

STOCHASTIC PERTURBATIONS IN GAME THEORY
AND APPLICATIONS TO NETWORKS

PANAYOTIS MERTIKOPOULOS



Department of Physics
National & Kapodistrian University of Athens

A thesis submitted for the degree of
Doctor of Philosophy

Athens, June 2010



*"But is mathematics not the sister, as well as the servant, of the arts?
And does she not share the same touch of genius and melancholy madness?"*

— Harold Marston Morse, topologist

*“The happiest moments of my life have been those few
which I have passed at home, in the bosom of my family.”*

— Thomas Jefferson

This thesis can be dedicated to none other but my beloved family, the warm home that instilled in me the love of knowledge.

STOCHASTIC PERTURBATIONS IN GAME THEORY
AND APPLICATIONS TO NETWORKS

Department of Physics
National & Kapodistrian University of Athens

This dissertation is accepted in its present form by the thesis committee of
Panayotis Mertikopoulos as satisfying the thesis requirement for the degree of
Doctor of Philosophy.

THE THESIS COMMITTEE

ARIS L. MOUSTAKAS (Chair)

Date

ANDREAS POLYDOROS

Date

DIMITRIS FRANTZESKAKIS

Date

ABSTRACT

As far as complex organisms go, the Internet is a textbook specimen: its billions of man-made components might be relatively simple to understand on an individual basis, but the complexity of their interactions far outstrips the deductive capabilities of their creators. As a result, a promising way to analyze these interactions is provided by game theory, a discipline devoted to the (often chimeric) task of making sense of an even more obscure and complex entity, human rationality.

Nevertheless, this inherent goal of game theory is also the source of a fundamental controversy: is it really possible to apply our understanding of human interactions and behavioral patterns to networks of machines, devoid as they are of the feelings and irrational drives that characterize the former?

The way out of this impasse is to recall that game theory encompasses the study of “rationality” in its most abstract form, even when confronted with entities that have little or no capacity for rational thought (e.g. animals, bacteria, or, in our case, computers). This line of reasoning has led to the emergence of *evolutionary game theory* (EGT), an offshoot of game theory which is concerned with large populations of barely sentient species that interact with each other in competitive habitats that promote natural selection. As such, evolutionary methods appear particularly attractive as a means to understand the inner workings of communication networks: after all, it is a rare occasion when Internet users do not compete with one another for the resources of the network they live in.

Perhaps the most well-studied model of population evolution in this context is that of the *replicator dynamics*, a dynamical system first introduced by Taylor and Jonker (1978) to study the interactions of different phenotypes within biological species. These studies quickly attracted the interest of game theorists and, after about a decade, eventually culminated in the “folk theorem” of EGT, a theorem which links evolution and rationality by showing that what appears as the consequence of rational thought is, in fact, the byproduct of natural selection favoring the survival of the “fittest”.

On the other hand, if rationality can be construed as the outcome of an evolutionary process, then, by providing appropriate selection criteria, evolution can also be steered to any state which is efficient with respect to these criteria. In this way, the replicator dynamics take on a different guise, that of a *learning mechanism* which the users of a network might employ in order to reach a “socially efficient” steady state; put differently, by manipulating the incentives of the “players”, the designers of the “game” can guide them to whatever state suits their purpose.

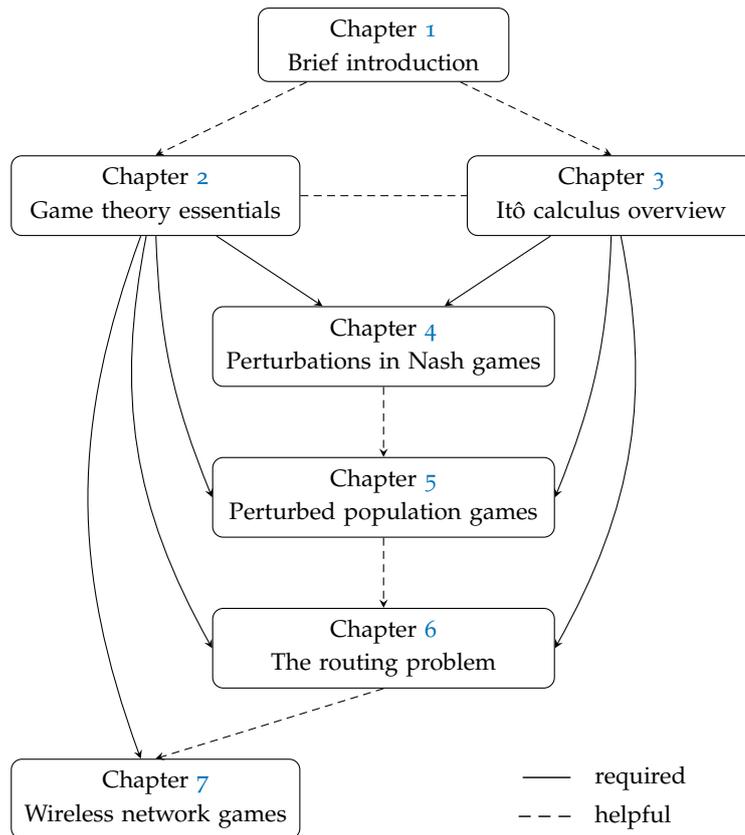
The main issue that this dissertation seeks to address is what happens if, in addition to the interactions between the players of a game (e.g. the users of a network), the situation is exacerbated by the interference of an assortment of exogenous and unpredictable factors, commonly referred to as “nature”. We find that this random interference differentiates crucially between the evolutionary and learning approaches, leading to different (stochastic) versions of the replicator dynamics.

