\beta q_i q_j$ and $\Psi_i(q_i) = \alpha q_i - \beta q_i^2 - c_i(q_i)$. Clearly, the Cournot oligopoly fulfils the conditions of Theorem 2.5 and, therefore, is an exact potential game. Furthermore, its potential function is given by

$$\Psi(q_1,\ldots,q_n) = \sum_{i,j\in N\colon i\neq j}\Psi_{ij}(q_i,q_j) + \sum_{i\in N}\Psi_i(q_i) = \alpha Q - \beta\sum_{i,j\in N} q_i q_j - \sum_{i\in N} c_i(q_i)$$

It can be concluded that Ui's characterisation is a very useful tool if one can easily decompose individual payoff functions in clearly delineated interaction potential elements. The exact potential then follows quite straightforwardly. ♦

The next example considers the public-good provision game. Such games may be highly non-continuous, so the Monderer-Shapley characterisation is unavailable—whereas Ui's characterisation applies without difficulty.

Example 2.8 Collective good provision game

Consider an association of n members that provide a collective good to the members through voluntary contributions. Let $x_i \geqslant 0$ be the contribution of member i and let $X = \sum_{i=1}^{n} x_i$ be the total contribution collected. Now we assume that the association provides the collective good to the members if and only if $X \geqslant \underline{X} \geqslant 0$, where $\underline{X}$ expresses a minimum level of contributions before provision is agreed.

Each member has individual preferences over the collective good provided, which are expressed in monetary terms, i.e., as a willingness to pay. Let the willingness to pay be a monotone function $w_i\colon \mathbb{R}_+ \to \mathbb{R}_+$ with $w_i(0) = 0$.

Now, each member's payoff from collective good provision can be described by

$$\pi_i(x_1,\ldots,x_n) = \begin{cases} w_i(X) - x_i & \text{if } X \geqslant \underline{X} \\ -x_i & \text{otherwise} \end{cases}$$

It is easy to see that this payoff function is decomposable as $\pi_i(x_1,\ldots,x_n) = \Psi_N(x_1,\ldots,x_n) + \Psi_i(x_i)$ *if and only if* $w_i = w_j$ for all $i,j \in N$. In that case, we can define $\Psi_i(x_i) = -x_i$ and

$$\Psi_N(x_1,\ldots,x_n) = \begin{cases} w(X) & \text{if } X \geqslant \underline{X} \\ 0 & \text{otherwise} \end{cases}$$

where $w = w_1 = \cdots = w_n$. In that case, an exact potential is given by

$$\Psi(x_1,\ldots,x_n) = \Psi_N(x_1,\ldots,x_n) + \sum_{i\in N}\Psi_i(x_i) = \begin{cases} w(X) - X & \text{if } X \geqslant \underline{X} \\ -X & \text{otherwise} \end{cases}$$

This example illustrates the usefulness of Ui's characterisation for games with non-continuous payoffs. ♦

2.2.3 Sandholm's characterisation

Sandholm (2010) introduced a characterisation of exact potential games that simultaneously provides a test whether a given game has an exact potential. The method was extended by Hwang and Rey-Bellet (2020); the resulting Sandholm-Hwang-Rey-Bellet approach—or, for short, Sandholm's characterisation—is the subject of this section. We follow the exposition of Hwang and Rey-Bellet (2020).

Considering a Euclidean pre-game $\langle N, \mathbf{A} \rangle$ with compact action sets $A_i \subset \mathbb{R}^{k_i}$ for all players $i \in N$, we might assume that for every player $i \in N$ the action set can be endowed with a standard measure. Given that A_i is compact, it follows that λ_i is a finite measure on the standard Lebesgue measure space restricted to A_i.

Now, if A_i is finite, this measure would be the counting measure, putting an equal weight on each action. Similarly, if the compact action set $A_i \subseteq \mathbb{R}^{k_i}$ is infinite, we can impose the standard Lebesgue measure on A_i. We denote this standard measure by λ_i for player $i \in N$.[5]

For every player $i \in N$, we introduce the *expectation operator* $\mathcal{E}_i \colon \mathbb{R}^{\mathbf{A}} \times \mathbf{A} \to \mathbb{R}$ as the function that assigns for every action profile $a \in \mathbf{A}$ the subjective (partial) expected value of a measurable function or stochastic variable $f \colon \mathbf{A} \to \mathbb{R}$ in a over the i-th dimension defined by

$$\mathcal{E}_i(f, a) = \frac{1}{\lambda(A_i)} \int_{A_i} f(\cdot; a_{-i})\, d\lambda_i \in \mathbb{R} \tag{2.11}$$

Note that this expectation operator is a linear functional or map on the Euclidean space $\mathbb{R}^{\mathbf{A}} \times \mathbf{A}$.

The following characterisation of Euclidean exact potential games follows from Proposition 2(i) as formulated and proved by Hwang and Rey-Bellet (2020). For this characterisation the standard identity operator $\mathcal{I}$ on $\mathbb{R}^{\mathbf{A}} \times \mathbf{A}$ is defined by $\mathcal{I}(f, a) = f(a)$.

Theorem 2.6 (Hwang and Rey-Bellet 2020, Proposition 2)

The Euclidean game $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ is an exact potential game if and only if for all players $i, j \in N$:

$$(\mathcal{I} - \mathcal{E}_i)(\mathcal{I} - \mathcal{E}_j)(\pi_i - \pi_j) = 0 \tag{2.12}$$ ♡

If $\mathbf{A}$ is finite and λ is the counting measure on $\mathbf{A}$, we can rewrite Theorem 2.6 to arrive at the condition of Sandholm (2010). We can express this as follows for a two-player game, using the formulation of Hwang and Rey-Bellet (2020).

Corollary 2.2 (Sandholm's characterisation for 2-player games)

Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be a finite game with $N = \{1, 2\}$, $|A_1| = k_1$, $|A_2| = k_2$, and payoff representation $\pi_1 = \{\alpha_{hl}\}_{h \in A_1, l \in A_2}$ and $\pi_2 = \{\beta_{hl}\}_{h \in A_1, l \in A_2}$.
Γ is an exact potential game if and only if for all $(h, l) \in A_1 \times A_2$:

$$\alpha_{hl} - \frac{1}{|A_1|} \sum_{m \in A_1} \alpha_{ml} - \frac{1}{|A_2|} \sum_{p \in A_2} \alpha_{hp} + \frac{1}{|A_1|} \frac{1}{|A_2|} \sum (m, p) \in \mathbf{A} \alpha_{mp}$$

$$= \beta_{hl} - \frac{1}{|A_1|} \sum_{m \in A_1} \beta_{ml} - \frac{1}{|A_2|} \sum_{p \in A_2} \beta_{hp} + \frac{1}{|A_1|} \frac{1}{|A_2|} \sum (m, p) \in \mathbf{A} \beta_{mp}$$ ♡

The Corollary follows immediately from rewriting (2.12) for two players as

[5] The different nature of an action set, even in a Euclidean game, necessitates the use of individualised measures on action spaces.

$$(\mathcal{I} - \mathcal{E}_1 - \mathcal{E}_2 + \mathcal{E}_1\mathcal{E}_2)\pi_1 = (\mathcal{I} - \mathcal{E}_1 - \mathcal{E}_2 + \mathcal{E}_1\mathcal{E}_2)\pi_2$$

The expression above makes it easy to test whether a finite two-player game is an exact potential game.

2.2.4 Application: Networked Cournot oligopolies

The characterisations developed above help us identify which interactive decision situations are potential games. In this section we apply them to an economically interesting class: oligopolies operating in competitive networks.

One of the most fruitful applications of exact potential games lies in industrial organisation, centred on the basic Cournot oligopoly model of Chapter 1 (Example 2.7), which serves as the quintessential exact potential game in the literature.

Following Ederer and Pellegrino (2025) and Pellegrino (2025), we extend the basic model to multi-interactional Cournot oligopolies combining market interactions with financial relationships. Since both interactions are naturally represented as networks, we refer to these situations as *networked Cournot oligopolies*. Cao, Li, and Zhu (2025) develop the essential form of this class and show that it is an exact potential game, so that the full machinery of potential game theory can be brought to bear on them.

A network Cournot oligopolistic model We develop the model as presented and analysed by Cao, Li, and Zhu (2025). We consider n firms that operate in a homogeneous good market and compete in this market through strategic selection of output quantities à la Cournot. For each firm $i \in N = \{1, \ldots, n\}$ we denote its output level by $q_i \in A_i = [0, \infty)$, resulting in a strategy tuple $q = (q_1, \ldots, q_n) \in \mathbf{A} = \mathbb{R}^n_+$.

The total output that is brought to market is now derived as $Q = \sum_{i \in N} q_i$. We use the standard linear market demand model, represented by a inverse demand function $p = \alpha - \beta Q$, where $\alpha, \beta > 0$.[6] We make the simplifying assumption that production does not result in costs. Hence, the cost functions in this oligopoly are set to zero. This, in turn, implies that the net earnings of firm $i \in N$ are given by

$$E_i(q) = (\alpha - \beta Q)q_i \tag{2.13}$$

We next introduce the second relationship between firms that is represented as an *ownership network* $\Delta = [\delta_{ij}]_{i,j \in N}$—also referred to as the "cross-holding network" by Cao, Li, and Zhu (2025). Clearly Δ is an $n \times n$-matrix in which $\delta_{ij} \in [0, 1]$ denotes the ownership of firm j in firm i. Excluding outside ownership, this necessarily implies that the full ownership of each firm is divided among all n firms: For every $i \in N$: $\sum_{j \in N} \delta_{ij} = 1$. Finally, we make the hypothesis that each firm retains a controlling ownership of their own operations, implying that $\delta_{ii} \geqslant \frac{1}{2}$ for all $i \in N$.

From the ownership network Δ we can now also derive an $n \times n$ *influence network* $W = [w_{ij}]_{i,j \in N}$ defined as $w_{ij} = \delta_{ii} + \delta_{ji}$ as the total fraction of shares held by firm $i \in N$ of itself as well as firm j. Note that $w_{ii} = 2\delta_{ii}$ and that in general $w_{ij} \neq w_{ji}$.

These modelling assumptions now result in the game-theoretic payoff structure. In particular, we implement the simple rule that ownership directly results in dividend payments. Thus, each firm pays out its full profits as dividends that are distributed according to the ownership of that particular firm. Therefore, firm j collects a fraction δ_{ij} of the earnings $E_i(q)$ of firm i. This is summarised in the resulting game-theoretic payoffs of firm

[6] We conclude from this that the effective range for the total market output is $0 \leqslant Q \leqslant \frac{\alpha}{\beta}$.

i as

$$\pi_i(q) = \sum_{j\in N} \delta_{ji}\, E_j(q) = (\alpha - \beta Q) \sum_{j\in N} \delta_{ji}\, q_j \tag{2.14}$$

We can rewrite the payoff or profit function in terms of the interaction matrix W instead of the ownership matrix Δ. This can be stated as follows.

Lemma 2.1

For every $i \in N$, π_i is quadratic in q, linear in W and can be written as

$$\pi_i(q) = (\alpha - \beta Q) \left(\sum_{j\in N} w_{ij}\, q_j - \tfrac{w_{ii}}{2} Q \right) \tag{2.15}$$

♡

Proof Consider that for any $i, j \in N$ we have that $w_{ij} = \delta_{ii} + \delta_{ji}$, implying that $\delta_{ii} = \frac{w_{ii}}{2}$. Therefore,

$$\begin{aligned}
\pi_i(q) &= (\alpha - \beta Q) \sum_{j\in N} \delta_{ji}\, q_j \\
&= (\alpha - \beta Q) \left[\sum_{j\in N} (\delta_{ii} + \delta_{ji})\; q_j - \sum_{j\in N} \delta_{ii}\, q_j \right] \\
&= (\alpha - \beta Q) \left[\sum_{j\in N} w_{ij}\, q_j - \tfrac{w_{ii}}{2}\, Q \right]
\end{aligned}$$

This shows the assertion of the lemma. ¶

The next result ties the model together as an exact potential game. The proof uses the Monderer-Shapley characterisation (Theorem 2.1) and checks the second derivative of the payoff function.

Lemma 2.2

The network Cournot oligopoly model is an exact potential game if and only if for all firms $i, j \in N$: $w_{ij} = w_{ji}$. In that case, a corresponding potential function is given by

$$P(q) = \alpha \sum_{i\in N} \frac{w_{ii}}{2} q_i - \beta \sum_{i\in N} \sum_{j\in N} w_{ij} q_i q_j \tag{2.16}$$

♡

Proof Consider $i, j \in N$. Using Theorem 2.1, we check that $\frac{\partial^2 \pi_i}{\partial q_i \partial q_j} = \frac{\partial^2 \pi_j}{\partial q_j \partial q_i}$. For that purpose, we compute the regular derivative:

$$\begin{aligned}
\frac{\partial \pi_i}{\partial q_i} &= -\beta \left(\sum_{h\in N} w_{ih} q_h - \tfrac{w_{ii}}{2} \sum_{h\in N} q_h \right) + \left(\alpha - \beta \sum_{h\in N} q_h \right) \left(w_{ii} - \frac{w_{ii}}{2} \right) \\
&= \alpha \frac{w_{ii}}{2} - \beta \sum_{h\in N} w_{ih} q_h
\end{aligned}$$

and, subsequently, we compute

$$\frac{\partial^2 \pi_i}{\partial q_i \partial q_j} = \beta \left(\frac{w_{ii}}{2} - w_{ij} \right) + \frac{w_{ii}}{2}(-\beta)$$
$$= -\beta w_{ij}$$

This implies that, indeed, the networked Cournot oligopoly model is an exact potential game if and only if $w_{ij} = w_{ji}$ for all firms $i, j \in N$.

To check that the proposed function P is indeed a potential function of the networked Cournot oligopoly model we only have to check that for all $i \in N$: $\frac{\partial P}{\partial q_i} = \frac{\partial \pi_i}{\partial q_i}$. This is quite easy to verify and we leave this task to the interested reader. ¶

2.3 Determining exact potential functions for games

The previous section considered various methods for verifying whether a game has an exact potential. Several of these characterisations are impractical as tests, and most are poorly suited to actually constructing the potential function for a given game.

The main exception is the method of Monderer and Shapley (1996) for smooth games. When payoff functions are smooth, the fundamental theorem of line integrals yields an integral formula for the exact potential, as follows.

