Lecture 09: Online Learning in Zero-Sum Games

Lecturer: Anas Barakat

October 14, 2025

Abstract

We introduce the fundamental class of two-player zero-sum games. We state and prove the celebrated min-max theorem in the case of finite normal-form games using the existence of no-regret learners. Then we show how to learn approximate Nash equilibria in two player-zero sum games via online learning. We conclude by a brief discussion on the limitations of the average-iterate convergence guarantee compared to last-iterate convergence.

1 Two Player Zero-Sum Games

Zero-sum games are two-player games where the payoff functions of the players sum up to zero. Therefore the interests of the players are conflicting and there is no cooperation. This is the archetype of the fully competitive setting in game theory.

Definition 1 (**Two-player zero-sum game**). *Consider a* 2*-player normal-form game*

$$\Gamma = (\mathcal{I} = \{1,2\}, \{\mathcal{A}_i\}_{i\in\{1,2\}}, \{u_i\}_{i\in\{1,2\}}) \,. \tag{1}$$

When $u_1 + u_2 = 0$*, the game is said to be zero-sum.*

We denote by $n = |\mathcal{A}_1|, m = |\mathcal{A}_2|$ the number of available strategies for Players 1 and 2, respectively.

In this case, it is enough to describe the payoffs of the game using the payoff function of the first player (which is the opposite of the payoff of the second player). The utility function can then be encoded using a single matrix M whose entries are defined as follows:

$$M_{i,j} := u_2(a_i, a_j) = -u_1(a_i, a_j), \quad \forall i \in [n], j \in [m], \quad \forall a_i \in \mathcal{A}_i, a_j \in \mathcal{A}_j \,, \tag{2}$$

i.e. $M_{i,j}$ is the payoff of player 1 (the row player) when player 1 chooses action $a_i \in \mathcal{A}_i$ and player 2 selects action $a_j \in \mathcal{A}_j$. We often also say for simplicity that player 1 selects action $i \in [n]$ and player 2 selects action $j \in [m]$. The goal of player 1 is to minimize the payoff function u_1 which depends on the actions of both players. The goal of player 2 is to maximize the utility u_1. Player 1 can only select their own strategy $x \in \Delta(\mathcal{A}_1)$ but not the strategy $y \in \Delta(\mathcal{A}_2)$ for player 2 and vice-versa for player 2.

Remark 2. Conversely, any real matrix induces as a finite zero-sum game. The terminology of *matrix game* is also commonly used.

The expected payoff of player 1 (row player) when player 1 chooses a mixed strategy $x \in \Delta(\mathcal{A}_1)$ and player 2 (column player) chooses a mixed strategy $y \in \Delta(\mathcal{A}_2)$ is given by:

$$u_1(x,y) = \sum_{a_i\in\mathcal{A}_1, a_j\in\mathcal{A}_2} x_{a_i} y_{a_j} u_1(a_i, a_j) = \sum_{a_i\in\mathcal{A}_1, a_j\in\mathcal{A}_2} x_{a_i} y_{a_j} M_{i,j} = x^\top M y \,. \tag{3}$$

Examples. See Matching Pennies, Rock-Paper-Scissors from last lectures. Beyond normal-form games, the class of zero-sum games (in continuous games) finds several applications in Machine Learning and beyond, and has attracted a lot of attention in the last few years. These include Generative Adversarial Networks (GANs, generator vs discriminator), robust ML and adversarial training (min-max optimization, e.g. classifier or ML model vs adversary) security games (attackers vs defenders), self-play (agent vs a copy of itself, i.e. using the same algorithm) and many others.

Definition 3 (Minmax and maxmin). *The maxmin of a finite 2-player game* Γ *is defined by:*

$$\underline{v} := \max_{y \in \Delta(A_2)} \min_{x \in \Delta(A_1)} x^\top M y \,. \tag{4}$$

The minmax of the game is defined by:

$$\bar{v} := \min_{x \in \Delta(A_1)} \max_{y \in \Delta(A_2)} x^\top M y \,. \tag{5}$$

Suppose that player 2 selects their strategy $a_j \in \mathcal{A}_j$ first and then player 1 selects their strategy $a_i \in \mathcal{A}_i$ knowing the strategy $a_j \in \mathcal{A}_j$ of player 2. The the maxmin is the best payoff that player 2 can guarantee for themselves in the worst case (i.e. if they move first, and hence player 1 best responds by selecting the min of their payoff knowing the strategy of player 2). Similarly, the minmax is the is the best payoff that player 1 can guarantee for themselves in the worst case (asuming that player 1 moves first and player 2 selects their strategy given the strategy of player 1).