Rather surprisingly, in the case of learning, we find that many aspects of rationality remain unperturbed by the effects of noise: regardless of the

fluctuations' magnitude, players are still able to identify suboptimal actions, something which is not always possible in the evolutionary setting. Even more to the point, the "strict (Nash) equilibria" of the game (an important class of steady states) turns out to be stochastically stable and attracting, again irrespective of the noise level.

From the viewpoint of network theory (where stochastic perturbations are, quite literally, omnipresent), the importance of these results is that they guarantee the *robustness* of the replicator dynamics against the advent of noise. In this way, if the users of a stochastically fluctuating network adhere to a replicator learning scheme and are patient enough, we show that the flow of traffic in the network converges to an invariant (stationary) distribution which is sharply concentrated in a neighborhood of the network's equilibrium point.

DIAGRAMMATIC OUTLINE OF THE THESIS



PUBLICATIONS

Some of the ideas presented in this thesis have appeared previously in the following publications:

1. P. Mertikopoulos and A. L. Moustakas: "Balancing traffic in networks: redundancy, learning and the effect of stochastic fluctuations", *submitted*. URL: <http://arxiv.org/abs/0912.4012>.
2. P. Mertikopoulos and A. L. Moustakas: "The emergence of rational behavior in the presence of stochastic perturbations", in *The Annals of Applied Probability*, vol. 20, no. 4, July 2010. URL: <http://arxiv.org/abs/0906.2094>.
3. P. Kazakopoulos, P. Mertikopoulos, A. L. Moustakas and G. Caire: "Living at the edge: a large deviations approach to the outage MIMO capacity", to appear in the *IEEE Transactions on Information Theory*. Available online: <http://arxiv.org/abs/0907.5024>.
4. P. Kazakopoulos, P. Mertikopoulos, A. L. Moustakas and G. Caire: "Distribution of MIMO mutual information: a large deviations approach", in *ITW '09: Proceedings of the 2009 IEEE Workshop on Networking and Information Theory*.
5. P. Mertikopoulos and A. L. Moustakas: "Learning in the presence of noise", in *GameNets '09: Proceedings of the 1st International Conference on Game Theory for Networks*, May 2009.
6. P. Mertikopoulos and A. L. Moustakas: "Correlated anarchy in overlapping wireless networks", in *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, special issue on the applications of game theory to communication networks, September 2008. URL: <http://arxiv.org/abs/0805.0963>.
7. P. Mertikopoulos, N. Dimitriou and A. L. Moustakas: "Vertical handover between service providers", in *WiOpt '08: Proceedings of the 6th International Symposium on Modelling and Optimization of Wireless Networks*, April 2008.
8. N. Dimitriou, P. Mertikopoulos and A. L. Moustakas: "Vertical handover between wireless standards", in *ICC'08: Proceedings of the 2008 IEEE International Conference on Communications*, May 2008.
9. P. Mertikopoulos and A. L. Moustakas: "The Simplex Game: can selfish users learn to operate efficiently in wireless networks?", in *GameComm '07: Proceedings of the 1st International Workshop on Game Theory for Communication Networks*, October 2007.

*“Science, my lad, is made up of mistakes,
but they are mistakes which it is useful to make,
because they lead little by little to the truth.”*

— Jules Verne, *A Journey to the Center of the Earth*

ACKNOWLEDGEMENTS

Towards the end of “*The Lord of the Rings*” trilogy, and after having spent the most perilous part of a three-book journey keeping his master Frodo out of harm’s way, Sam Gamgee was confronted in the safety of the Shire by an angry Jolly Cotton who asked him why he is not “looking out for Frodo now that things are looking dangerous”. To that question, Sam surmised that he can either give a two-week answer or none at all, and, being pressed for time, chose to give the latter.

In writing this acknowledgements section, I found myself in a similar dilemma because, for a long time, I had contemplated writing only two words:

“Ari, thanks!”

The reason that I chose not to take Sam’s approach and let two words carry the weight of many has more to do with my advisor’s character and modesty than with anything else: I doubt that he would find it fair or fitting to be favored in this way over the large number of people who have helped shape this thesis and its author. Nevertheless, this thesis would never have been written if it weren’t for him, so the credit is his, whether he ever claims it or not.

The journey that I have Aris to thank for is a long one and I can only hope that it does not end with this thesis. Without a doubt, the going has not always been pleasant, but he was always there to carve a path where I failed. His door was always open, and he opened it as if I were an equal, not a student; he gave me the freedom to pursue what I wanted and was patient enough to see my choices through, even the failed ones (and there were many); whenever I despaired with a proof or a calculation, Aris was always there to lend a helping hand and a word of support; and in my meandering monologues of technical details, he would always identify the essence and the correct path to proceed. But, above all else, Aris has been a friend, a wise friend who gave advice carefully but freely, and on whose experience and kindness I could always draw when needed.