Theorem 2.7 (Monderer and Shapley (1996, Theorem 4.5)) *Let Γ be a smooth game such that characterisation (2.5) stated in Theorem 2.2 holds, then for every fixed action profile $a^0 \in \mathbf{A}$, we can construct an exact potential function for Γ by*

$$\Psi(a) = \int_0^1 \sum_{i \in N} \left[\frac{\partial \pi_i}{\partial a_i}(\gamma(t)) \, |\gamma_i'(t)| \right] dt \tag{2.17}$$

where $\gamma \colon [0,1] \to \mathbf{A}$ is a piecewise continuously differentiable path with $\gamma(0) = a^0$ and $\gamma(1) = a$. ♡

The formula in Theorem 2.7 remains rather opaque and is not always easy to apply, even for smooth games satisfying Theorems 2.2 or 2.4. The next example illustrates the computation on a simple smooth exact potential game.

Example 2.9 Consider a simple two-player game with $A_1 = A_2 = \mathbb{R}_+ = [0, \infty)$ and payoff functions given by

$$\pi_1(x, y) = x^2y^2 - 3xy + x^2 + y^2$$
$$\pi_2(x, y) = x^2y^2 - 3xy$$

We first establish that this game indeed is an exact potential game using Theorem 2.2. For that purpose we compute the required derivatives:

$$\frac{\partial \pi_1}{\partial x} = 2xy^2 - 3y + 2x \qquad \frac{\partial^2 \pi_1}{\partial y \partial x} = 4xy - 3$$
$$\frac{\partial \pi_2}{\partial y} = 2x^2y - 3x \qquad \frac{\partial^2 \pi_2}{\partial x \partial y} = 4xy - 3$$

This implies that this game indeed has an exact potential according to the Monderer-Shapley criterion. Next we compute the exact potential function by using the formula stated in Theorem 2.7. It is usually the easiest to take the linear path from standard anchor point $(0,0)$ and (x,y). Hence, we select $\gamma(t) = (tx, ty)$ for $t \in [0,1]$. Therefore, $\gamma'(t) = (x,y)$. Now, we know that

$$\begin{aligned}
\Psi(x,y) &= \int_0^1 \left(\frac{\partial \pi_1}{\partial x} + \frac{\partial \pi_2}{\partial y} \right)(tx,ty) \cdot (x,y)\, dt \\
&= x \int_0^1 \frac{\partial \pi_1}{\partial x}(tx,ty)\, dt + y \int_0^1 \frac{\partial \pi_2}{\partial y}(tx,ty)\, dt \\
&= x \int_0^1 \left[2xy^2t^3 - 3yt + 2xt \right] dt + y \int_0^1 \left[2x^2yt^3 - 3xt \right] dt \\
&= x \left[\tfrac{1}{2}xy^2t^4 - \tfrac{3}{2}yt^2 + xt^2 \right]_0^1 + y \left[\tfrac{1}{2}x^2yt^4 - \tfrac{3}{2}xt^2 \right]_0^1 \\
&= x \left(\tfrac{1}{2}xy^2 - \tfrac{3}{2}y + x \right) + y \left(\tfrac{1}{2}xy^2 - \tfrac{3}{2}y \right) \\
&= x^2y^2 - 3xy + x^2
\end{aligned}$$

It is indeed easy to verify that the constructed function Ψ is an exact potential function for this game. ♦

An alternative approach to the issue of determining the exact potential function for a game has been developed by Arefizadeh, Nedić, and Dasarathy (2024). Throughout the discussion in this section we let again $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be a Euclidean game with $A_i \subset \mathbb{R}^{k_i}$ and $\bar{k} = \sum_{i \in N} k_i$.

Definition 2.2

*Let $a, b \in \mathbf{A}$ be two action tuples in the game Γ. We define the **path potential** of the Euclidean game Γ from a to b to be given by the difference function $\varphi \colon \mathbf{A}^2 \to \mathbb{R}$ as the difference value of the payoffs on the path from a to b defined by*

$$\varphi(a,b) = \sum_{i=1}^{n} \left[\pi_i(b_1, \dots, b_i, a_{i+1}, \dots, a_n) - \pi_i(b_1, \dots, b_{i-1}, a_i, \dots, a_n) \right] \tag{2.18}$$

The path potential φ of Γ is also denoted as φ_Γ. ♣

The path potential of a game Γ refers to the total of the payoff differences along a fixed path from a given action tuple $a \in \mathbf{A}$ and an action tuple $b \in \mathbf{A}$. In this particular path, each player $i \in N$ is sequentially moving from a_i to b_i in the order of the player set N. The path potential is non-trivial for each game that is not a dummy game.

For the next proposition, recall that a player $i \in N$ in the game Γ is *non-effective* if for every action tuple $a \in \mathbf{A}\colon \pi_i(a) = f_i(a_{-i})$, i.e., the payoff of player i is independent of i's own strategic choice. The game Γ is a *dummy game* if every player $i \in N$ is non-effective, and *effective* otherwise—i.e., if at least one player's payoff depends on her own action.

Proposition 2.2

Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be an effective Euclidean game such that $0 \in \mathbf{A} \subseteq \mathbb{R}^{\bar{k}}$. Then the path potential φ of Γ is non-trivial in the sense that there is at least one action tuple $a \in \mathbf{A}$ such that $\varphi(0,a) \neq 0$. ♠

The path potential is straightforward to compute. The next result links it to the exact potential of a Euclidean game under mild conditions and thus provides a workable route to the exact potential function. It slightly extends Arefizadeh, Nedić, and Dasarathy (2024, Theorem 6) by removing the requirement that all action sets are symmetric.

Theorem 2.8

Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be an effective Euclidean game such that $0 \in \operatorname{int} \mathbf{A} \subseteq \mathbb{R}^{\bar{k}}$.
Then Γ is an exact potential game if and only if for Γ's path potential φ it holds that for all action tuples $a, b \in \mathbf{A}$:

$$\varphi(a,b) = \varphi(a,0) + \varphi(0,b) \qquad \textit{as well as} \qquad \varphi(a,0) = -\varphi(0,a). \tag{2.19}$$

Furthermore, an exact potential Ψ of Γ is given by $\Psi(a) = \varphi(0,a) = -\varphi(a,0)$. ♡

We explore next the effectiveness of the approach founded on the path potential to identify an exact potential function for some given game. The next example takes the game considered in Example 2.9 and applies the path potential approach to it.

Example 2.10 Consider the two-player game from Example 2.9 with $A_1 = A_2 = \mathbb{R}_+$ and

$$\begin{aligned}\pi_1(x,y) &= x^2y^2 - 3xy + x^2 + y^2\\ \pi_2(x,y) &= x^2y^2 - 3xy\end{aligned}$$

We can now establish that this game's path potential is given by

$$\begin{aligned}\varphi((x,y);(v,w)) &= \pi_1(v,y) - \pi_1(x,y) + \pi_2(v,w) - \pi_2(v,y)\\ &= (\,v^2y^2 - 3vy + v^2 + y^2\,) - (\,x^2y^2 - 3xy + x^2 + y^2\,) +\\ &\qquad + (\,v^2w^2 - 3vw\,) - (\,v^2y^2 - 3vy\,)\\ &= v^2w^2 + v^2 - 3vw - x^2y^2 - x^2 + 3xy\\ &= v^2(1+w^2) - 3vw - x^2(1+y^2) + 3xy\end{aligned}$$

This implies that

$$\begin{aligned}\varphi((x,y);0) &= -x^2(1+y^2) + 3xy\\ \varphi(0;(v,w)) &= v^2(1+w^2) - 3vw\end{aligned}$$

It is easy to see that $\varphi((x,y);(v,w)) = \varphi((x,y);0) + \varphi(0;(v,w))$ and that $\Psi(x,y) = -\varphi((x,y);0) = x^2(1+y^2) - 3xy$. It is clear that this approach is easier to work through than the Monderer-Shapley path integral formula. ♦

2.4 The potentialness of games

The preceding sections developed several characterisations of exact potential games, each of which provides a sharp test: a game either possesses an exact potential or it does not. In practice, however, this dichotomy is rather limiting. Many games of considerable importance in economic applications—such as auctions and contests—are not exact potential games, yet experimental evidence shows that standard learning dynamics frequently converge to Nash equilibrium in these games all the same. This raises a natural question: can we measure how "close" a given game is to being a potential game, and does such a measure help predict whether convergence will occur?

I now turn to a framework that addresses this question for finite normal form games. The approach rests on a combinatorial decomposition of games seminally introduced by Candogan, Menache, et al. (2011), who showed that any finite game can be resolved into a *potential* component, a *harmonic* component, and a *non-strategic* component. The potential component captures the part of the game's strategic structure that behaves like a potential game; the harmonic component captures the part that is, in a precise sense, maximally opposed to potential behaviour—exhibiting cyclical strategic incentives. The non-strategic component affects neither the game's equilibrium structure nor the players' payoff differentials, and is therefore strategically irrelevant.

Building on this decomposition, Candogan, Ozdaglar, and Parrilo (2013a,b) studied the long-run behaviour of dynamics in near-potential games, and Bichler et al. (2025) introduced a scalar measure of *potentialness* that quantifies the relative weight of the potential component. I present this measure here and establish its basic properties. The development requires some combinatorial and algebraic tools, which are collected in the next mathematical notes.

Mathematical notes *The decomposition of Candogan, Menache, et al. (2011) is grounded in combinatorial Hodge theory, a branch of algebraic topology concerned with decomposing functions on graphs and simplicial complexes into gradient, harmonic, and divergence-free parts. I restrict myself here to the minimal apparatus needed for finite games; for a comprehensive treatment of the underlying mathematics, I refer to Munkres (1984) and Jiang et al. (2011).*

The response graph of a finite game *Consider a finite normal form game* $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ *with* $N = \{1, \dots, n\}$ *players and finite action sets* A_i *for each player* $i \in N$*. Recall that* $\mathbf{A} = \prod_{i \in N} A_i$ *denotes the set of all action profiles, and write* $|\mathbf{A}|$ *for the total number of action profiles.*
A pair of action profiles (a, a') *with* $a, a' \in \mathbf{A}$ *is called a* unilateral deviation *if* a *and* a' *differ in the action of exactly one player: there exists a unique* $i \in N$ *such that* $a_i \neq a'_i$ *and* $a_j = a'_j$ *for all* $j \neq i$*. Denote the set of all unilateral deviations by* $\mathcal{E}$*.*
The response graph *of the game is the graph with one node for each action profile* $a \in \mathbf{A}$ *and one edge for each unilateral deviation* $(a, a') \in \mathcal{E}$*.*[7] *Edges are oriented: the deviation* (a, a') *is directed from* a *to* a'*, and reversing the orientation negates any associated quantity.*

Chain groups and flows *On the response graph, one defines two vector spaces that serve as the algebraic foundation for the decomposition. The* vertex space $C_0 \cong \mathbb{R}^{|\mathbf{A}|}$ *consists of all assignments of a real number to each node (action profile) of the response graph. The* edge space $C_1 \cong \mathbb{R}^{|\mathcal{E}|}$ *consists of all assignments of a real number to each oriented edge (unilateral deviation); an element of* C_1 *is called a* flow *on the response graph.*[8] *Both* C_0 *and* C_1 *are endowed with the standard Euclidean inner product, denoted* $\langle \cdot, \cdot \rangle_0$ *and* $\langle \cdot, \cdot \rangle_1$*, respectively. The induced norms are written as* $\| \cdot \|_0$ *and* $\| \cdot \|_1$*, or simply* $\| \cdot \|$ *when the context is clear.*

[7] The response graph and its properties are studied in detail in Biggar and Shames (2023).

[8] In the language of algebraic topology, C_0 and C_1 are the chain groups of the simplicial complex formed by the response graph; see Munkres (1984) for further details.

The gradient and divergence maps *Two linear maps between C_0 and C_1 play a central role.*
The gradient map $d_0 \colon C_0 \to C_1$ *takes a vertex assignment $\Psi \in C_0$ and produces the flow $d_0\Psi \in C_1$ defined by*

$$(d_0\Psi)(a, a') = \Psi(a') - \Psi(a)$$

for every oriented edge $(a, a') \in \mathcal{E}$. Thus, d_0 computes the difference of a vertex function along each edge, exactly as a discrete gradient does.[9]
The divergence map $\partial_1 \colon C_1 \to C_0$ *is the adjoint of the gradient map with respect to the inner products on C_0 and C_1, meaning that*

$$\langle X, d_0\Psi\rangle_1 = \langle \partial_1 X, \Psi\rangle_0$$

for all flows $X \in C_1$ and all vertex assignments $\Psi \in C_0$. Since both spaces are finite-dimensional and endowed with the standard Euclidean inner product, the divergence map is represented by the transpose of the matrix representing d_0.

Orthogonal projection onto gradient flows *A flow $X \in C_1$ that lies in the image $\operatorname{Im} d_0$ of the gradient map—that is, $X = d_0\Psi$ for some $\Psi \in C_0$—is called a* potential flow *or* gradient flow. *A flow $X \in C_1$ that lies in the kernel $\ker \partial_1$ of the divergence map is called a* harmonic flow.[10]
Since $\operatorname{Im} d_0$ and $\ker \partial_1$ are linear subspaces of C_1, one can form the orthogonal projection of any flow onto each subspace. The orthogonal projection onto $\operatorname{Im} d_0$ is the linear map $e \colon C_1 \to C_1$ given by $e = d_0\tilde{d}_0$, where $\tilde{d}_0 \colon C_1 \to C_0$ denotes the Moore–Penrose pseudo-inverse of the gradient map.[11] □

With the algebraic apparatus in hand, I now develop the game-theoretic content of the decomposition. The central idea is to represent a game not by its payoff functions directly, but by the payoff *differences* associated with unilateral deviations. This representation captures all strategically relevant information: two games with identical payoff differences across all unilateral deviations share the same set of Nash equilibria, even if their payoff levels differ Candogan, Menache, et al. 2011.

Definition 2.3

Let $\Gamma = \langle N, \mathbf{A}, \pi\rangle$ be a finite normal form game. The ***deviation map*** *of Γ is the linear map $D \colon \mathcal{U} \to C_1$ from the space $\mathcal{U}$ of payoff profiles to the edge space C_1, defined by*

$$(D\pi)(a, a') = \pi_i(a') - \pi_i(a) \in \mathbb{R}^{|\mathcal{E}|} \tag{2.20}$$

for every unilateral deviation $(a, a') \in \mathcal{E}$, where $i \in N$ is the unique player whose action differs between a and a'. The image $D\pi \in C_1$ is called the ***deviation flow*** *of the game.* ♣

The deviation flow assigns to each edge of the response graph the payoff difference experienced by the deviating player. It is through this flow—rather than through the payoff functions themselves—that the strategic structure of a game is most transparently expressed. Indeed, games whose deviation flows coincide are said to be *strategically equivalent*, since they share the same best-response structure and Nash equilibria.