The next lemma shows that the latter situation is more favorable for the second player than the first one.

Lemma 4. *The maxmin is smaller than the minmax of the game, i.e.* $\underline{v} \leq \bar{v}$ *.*

The first mover seems to face a disadvantage in pure antagonism.

Proof. Using the definitions of min and max, we can write:

$$\begin{aligned}
\min_{x \in \Delta(\mathcal{A}_1)} x^\top M y' &\leq x'^\top M y' \leq \max_{y \in \Delta(\mathcal{A}_2)} x'^\top M y, \quad \forall x' \in \Delta(\mathcal{A}_1), y' \in \Delta(\mathcal{A}_2)\,, \\
\min_{x \in \Delta(\mathcal{A}_1)} x^\top M y' &\leq \min_{x \in \Delta(\mathcal{A}_1)} \max_{y \in \Delta(\mathcal{A}_2)} x^\top M y\,, \quad \forall y' \in \Delta(\mathcal{A}_2)\,, \\
\max_{y \in \Delta(\mathcal{A}_2)} \min_{x \in \Delta(\mathcal{A}_1)} x^\top M y &\leq \min_{x \in \Delta(\mathcal{A}_1)} \max_{y \in \Delta(\mathcal{A}_2)} x^\top M y\,.
\end{aligned} \tag{6}$$

□

The difference $\bar{v} - \underline{v}$ is called the duality gap.

Definition 5 (Value of a game). *The game* Γ *has a value if* $\underline{v} = \bar{v}$*.*

The value of a game denoted by v refers to the rational outcome of the interaction between the two rational players. Intuitively, the value v is the fair price of the contest: If both players are perfectly rational, they can guarantee this same number and no one can force a better or worse outcome.

For the minimizing player, v is the loss they can ensure never to exceed, no matter what the opponent does. For the maximizing player, v is the gain they can ensure to obtain, no matter what the opponent does.

2 Learning NE in Zero-Sum Games

Recall first the definition of Nash Equilibria (NE) in our 2-player setting. No unilateral strategy deviation for any player is profitable.

Definition 6 ((Mixed) Nash Equilibrium). *A strategy profile* $(x^\star, y^\star) \in \Delta(\mathcal{A}_1) \times \Delta(\mathcal{A}_2)$ *is a NE of the two-player normal-form game* Γ *defined in* (1) *if:*

$$\forall x \in \Delta(\mathcal{A}_1), \quad x^\top (-A) y^\star \leq (x^\star)^\top (-A) y^\star\,, \tag{7}$$

$$\forall y \in \Delta(\mathcal{A}_2), \quad (x^\star)^\top A y \leq (x^\star)^\top A y^\star\,. \tag{8}$$

One of the first and most important results in game theory is the minmax theorem which was proved by the John von Neumann in 1928. The theorem states that neither player benefits from deviating: each has secured the best guaranteed outcome against all possible counter-moves.

Theorem 7 (Von-Neumann's Minmax Theorem (Neumann, 1928)). *Let Γ be a 2-player zero-sum normal-form game as defined in Definition 1 with payoff matrix M. Then we have:*

$$\min_{x\in\Delta(A_1)}\max_{y\in\Delta(A_2)} x^\top My = \max_{y\in\Delta(A_2)}\min_{x\in\Delta(A_1)} x^\top My. \tag{9}$$

Extensions of this theorem to continuous settings (e.g. with convex compact sets and a convex-concave continuous function) has been established by Sion (1958) (see also the references therein). For a detailed discussion for this more general setting, see e.g. (Orabona, 2019, Chapter 12).

Proof. There exist different proofs for this theorem. A classical proof is based on duality in linear programming in finite dimensions (maximization of a linear form under linear constraints), see e.g. section 2.3, chapter 2 in Laraki et al. (2019). In the following, given our focus on online learning, we rather present a different constructive proof based on the existence of a no-regret learner, provided in Freund and Schapire (1996).