For all these reasons I am truly indebted to Aris and my most sincere wish is that I will be able to pass on this privilege to my students by being for them the advisor and mentor that Aris was for me. So, *Ari, thanks!*

There are, of course, many others who have had a profound impact on this thesis. First off, I was extremely lucky and honored to have Professors Andreas Polydoros and Dimitris Frantzeskakis in my thesis committee. They are both wonderful people and I am deeply indebted to them for their kind support at a time when it had become doubtful whether this thesis would ever see the light of day. For that and for the warm smile that they greeted me with, they have my most sincere thanks and gratitude.

This is also the place to thank my undergraduate mentors at the University of Athens: Christos Ktorides for encouraging me to study mathematics at Brown and for supporting me at every step of the way, ever since we first met

in his quantum mechanics course; Alekos Karanikas for showing me how not to give up on calculations when the going gets rough and for being one of the kindest persons I have ever met; Haris Apostolatos for being a role model of clarity of thought; Professor Yannis Stavrakakis from the CS department for his help and understanding; and, last but not least, my undergraduate advisor, Petros Ioannou, for teaching me that audacity sometimes serves science just as well as humility.

I also cannot thank enough my officemate, Pavlos Kazakopoulos, for putting up with me for a bit more than three years. The whiteboard in our office always had a “Problem of the Week” scribbled at a corner somewhere, and these problems led to endless hours of excitement and scintillating discussions on mathematics, physics, and life in general. Sofia Zarbouti provided invaluable support with the countless bureaucratic minutiae that came up during my years in the University of Athens; without her support, I doubt that I could have made it through the administrative mazes of the University of Athens.

Indeed, the University of Athens is not always a very friendly place. Even though it is my official *alma mater* (and, in fact, doubly so since I did my undergraduate studies there as well), I often found it to be a cold and uncaring place, stifling its many excellent teachers and its students with concerns and worries that do not belong in the realm of science or education. Without my friends and teachers in the Physics Department to support me, I would have been consumed in a pointless rebellion against a sterile status quo that I would not have been able to change; for their kindness and endurance in this often hostile environment, I thank them all from the bottom of my heart.

On the other hand (and setting official titles aside), I can hardly imagine a better place to do research than what I consider to be my true “nourishing mother”, Brown University. I studied mathematics there for three years, and I owe every bit of my mathematical expertise and taste for mathematical beauty to these years and to my teachers there. The Mathematics Department was a small place, with teachers and students alike living like a family under the same roof, doing mathematics simply for the joy of it. Though I do not remember the harsh New England winters with any fondness, Kassar House will always occupy one of the warmest spots in my heart.

The list of the people at Brown that I would like to thank seems inexhaustible, but first and foremost is my advisor there, George Daskalopoulos. I still remember our first conversation over the phone when I was considering graduate schools in the States: we talked for a few minutes in Greek and then the conversation drifted to his research interests, Teichmüller spaces and Higgs bundles, topics that I knew nothing of at the time, but which took on a magical sheen when I heard them explained to me. When we hung up, I immediately decided that Brown was the ideal place for me, and, while there, George helped me immensely: he was never bothered by my interminable list of questions and he guided me patiently until I had found a thesis topic to my liking. Unfortunately, my research interests drifted away from differential geometry after a while and I never completed my thesis, but I will always remember fondly the myriad afternoons that we spent discussing generalized complex geometry and Hyperkähler quotients.

I have a similar debt of gratitude to my other teachers at Brown: to Tom Goodwillie and the always smiling Jeff Brock for their topology courses; to Nikos Kapouleas and Justin Corvino for teaching me Riemannian geometry; and to Sergei Treil for teaching me analysis the old-school Russian way.

To my friends at Brown, I owe a different kind of debt, one of companionship and of making the winters in Providence bearable. Jonathan was *always* eager to “talk some math” at the department, whether it was 1 AM on a Thursday night, or over a game of table tennis; Kate, Michelle, Graeme, Henning and Mike were always around as well, and I cannot imagine how life in Providence would have been if it hadn’t been for this creative and downright *fun* bunch of people to share it with. I miss you all!

There is also the Greeks and the Turks to thank: Dimitris the Lesbian, my roommate Socrates (with or without his hood) and Yannis (when not “zeroing out” everything) were my best friends in Providence, and our 3000 mile road-trip in the Canyonlands and the Pacific Southwest is my fondest memory of the States. Then there is (pauvre) Babis, Vassilis, Nikos, Olga, Ioanna G. and Menia; Duygu the social butterfly, Murat and Lüfo, Nathalie, Ebru and Avram. Thank you all!

Aurélie, merci pour tout et surtout pour ton hospitalité à Lyon; tu me manques toujours.

Last but (definitely) not least of my friends from Brown, there is Daniel. There is really no point in recounting here the times that we spent together at Brown, whether we were stooped over a math book trying to work out an estimate for a parabolic PDE, or having a couple of beers at the GCB. I cannot think of a better friend than Daniel, and his lucky arrival at Brown in my third (and hardest) year there was a true blessing; without Daniel I wouldn’t have been able to sort out my feelings on leaving, and that would have made a hard year a good deal worse. For all that, I thank you dearly my good friend.