[9] The gradient map is an instance of a *coboundary map* in the theory of simplicial complexes; see Munkres (1984).

[10] The use of the term “harmonic” originates in combinatorial Hodge theory; see Jiang et al. (2011) for a concise development.

[11] I refer to Golan (1992, Chapter 3) for a treatment of the Moore–Penrose pseudo-inverse in the context of linear algebra.

The image of the deviation map, $\operatorname{Im} D \subset C_1$, defines the subspace of *feasible flows*: flows that can arise as the deviation flow of some game. Not every flow in C_1 is feasible; feasibility imposes structural constraints reflecting the combinatorics of multi-player interaction.

Having introduced the deviation map, we can now provide a concise characterisation of potential games in the language of flows. Recall from the definition of an exact potential game—equation (2.1)—that a game $\Gamma = \langle N, \mathbf{A}, \pi\rangle$ is an exact potential game with potential function $\Psi\colon \mathbf{A} \to \mathbb{R}$ if every player's payoff differences from unilateral deviations coincide with the corresponding differences in Ψ. Comparing this with the definitions of the deviation map D and the gradient map d_0, we arrive at the next result.

Proposition 2.3

A finite game $\Gamma = \langle N, \mathbf{A}, \pi\rangle$ is an exact potential game with potential function Ψ if and only if $D\pi = d_0\Psi$. ♠

Proof Let $\Gamma = \langle N, \mathbf{A}, \pi\rangle$ be an exact potential game with potential function Ψ. Then for every unilateral deviation $(a, a') \in \mathcal{E}$ with deviating player $i \in N$, the defining property (2.1) gives

$$(D\pi)(a, a') = \pi_i(a') - \pi_i(a) = \Psi(a') - \Psi(a) = (d_0\Psi)(a, a')$$

so that $D\pi = d_0\Psi$.

Conversely, suppose that $D\pi = d_0\Psi$ for some $\Psi \in C_0$. Then for every player $i \in N$ and every unilateral deviation $(a, a') \in \mathcal{E}$ acted by i,

$$\pi_i(a') - \pi_i(a) = (D\pi)(a, a') = (d_0\Psi)(a, a') = \Psi(a') - \Psi(a)$$

which is exactly the defining property of an exact potential game. ¶

Proposition 2.3 tells us that a game is an exact potential game precisely when its deviation flow is a *gradient flow*—that is, when $D\pi \in \operatorname{Im} d_0$. The space of potential games is therefore characterised by the linear subspace $D^{-1}(\operatorname{Im} d_0) \subset \mathcal{U}$.

2.4.1 Harmonic games

In contrast, games whose deviation flow is *harmonic*—that is, $D\pi \in \ker \partial_1$—are called *harmonic games*. These games represent the strategic opposite of potential games. In a potential game, individual incentives are perfectly aligned with a single global objective. In a harmonic game, incentives are purely cyclical: the payoff differences along unilateral deviations circulate without admitting any consistent ranking of action profiles.

The condition $D\pi \in \ker \partial_1$ admits a revealing restatement directly in terms of payoff increments at each action profile. Recall that the divergence ∂_1 is defined as the adjoint of the gradient map d_0. Applied to a deviation flow $D\pi$, it produces a vertex function $\partial_1 D\pi \in C_0$ whose value at each action profile $a \in \mathbf{A}$ aggregates—with signs determined by the adjoint relationship—the deviation flow along all edges incident to a in the response graph. Unpacking this adjoint yields a concrete, finite-dimensional condition on payoffs, as the next proposition shows.[12]

[12] The result is due to Candogan, Menache, Ozdaglar and Parrilo (2011); see also Abdou, Pnevmatikos, Scarsini and Venel (2022) for a complementary development of game decompositions and their strategic interpretation.

Proposition 2.4 (Candogan, Menache, Ozdaglar and Parrilo, 2011)

A finite normal form game $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ *is harmonic if and only if for every action profile* $a \in \mathbf{A}$*:*

$$\sum_{i \in N} \sum_{a'_i \in A_i} \left[\pi_i(a'_i, a_{-i}) - \pi_i(a) \right] = 0. \tag{2.21}$$

♠

Proof Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be a finite game. By definition, Γ is harmonic if and only if $\partial_1 D\pi = 0$ in C_0, which holds if and only if $(\partial_1 D\pi)(a) = 0$ for every action profile $a \in \mathbf{A}$. It therefore suffices to compute $(\partial_1 D\pi)(a)$ explicitly.

Denote by $e_a \in C_0$ the indicator function of a—the vertex assignment that takes value 1 at a and 0 at every other profile. Using the adjoint characterisation of ∂_1:

$$(\partial_1 D\pi)(a) = \langle \partial_1 D\pi, e_a \rangle_0 = \langle D\pi, d_0 e_a \rangle_1$$

Now $(d_0 e_a)_{(b,b')} = e_a(b') - e_a(b)$ vanishes unless exactly one endpoint of the edge (b, b') equals a, in which case it is ± 1. Hence the inner product $\langle D\pi, d_0 e_a \rangle_1$ reduces to a signed sum of the deviation flow along the edges incident to a. Substituting $(D\pi)_{(a,a')} = \pi_i(a') - \pi_i(a)$ for the unique deviating player i and absorbing the overall sign into the vanishing condition, we obtain

$$(\partial_1 D\pi)(a) = \sum_{i \in N} \sum_{a'_i \in A_i,\, a'_i \neq a_i} \left[\pi_i(a'_i, a_{-i}) - \pi_i(a) \right]$$

The restriction to $a'_i \neq a_i$ may be dropped without affecting the sum, since the term with $a'_i = a_i$ contributes zero. Therefore $(\partial_1 D\pi)(a) = 0$ for every $a \in \mathbf{A}$ if and only if condition (2.21) holds. ¶

Proposition 2.4 provides a concrete, payoff-level characterisation of harmonic games: at every action profile, the *total* unilateral improvement summed over all players and all alternative actions is exactly zero. The gains realisable through some deviations are precisely offset by the losses entailed by others. This is an extremely stringent condition, and in practice only games exhibiting substantial symmetry in their payoff structure satisfy it.

Example 2.11 Consider the classical *Matching Pennies* game with $N = \{1, 2\}$, $A_1 = A_2 = \{\mathbf{H}, \mathbf{T}\}$, and the zero-sum payoff structure

	H	**T**
H	1, −1	−1, 1
T	−1, 1	1, −1

Player 1 receives payoff 1 when both players select the same action and payoff -1 when their actions differ; player 2 receives the opposite payoffs. The game is zero-sum, admits no pure Nash equilibrium, and possesses a unique completely mixed Nash equilibrium in which each player randomises uniformly over $\{\mathbf{H}, \mathbf{T}\}$.

We verify that Matching Pennies is a harmonic game by checking the characterisation (2.21) at each of the four action profiles. Consider first $(\mathbf{H}, \mathbf{H})$. Each player has only one alternative action available, so the sum in (2.21) contains exactly two non-zero terms:

$$\text{Player 1 deviates to } \mathbf{T}: \quad \pi_1(\mathbf{T},\mathbf{H}) - \pi_1(\mathbf{H},\mathbf{H}) = -1 - 1 = -2$$
$$\text{Player 2 deviates to } \mathbf{T}: \quad \pi_2(\mathbf{H},\mathbf{T}) - \pi_2(\mathbf{H},\mathbf{H}) = 1 - (-1) = +2$$

The two increments sum to zero, verifying the harmonic condition at $(\mathbf{H},\mathbf{H})$. Analogous computations at each of the three remaining profiles yield zero as well; this is easy to check. Hence $\partial_1 D\pi = 0$, so Matching Pennies is a harmonic game. By part (c) of Proposition 2.5 we conclude that $P(\Gamma) = 0$.

The harmonic nature of Matching Pennies is accompanied by the cyclic structure of its deviation flow. Arrange the four action profiles into the closed cycle

$$(\mathbf{H},\mathbf{H}) \longrightarrow (\mathbf{T},\mathbf{H}) \longrightarrow (\mathbf{T},\mathbf{T}) \longrightarrow (\mathbf{H},\mathbf{T}) \longrightarrow (\mathbf{H},\mathbf{H})$$

in which each arrow represents a unilateral deviation by a single player. The deviation flow along the four edges of this cycle takes the values

$$-2, \quad -2, \quad -2, \quad -2$$

respectively, summing to $-8 \neq 0$ around the loop. This non-vanishing cycle sum rules out the existence of an exact potential function: were Γ an exact potential game with potential Ψ, then by Proposition 2.3 the flow values along any closed cycle would telescope to zero, since each edge value $\Psi(a') - \Psi(a)$ would be cancelled by the corresponding value along the return path. The example thus exhibits the characteristic dichotomy of harmonic games: local conservation of payoff increments at every vertex coexists with global circulation around closed cycles, leaving no room for a consistent global potential. ♦ ♦

Legacci, Mertikopoulos and Pradelski (2024) show that exponential weight dynamics—perhaps the most widely studied family of no-regret learning algorithms—exhibit quasi-periodic, non-convergent behaviour (Poincaré recurrence) in every harmonic game.

2.4.2 The Potential-Harmonic Decomposition Theorem

The fundamental result that connects these two classes of games—exact potential games and harmonic games—is the combinatorial Hodge decomposition of feasible flows, established by Candogan, Menache, et al. (2011):

Theorem 2.9 (Candogan, Menache, et al. 2011)

Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be a finite normal form game. The space of feasible flows decomposes as the orthogonal direct sum

$$\operatorname{Im} D = \operatorname{Im} d_0 \oplus \ker \partial_1 \tag{2.22}$$

Equivalently, the deviation flow $D\pi$ of any finite game admits a unique decomposition

$$D\pi = D\pi_P + D\pi_H \tag{2.23}$$

where $D\pi_P = e\, D\pi \in \operatorname{Im} d_0$ is the orthogonal projection of $D\pi$ onto the space of potential flows, and $D\pi_H \in \ker \partial_1$ is the orthogonal projection onto the space of harmonic flows. ♡

The decomposition of flows in (2.22) induces a corresponding decomposition in the space of payoff profiles. Candogan, Menache, et al. (2011) show that the payoff function π of any finite game can be uniquely written as

$$\pi = \pi_P + \pi_H + \pi_K \tag{2.24}$$

where π_P is a *normalised potential game*, π_H is a *normalised harmonic game*, and π_K is a *non-strategic or "dummy" game*.[13] The non-strategic component π_K has zero deviation flow and therefore affects neither the Nash equilibria nor the unilateral payoff differences of the original game. The potential component π_P is, moreover, the potential game that is closest to π in the Euclidean sense Candogan, Menache, et al. 2011.

Example 2.12 Consider Shapley's game, a two-player game in which each player has three actions, $A_1 = A_2 = \{1, 2, 3\}$. The payoff matrices are given by

	1	**2**	**3**
1	$1, 0$	$0, 0$	$0, 1$
2	$0, 1$	$1, 0$	$0, 0$
3	$0, 0$	$0, 1$	$1, 0$

Player 1 receives a payoff of 1 when both players choose the same action, and 0 otherwise. Player 2 receives a payoff of 1 at the action profiles $(1, 3)$, $(2, 1)$, and $(3, 2)$, and 0 otherwise—reflecting a cyclical mismatch incentive. This game has no pure Nash equilibrium and admits only a unique completely mixed Nash equilibrium. Applying the decomposition (2.24), we obtain the following components. The potential component is:

$\pi_P =$

	1	**2**	**3**
1	$0.17,\ 0.17$	$-0.33,\ -0.33$	$0.17,\ 0.17$
2	$0.17,\ 0.17$	$0.17,\ 0.17$	$-0.33,\ -0.33$
3	$-0.33,\ -0.33$	$0.17,\ 0.17$	$0.17,\ 0.17$

We note that the payoffs of the two players are identical in every cell, confirming that π_P is indeed a potential game. The harmonic component is:

$\pi_H =$

	1	**2**	**3**
1	$0.50,\ -0.50$	$0.00,\ 0.00$	$-0.50,\ 0.50$
2	$-0.50,\ 0.50$	$0.50,\ -0.50$	$0.00,\ 0.00$
3	$0.00,\ 0.00$	$-0.50,\ 0.50$	$0.50,\ -0.50$

The harmonic component exhibits the antisymmetric payoff structure characteristic of a zero-sum-like game: what one player gains, the other loses. The remaining non-strategic component assigns the constant payoff $\frac{1}{3}$ to each player in every cell, contributing nothing to the strategic structure of the game.

Shapley's game is neither potential nor harmonic, but it is much closer to being harmonic than to being potential. The potentialness of this game, as defined below, is $P(\Gamma) \approx 0.36$. ♦

[13] A game $\langle N, \mathbf{A}, \pi \rangle$ is *non-strategic* if $D\pi = 0$, which holds if and only if every player is indifferent among all of her available actions for every fixed choice of the opponents; see Candogan, Menache, et al. (2011) for a full development. The normalisation selects a canonical representative from each class of strategically equivalent games.

The decomposition of Theorem 2.9 makes it possible to quantify the relative weight of the potential component of any finite game. The next definition, due to Bichler et al. (2025), introduces a normalised measure that maps this relative weight to the unit interval.[14]

Definition 2.4

Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ *be a finite normal form game with deviation flow decomposition* $D\pi = D\pi_P + D\pi_H$ *as in (2.23). The* ***potentialness*** *of* Γ *is the real number*

$$P(\Gamma) := \frac{\|D\pi_P\|}{\|D\pi_P\| + \|D\pi_H\|} \tag{2.25}$$

where $\|\cdot\|$ *denotes the Euclidean norm on* C_1*.*[a]

[a] The definition presupposes that the deviation flow $D\pi$ is non-zero, so that the denominator $\|D\pi_P\| + \|D\pi_H\|$ is positive. If $D\pi = 0$, the game is non-strategic and every player is indifferent among all available actions; in this degenerate case the potentialness is not defined, since there is no strategic content to decompose.

♣

The potentialness of a game measures the fraction of its strategic content—as captured by the deviation flow—that is attributable to the potential component. A game whose strategic structure is entirely potential-like has potentialness equal to one; a game whose strategic structure is entirely harmonic has potentialness equal to zero. The next proposition, due to Bichler et al. (2025), formalises these boundary cases.