We prove the two different inequalities to obtain the equality of the minmax theorem. The first inequality has been proved in Lemma 4. We now show the remaining inequality. We consider a repeated game setting where a no-regret minimizer selects a mixed strategy $x_t \in \Delta(\mathcal{A}_1)$ and we assume that the strategy $y_t \in \Delta(\mathcal{A}_2)$ is chosen by the environment as a best response to x_t, i.e.

$$y_t \in \operatorname*{argmax}_{y\in\Delta(\mathcal{A}_2)} x_t^\top My\,. \tag{10}$$

The loss function observed by the regret minimizer at time t is the linear function: u^t defined for all $x \in \Delta(\mathcal{A}_1)$ by $u^t(x) = x^\top My_t$. We also define the average strategies:

$$\bar{x}_T := \frac{1}{T}\sum_{t=1}^T x_t \in \Delta(\mathcal{A}_1)\,, \quad \bar{y}_T := \frac{1}{T}\sum_{t=1}^T y_t \in \Delta(\mathcal{A}_2)\,. \tag{11}$$

We now have the following inequality:

$$\min_{x\in\Delta(A_1)}\max_{y\in\Delta(A_2)} x^\top My \le \max_{y\in\Delta(A_2)} \bar{x}_T^\top My = \frac{1}{T}\max_{y\in\Delta(A_2)}\sum_{t=1}^T x_t^\top My \le \frac{1}{T}\sum_{t=1}^T x_t^\top My_t\,, \tag{12}$$

where the last inequality follows fron the fact that y_t is a best response to x_t for any time t.

Recall now the definition of regret:

$$\mathrm{Reg}_x^T := \sum_{t=1}^T x_t^\top My_t - \min_{x\in\Delta(A_1)}\sum_{t=1}^T x^\top My_t\,, \tag{13}$$

which implies by rearranging and dividing by T that

$$\frac{1}{T}\sum_{t=1}^T x_t^\top My_t = \min_{x\in\Delta(A_1)} x^\top M\bar{y}_T + \frac{\mathrm{Reg}_x^T}{T}\,. \tag{14}$$

Using the above identity in (12) yields:

$$\min_{x\in\Delta(A_1)}\max_{y\in\Delta(A_2)} x^\top My \le \min_{x\in\Delta(A_1)} x^\top M\bar{y}_T + \frac{\mathrm{Reg}_x^T}{T} \le \max_{y\in\Delta(A_2)}\min_{x\in\Delta(A_1)} x^\top My + \frac{\mathrm{Reg}_x^T}{T}\,. \tag{15}$$

Taking the limit when $T \to \infty$ in the last inequality above yields:

$$\min_{x\in\Delta(A_1)} \max_{y\in\Delta(A_2)} x^\top My \leq \max_{y\in\Delta(A_2)} \min_{x\in\Delta(A_1)} x^\top My\,, \tag{16}$$

which concludes the proof. □

Corollary 8. *A strategy profile $(x^\star, y^\star) \in \Delta(\mathcal{A}_1) \times \Delta(\mathcal{A}_2)$ is a NE of the two-player zero-sum normal-form game Γ defined in (1) if and only if $(x^\star, y^\star)$ is a max-min strategy, i.e. if and only if*

$$x^\star \in \operatorname*{argmin}_{x\in\Delta(\mathcal{A}_1)} \max_{y\in\Delta(\mathcal{A}_2)} x^\top My\,, \quad y^\star \in \operatorname*{argmax}_{y\in\Delta(\mathcal{A}_2)} \min_{x\in\Delta(\mathcal{A}_1)} x^\top My\,. \tag{17}$$

Proof. ($\Longleftarrow$) Suppose $(x^\star, y^\star)$ is a max-min strategy. Then we have

$$\begin{aligned}\forall x \in \Delta(\mathcal{A}_1), x^\top My^\star \geq \min_{x\in\Delta(\mathcal{A}_1)} x^\top Ay^\star = \max_{y\in\Delta(\mathcal{A}_2)} \min_{x\in\Delta(\mathcal{A}_1)} x^\top Ay^\star = \min_{x\in\Delta(\mathcal{A}_1)} \max_{y\in\Delta(\mathcal{A}_2)} x^\top My \\ = \max_{y\in\Delta(\mathcal{A}_2)} (x^\star)^\top My \geq (x^\star)^\top My^\star\,.\end{aligned} \tag{18}$$