The decision to leave Brown was, without a doubt, the hardest one that I have made in my entire life. Of course, this is hardly the place to get into the reasons behind this decision, except to mention that one of the most important ones had to do with the friends that I had made in Greece and who had remained an ocean and several time zones away. Athina and Alexandros were at Imperial at the time, and my London trip and their continued company during all these years kept me from losing my nerve. Along with Marios, Ilias and Lenja, these are the people outside my family that have been in my life the longest, and my life is all the richer for it.

Without Rudolf Boris, my first true teacher of mathematical analysis, I might never have decided to embark upon a life of mathematical research. It was he that ignited in me the spark of mathematics and he has fuelled it ever since. I hope he enjoys this thesis and that it is fitting tribute to his quality as a teacher and as a human being.

Tristan Tomala of HEC, Marco Scarsini of LUISS, Rida Laraki of the École Polytechnique and the rest of the audience of the “*Séminaire Parisien de Théorie des Jeux*” provided invaluable assistance in shaping some of the results of this thesis with their insightful comments and remarks – I would never have spotted a subtle mistake in an equilibrium characterization had it not been for an after-seminar discussion with Marco. Tristan’s interest in particular was especially touching right from the beginning, and I hope that I will be able to live up to his faith and trust in me.

Captain Nikos Soulias, long a friend of my parents’, was the equivalent of family when I was in the States. Whenever he visited New York, he would travel 180 miles just to come to Providence and take me out to lunch or dinner at the Capital Grille and to see how I was doing. I will never forget that, and I only hope that I will be able to one day repay his kindness and care.

Tonia, my love, has spent countless hours encouraging me when it seemed that this thesis would never finish. She can always bring a smile to my face (even when there is nothing to smile about), so, along with a big hug, I also give her this thesis; it is as much the product of her toils as it is of mine.

Finally, in trying to find the words to thank my parents, Vlassis and Kallia, my sister, Victoria, and the rest of my family for their unconditional love and support, I find myself in the same spot as King Lear's daughter, Cordelia: "Unhappy that I am, I cannot heave my heart into my mouth. I love [my family] according to my bond; no more nor less.". So, with all my love and gratitude for arousing in me the passion for the pursuit of knowledge, this thesis is for them.

FINANCIAL SUPPORT In terms of financial assistance, this thesis was supported by the European Union research grants Net-ReFound (EU-IST-FP6-034413) and URANUS (EU-IST-2006-2790). I would also like to take this opportunity to gratefully acknowledge the generous scholarship and support that I received for the duration of my studies from the Empirikeion Foundation of Athens, Greece.

CONTENTS

1	INTRODUCTION	1
1.1	The Natural Habitat of Game Theory	2
1.2	Evolution, Learning, and the Replicator Dynamics	4
1.3	The Effect of Stochastic Fluctuations	6
1.4	Thesis Outline and Overview of Results	8
	Notational Conventions	10
I	THEORETICAL PRELIMINARIES	13
2	A BRIEF EXCURSION IN GAME THEORY	15
2.1	Games in Normal Form	15
2.1.1	N -person Games	17
2.1.2	Populations and Evolutionary Games	17
2.1.3	Potential Games	20
2.1.4	Congestion Games	21
2.2	Solution Concepts in Game Theory	23
2.2.1	Dominated Strategies	23
2.2.2	Nash Equilibrium	24
2.2.3	Justifying Equilibrial Play: Correlation and Evolution	26
2.3	Evolution, Learning, and the Replicator Dynamics	33
2.3.1	The Replicator Dynamics	33
2.3.2	Entropy and Rationality	36
3	ELEMENTS OF STOCHASTIC ANALYSIS	39
3.1	Brownian Motion	40
3.2	Stochastic Integration	43
3.3	Itô Calculus and Stochastic Differential Equations	46
3.4	Diffusions and their Generators	49
II	STOCHASTIC FLUCTUATIONS IN NON-COOPERATIVE GAMES	53
4	STOCHASTIC PERTURBATIONS IN NASH GAMES	55
4.1	Exponential Learning	56
4.2	Learning in the Presence of Noise	59
4.3	Extinction of Dominated Strategies	62
4.4	Congestion Games: A Suggestive Digression	66
4.5	Stability of Equilibrial Play	70
5	STOCHASTIC PERTURBATIONS IN POPULATION GAMES	75
5.1	Deterministic Evolution in Population Games	76
5.1.1	Exponential Learning and Rational Behavior	77
5.1.2	Evolution in Potential Games	81
5.2	The Effect of Stochastic Fluctuations	88
III	APPLICATIONS TO NETWORKS	95
6	THE ROUTING PROBLEM	97
6.1	Networks and Congestion Models	100
6.1.1	Networks and Flows	100
6.1.2	Congestion Models and Equilibrium	102
6.2	Learning, Evolution and Rational Behavior	106