Proposition 2.5

Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ *be a finite normal form game. The potentialness* $P(\Gamma)$ *satisfies:*

(a) $P(\Gamma) \in [0, 1]$*;*

(b) $P(\Gamma) = 1$ *if and only if* Γ *is an exact potential game;*

(c) $P(\Gamma) = 0$ *if and only if* Γ *is a harmonic game.*

♠

Proof Consider the game $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ and its decomposition $D\pi = D\pi_P + D\pi_H$.

Proof of (a). Since norms are non-negative, both $\|D\pi_P\|$ and $\|D\pi_H\|$ are non-negative. Hence the ratio in (2.25) lies in $[0, 1]$.

Proof of (b). We have $P(\Gamma) = 1$ if and only if $\|D\pi_H\| = 0$, which holds if and only if $D\pi_H = 0$. In that case, $D\pi = D\pi_P + D\pi_H = D\pi_P \in \operatorname{Im} d_0$, so that $D\pi = d_0\Psi$ for some $\Psi \in C_0$ and the game is an exact potential game by Proposition 2.3.

Conversely, suppose Γ is an exact potential game with potential Ψ. Then $D\pi = d_0\Psi \in \operatorname{Im} d_0$. Since $D\pi_P$ is the orthogonal projection of $D\pi$ onto $\operatorname{Im} d_0$, and $D\pi$ already belongs to $\operatorname{Im} d_0$, the projection leaves it invariant: $D\pi_P = D\pi$. Hence $D\pi_H = D\pi - D\pi_P = 0$, so $\|D\pi_H\| = 0$ and $P(\Gamma) = 1$.

Proof of (c). We have $P(\Gamma) = 0$ if and only if $\|D\pi_P\| = 0$, which holds if and only if $D\pi_P = 0$. In that case, $D\pi = D\pi_H \in \ker \partial_1$ and the game is harmonic by definition.

Conversely, if Γ is a harmonic game then $D\pi \in \ker \partial_1$. By an argument analogous to the one in assertion (b), the projection onto the harmonic subspace leaves $D\pi$ invariant, giving $D\pi_H = D\pi$ and hence $D\pi_P = 0$. ¶

[14] The definition is closely related to the *maximum pairwise difference* $\delta(\pi, \pi_P) = \|D\pi - D\pi_P\| = \|D\pi_H\|$ introduced by Candogan, Ozdaglar, and Parrilo (2013b) as a measure of distance from a game to its nearest potential game. The potentialness $P(\Gamma)$ normalises this distance into a dimensionless ratio.

The properties stated in Proposition 2.5 confirm that potentialness provides a meaningful and well-calibrated measure of proximity to the class of exact potential games. In particular, potentialness depends only on the deviation flow $D\pi$ and is therefore invariant under the addition of non-strategic components: strategically equivalent games have identical potentialness.

To illustrate the discriminating power of this measure, I note that the Prisoners' Dilemma—an exact potential game, as demonstrated earlier in this chapter—has potentialness $P(\Gamma) = 1$. The Battle of the Sexes, which is not an exact potential game but possesses pure Nash equilibria and exhibits convergent learning dynamics, has a potentialness of approximately 0.94. Matching Pennies, a zero-sum game with only a mixed Nash equilibrium and non-convergent dynamics, is a harmonic game with $P(\Gamma) = 0$. Shapley's game, as computed in Example 2.12 above, has $P(\Gamma) \approx 0.36$.

These values suggest that potentialness captures meaningful structural differences between games. Bichler et al. (2025) provide extensive numerical evidence that potentialness serves as a strong predictor both for the existence of strict pure Nash equilibria and for the convergence of no-regret learning dynamics in finite games. In random matrix games, they find that games with potentialness below 0.4 rarely exhibit convergent dynamics, while games with potentialness exceeding 0.6 mostly do. In economically motivated games—such as first- and second-price auctions, Tullock contests, and all-pay auctions—potentialness characterises the observed differences in convergence behaviour: all-pay auctions, which have very low potentialness and only mixed equilibria, are resistant to learning, whereas first- and second-price auctions and contests, which have substantially higher potentialness, admit convergent dynamics.[15]

The potentialness metric thus provides an ex-ante diagnostic for the learnability of a game under standard dynamics, complementing the exact characterisations developed in the earlier sections of this chapter and offering a more graded perspective on the strategic landscape of finite games.

2.4.3 A parametric family of Cournot-like games with prescribed potentialness

The decomposition of a finite game into potential, harmonic, and non-strategic components provides a constructive means of building games with prescribed levels of potentialness. Starting from three "building block" games—one potential, one harmonic, and one non-strategic—a weighted sum produces a game whose deviation flow is itself a weighted sum of the three components' deviation flows. Since the deviation flow of a non-strategic game vanishes by definition, and the potential and harmonic components of the sum are orthogonal in C_1, the norms in the definition of potentialness (Definition 2.10 in the textbook) satisfy a simple algebraic relationship that allows the potentialness to be read off directly from the weights.

To illustrate this construction with an economically motivated example, I assemble a two-player, two-action game in which each component carries concrete economic meaning. The two players are firms operating in a common market; each firm chooses between a low output level $q_L = 1$ and a high output level $q_H = 3$, so that $A_1 = A_2 = \{L, H\}$.

The potential component: a Cournot duopoly The potential component π_P describes the quantity competition between the two firms in a linear market. The inverse demand function is

$$p(Q) = a - b\,Q \qquad \text{with } a = 10,\ b = 1,$$

[15] For a detailed analysis of potentialness in auction games and its relationship to both complete-information and Bayesian settings, I refer to Bichler et al. (2025).

where $Q = q_1 + q_2$ denotes aggregate output. Each firm faces constant marginal production cost $c = 2$. The resulting Cournot payoffs are $\pi_i^C(q_i, q_j) = [p(q_i + q_j) - c]\, q_i$, yielding the normal form

π_P:

	L_2	H_2
L_1	6, 6	4, 12
H_1	12, 4	6, 6

This Cournot duopoly is an exact potential game, with potential function

$$\Psi(q_1, q_2) = (a - c)(q_1 + q_2) - b(q_1^2 + q_2^2) - b\, q_1 q_2,$$

as established by Monderer and Shapley (1996). On the action grid $\{L, H\}^2$, the potential takes the values

$$\Psi(L, L) = 13, \quad \Psi(L, H) = \Psi(H, L) = 19, \quad \Psi(H, H) = 21.$$

The unique maximiser of Ψ on $\{L, H\}^2$ is the profile (H, H), which coincides with the unique strict pure Nash equilibrium of the Cournot duopoly.[16]

The harmonic component: market interference The harmonic component π_H captures a pattern of strategic interference that disrupts the coordinated quantity competition of the Cournot game. I model this interference as a Matching Pennies-style interaction, in which firm 1 benefits from matching firm 2's output level and firm 2 benefits from mismatching firm 1's output level:

π_H:

	L_2	H_2
L_1	$+1, -1$	$-1, +1$
H_1	$-1, +1$	$+1, -1$

Economically, this component can be interpreted as capturing an antagonistic market-timing effect: firm 1 prefers to coordinate its output scale with firm 2 (e.g., matching a market leader's pace), while firm 2 derives advantage from doing the opposite (e.g., avoiding direct head-to-head scale competition). The zero-sum structure reflects that this "rivalry" redistributes a fixed quantity of market surplus rather than creating or destroying it. As established in Example 2.9 of the chapter, this is a harmonic game with $P(\pi_H) = 0$.

The non-strategic component: a pollution externality The non-strategic component π_K represents a commons-type externality: each firm's production pollutes a shared resource input, and the pollution imposes a cost on the *other* firm's operations. Formally, each firm i suffers a loss $e\, q_j$ proportional to the output of its rival, with $e = 3$:

$$\pi_i^K(q_i, q_j) = -e\, q_j \qquad (i, j \in \{1, 2\},\ i \neq j).$$

This gives the normal form

π_K:

	L_2	H_2
L_1	$-3, -3$	$-9, -3$
H_1	$-3, -9$	$-9, -9$

[16] This is the discrete analogue of the familiar Cournot–Nash equilibrium of the continuous game, in which each firm produces its best response to the other's output. For the continuous Cournot duopoly with the parameters above, the interior Nash equilibrium quantities are $q_1^* = q_2^* = (a - c)/(3b) = 8/3$, which lies between q_L and q_H.

Each firm's payoff in this game depends only on the *other* firm's action; consequently, every firm is indifferent among its own available actions for every fixed choice of the rival. This is precisely the defining property of a non-strategic game: $D\pi_K = 0$. The externality is economically meaningful—it imposes real welfare losses—but it has no strategic bite, because no firm can affect its own payoff through its own action choice.

The weighted family For non-negative weights $\alpha, \beta, \gamma \geqslant 0$ with $\alpha + \beta > 0$, define the weighted sum

$$\pi_{\alpha,\beta,\gamma} := \alpha\, \pi_P + \beta\, \pi_H + \gamma\, \pi_K. \tag{2.26}$$

The deviation flow of $\pi_{\alpha,\beta,\gamma}$ decomposes as

$$D\pi_{\alpha,\beta,\gamma} = \alpha\, D\pi_P + \beta\, D\pi_H + \gamma\, D\pi_K = \alpha\, D\pi_P + \beta\, D\pi_H,$$

since $D\pi_K = 0$. Moreover, $D\pi_P \in \operatorname{Im} d_0$ and $D\pi_H \in \ker \partial_1$ are orthogonal in C_1 by Lemma 2.7 in the chapter. It follows that the orthogonal projection of $D\pi_{\alpha,\beta,\gamma}$ onto $\operatorname{Im} d_0$ is precisely $\alpha\, D\pi_P$, and the projection onto $\ker \partial_1$ is $\beta\, D\pi_H$. The potentialness of the weighted game is therefore

$$P(\pi_{\alpha,\beta,\gamma}) = \frac{\alpha\, \|D\pi_P\|}{\alpha\, \|D\pi_P\| + \beta\, \|D\pi_H\|}, \tag{2.27}$$

independent of the weight γ on the non-strategic component. Direct computation gives

$$\|D\pi_P\| = \sqrt{6^2 + 2^2 + 6^2 + 2^2} = \sqrt{80} = 4\sqrt{5} \approx 8.944$$

and

$$\|D\pi_H\| = \sqrt{(-2)^2 + 2^2 + 2^2 + (-2)^2} = \sqrt{16} = 4.$$

Consequently, for the family $\pi_{\alpha,\beta,\gamma}$ the potentialness reduces to

$$P(\pi_{\alpha,\beta,\gamma}) = \frac{\alpha\sqrt{5}}{\alpha\sqrt{5} + \beta}. \tag{2.28}$$

The weight γ drops out entirely—as it must, since potentialness is invariant under the addition of a non-strategic game—while the ratio β/α serves as a single continuous parameter that tunes the potentialness from 1 (when $\beta = 0$) to 0 (as $\alpha \to 0$ with β fixed).

Numerical illustration Table 2.1 reports the potentialness and the set of pure Nash equilibria of the weighted game $\pi_{\alpha,\beta,1}$ for several representative choices of (α, β).

Three observations emerge from the table.

First, at $(\alpha, \beta) = (1, 0)$ the weighted game reduces to the Cournot duopoly (up to the non-strategic pollution component), which is an exact potential game with $P = 1$ and a unique strict pure Nash equilibrium at (H_1, H_2). As the harmonic weight β is increased while α remains fixed, the potentialness declines monotonically, and the strict Nash equilibrium eventually ceases to exist.

Second, the transition from existence to non-existence of a pure Nash equilibrium occurs around $P \approx 0.4$. For $(\alpha, \beta) = (1, 3)$, where $P \approx 0.427$, a unique pure Nash equilibrium (H_1, L_2) persists; but at $(\alpha, \beta) = (1, 4)$, with $P \approx 0.359$, no pure Nash equilibrium exists at all. This threshold behaviour exemplifies the empirical finding of Bichler, Legacci, Mertikopoulos, Oberlechner and Pradelski (2025) that potentialness below roughly 0.4 is rarely compatible with the existence of strict pure Nash equilibria in random matrix games.

Third, the pure Nash equilibrium *shifts* as β increases. In the unweighted Cournot game (i.e., $\beta = 0$), the equilibrium is the Cournot–Nash profile (H_1, H_2). As the harmonic rivalry strengthens, firm 2's best response

α	β	$P(\pi_{\alpha,\beta,1})$	Pure Nash equilibria	Strategic character
1	0	1.000	$\{(H_1, H_2)\}$	pure Cournot; exact potential game
3	1	0.870	$\{(H_1, H_2)\}$	Cournot-dominated; SPNE preserved
2	1	0.817	$\{(H_1, H_2)\}$	Cournot-dominated; SPNE preserved
1	1	0.691	$\{(H_1, L_2), (H_1, H_2)\}$	balanced; two SPNE
1	2	0.528	$\{(H_1, L_2)\}$	harmonic-dominated; NE shifts
1	3	0.427	$\{(H_1, L_2)\}$	harmonic-dominated; NE shifts
1	4	0.359	$\varnothing$	no pure NE; only mixed
0	1	0.000	$\varnothing$	pure Matching Pennies

Table 2.1: Potentialness and pure Nash equilibrium structure of the weighted game $\pi_{\alpha,\beta,1}$ for representative weights. The empirical threshold at $P \approx 0.4$, below which pure Nash equilibria cease to exist, matches the general pattern documented by Bichler, Legacci, Mertikopoulos, Oberlechner and Pradelski (2025) for random and economically structured games.

changes: at $(\alpha, \beta) = (1, 2)$ the unique equilibrium becomes (H_1, L_2), reflecting that firm 2 now benefits from mismatching firm 1's output (as per the harmonic component). For β sufficiently large, no profile is mutually stable, and convergence to any pure outcome is lost entirely.

Concretely, the fully assembled game $\pi_{1,1,1}$—equal weights on all three components—has the payoff matrix

$\pi_{1,1,1}$:

	L_2	H_2
L_1	$+4,\ +2$	$-6,\ +10$
H_1	$+8,\ -4$	$-2,\ -4$

with potentialness $P \approx 0.691$ and two pure Nash equilibria, at (H_1, L_2) and (H_1, H_2). The economic interpretation is intuitive: moderate harmonic interference creates strategic ambiguity—both the Cournot outcome and the mismatch outcome can sustain themselves—while the pollution externality redistributes welfare without altering the equilibrium structure.