Similarly, for the second player, we have

$$\begin{aligned}\forall y \in \Delta(\mathcal{A}_2), (x^\star)^\top My \leq \max_{y\in\Delta(\mathcal{A}_2)} (x^\star)^\top My = \min_{x\in\Delta(\mathcal{A}_1)} \max_{y\in\Delta(\mathcal{A}_2)} x^\top My = \max_{y\in\Delta(\mathcal{A}_2)} \min_{x\in\Delta(\mathcal{A}_1)} x^\top My \\ = \min_{x\in\Delta(\mathcal{A}_1)} x^\top My^\star \geq (x^\star)^\top My^\star\,.\end{aligned} \tag{19}$$

We have proved that $(x^\star, y^\star)$ is a NE.

($\Longrightarrow$) Suppose now that $(x^\star, y^\star)$ is a NE. By definition, we get:

$$\max_{y\in\Delta(\mathcal{A}_2)} (x^\star)^\top My \leq (x^\star)^\top My^\star \leq \min_{x\in\Delta(\mathcal{A}_1)} x^\top My^\star\,. \tag{20}$$

This inequality implies that:

$$\min_{x\in\Delta(\mathcal{A}_1)} \max_{y\in\Delta(\mathcal{A}_2)} x^\top My \leq \max_{y\in\Delta(\mathcal{A}_2)} (x^\star)^\top My \leq (x^\star)^\top My^\star \leq \min_{x\in\Delta(\mathcal{A}_1)} x^\top My^\star \leq \max_{y\in\Delta(\mathcal{A}_2)} \min_{x\in\Delta(\mathcal{A}_1)} x^\top My\,. \tag{21}$$

However, by the minmax theorem, $\min_{x\in\Delta(\mathcal{A}_1)} \max_{y\in\Delta(\mathcal{A}_2)} x^\top My = \max_{y\in\Delta(\mathcal{A}_2)} \min_{x\in\Delta(\mathcal{A}_1)} x^\top My$. This means that all the inequalities in (21) are actually equalities. Hence,

$$\min_{x\in\Delta(\mathcal{A}_1)} \max_{y\in\Delta(\mathcal{A}_2)} x^\top My = \max_{y\in\Delta(\mathcal{A}_2)} (x^\star)^\top My \leq (x^\star)^\top My^\star\,, \quad \min_{x\in\Delta(\mathcal{A}_1)} x^\top My^\star = \max_{y\in\Delta(\mathcal{A}_2)} \min_{x\in\Delta(\mathcal{A}_1)} x^\top My\,, \tag{22}$$

which means that $(x^\star, y^\star)$ is a max-min strategy in the sense of (17) and this concludes the proof. □

Definition 9 (Duality gap function). *The duality gap function $\Delta : \Delta(\mathcal{A}_1) \times \Delta(\mathcal{A}_2) \to \mathbb{R}$ is defined for every $(x', y') \in \Delta(\mathcal{A}_1) \times \Delta(\mathcal{A}_2)$ by:*

$$\Delta(x', y') := \max_{y\in\Delta(\mathcal{A}_2)} x'^\top My - \min_{x\in\Delta(\mathcal{A}_1)} x^\top My'\,. \tag{23}$$

Lemma 10 (Nonnegativity of duality function). *For all $(x', y') \in \Delta(\mathcal{A}_1) \times \Delta(\mathcal{A}_2), \Delta(x', y') \geq 0$.*

Proof. For any $(x', y') \in \Delta(\mathcal{A}_1) \times \Delta(\mathcal{A}_2), \Delta(x', y')$,

$$\max_{y\in\Delta(\mathcal{A}_2)} x'^\top My \geq x'^\top My' \geq \min_{x\in\Delta(\mathcal{A}_1)} x^\top My'\,. \tag{24}$$

□

Proposition 11 (NE characterization via duality gap). *A strategy profile $(x^\star, y^\star) \in \Delta(\mathcal{A}_1) \times \Delta(\mathcal{A}_2)$ is a NE of the two-player zero-sum normal-form game Γ defined in (1) if and only if $\Delta(x^\star, y^\star) = 0$.*