6.2.1	Learning and the Replicator Dynamics	106
6.2.2	Entropy and Rationality	108
6.3	Stochastically Perturbed Congestion Models	112
6.3.1	The Stochastic Replicator Dynamics	113
6.3.2	Stochastic Fluctuations and Rationality	115
6.4	Some Loose Ends	123
7	APPLICATIONS TO WIRELESS NETWORKS	127
7.1	The Simplex Game	129
7.2	Selfishness and Efficiency	134
7.2.1	Notions of Equilibrium	134
7.2.2	Harvesting the Equilibria	135
7.2.3	Anarchy and Efficiency	136
7.3	Iteration and Evolution	137
7.3.1	Evolutionary Dynamics	137
7.3.2	The Steady States	139
7.4	Measuring the Price of Anarchy	139
7.5	A Short Discussion	144
8	CONCLUSIONS	147
8.1	Some Open Questions	148
IV	APPENDIX	151
A	REPLICA-THEORETIC CALCULATIONS	153
A.1	Irrelevance of Node Strength and Reduction of Choices	153
A.2	An Analytic Expression for the Price of Anarchy	154
	BIBLIOGRAPHY	159

LIST OF FIGURES

Frontispiece	“ <i>Melencolia I</i> ”, by Albrecht Dürer (engraving, 1514).	iii
Figure 2.1	Rock-Paper-Scissors in strategic form	17
Figure 5.1	An attracting state which is not evolutionarily stable	81
Figure 5.2	The quadratic estimates of Lemma 5.7.	85
Figure 6.1	Reducible and irreducible networks.	102
Figure 6.2	Braess’s paradox.	106
Figure 6.3	The various sets in the proof of Theorem 6.12.	111
Figure 6.4	Learning in reducible and irreducible networks.	120
Figure 7.1	Convergence of a simplex game to equilibrium.	140
Figure 7.2	Independence on Node Strength.	143
Figure 7.3	Choice, sophistication, and the price of anarchy.	144
Figure 7.4	The (ir)relevance of asymptotics.	145

LIST OF TABLES

Table 2.1	Some game-theoretic prefixes and adjectives.	22
-----------	--	----

ACRONYMS

BNN	Brown-von Neumann-Nash dynamics
CDF	cumulative distribution function
CEQ	correlated equilibrium
EGT	evolutionary game theory
ESS	evolutionarily stable strategy
IEEE	Institute of Electrical and Electronics Engineers
KBE	Kolmogorov backward equation
KFE	Kolmogorov forward equation
MIMO	multiple-input, multiple-output
NEQ	Nash equilibrium
PDF	probability density function

RSB	replica symmetry breaking
SDE	stochastic differential equation
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
WEQ	Wardrop equilibrium
WLAN	Wireless Local Area Network

INTRODUCTION

WITHOUT a doubt, the Internet is a beast, and a web-spinning one at that. As of the latest census of December 2009, the number of actual human users of the Internet has reached the hallmark number of 1.8 billion, just over a fourth of the Earth's current population. In their turn, as these human users simultaneously surf the Web, check their email and share their files over RapidShare, they give rise to vast hordes of virtual "application users" whose number is downright impossible to estimate. Accordingly, to serve the demands issued by these users, the Internet has been spread over approximately 27,000 autonomous systems (effectively stand-alone local subnetworks), which are managed by more than 150,000 routers operating in the massive 100 Gbps range.¹ And still, there are times when this gargantuan infrastructure simply grounds to a halt, unable to process the billions of service requests that compete with one another in this grand scale.

So, is there any hope to tame this formidable web of users and their typically conflicting demands? On the one hand, given that our email clients and web browsers really *do* work (most of the time, at least), one might be tempted to say that we have already done so. On the other hand, as Costis Daskalakis points out in his celebrated thesis (Daskalakis, 2008), this does not mean that we really understand the inner workings of the Internet: if a web page fails to load, do we really have the means to backtrack this failure all the way to its true source (which might be a protocol failure that occurred at the other end of the globe)? And, even more importantly, do we have any sort of predictive power over such failures?

The short answer to both these questions is "no", largely owing to the immense and multi-layered complexity of the Internet: it would take an Internet-sized supercomputer simply to monitor its operation, let alone understand it. However, if we turn this last proposition on its head, we see that the Internet itself may be used to understand the Internet: just like molecules in a gas, the driving dynamics of each user are relatively simple to understand, so even though one cannot foresee the actions of a particular user, the laws of statistics may be used to infer the general flow of future events.

This is precisely the seminal idea that has brought game theory to the forefront of theoretical investigations into the nature of the Internet – and, in so doing, supplied renewed applicational impetus to an otherwise unrelated branch of applied mathematics. By shedding light on the user-user interactions at a local scale (or, conversely, by *controlling* them), game theory might well provide us with the tools needed to analyze and steer the evolution of the Internet.

¹ These Internet usage statistics are only estimates of the true numbers. The most accurate sources for this kind of information are the US census bureau (<http://www.census.gov>) and the Internet World Stats consortium (<http://www.internetworldstats.com/stats.htm>).