Interpretation The weighted family constructed in this subsection demonstrates three features of potentialness that are central to its use as a diagnostic tool.

First, potentialness is invariant under non-strategic perturbations: the weight γ on the pollution externality does not enter formula (2.28). Non-strategic externalities can substantially shift the distribution of welfare without altering the underlying strategic structure, and this economic fact is reflected in the invariance of potentialness under such perturbations.

Second, potentialness provides a smooth, continuously varying measure of distance from the class of exact potential games. The transition from a pure potential game ($\alpha = 1, \beta = 0, P = 1$) to a pure harmonic game ($\alpha = 0, \beta = 1, P = 0$) traces out a one-parameter family whose behaviour interpolates monotonically between the two extremes.

Third, and most importantly, the threshold behaviour of pure Nash equilibrium existence at $P \approx 0.4$ aligns with the general empirical pattern documented by Bichler, Legacci, Mertikopoulos, Oberlechner and Pradelski (2025). In this carefully controlled family of games, potentialness acts as a precise predictor of the equilibrium structure: for $P \gtrsim 0.4$, a strict pure Nash equilibrium exists; for $P \lesssim 0.4$, it does not. Although the specific threshold is tied to the particular components used, the qualitative phenomenon—that low potentialness is incompatible with pure-strategy learnability—is remarkably robust.

2.5 Notes

The notion of exact potential developed in Section 2.1 is the strongest of the several flavours of potential game studied in this book. It delivers the most complete catalogue of consequences — existence of strict pure Nash equilibria, convergence of best- and better-response dynamics, a direct welfare reading of the potential landscape, and integrability of the payoff flow — but it is also the hardest property for a game to possess. In these notes I collect three clarifying strands: Subsection 2.5.1 situates the exact notion within the hierarchy of potential concepts encountered in the book and articulates the power–restrictiveness trade-off; Subsection 2.5.2 organises the several characterisations of Section 2.2 into three methodological families; and Subsection 2.5.3 surveys the principal analytic methods for constructing an exact potential function in practice, supplementing the two methods presented in Section 2.3.

2.5.1 Exact potentials: most powerful, most restrictive

The defining condition of Section 2.1 requires that a scalar function $\Psi\colon \mathbf{A} \to \mathbb{R}$ reproduce, with sign and magnitude, every unilateral payoff increment of every player. This is demanding. It fixes not only the direction of improving deviations — the ordinal content of the game — but also their cardinal magnitudes, and it does so uniformly across the whole player set.

The remaining chapters of this book will progressively relax this condition. The hierarchy that emerges is

$$\text{exact} \subset \text{weighted} \subset \text{ordinal} \subset \text{generalised-ordinal} \subset \text{best-response}.$$

Weighted potentials, already introduced by Monderer and Shapley (1996), replace cardinal matching by matching up to player-specific positive multipliers. Ordinal potentials require only sign matching; generalised-ordinal potentials require sign matching only for payoff-improving deviations; best-response potentials, due to Voorneveld (2000), retain only what is needed to preserve the best-response correspondence. Further relaxations — pseudo- and quasi-potentials — appear in Dubey, Haimanko, and Zapechelnyuk (2006) and in the semi-tensor-product literature surveyed in Chapter 4 (see Cheng and Ji 2021). The near-potential games of Candogan, Ozdaglar, and Parrilo (2013b) and Candogan, Ozdaglar, and Parrilo (2013a) form a continuous relaxation that is in turn the direct antecedent of the potentialness measure of Section 2.4.

The trade-off is clear in substance. All the flavours above preserve the existence of pure Nash equilibria on finite action sets, but progressively shed the cardinal welfare reading of the potential, the quantitative rate estimates for best- and better-response dynamics, the continuity of the potential under payoff perturbations, and — decisively — the closure under summation. Exact potential games form a linear subspace of the space of all finite games, a fact on which the orthogonal decomposition of Section 2.4 explicitly relies; ordinal potential games do not. The restrictiveness of exactness is, in other words, precisely what makes the quantitative theory of potentialness possible.

2.5.2 A classification of the characterisations of Section 2.2

The characterisations presented in Section 2.2 are often stated successively, as if they were alternative proofs of a single equivalence. They are equivalent, of course, but they are not of the same kind. Each belongs to a distinct methodological family with its own natural domain of application, and the distinction is practically decisive in applied work.

Family A: differential and integrability-based characterisations.

This family applies to games with smooth payoffs on an open, simply connected subset of a Euclidean action space. Its two representatives in Section 2.2 are the classical cross-derivative condition

$$\frac{\partial^2 \pi_i}{\partial a_i\, \partial a_j} = \frac{\partial^2 \pi_j}{\partial a_j\, \partial a_i} \qquad (i \neq j)$$

of Monderer and Shapley (1996), and its multi-dimensional Euclidean generalisation by Arefizadeh, Nedić, and Dasarathy (2024), which matches off-diagonal Jacobians. Both are restatements of the closedness of the one-form built from the players' gradient payoffs, and both reduce via Poincaré's Lemma to its exactness on the action space. The strength of this family is that it suggests its own constructive companion — the line-integral formula of Section 2.3 — but it requires $\mathcal{C}^2$ regularity and simple connectedness, and is silent for finite games.

Family B: combinatorial and decompositional characterisations.

This family applies to finite games without regularity assumptions. Its principal representative in Section 2.2 is the characterisation of Ui (2000): a finite game is an exact potential game if and only if every player's payoff function admits an expansion

$$\pi_i(a) = \sum_{\substack{T \subseteq N \\ i \in T}} \phi_T(a_T)$$

in terms of interaction potentials $\phi_T \colon \mathbf{A}_T \to \mathbb{R}$ shared across all players in T; the potential is then $\Psi(a) = \sum_{T \subseteq N} \phi_T(a_T)$. The standard decomposition via dummy games, used repeatedly in Section 2.2, is a pragmatic instance of the same principle, restricted to singletons and pairs. Interaction potentials are recoverable by Möbius inversion on the coalition lattice, and for games of low coalition order the expansion collapses to a short closed-form expression. The number of coalitions grows as $2^{|N|}$, which limits the method on large games unless the coalition structure is itself sparse.

Family C: linear-operator and expectation-based characterisations.

This family casts exactness as a vanishing condition in a Hilbert space of payoff flows. The first representative in Section 2.2 is the characterisation of Sandholm (2010), reformulated operationally by Hwang and Rey-Bellet (2020): exactness is equivalent to the annihilation of $D\pi$ by a specific expectation operator. The Hodge-theoretic characterisation of Candogan, Menache, et al. (2011) — a finite game is exact-potential precisely when its deviation flow $D\pi$ lies in the image of the coboundary operator d_0 — belongs to the same family, and it is the characterisation on which the decomposition theorem of Section 2.4 rests. Family C bridges the combinatorial and analytic viewpoints and yields the quantitative potentialness diagnostic of Section 2.4 as a by-product, but its full power is visible only once the flow-space formalism is in place.

A unifying thread Each family is, in fact, a restatement of a single topological trait: the deviation flow $D\pi$ has trivial circulation around every closed cycle in the action profile space. Family A expresses this through differential forms, Family B through coalition-indexed inclusion–exclusion, and Family C through Hilbert-space orthogonality. The choice of language can be tuned to the data at hand. Smooth oligopoly models on convex Euclidean domains yield most cheaply to Family A; congestion and networked games, where coalition locality is the key structural feature, reduce with minimal labour under Family B; population-level and large random models, in which the flow-space decomposition of Section 2.4 is the natural setting, call for Family C.

2.5.3 Additional methods for computing an exact potential

Section 2.3 presents two constructive methods. The first is the classical line-integral formula of Monderer and Shapley (1996) for differentiable games on a simply connected action domain: once closedness of the gradient one-form has been verified, the potential Ψ is obtained by integrating that one-form along any convenient path from a reference profile to the profile of interest, with path-independence guaranteed by exactness. The second is the path-potential construction of Arefizadeh, Nedić, and Dasarathy (2024), which adapts the line-integral idea to a multi-dimensional Euclidean action space by evaluating the line integral along a coordinate-wise path. Both methods are, structurally, representatives of a single *integrability family*: they fix a reference profile, sum payoff increments along a path, and appeal to exactness to secure path-independence. In the remainder of this subsection I outline five additional methods that are either of a different type altogether or are specialisations adapted to structural features of particular classes of games. The treatment is deliberately brief; the aim is to provide the reader with a map of what is available, together with pointers to the literature, rather than a full exposition. Purely numerical algorithms are not discussed; the focus throughout is on constructive analytic methods.

Method 1: coalitional construction via Möbius inversion.

For a finite game satisfying Ui's coalitional decomposition of Family B above, the interaction potentials ϕ_T are recoverable from the payoff functions by Möbius inversion on the coalition lattice. With a fixed reference action $0 \in A_i$ for each player, one obtains

$$\phi_T(a_T) \;=\; \sum_{S \subseteq T} (-1)^{|T \setminus S|} \, \pi_{\min(T)}\big(a_S, 0_{N \setminus S}\big),$$

and the potential follows as $\Psi(a) \;=\; \sum_{T \subseteq N} \phi_T(a_T)$. For games with low-order coalition interaction — additively separable payoffs, pairwise interactions, networked oligopolies — the sum collapses to a short closed-form expression and the method produces the potential directly, without integration. I refer to Ui (2000) for the proof of correctness of this inversion and to Chapter 3 for its extension to weighted potentials.

Method 2: Hodge projection onto the potential subspace.

For any finite game, Theorem 2.9 of Section 2.4 gives the Hodge decomposition $D\pi = D\pi_P + D\pi_H$. When the game is already exact-potential the harmonic component $D\pi_H$ vanishes, and the potential is recovered in closed form as

$$\Psi \;=\; d_0^{\dagger}(D\pi),$$

where $d_0^{\dagger}$ denotes the Moore–Penrose pseudo-inverse of the coboundary operator $d_0 \colon C_0 \to C_1$ introduced in Section 2.4. Equivalently, Ψ is the unique element of C_0, up to an additive constant, that satisfies $d_0\Psi = D\pi$. The construction is coordinate free and, importantly, extends to games that are not exactly potential: the projection $D\pi_P = e\, D\pi$ then returns an approximate potential function, corresponding to the nearest exact potential game in the Euclidean sense (Candogan, Ozdaglar, and Parrilo 2013b). This method is particularly well-suited to empirical settings in which the deviation flow is directly measurable, for example auction experiments of the kind analysed by Bichler et al. (2025).

Method 3: semi-tensor-product potential equation.

In a parallel development largely independent of the Hodge-theoretic tradition, Cheng, Liu, et al. (2016) and collaborators reformulated the exact-potential condition as a linear algebraic equation within the semi-tensor-product (STP) algebra. When the payoff functions of a finite game are encoded as Kronecker structured vectors, the exact-potential condition becomes a linear system in the STP algebra — the *potential equation* — and its solution, when it exists, is the potential vector. The approach reduces the twin tasks of detecting exact-potentiality and of constructing the potential function to the analysis of a single structured linear system. Cheng and Ji (2021) and Yuanhua Wang and Cheng (2017) extend the machinery to weighted, ordinal, and vector-payoff generalisations, and to Bayesian games; I discuss the STP approach at length in Chapter 4. For present purposes I note two attractive features. First, the method yields a complete answer — a potential function if one exists, a precise obstruction if it does not — without requiring any path integration. Second, the STP potential equation is intrinsically computable, and the Cheng school has developed a corresponding algorithmic theory that is of interest to the applied literature; see Yanfei Wang, Hu, and Cheng (2024) for a typical treatment and the review in Cheng and Ji (2021).

Method 4: direct construction from the standard decomposition.

The standard decomposition of Section 2.2 writes each player's payoff as $\pi_i = \pi_i' + \xi_i$, where π_i' is common to all players up to the dummy residual ξ_i. When the common components π_i' in fact coincide across players — as they do in additively coupled payoff structures, in networked Cournot duopolies with symmetric cost and demand, and in public-good games with linear spillovers — the shared component π' is itself the exact potential, up to an additive constant. This method is the simplest of all available routes to the potential when it applies, and it is the working method behind the treatment of additively coupled games in Section 2.1 and the networked Cournot application in Section 2.2. It is best understood as a special case of Method 1 in which Möbius inversion is trivial.

Method 5: symmetry-exploiting constructions.

When a game admits a group action — a permutation symmetry of the player set, a graph automorphism of the interaction network, or a translation symmetry of the common action set — the exact potential inherits that symmetry, and it is enough to construct the potential on a fundamental domain and then extend it by invariance. The canonical illustration is Rosenthal (1973)'s congestion game, in which the player-wise symmetry of congestion costs reduces the potential to the single aggregate sum $\Psi(a) = \sum_{r\in R}\sum_{k=1}^{n_r(a)} c_r(k)$, where $n_r(a)$ denotes the number of players using resource r at profile a and c_r the congestion cost on that resource. Marden and collaborators exploit the same principle extensively in distributed engineered systems; see Marden (2012), Marden, Arslan, and Shamma (2009), and the survey in Marden and Shamma (2015). Strictly speaking, the symmetry method is a constructive shortcut rather than an independent characterisation: whichever of Methods 1–4 is in principle available will, when symmetry is present, reduce under the symmetry group to a far shorter computation. I include it here as a distinct entry because, in applied work, it is the method most frequently used to handle genuinely large games.

Putting the methods to work The seven methods — the two of Section 2.3 and the five above — are not independent. Methods 1, 4, and 5 are specialisations within Family B of Subsection 2.5.2; Method 2 is the Family C construction in its most concrete form and Method 3 its STP-coordinate specialisation; the two methods of Section 2.3 are the Family A representatives. In practice, the choice is dictated by the structure of the game. Differentiable games on a simply connected domain default to the line integral of Section 2.3; finite games of low coalition order to coalitional decomposition (Method 1) or the standard decomposition (Method 4); games whose natural object is the deviation flow to the Hodge projection (Method 2); games approached algebraically via the STP framework of Chapter 4 to the potential equation (Method 3); and genuinely large games with transparent symmetry to the invariance shortcut of Method 5.