Proof. The proof follows from the following equivalences:

$$\begin{aligned}(x^\star, y^\star) \text{ is a NE} &\iff \min_{x\in\Delta(A_1)} x^\top M y^\star \geq (x^\star)^\top M y^\star \geq \max_{y\in\Delta(A_2)} (x^\star)^\top M y \\ &\iff \min_{x\in\Delta(A_1)} x^\top M y^\star = \max_{y\in\Delta(A_2)} (x^\star)^\top M y \\ &\iff \Delta(x^\star, y^\star) = 0\,, \end{aligned} \tag{25}$$

where the second equivalence follows from using the fact that we always have the inequality $\min_{x\in\Delta(A_1)} x^\top M y^\star \leq \max_{y\in\Delta(A_2)} (x^\star)^\top M y$ and using the definitions of min and max . □

Proposition 12 (Approximate NE and duality gap). *For any desired accuracy $\varepsilon \geq 0$,*

(i) if $\Delta(x^\star, y^\star) \leq \varepsilon$ then $(x^\star, y^\star)$ is an ε-NE.

(ii) if $(x^\star, y^\star)$ is an ε-NE, then $\Delta(x^\star, y^\star) \leq 2\varepsilon$.

Proof. The proof of this result is left as an exercise. □

We consider now a repeated game setting similarly to the previous lecture. Recall the definitions of regret for both players:

$$\mathrm{Reg}_x^T := \sum_{t=1}^T x_t^\top M y_t - \min_{x\in\Delta(A_1)} \sum_{t=1}^T x^\top M y_t\,, \tag{26}$$

$$\mathrm{Reg}_y^T := \max_{y\in\Delta(A_2)} \sum_{t=1}^T x_t^\top M y - \sum_{t=1}^T x_t^\top M y_t\,. \tag{27}$$

The duality gap function and the regrets of both players are connected by the following result.

Proposition 13. *Let $\bar{x}_T, \bar{y}_T$ be the average of the sequence of iterates $(x_t), (y_t)$ respectively, i.e. $\bar{x}_T = \frac{1}{T}\sum_{t=1}^T x_t, \bar{y}_T = \frac{1}{T}\sum_{t=1}^T y_t$. Then we have the following link between the duality function evaluated at the average iterates and the sum of regrets of players in a 2-player zero-sum game:*

$$\Delta(\bar{x}_T, \bar{y}_T) = \frac{Reg_x^T + Reg_y^T}{T}\,. \tag{28}$$

Proof. The proof is immediate from summing up (26)-(27) and using the definitions of the average iterates $\bar{x}_T, \bar{y}_T$. □

We are now ready to state our NE learning result. Regret minimizing dynamics converge in a time average sense to NE.

Corollary 14. *If the two sequences $(x_t), (y_t)$ are generated using two no-regret algorithms (not necessarily the same) satisfying $Reg_x^T = \mathcal{O}(\sqrt{T})$ and $Reg_y^T = \mathcal{O}(\sqrt{T})$ in the two-player zero-sum game Γ, then*

$$\Delta(\bar{x}_T, \bar{y}_T) = \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)\,. \tag{29}$$

Hence if $T = \mathcal{O}(\varepsilon^{-2})$, then $(\bar{x}_T, \bar{y}_T)$ is an ε-NE.

Proof. Use Proposition 13 for the first statement and then Proposition 12 for the second. □

Faster rates? Observe that improved convergence rates can be obtained if regret bounds can be improved. Note that the above result uses a no-regret algorithm in the case where the losses can be adversarial, which is a worst-case. In our setting, we deal with a game setting in which the loss function is not arbitrary and adversarial but can rather be *predictable* if both players use a no-regret algorithm. We will discuss the idea of optimism in lecture 11 which allows to get improved regret rates and hence faster convergence to approximate NE.

Average vs Last-iterate convergence. Note that the above result shows a convergence result of the *average iterates* rather than the *last iterates* (x_T, y_T). In general these two modes of convergence are distinct. It could be that last iterates are not converging whereas the average iterates are. We provide a simple example below in a simple continuous game setting.