1.1 THE NATURAL HABITAT OF GAME THEORY

The purpose of Game Theory,

Game theory, like its fictional counterpart, psychohistory, has a very ambitious goal: to explain and predict the behavior of selfishly minded individuals who interact with each other and with their environment. But, whereas Asimov's pen lent the brilliance of its wielder to the psychohistorians of the Foundation universe, psychology, cognitive science, or even philosophy might seem better suited to tackle the difficult task of understanding human rationality.

its applications,

To be sure, trying to quantify the aspects of human behavior with the mathematical tools currently at the disposal of game theory is a fool's errand, threatening to take away much of the discipline's credibility. For instance, consider the (over-advertised but, ultimately, ill-understood) Prisoner's Dilemma, proposed by Merrill Flood while working at RAND (Flood, 1951): two outlaws are apprehended by the authorities and, because of insufficient evidence for a conviction, each is offered the same, secret deal. The prisoner who gives up incriminating evidence on his partner goes free, while said partner gets sentenced to 10 years in prison. If both betray each other, they will both go to jail for 5 years, and if they both evoke their Miranda right to remain silent, they will only serve 6 months on a minor offense.

What is the prisoners' way out of this quandary? Contrary to what might be expected, game-theoretic analysis suggests that the (Nash) "equilibrium" of this dilemma is for both prisoners to betray each other – to "defect" in the parlance of game theorists. One justification behind this solution concept is that the gains of defecting dominate the gains of remaining silent. In other words, given the choice of one's opponent (in fact, independently of it), one is better off to defect than not to: being set free is preferable to serving 6 months, and 5 years in prison are preferable to 10.

and the inherent difficulties.

Needless to say, this solution is highly debatable – after all, is it not better to aim for the common good and serve a measly 6 months in jail than to risk a 5-year sentence? Well, is it, really? There are those who would selflessly take the risk of trusting their fellow outlaw in the pursuit of a mutually beneficial outcome, and there are those who would selfishly jump at the opportunity of being set free without a second thought.² However, what if the penalty for mutual silence is not 6 months but only 6 days? Or, on the contrary, what if a conviction carries the death penalty instead?

Albeit reasonable, questions such as these invariably take us away from science and plunge us deep into the realm of speculation, best left to the literary masters of our time. As long as vast areas of the human decision process remain uncharted, there is no mathematical way to quantify "rational" human responses to a given stimulus, even when the stimulus itself is readily quantifiable (which, in the Prisoner's Dilemma, it is not). Consequently, the explanation (or, more to the point, the prediction) of human behavior by game theoretic means does not seem to be a realistic prospect at this time.

Subjectivity and refinements.

There are two ways to circumvent this impasse. One is to decrease the complexity of the strategic scenaria considered, so as to at least be able to make verifiable predictions regarding human players in simple decision problems. For example, an idea that was eagerly pursued in the 1960's was to incorporate the effect that the (subjective) beliefs of an individual have on his decision processes. But, even at a rudimentary level, this led to an interminable cascade of subjective beliefs about another person's subjective beliefs, their beliefs

² Interestingly enough, in a psychological experiment mimicking the conditions of the Prisoner's Dilemma, over 40% of the participants chose to remain silent (Tversky, 2004).

about those beliefs, and so on ad nauseam. This discussion prompted such magnificent set-theoretic structures as Harsányi's "tower of beliefs" (also known as a "belief hierarchy"; see Harsányi, 1967-1968) but, like its biblical predecessor in Babylon, the tower's original objective was not met. Belief hierarchies comprise a huge step forward in our understanding of Bayesian rationality – the seminal work of Aumann (1974) on subjectivity and correlation being another – but the vast majority of the intricacies of human rationality still elude us (and will probably keep doing so for some time to come).

Other attempts to justify the predictions of game theory by refining the concept of an equilibrium were not as successful. As Ken Binmore reproachingly points out, "various bells and whistles were appended to the idea of rationality", each one an objective reflection of its creator's subjective interpretation of what constitutes rational behavior. Thus, as one would expect, this rather disordered approach quickly led to the downright embarrassing situation where "almost any Nash equilibrium could be justified in terms of one or other's refinement".³

The other way around the mathematical intractability of the human condition is more radical. Indeed, as was already mentioned, the goal that game theorists set out to achieve is to understand the behavior of selfish individuals that interact with one another. But is the notion of a "selfish individual" truly synonymous to that of a "selfish human"? Surely, "selfishness" is not a plague exclusive to mankind: it can be found in the majestic sequoias that race to break through the canopy of their own groves in their quest for sunlight; a more savage display is when the new alpha lion attacks the cubs of the previous alpha in order to safeguard his standing; and even in the strictly regulated realm of technology, there are few entities as avaricious as a computer, constantly vying for more resources in order to complete its appointed tasks. So, instead of examining the behavior of humans (perhaps the most complex entities in the known cosmos), the idea is to begin the study of rationality more humbly, by looking at individuals whose motives and responses can be quantified in an objective manner.

This idea was first put forth by Maynard Smith and Price (1973), the founding fathers of *evolutionary game theory*. As above, their fundamental tenet was disarmingly simple: they just focused on species whose stimulus-response characteristics can be fully determined by their genetic programming (e.g. the growth-reproduction properties of bacteria). In this way, no longer thwarted by the complexity of the high-level cognition functions of humans, Maynard Smith and Price paved the way for a multitude of concrete applications of game theory in population biology. To wit, whereas the retreat of the weaker animal from a contest of strength before deadly injuries have been incurred had previously been attributed to an altruistic attitude serving the "good of the species", it was shown that this is, in fact, the result of animals playing out an (inherently selfish) *evolutionarily stable strategy* (ESS).