Forward pointers The three families of Subsection 2.5.2 reappear repeatedly in the remaining chapters of this book. Family B underlies the weighted-potential theory of Chapter 3, in which Ui's interaction-potential construction extends with player-specific multipliers. Family A reappears in Chapter 4 on potential games with continuous action spaces. Family C underlies the near-potential and potentialness analyses of Chapter 6, where the diagnostic of Section 2.4 is developed into a quantitative theory of almost-potential behaviour. The semi-tensor-product apparatus that underpins Method 3 is the subject of the second half of Chapter 4, where its relationship to the line-integral and Hodge constructions is made precise.

2.6 Proofs of Theorems and Propositions

Some of the proofs of the results presented in this chapter rely on tools from differential multi-variable calculus. These tools are presented here first.

2.6.1 Preamble: Some mathematical concepts

In these auxiliary mathematical notes, we restrict ourselves here to the discussion of so-called one-forms only. This is presented in a general context, without explicit reference to a game theoretic structure.

Differential calculus Consider a function $f\colon \mathbf{A} \to \mathbb{R}^k$. Then the *differential* of f is defined as

$$df = \sum_{i=1}^{k} \frac{\partial f}{\partial a_i}\, da_i$$

On the other hand, a *differential one-form* is defined as

$$\omega = \sum_{i=1}^{n} g_i(a)\, da_i \qquad \text{where } g_i\colon \mathbf{A} \to \mathbb{R}^n \quad \text{for } i \in \{1, \ldots, k\}.$$

It should be clear that the differential of a function is a one-form, but the other way is not true. The one-forms that have this property are called exact: A one-form ω is *exact* if there exists some function $f\colon \mathbf{A} \to \mathbb{R}^k$ such that $\omega = df$.

Exact one-forms are natural and easy to work with. Therefore, it is a primary objective of the theory to present a relatively easy way to check when a one-form is exact or not. Exactness of a differential one-form can indeed be established through testing of a property that is referred to as closedness: A one-form $\omega = \sum_{i=1}^{k} g_i(a)\, da_i$ is *closed* if for all $i, j \in \{1, \ldots, k\}$: $\frac{\partial g_i}{\partial a_j} = \frac{\partial g_j}{\partial a_i}$.

It is relatively easy to establish that exact one-forms are closed. The reverse is less obvious.

Lemma 2.3 (Poincaré's Lemma)

Let ω be a one-form that is defined on a open, convex set $\mathbf{A}^e$, where $\mathbf{A} \subset \mathbf{A}^e$. Then ω is exact if and only if it is closed. ♡

I refer to Bouchard (2024, Theorem 3.6.4) and its proof to establish this important result.[17]

The role of exact one-forms is important to understand potential games. If $\omega = df$, then the function f acts as a potential function with regard to the one-form ω.[18] It should be clear that the mathematical theory of one-forms gives rise to the use of the term "potential" to describe these structures.

It is clear that exact one-forms have special properties regarding their integration. For that, we introduce a parametric curve in $\mathbb{R}^k$ as a vector-valued function $\alpha\colon [0,1] \to \mathbb{R}^k$ with $\alpha(t) = (x_1(t), x_2(t), \ldots, x_k(t))$ such that α can be extended to a $\mathcal{C}^1$ function on an open set containing $[0,1]$ with $\alpha'(t) \neq 0$ for all $t \in [0,1]$ and, if $\alpha(t) = \alpha(s)$ for $t \neq s$, it has to hold that $s, t \in \{0,1\}$. A parametric curve is *closed* if $\alpha(0) = \alpha(1)$.

Note that a parametric curve as introduced here has a natural orientation given by the trajectory from $\alpha(0)$ to $\alpha(1)$.

Line integrals Let $\omega = \sum_{i=1}^{k} g_i(a)\, da_i$ be a one-form and let $\alpha\colon [0,1] \to \mathbb{R}^k$ be a parametric curve. Then we write

$$\alpha^*\omega = \left(\sum_{i=1}^{k} g_i(\alpha(t)) \frac{da_i}{dt} \right) dt$$

as the associated one-form. The line integral of ω over the parametric curve α is now defined as

$$\oint_\alpha \omega = \int_0^1 \alpha^*\omega = \int_0^1 \left(\sum_{i=1}^{k} g_i(\alpha(t)) \frac{da_i}{dt} \right) dt$$

Lemma 2.4 (Fundamental Theorem of Line Integrals)

Let $\omega = df$ be an exact one-form on an open set $U \subseteq \mathbb{R}^k$ and let α be a parametric curve with $\alpha([0,1]) \subset U$. Then

$$\oint_\alpha \omega = \oint_\alpha df = f(\alpha(1)) - f(\alpha(0))$$

♡

Thus, the line integral of a one-form only depends on the starting and ending points of the curve over which the integral is executed. An easy corollary of this fundamental insight is that the line integral of an exact one-form over a closed curve is always zero. This gives rise to the following formulation of the fundamental characteristics of an exact one-form $\omega = df$ defined on a convex, open set in $\mathbb{R}^k$. Indeed, the following four statements are equivalent:

(i) ω is an exact one-form;

(ii) ω is a closed one-form;

(iii) For any closed parametric curve α it holds that $\oint_\alpha \omega = 0$, and

(iv) All line integrals of ω are independent and given by $\oint_\alpha \omega = f(\alpha(1)) - f(\alpha(0))$.

[17] Note here that the Lemma is there formulated for a simply connected set, but that the version of the Lemma for convex sets results from that by noting that convex sets are simply connected. For a direct proof of this Lemma, I also refer to Rudin (1976, Theorem 10.39).

[18] In fact, the vector field generated by ω is conservative and its potential is exactly given by the vector form of f. I refer to Bouchard (2024, Section 2.2.2) for some details.

The proofs of Theorem 2.3 as well as Theorem 2.4 rely heavily on these concepts from vector calculus. It is also clear that Monderer-Shapley's formulation of the potential as stated in Theorem 2.2 is a simple application of the Fundamental Theorem of Line Integrals.

2.6.2 Proof of Theorem 2.3

The proof of the generalisation of the Monderer-Shapley characterisations rests on multivariable calculus.
Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be a Euclidean game with, for every player $i \in N$, action set $A_i \subset \mathbb{R}^{k_i}$ with $k_i \geqslant 1$, and payoff function π_i continuously differentiable on an open set $\mathbf{A}^e \subset \mathbb{R}^{\bar{k}}$ containing $\mathbf{A} \subset \mathbb{R}^{\bar{k}}$, where $\bar{k} = \sum_{i \in N} k_i \geqslant n$.
We can re-represent the extended open strategy space $\mathbf{A}^e$ and strategy tuples $a \in \mathbf{A}^e$ using the multi-dimensional structure of these strategy spaces. Indeed, we can write

$$a = (a_1, \dots, a_n) = (a_{1,1}, \dots, a_{1,k_1}, a_{2,1}, \dots, a_{2,k_2}, a_{3,1}, \dots, a_{n,k_n}) \in \mathbb{R}^{\bar{k}}$$

where $a_i = (a_{i,1}, \dots, a_{i,k_i}) \in A_i^e \subset \mathbb{R}^{k_i}$, where A_i^e is the projection of $\mathbf{A}^e$ on $\mathbb{R}^{k_i}$.
It should be clear that Γ is an exact potential game if and only if there exists $\Psi \colon \mathbf{A} \to \mathbb{R}$ such that for every $a \in \mathbf{A}$, $i \in N$ and $k \in \{1, \dots, k_i\}$ it holds that $\pi_i(a) - \pi_i(a_{-(i,k)}, a_{i,k} + \delta) = \Psi(a) - \Psi(a_{-(i,k)}, a_{i,k} + \delta)$ for $\delta \in (-1, 1)$ sufficiently small such that $(a_{i,1}, \dots, a_{i,k-1}, a_{i,k} + \delta, a_{i,k+1}, \dots, a_{i,k_i}) \in A_i^e$.
In particular, this implies by the application of the standard definition of the partial derivative regarding dimension (i, k) that Γ is an exact potential game if and only if there exists $\Psi \colon \mathbf{A} \to \mathbb{R}$ such that for every $a \in \mathbf{A}$, $i \in N$ and $k \in \{1, \dots, k_i\}$ it holds that

$$\frac{\partial \pi_i}{\partial a_{i,k}}(a) = \frac{\partial \Psi}{\partial a_{i,k}}(a).$$

From this it immediately follows that Γ is an exact potential game if and only if there exists some differentiable function $\Psi \colon \mathbf{A} \to \mathbb{R}$ such that for every $i \in N$ and $a \in \mathbf{A}$:

$$D_i\, \pi_i(a) = \left(\frac{\partial \pi_i}{\partial a_{i,1}}(a), \dots, \frac{\partial \pi_i}{\partial a_{i,k_i}}(a) \right) = \left(\frac{\partial \Psi}{\partial a_{i,1}}(a), \dots, \frac{\partial \Psi}{\partial a_{i,k_i}}(a) \right) = D_i\, \Psi(a)$$

This concludes the proof of the assertion of Theorem 2.3.

2.6.3 Proof of Theorem 2.4

Consider a game $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ that is a Euclidean game with for every player $i \in N$ the action set $A_i \subset \mathbb{R}^{k_i}$ is convex with $k_i \geqslant 1$ and the payoff function π_i is twice continuously differentiable on an open, convex set $\mathbf{A}^e \subset \mathbb{R}^{\bar{k}}$ containing $\mathbf{A} \subset \mathbb{R}^{\bar{k}}$ with $\bar{k} = \sum_{i \in N} k_i \geqslant n$.
From the game Γ we can construct a one-form defined by

$$\omega = \sum^{(n,k_n)} (i,k) = (1,1)\, \frac{\partial \pi_i}{\partial a_{i,k}}(a)\, da_{i,k} \qquad \text{for every } a \in \mathbf{A}^e.$$

From Theorem 2.3, it now follows that the game Γ is an exact potential game if and only if the one-form ω is exact in the sense that there exists a function $\Psi \colon \mathbf{A}^e \to \mathbb{R}$ such that

$$\omega = df = \sum_{(i,k)=(1,1)}^{(n,k_n)} \frac{\partial \Psi}{\partial a_{i,k}}(a)\, da_{i,k} \qquad \text{for every } a \in \mathbf{A}^e.$$

Furthermore, from Poincaré's Lemma we know that the one-form ω is exact on $\mathbf{A}^e$ if and only if it is closed in the sense that

$$\frac{\partial^2 \pi_i}{\partial a_{j,m}\, \partial a_{i,k}}(a) = \frac{\partial}{\partial a_{j,m}} \left(\frac{\partial \pi_i}{\partial a_{i,k}} \right)(a) = \frac{\partial}{\partial a_{i,k}} \left(\frac{\partial \pi_j}{\partial a_{j,m}} \right)(a) = \frac{\partial^2 \pi_j}{\partial a_{i,k}\, \partial a_{j,m}}(a)$$

for all $i, j \in N$, $k \in \{1, \ldots, k_i\}$ and $m \in \{1, \ldots, k_j\}$.
From this we conclude immediately the assertion of Theorem 2.4 by rewriting the second partial derivatives in matrix form to get $D^2_{i,j}\pi_i(a) = D^2_{j,i}\pi_j(a)$ for all $i, j \in N$.

2.6.4 Proof of Theorem 2.5

Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be some game.

First, assume that Γ has an interaction potential, in the sense that for every coalition of players $S \subseteq N$ there exists an interaction potential function $\Psi_S \colon \mathbf{A}_S \to \mathbb{R}$ such that for every $a \in \mathbf{A}$ and every player $i \in N$:

$$\pi_i(a) = \sum_{S \subseteq N \colon i \in S} \Psi_S(a_S).$$

Now, define for every $a \in \mathbf{A}$, $\Psi(a) = \sum_{S \subseteq N} \Psi_S(a_S)$. Then we derive for every $i \in N$, $a_i, b_i \in N$ and $\bar{a}_{-i} \in \mathbf{A}_{-i}$ that

$$\begin{aligned}
\Psi(a_i, \bar{a}_{-i}) - \Psi(b_i, \bar{a}_{-i}) &= \sum_{S \subseteq N} \Psi_S((a_i, \bar{a}_{-i})_S) - \sum_{S \subseteq N} \Psi_S((b_i, \bar{a}_{-i})_S) \\
&= \sum_{S \subseteq N \colon i \in S} \Psi_S((a_i, \bar{a}_{-i})_S) - \sum_{S \subseteq N \colon i \in S} \Psi_S((b_i, \bar{a}_{-i})_S) \\
&= \pi_i(a_i, \bar{a}_{-i}) - \pi_i(b_i, \bar{a}_{-i})
\end{aligned}$$

This shows that Γ is indeed an exact potential game for the defined function Ψ.

Conversely, suppose that the game Γ has an exact potential given by Ψ. Recall that $\delta_i(a_{-i}) = \pi_i(a) - \Psi(a)$ for every $a \in \mathbf{A}$ defines a dummy game for every $i \in N$. For every coalition $S \subseteq N$ we now introduce $\Psi_S \colon \mathbf{A}_S \to \mathbb{R}$ as the function defined by

$$\Psi_S(a_S) = \begin{cases} \Psi(a) + \sum_{i \in N} \delta_i(a_{-i}) & \text{if } S = N \\ -\delta_i(a_{-i}) & \text{for } S = N - i \text{ if some } i \in N \\ 0 & \text{if } |S| \leqslant n - 2 \end{cases}$$

where we recall that $n = |N|$.
Now, let $i \in N$ and $a \in \mathbf{A}$. Then we derive that

$$\begin{aligned}\sum_{S\subseteq N:\, i\in S} \Psi_S(a_S) &= \sum_{j\in N-i} \Psi_{N-j}(a_{N-j}) + \Psi_N(a)\\ &= -\sum_{j\in N-i} \delta_j(a_{-j}) + \Psi(a) + \sum_{h\in N} \delta_h(a_{-h})\\ &= \Psi(a) + \delta_i(a_{-i}) = \pi_i(a)\end{aligned}$$

We conclude the proof by checking whether the sum of the interaction potential functions Ψ_S is exactly the potential function Ψ. Indeed, for every $a \in \mathbf{A}$:

$$\begin{aligned}\sum_{S\subseteq N} \Psi_S(a_S) &= \sum_{j\in N} \Psi_{N-j}(a_{N-j}) + \Psi_N(a)\\ &= -\sum_{j\in N} \delta_j(a_{-j}) + \Psi(a) + \sum_{h\in N} \delta_h(a_{-h}) = \Psi(a)\end{aligned}$$

showing that the interaction potential construction is indeed equivalent to a game having an exact potential.

2.6.5 Proof of Theorem 2.6

To show the assertion of Theorem 2.6, we first show that if Γ is a Euclidean potential game, then property (2.12) holds.