Example 15 (Continuous bilinear zero-sum game). Consider the bilinear zero-sum game:

$$\min_{x\in[0,1]} \max_{y\in[0,1]} xy\,. \tag{30}$$

It can be verified that this game has a unique NE $(0,0)$. Consider now the gradient descent algorithm as a no-regret learner used by both players:

$$x_{t+1} = x_t - \eta y_t \tag{31}$$
$$y_{t+1} = y_t + \eta x_t\,, \tag{32}$$

where $\eta > 0$ is a step size. It is straightforward to check that:

$$\|(x_{t+1}, y_{t+1}) - (0,0)\|_2^2 = (1+\eta^2)\|(x_t, y_t) - (0,0)\|_2^2\,. \tag{33}$$

Since $1+\eta^2 > 1$, the above identity that the last iterates do not converge to the unique NE. However, the average iterates do converge using an extension of our results to continuous 2-player zero-sum games.

This question of average vs last-iterate convergence is an active research topic (Hsieh et al., 2019; Golowich et al., 2020; Lin et al., 2020; Cai et al., 2022, 2024, 2025b,a). For zero-sum normal-form games, similar non-convergence behavior has been observed for no-regret learning algorithms such as the Multiplicative Weights Update algorithm (MWU), see e.g. Bailey and Piliouras (2018). Last iterate convergence can be achieved using modified algorithms based on optimism, see e.g. Daskalakis and Panageas (2018, 2019) for asymptotic last-iterate convergence and Wei et al. (2021) for last-iterate convergence rates. We will discuss some of these ideas in lecture 11.

3 Next Lecture

We will discuss alternative (relaxed) equilibrium concepts in general-sum games beyond NE, namely (coarse) correlated equlibria and we will show how we can approximate them using no-regret learning.

References

James P Bailey and Georgios Piliouras. Multiplicative weights update in zero-sum games. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 321–338, 2018.

Yang Cai, Argyris Oikonomou, and Weiqiang Zheng. Finite-time last-iterate convergence for learning in multi-player games. *Advances in Neural Information Processing Systems*, 35: 33904–33919, 2022.

Yang Cai, Gabriele Farina, Julien Grand-Clément, Christian Kroer, Chung-Wei Lee, Haipeng Luo, and Weiqiang Zheng. Fast last-iterate convergence of learning in games requires forgetful algorithms. *Advances in Neural Information Processing Systems*, 37:23406–23434, 2024.

Yang Cai, Gabriele Farina, Julien Grand-Clément, Christian Kroer, Chung-Wei Lee, Haipeng Luo, and Weiqiang Zheng. On separation between best-iterate, random-iterate, and last-iterate convergence of learning in games. *arXiv preprint arXiv:2503.02825*, 2025a.

Yang Cai, Haipeng Luo, Chen-Yu Wei, and Weiqiang Zheng. From average-iterate to last-iterate convergence in games: A reduction and its applications. *arXiv preprint arXiv:2506.03464*, 2025b.

Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

Constantinos Daskalakis and Ioannis Panageas. The limit points of (optimistic) gradient descent in min-max optimization. *Advances in neural information processing systems*, 31, 2018.

Constantinos Daskalakis and Ioannis Panageas. Last-iterate convergence: Zero-sum games and constrained min-max optimization. In *Innovations in Theoretical Computer Science*, 2019.

Yoav Freund and Robert E Schapire. Game theory, on-line prediction and boosting. In *Proceedings of the ninth annual conference on Computational learning theory*, pages 325–332, 1996.

Noah Golowich, Sarath Pattathil, and Constantinos Daskalakis. Tight last-iterate convergence rates for no-regret learning in multi-player games. *Advances in neural information processing systems*, 33:20766–20778, 2020.

Yu-Guan Hsieh, Franck Iutzeler, Jérôme Malick, and Panayotis Mertikopoulos. On the convergence of single-call stochastic extra-gradient methods. *Advances in Neural Information Processing Systems*, 32, 2019.

Rida Laraki, Jérôme Renault, and Sylvain Sorin. *Mathematical foundations of game theory*, volume 202. Springer, 2019.

Tianyi Lin, Zhengyuan Zhou, Panayotis Mertikopoulos, and Michael Jordan. Finite-time last-iterate convergence for multi-agent learning in games. In *International Conference on Machine Learning*, pages 6161–6171. PMLR, 2020.

J. von Neumann. Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 100:295–320, 1928. URL http://eudml.org/doc/159291.

Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*, 2019.

Maurice Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8(1):171 – 176, 1958.

Chen-Yu Wei, Chung-Wei Lee, Mengxiao Zhang, and Haipeng Luo. Linear last-iterate convergence in constrained saddle-point optimization. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=dx11_7vm5_r.