Rather ironically, by taking this premise to its logical extreme, we see that a most fertile playground for evolutionary games of this sort is one where biological organisms are altogether absent: the realm of computer networks. Indeed, barring malfunctions and similar (relatively rare) errors, a computer will react to a given stimulus precisely as it was programmed to react; hence, computer interactions will be governed by a set of very specific rules and

Turning to simpler entities.

Evolutionary Game Theory

and network applications.

³ In Binmore's foreword to Weibull (1995, p. ix). I cannot help but wonder whether this sentiment is shared by those physicists and mathematicians working in string theory, where almost every string vacuum is sufficiently diverse to explain any phenomenon observed at lower energies.

protocols which allow no room for doubt or uncertainty. Therefore, given that the mindless (but also selfish) purpose behind the programming of every computer is to carry out its appointed task heedless of its surroundings (anything not fed to it as input), we see that such networks constitute the perfect subjects of game-theoretic considerations.⁴

1.2 EVOLUTION, LEARNING, AND THE REPLICATOR DYNAMICS

Evolution, Network applications of evolutionary game theory are characterized by a certain duality. First off, just as in population biology, one aspect is to understand the evolution of a network by examining the selfish interactions of its components and by determining those states which are “stable” with respect to these interactions. As such, this direction is usually described by the term “evolution” and it has been at the forefront of game-theoretic investigations ever since the seminal work of [Maynard Smith and Price \(1973\)](#) – see also the masterful surveys by [Weibull \(1995\)](#) and by [Hofbauer and Sigmund \(1998, 2003\)](#) for more information.

learning, Then again, there is also the flip side of this coin, summed up in the following question: if we want a system of interacting entities (e.g. a network) to operate in some “optimal” state, can we somehow educate the competing entities so that their selfish interests will eventually lead them to this state? Quite aptly, this aspect of evolutionary game theory is referred to as “learning” and it has been studied just as extensively – the book by [Fudenberg and Levine \(1998\)](#) is an excellent starting point, while the article by [Hart \(2005\)](#) provides a more recent account.

In both questions, there are two cardinal components. The first one is to specify the actual “strategic game” being played, i.e. the specific framework which will be used to model and quantify the interactions that occur, defining such things as the players’ actions (e.g. to transmit or to remain silent) and the gains that correspond to them. The second one consists of the selection mechanism by which the players rate their choices and their reactions – for instance, by imitating the most successful players or by learning to avoid their past mistakes.⁵ Clearly, the first of these components (the choice of game) depends on the particulars of the application considered, so it is best left to be addressed on an individual basis. However, the natural selection process that forms the core of the latter component can be studied independently of the former, by examining the convergence and stability properties of the evolutionary (or learning) mechanism with respect to the “equilibria” (or any other solution concept) of the underlying game.

and the replicator dynamics. One of the most widely used models for evolution and learning in this context is that of the *replicator dynamics*, a dynamical system that was first introduced by [Taylor and Jonker \(1978\)](#) to model large populations which interact with one another by means of random matchings in some contest of reproductive fitness (such as sparring duels between deer). In a nutshell, these dynamics arise as the byproduct of an “imitation of the fittest” mechanism which reflects the outcome of these matches; more precisely, the per capita growth rate of a phenotype (i.e. the number of offspring in the unit of time) is

⁴ For a comprehensive survey, see [Altman et al. \(2006\)](#) and references therein.

⁵ We see here that the “simple organisms” maxim of evolutionary game theory again turns out to be of paramount importance: modelling human interactions is already a Sisyphean undertaking by itself; trying to delve even deeper into the human thought process by proposing a mathematical model for human learning certainly bears little ties to reality.

assumed proportional to the rewards gained therein by the individuals of the phenotype in question.⁶

So, if $z_{i\alpha}$ denotes the population of the α -th phenotype of species i and $u_{i\alpha}$ denotes its reproductive fitness (as determined by the state of the various species and the evolutionary game in question), the evolutionary postulate that we just described gives:

$$\frac{dz_{i\alpha}}{dt} = z_{i\alpha} u_{i\alpha}. \quad (1.1)$$

Thus, if $z_i = \sum_{\beta} z_{i\beta}$ is the total population of the i -th species and $x_{i\alpha} = z_{i\alpha}/z_i$ denotes the relative population share of phenotype α , a simple application of the chain rule yields the *replicator equation*:

$$\frac{dx_{i\alpha}}{dt} = x_{i\alpha} (u_{i\alpha} - u_i), \quad (1.2)$$

where u_i denotes the population average of the i -th species: $u_i = \sum_{\beta} x_{i\beta} u_{i\beta}$.

Now, given that this evolutionary approach concerns rapidly evolving species and populations which are rarely encountered outside population biology, the “learning” value of (1.2) is not immediately apparent. However, if the index α instead refers to one of the possible actions of some individual i (say, which path to choose in a network) and $u_{i\alpha}$ represents the gains associated to this action, then (1.2) can be viewed as a learning process, simply by reinterpreting $x_{i\alpha}$ as the probability with which player i employs his α -th action.⁷ In this manner, the replicator dynamics also represent an adaptive learning scheme with the payoff differences $u_{i\alpha} - u_i$ measuring the propensity of player i choosing action α .