Indeed, if Γ has an exact potential, it follows from Proposition 2.1 that for all players $i, j \in N$ and for all $a \in \mathbf{A}$: $\pi_i(a) - \delta_i(a_{-i}) = \pi_j(a) - \delta_j(a_{-j})$, or $\pi_i(a) - \pi_j(a) = \delta_i(a_{-i}) - \delta_j(a_{-j})$. We apply this to the lefthand side of (2.12). We build the lefthand side in two steps:

$$\begin{aligned}(\mathcal{I} - \mathcal{E}_j)(\pi_i - \pi_j, a) &= (\mathcal{I} - \mathcal{E}_j)(\delta_i - \delta_j, a\,)\\ &= \delta_i(a_{-i}) - \mathcal{E}_j(\delta_i, a) - \delta_j(a_{-j}) + \mathcal{E}_j(\delta_j, a)\\ &= \delta_i(a_{-i}) - \mathcal{E}_j(\delta_i, a) - \delta_j(a_{-j}) + \delta_j(a_{-j}) = \delta_i(a_{-i}) - \mathcal{E}_j(\delta_i, a)\end{aligned}$$

Hence, this further implies that

$$\begin{aligned}(\mathcal{I} - \mathcal{E}_i)(\mathcal{I} - \mathcal{E}_j)(\pi_i - \pi_j)(a) &= (\mathcal{I} - \mathcal{E}_i)(\delta_i(a_{-i}) - \mathcal{E}_j\delta_i(a_{-i})\,)\\ &= \delta_i(a_{-i}) - \mathcal{E}_j\delta_i(a_{-i}) - (\mathcal{E}_i\delta_i(a_{-i}) - \mathcal{E}_i\,\mathcal{E}_j\delta_i(a_{-i})\,)\\ &= \delta_i(a_{-i}) - \mathcal{E}_j\delta_i(a_{-i}) - (\delta_i(a_{-i}) - \mathcal{E}_j\delta_i(a_{-i})\,) = 0\end{aligned}$$

This shows that an exact potential game satisfies (2.12).

Next suppose a game Γ satisfies (2.12). First we note that the expectation operators $\{\mathcal{E}_i \mid i \in N\}$ are well-defined on the space of all measurable functions on all action sets of these players. These operators are commutative and, therefore, it holds that for all $i, j \in N$:

$$(\mathcal{I} - \mathcal{E}_i)\ (\mathcal{I} - \mathcal{E}_j) = \mathcal{I} - (\mathcal{E}_i + (\mathcal{I} - \mathcal{E}_i)\ \mathcal{E}_j) \tag{2.29}$$

By induction we arrive at the expression

$$\prod_{i\in N}(\mathcal{I}-\mathcal{E}_i)=\mathcal{I}-\left(\mathcal{E}_1+\sum_{j=2}^{n}\prod_{k=1}^{j-1}(\mathcal{I}-\mathcal{E}_k)\ \mathcal{E}_j\right)$$

Next, let $i\in N$. Then

$$\mathcal{I}=(\mathcal{I}-\mathcal{E}_i)+\mathcal{E}_i=\sum_{S\subseteq N\colon i\in S}\prod_{j\notin S}\prod_{k\in S}\mathcal{E}_j(\mathcal{I}-\mathcal{E}_k)+\mathcal{E}_i$$

Therefore,

$$\pi_i=\sum_{S\subseteq N\colon i\in S}\prod_{j\notin S}\prod_{k\in S}\mathcal{E}_j(\mathcal{I}-\mathcal{E}_k)\pi_i+\mathcal{E}_i\pi_i \tag{2.30}$$

Together with the above, this results in the conclusion that for all $S\subseteq N$ and all $i,h\in S$:

$$\begin{aligned}\prod_{j\notin S}\prod_{k\in S}\mathcal{E}_j(\mathcal{I}-\mathcal{E}_k)\pi_i&=\prod_{j\notin S}\prod_{k\in S\colon k\neq i,h}\mathcal{E}_j(\mathcal{I}-\mathcal{E}_k)(\mathcal{I}-\mathcal{E}_i)(\mathcal{I}-\mathcal{E}_h)\pi_i\\&=\prod_{j\notin S}\prod_{k\in S\colon k\neq i,h}\mathcal{E}_j(\mathcal{I}-\mathcal{E}_k)(\mathcal{I}-\mathcal{E}_i)(\mathcal{I}-\mathcal{E}_h)\pi_h\\&=\prod_{j\notin S}\prod_{k\in S}\mathcal{E}_j(\mathcal{I}-\mathcal{E}_k)\pi_h\end{aligned}$$

Introducing $\xi_S=\prod_{j\notin S}\prod_{k\in S}\mathcal{E}_j(\mathcal{I}-\mathcal{E}_k)\pi_i$ as the common factor in the expression above, we can now rewrite (2.30) as

$$\pi_i=\sum_{S\subseteq N\colon i\in S}\xi_S+\mathcal{E}_i\pi_i$$

Hence,

$$\pi_i=\sum_{S\subseteq N\colon S\neq\varnothing}\xi_S-\sum_{S\subseteq N\colon S\neq\varnothing,\ i\neq S}\xi_S+\mathcal{E}_i\pi_i$$

In this expression, $\sum_{S\subseteq N\colon S\neq\varnothing,\ i\neq S}\xi_S$ does not depend on a_i. Furthermore, $\mathcal{E}_i\pi_i$ does not depend on a_i either. This shows that we have decomposed π_i in line with Proposition 2.1, showing that $\Gamma=\langle N,\mathbf{A},\pi\rangle$ indeed has to be an exact potential game.

This shows the assertion of Theorem 2.6.

2.6.6 Proof of Theorem 2.7

Consider a smooth Euclidean game $\Gamma=\langle N,\mathbf{A},\pi\rangle$ with $A_i=[\underline{a}_i,\bar{a}_i]\subset\mathbb{R}$ and $\pi_i\in\mathcal{C}^2(\mathbf{A}^e)$ at least twice continuously differentiable on an open set $\mathbf{A}^e\subset\mathbb{R}^N$ with $\mathbf{A}\subset\mathbf{A}^e$.
We refer to Theorem 2.2 and the proof of Theorem 2.4 to determine that a smooth game has an exact potential if and only if the corresponding one-form $\omega=\sum_{i\in N}\frac{\partial\pi_i}{\partial a_i}\,da_i$ is exact. In that case there exists a smooth potential function $\Psi\colon\mathbf{A}\to\mathbb{R}$ such that $\omega=d\Psi$.
From the mathematical notes about the fundamental theorem of line integrals it is immediately determined that exactness of ω implies that for any parametric curve $\gamma\colon[0,1]\to\mathbf{A}$ it holds that

$$\oint_\gamma \omega = \oint_\gamma d\Psi = \Psi(\gamma(1)) - \Psi(\gamma(0)) \tag{2.31}$$

Consider some fixed action profile $a^0 \in \mathbf{A}$ and let $a \in \mathbf{A}$ be arbitrary. Now consider any path γ such that $\gamma(0) = a^0$ and $\gamma(1) = a$. Then (2.31) can be rewritten as

$$\Psi(a) = \Psi(\gamma(1)) = \oint_\gamma \omega + \Psi(\gamma(0)) = \oint_\gamma \sum_{i\in N} \frac{\partial \pi_i}{\partial a_i} da_i + \Psi(a^0) \tag{2.32}$$

Normalising the function Ψ such that $\Psi(a^0)$ and rewriting the line integral with the standard formulation we arrive at the desired formula:

$$\Psi(a) = \int_0^1 \sum_{i\in N} \left[\frac{\partial \pi_i}{\partial a_i}(\gamma(t))\, |\gamma_i'(t)| \right] dt$$

2.6.7 Proof of Proposition 2.2

Let $\Gamma = \langle N, \mathbf{A}, \pi\rangle$ be an effective Euclidean game such that $0 \in \mathbf{A} \subseteq \mathbb{R}^{\bar{k}}$.
Next, assume by contradiction that $\varphi(0, a) = 0$ for every action tuple $a \in \mathbf{A}$. First, we note that for every $i \in N$ from this by selecting $a(i) = (0, \dots, 0, a_i, 0 \dots, 0)$ that

$$\varphi(0, a(i)) = \pi_i(a(i)) - \pi_i(0) = 0 \qquad \text{or} \qquad \pi_i(a(i)) = \pi_i(0).$$

Now, take $b(n-1) = (0, \dots, 0, a_{n-1}, a_n) \in \mathbf{A}$ with $a_{n-1} \in A_{n-1}$ and $a_n \in A_n$. Then it follows from the hypothesis that

$$\varphi(0, b(n-1)) = \pi_{n-1}(0, \dots, 0, a_{n-1}, 0) - \pi_{n-1}(0) + \pi_n(b(n-1)) - \pi_n(0, \dots, 0, a_{n-1}, 0) = 0$$

We already have established that $\pi_{n-1}(0, \dots, 0, a_{n-1}, 0) = \pi_{n-1}(0)$. This, in turn, implies that

$$\varphi(0, b(n-1)) = \pi_n(b(n-1)) - \pi_n(0, \dots, a_{n-1}, 0) = 0 \qquad \text{or} \qquad \pi_n(b(n-1)) = \pi_n(0, \dots, 0, a_{n-1}, 0).$$

Reasoning in this fashion repeatedly for players $n-2$ through 1, we arrive at the conclusion that for any action tuple $a \in \mathbf{A}$ it holds that $\pi_n(a) = \pi_n(a_{-i}, 0)$. Hence, this would mean that player n is a dummy player in Γ. Thus, Γ is non-effective, contradicting the assumption of the assertion stated in Proposition 2.2.

2.6.8 Proof of Theorem 2.8

Let $\Gamma = \langle N, \mathbf{A}, \pi\rangle$ be an effective Euclidean game such that $0 \in \mathbf{A} \subseteq \mathbb{R}^{\bar{k}}$.
We first establish the relationship between the path potential and the exact potential function if Γ is an exact potential game. We can rewrite the definition of the path potential function to show that Γ is an exact potential game if and only if $\varphi(a, b) = \Psi(b) - \Psi(a)$ for all $a, b \in \mathbf{A}$. Indeed, we have that Γ has an exact potential Ψ if and only if

$$\begin{aligned}\varphi(a,b) &= \sum_{i=1}^{n} [\pi_i(b_1,\dots,b_i,a_{i+1},\dots,a_n) - \pi_i(b_1,\dots,b_{i-1},a_i,\dots,a_n)] \\ &= \sum_{i=1}^{n} [\Psi(b_1,\dots,b_i,a_{i+1},\dots,a_n) - \Psi(b_1,\dots,b_{i-1},a_i,\dots,a_n)] \\ &= \Psi(b) - \Psi(a)\end{aligned}$$

The above is based on the insight that the path potential value $\varphi(a,b)$ captures the total payoff difference summed over all players $i \in N$ that transition from a_i to b_i. This total payoff difference sum is the same as expressed through the exact potential function.
From the fact that $0 \in \mathbf{A}$, we derive now that

$$\varphi(a,0) = \Psi(0) - \Psi(a) \quad \text{as well as} \quad \varphi(0,a) = \Psi(a) - \Psi(0).$$

Normalisation of the potential function by selecting $\Psi(0) = 0$, implies now that $\Psi(a) = \varphi(0,a) = -\varphi(a,0)$.

Next, we show the stated characterisation of an exact potential game in the assertion. Using the above, we determined that Γ is an exact potential game if and only if $\varphi(a,b) = \Psi(b) - \Psi(a)$. Thus, we arrive at the conclusion that for all $a, b \in \mathbf{A}$:

$$\varphi(a,b) = \Psi(b) - \Psi(a) = \varphi(a,0) - \varphi(b,0) = \varphi(0,b) - \varphi(0,a).$$

Letting $a = b$ shows that $\varphi(a,a) = 0$ for all $a \in \mathbf{A}$. Thus,

$$\varphi(0,a) = \varphi(0,0) - \varphi(a,0) = -\varphi(a,0)$$

Hence, we have indeed shown that Γ is an exact potential game if and only if the path potential φ is decomposable in the sense that for all $a, b \in \mathbf{A}$: $\varphi(a,b) = \varphi(a,0) + \varphi(0,b)$ as well as $\varphi(a,0) = -\varphi(0,a)$.

This shows the assertion of Theorem 2.8.

2.6.9 Proof of Theorem 2.9

Theorem 2.9 asserts that every feasible flow on the response graph of a finite game decomposes uniquely into a potential and a harmonic component. The original proof by Candogan, Menache, Ozdaglar and Parrilo (2011) draws on the machinery of combinatorial Hodge theory on simplicial complexes, including the curl operator d_1, the Hodge Laplacian Δ_1, and the identification of two-cells in the response graph. This apparatus can be avoided: the decomposition admits a short, self-contained proof that uses only the definitions introduced in Section 2.4 together with two elementary facts from finite-dimensional linear algebra.

I briefly recall the notation established in Section 2.4 of the chapter. Consider a finite normal form game $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ with $N = \{1, \dots, n\}$ players, finite action sets A_i for each player $i \in N$, and action profile space $\mathbf{A} = \prod_{i \in N} A_i$. The set of unilateral deviations between action profiles is denoted by $\mathcal{E}$.

The *vertex space* $C_0 \cong \mathbb{R}^{|\mathbf{A}|}$ and the *edge space* $C_1 \cong \mathbb{R}^{|\mathcal{E}|}$ are endowed with the standard Euclidean inner product, written $\langle \cdot, \cdot \rangle_0$ and $\langle \cdot, \cdot \rangle_1$ respectively. The *gradient map* $d_0 \colon C_0 \to C_1$ is defined by

$$(d_0\Psi)(a,a') = \Psi(a') - \Psi(a)$$

for every oriented edge $(a,a') \in \mathcal{E}$. The *divergence map* $\partial_1 \colon C_1 \to C_0$ is the adjoint of d_0 with respect to the inner products on C_0 and C_1, characterised by

$$\langle X, d_0\Psi\rangle_1 = \langle \partial_1 X, \Psi\rangle_0 \tag{2.33}$$

for all $X \in C_1$ and all $\Psi \in C_0$. The *deviation map* $D \colon \mathcal{U} \to C_1$ sends a payoff profile $\pi \in \mathcal{U}$ to the flow $D\pi \in C_1$ given by

$$(D\pi)(a,a') = \pi_i(a') - \pi_i(a)$$

where $i \in N$ is the unique player whose action differs between a and a'. The *space of feasible flows* is the image $\operatorname{Im} D \subseteq C_1$ of the deviation map. A flow in $\operatorname{Im} d_0$ is called a *potential flow*; a flow in $\ker \partial_1$ is called a *harmonic flow*.

The proof rests on two simple observations. The first is a direct consequence of the adjoint relationship (2.33).

Lemma 2.5

The subspaces $\operatorname{Im} d_0$ *and* $\ker \partial_1$ *of* C_1 *are mutually orthogonal:*

$$\operatorname{Im} d_0 \perp \ker \partial_1 \qquad \text{in } C_1.$$

♡

Proof Let $Y \in \operatorname{Im} d_0$ and $X \in \ker \partial_1$. Since $Y \in \operatorname{Im} d_0$, there exists $\Psi \in C_0$ with $Y = d_0\Psi$. Applying the adjoint relationship (2.33):

$$\langle X, Y\rangle_1 = \langle X, d_0\Psi\rangle_1 = \langle \partial_1 X, \Psi\rangle_0 = \langle 0, \Psi\rangle_0 = 0$$

where the third equality uses $X \in \ker \partial_1$. Hence $X \perp Y$, as required. ¶

The second observation is that every potential flow is realised as the deviation flow of some game—specifically, a common-interest game in which all players share the same payoff function. This shows that the subspace of potential flows is contained in the subspace of feasible flows.

Lemma 2.6

Every potential flow is a feasible flow: $\operatorname{Im} d_0 \subseteq \operatorname{Im} D$.

♡

Proof Let $Y \in \operatorname{Im} d_0$ and choose $\Psi \in C_0$ such that $Y = d_0\Psi$. Define the payoff profile $\pi \in \mathcal{U}$ by

$$\pi_i(a) := \Psi(a) \qquad \text{for every player } i \in N \text{ and every action profile } a \in \mathbf{A}.$$

This payoff profile describes a game in which all players share the common payoff function Ψ. Its deviation flow at any unilateral deviation $(a,a') \in \mathcal{E}$—with deviating player $i \in N$—is given by

$$(D\pi)(a,a') = \pi_i(a') - \pi_i(a) = \Psi(a') - \Psi(a) = (d_0\Psi)(a,a') = Y(a,a').$$

Hence $D\pi = Y$, so that $Y \in \operatorname{Im} D$. Since $Y \in \operatorname{Im} d_0$ was arbitrary, we conclude that $\operatorname{Im} d_0 \subseteq \operatorname{Im} D$. ¶

The decomposition theorem With these two lemmas in hand, we can now state and prove the combinatorial Hodge decomposition of feasible flows. The argument combines Lemmas 2.5 and 2.6 with the elementary fact that, in a finite-dimensional Euclidean space, any element admits a unique orthogonal decomposition with respect to a subspace and its orthogonal complement.

I restate Theorem 2.9 for completeness:

Theorem 2.10 (Combinatorial Hodge decomposition of feasible flows)

Let $\Gamma = \langle N, \mathbf{A}, \pi \rangle$ be a finite normal form game. Every feasible flow $D\pi \in \operatorname{Im} D$ admits a unique decomposition

$$D\pi = D\pi_P + D\pi_H \tag{2.34}$$

with $D\pi_P \in \operatorname{Im} d_0$ and $D\pi_H \in \ker \partial_1$. The decomposition is orthogonal in C_1, and both components $D\pi_P$ and $D\pi_H$ are themselves feasible flows. ♡

Proof The proof proceeds in three steps: existence, uniqueness, and feasibility of the two components.

Step 1: Existence of the decomposition. The subspace $\operatorname{Im} d_0$ is a linear subspace of the finite-dimensional Euclidean space $(C_1, \langle\cdot,\cdot\rangle_1)$. The orthogonal complement theorem for finite-dimensional inner product spaces guarantees that every element $X \in C_1$ admits a unique orthogonal decomposition

$$X = X_P + X_H \qquad \text{with } X_P \in \operatorname{Im} d_0 \text{ and } X_H \in (\operatorname{Im} d_0)^{\perp}.$$

Here X_P is the orthogonal projection of X onto $\operatorname{Im} d_0$, and $X_H = X - X_P$ lies in the orthogonal complement $(\operatorname{Im} d_0)^{\perp}$. We claim that this orthogonal complement coincides with $\ker \partial_1$:

$$(\operatorname{Im} d_0)^{\perp} = \ker \partial_1. \tag{2.35}$$

The inclusion $\ker \partial_1 \subseteq (\operatorname{Im} d_0)^{\perp}$ is Lemma 2.5. For the converse inclusion, suppose $X \in (\operatorname{Im} d_0)^{\perp}$. Then for every $\Psi \in C_0$,

$$0 = \langle X, d_0\Psi \rangle_1 = \langle \partial_1 X, \Psi \rangle_0$$

by the adjoint relationship (2.33). Since this holds for every $\Psi \in C_0$, the element $\partial_1 X \in C_0$ is orthogonal to every vector in C_0, and therefore $\partial_1 X = 0$. Hence $X \in \ker \partial_1$, establishing (2.35).

Applying this decomposition to $X = D\pi$, we obtain $D\pi = D\pi_P + D\pi_H$ with $D\pi_P \in \operatorname{Im} d_0$ and $D\pi_H \in \ker \partial_1$. The decomposition is orthogonal in C_1 by construction.

Step 2: Uniqueness of the decomposition. Suppose $D\pi = X_P + X_H = X'_P + X'_H$ are two decompositions with $X_P, X'_P \in \operatorname{Im} d_0$ and $X_H, X'_H \in \ker \partial_1$. Rearranging,

$$X_P - X'_P = X'_H - X_H.$$

The left-hand side lies in $\operatorname{Im} d_0$; the right-hand side lies in $\ker \partial_1$. By Lemma 2.5, $\operatorname{Im} d_0 \cap \ker \partial_1 = \{0\}$—since any element in the intersection is orthogonal to itself and therefore zero. Hence $X_P = X'_P$ and $X_H = X'_H$, proving uniqueness.

Step 3: Both components are feasible flows. By Lemma 2.6, $\operatorname{Im} d_0 \subseteq \operatorname{Im} D$, so $D\pi_P \in \operatorname{Im} d_0 \subseteq \operatorname{Im} D$ is a feasible flow. Since both $D\pi \in \operatorname{Im} D$ and $D\pi_P \in \operatorname{Im} D$, and $\operatorname{Im} D$ is a linear subspace of C_1, the difference

$$D\pi_H = D\pi - D\pi_P \in \operatorname{Im} D$$

is also a feasible flow. This completes the proof of Theorem 2.9.

Remark The theorem allows us to define the orthogonal projections unambiguously: the linear map $e\colon C_1 \to C_1$ given by $X \mapsto X_P$, the projection onto $\operatorname{Im} d_0$, restricts to a well-defined map from $\operatorname{Im} D$ to $\operatorname{Im} d_0$ extracting the potential component of any feasible flow. The harmonic component is then $X - e(X) \in \ker \partial_1$.

The proof above does not require the introduction of higher-order simplicial structure—specifically, the curl operator $d_1\colon C_1 \to C_2$ defined on a two-cell extension of the response graph. This is a pedagogical simplification: the original argument of Candogan, Menache, Ozdaglar and Parrilo (2011) uses the full three-term combinatorial Hodge decomposition

$$C_1 = \operatorname{Im} d_0 \oplus \ker \Delta_1 \oplus \operatorname{Im} \partial_2$$

where $\Delta_1 = d_0\partial_1 + \partial_2 d_1$ is the Hodge Laplacian, to derive the same conclusion. The shortcut presented here works because we do not need to separately identify the "harmonic in the strict sense" subspace $\ker \Delta_1$ from the "divergence-free" subspace $\ker \partial_1$; for feasible flows, these identifications are not required to establish the unique decomposition. The two notions of harmonic flow coincide on $\operatorname{Im} D$, as explained in Section 2.4 of the chapter.

Bibliography

Arefizadeh, Sina, Angelia Nedić, and Gautam Dasarathy (2024). "On characterizations of potential and ordinal potential games". In: *arXiv* 2405.06253.

Balder, Erik (1997). "Remarks on Nash equilibria for games with additively coupled payoffs". In: *Economic Theory* 9, pp. 161–167.

Baye, Michael R, Dan Kovenock, and Casper G de Vries (1994). "The Solution of the Tolluck Rent Seeking game when $R > 2$: Mixed strategy equilibria and mean dissipation rates". In: *Public Choice* 81, pp. 363–380.

Bichler, Martin et al. (2025). "Characterizing the Convergence of Game Dynamics via Potentialness". In: *Transactions on Machine Learning Research* 2503.16285.

Biggar, Oliver and Iman Shames (2023). "The Graph Structure of Two-Player Games". In: *Scientific Reports* 13, p. 1833.

Bouchard, Vincent (2024). *Calculus IV online lecture notes*.

Buchanan, James M (1980). "Rent Seeking and Profit Seeking". In: *Toward a Theory of the Rent-Seeking Society*. Ed. by James M Buchanan, Robert D Tollison, and Gordon Tullock. Texas A & M University Press, pp. 3–15.

Candogan, Ozan, Ishai Menache, et al. (2011). "Flows and Decompositions of Games: Harmonic and Potential Games". In: *Mathematics of Operations Research* 36.3, pp. 474–503.

Candogan, Ozan, Asuman Ozdaglar, and Pablo A. Parrilo (2013a). "Dynamics in Near-Potential Games". In: *Games and Economic Behavior* 82, pp. 66–90. DOI: 10.1016/j.geb.2013.07.001.

— (May 2013b). "Near-Potential Games: Geometry and Dynamics". In: *ACM Transactions on Economics and Computation* 1.2, pp. 1–32. DOI: 10.1145/2465769.2465776.

Cao, Zhigang, Guopeng Li, and Feng Zhu (2025). "Potentials in quadratic Cournot cross-holding games". In: *Journal of Mathematical Economics* 119, p. 103135.

Cheng, Daizhan and Zhengping Ji (2021). "Weighted and Near Weighted Potential Games with Application to Game Theoretic Control". arXiv preprint arXiv:2106.12335.

Cheng, Daizhan, Ting Liu, et al. (2016). "On Decomposed Subspaces of Finite Games". In: *Automatica* 65, pp. 162–172.

Dubey, Pradeep, Ori Haimanko, and Andriy Zapechelnyuk (2006). "Strategic complements and substitutes, and potential games". In: *Games and Economic Behavior* 54, pp. 77–94.

Ederer, Florian and Bruno Pellegrino (2025). "A Tale of Two Networks: Common Ownership and Product Market Rivalry". In: *Review of Economic Studies* forthcoming.

Facchini, Giovanni et al. (1997). "Congestion Models and Weighted Bayesian Potential Games". In: *Theory and Decision* 42, pp. 193–206.

Gilles, Robert P (2010). *The Cooperative Game Theory of Networks and Hierarchies*. Theory and Decision Library. Berlin, Germany: Springer Verlag.

Golan, Jonathan S (1992). *Foundations of Linear Algebra*. Springer Verlag.

Hwang, Sung-Ha and Luc Rey-Bellet (2020). "Simple Characterizations of Potential Games and Zero-sum Equivalent Games". In: *Journal of Economic Theory and Econometrics* 31, pp. 1–13.

Jiang, Xiaoye et al. (2011). "Statistical Ranking and Combinatorial Hodge Theory". In: *Mathematical Programming* 127, pp. 203–244.

Marden, Jason R. (2012). "State Based Potential Games". In: *Automatica* 48, pp. 3075–3088.

Marden, Jason R., Gürdal Arslan, and Jeff S. Shamma (2009). "Cooperative Control and Potential Games". In: *SIAM Journal on Control and Optimization* 47.3, pp. 1393–1407. DOI: 10.1137/060676386.

Marden, Jason R. and Jeff S. Shamma (2015). "Game Theory and Distributed Control". In: *Handbook of Game Theory with Economic Applications*. Ed. by H. Peyton Young and Shmuel Zamir. Vol. 4. Elsevier, pp. 861–899. DOI: 10.1016/B978-0-444-53766-9.00016-1.

Monderer, Dov and Lloyd S Shapley (1996). "Potential Games". In: *Games and Economic Behavior* 14, pp. 124–143.

Munkres, James R (1984). *Elements of Algebraic Topology*. Cambridge, MA: Perseus Press.

Pellegrino, Bruno (2025). “Product Differentiation and Oligopoly: A Network Approach”. In: *American Economic Review* 115.4, pp. 1170–1225.

Rosenthal, Robert W. (1973). “A Class of Games Possessing Pure-Strategy Nash Equilibria”. In: *International Journal of Game Theory* 2.1, pp. 65–67. DOI: 10.1007/BF01737559.

Rudin, Walter (1976). *Principles of Mathematical Analysis*. Third. New York, NY: McGraw-Hill.

Sandholm, William H (2010). “Decompositions and potentials for normal form games”. In: *Games and Economic Behavior* 70.2, pp. 446–456.

Shapley, Lloyd S (1953). “A Value for n-Person Games”. In: *Contributions to the Theory of Games*. Ed. by R.D. Luce and A.W. Tucker. Vol. II. Princeton, NJ: Princeton University Press.

Tullock, Gordon (1980). “Efficient Rent Seeking”. In: *Toward a Theory of the Rent-Seeking Society*. Ed. by James M Buchanan, Robert D Tollison, and Gordon Tolluck. Texas A & M University Press, pp. 97–112.

Ui, Takashi (2000). “A Shapley Value Representation of Potential Games”. In: *Games and Economic Behavior* 31, pp. 121–135.

Voorneveld, Mark (1999). “Potential Games and Interactive Decisions with Multi-Criteria”. PhD thesis. Tilburg, The Netherlands: Tilburg University.

— (2000). “Best-Response Potential Games”. In: *Economics Letters* 66.3, pp. 289–295. DOI: 10.1016/S0165-1765(99)00196-2.

Wang, Yanfei, Xiangyu Hu, and Daizhan Cheng (2024). “On Potential Equations of Finite Symmetric Games”. In: *Asian Journal of Control*. Cheng-school characterisation of finite symmetric potential games via the semi-tensor product potential equation.

Wang, Yuanhua and Daizhan Cheng (2017). “From Weighted Potential Game to Weighted Harmonic Game”. In: *IET Control Theory and Applications*. Extends the Candogan-Menache-Ozdaglar-Parrilo decomposition to weighted finite games and introduces weighted harmonic games.