A learning interpretation.

There is an alternative “learning” interpretation of the replicator dynamics which, while not as straightforward, will be equally important for our purposes. To see how it works, assume that every player of the game keeps a cumulative *score* of his actions by means of the differential equation:

$$dU_{i\alpha} = u_{i\alpha} dt. \quad (1.3)$$

Clearly, the higher the score $U_{i\alpha}$ of an action α , the better this action will have performed over time for player i . Taking note of this fact, the players update the probabilities $x_{i\alpha}$ with which they pick their actions according to the Boltzmann distribution

Exponential learning.

$$x_{i\alpha} = e^{U_{i\alpha}} / \sum_{\beta} e^{U_{i\beta}} \quad (1.4)$$

which, after a simple differentiation, again yields the replicator dynamics (1.2). This learning scheme is called “exponential learning”⁸ and, together with the direct learning interpretation of (1.2), it shows that the replicator dynamics emerge as the common denominator of both evolution and learning.

⁶ We should also mention here the important distinction between an organism’s “genotype” (the full hereditary information encoded in its genetic makeup) and its “phenotype”, i.e. the whole of its observed properties, such as its morphology and behavioral patterns (Johanssen, 1903; Lewontin, 1974). In the theory of evolutionary synthesis, natural selection is based on an organism’s phenotype (which determines the interaction of the organism with other organisms), while heredity is determined by its genotype. Unfortunately, the epigenetic process which maps genotypes to phenotypes is not very well understood, so we will take the biometric approach of population genetics and restrict ourselves to the phenotype space.

⁷ Of course, this presumes that players’ actions are drawn from a finite set and that their “rewards” are a function of the probability with which they pick an action; for more details, see Chapter 2.

⁸ Or “logit” learning, though “logistic” would be a better fit – see Chapter 4 for more details.

Of course, it should be noted here that the replicator dynamics are not the only evolutionary dynamics considered in connection with game theory.⁹ However, they do possess three unique qualities which have ever drawn mathematicians like moths to the proverbial flame: *a*) they are simple and straightforward to state;¹⁰ *b*) they are naturally derived from first principles; and *c*) they have this intangible trait which mathematicians call “elegance” and which quickly becomes apparent to any one who spends some time investigating their properties. For all these reasons, evolution will be, for us, synonymous to evolution under the replicator dynamics, and questions about evolutionary justification will always be phrased in terms of (1.2).¹¹

*Rationality
and evolution.*

Accordingly, these dynamics provide an excellent way to attack some of the most fundamental questions of evolutionary game theory: for instance, do ill-adapted (suboptimal) phenotypes become extinct over time? Are the “equilibria” predicted by game theory steady states of (1.2)? And, assuming they are, are they also stable? Are they attracting? Conversely, can evolution ever lead a species to a state which is game-theoretically “irrational”? In short: *what are the rationality properties of the replicator dynamics (1.2)?*

A revealing, if incomplete, list of answers to these questions is this: yes, ill-adapted phenotypes (“dominated strategies” in game-theoretic jargon) do become extinct, the only surviving phenotypes being the “rationally admissible” ones (Samuelson and Zhang, 1992); game-theoretic equilibria are indeed steady states of (1.2), but they are not always attracting, nor are they always stable (Weibull, 1995); in fact, they are not even the *only* steady states of (1.2), but if a steady state is evolutionarily attracting, then it is necessarily an equilibrium of the game (Nachbar, 1990).

*The Folk
Theorem.*

These properties are usually grouped together under the so-called *folk theorem* of evolutionary game theory, a term coined to reflect both a mildly vague statement (due to certain details being omitted when they depend on the class of games considered) and a rather muddled origin.¹² Nevertheless, despite its somewhat nebulous nature, the folk theorem shows that game-theoretic predictions are fully justified from an evolutionary/learning perspective: dubious as the postulates of rationality might appear when applied to entities with higher-level cognitive functions, they emerge triumphant when genetic (or ordinary) programming purges the particulars of individuality.

1.3 THE EFFECT OF STOCHASTIC FLUCTUATIONS

*Nature's
interference.*

Admittedly, the folk theorem rests on an implicit (but crucial) assumption: that the payoffs $u_{i\alpha}$ which measure the reproductive fitness of a specific phenotype depend on the state of the various interacting species alone, i.e. that they are independent of external environmental conditions. Unfortunately, this requirement is not always met: nature always has a say on reproductive matters, and there are many births and deaths which occur as the result of unpredictable external factors (commonly referred to as the “weather”, even though the actual climate might have little relevance to the specific problem).

⁹ A notable alternative is the Brown-von Neumann-Nash dynamics (Fudenberg and Levine, 1998).

¹⁰ In a certain sense, we will see that they are the *simplest* possible evolutionary dynamics.

¹¹ The fact that the Lotka-Volterra predator-prey model and the Price equation can both be recast in terms of the replicator dynamics further reinforces the prominent position of the latter – see also the unifying expositions by Nowak and Page (2002), Cressman (2003), and Hofbauer and Sigmund (2003).

¹² The first appearance of the theorem is often traced back to Friedman (1971), but Weibull (1995) gives a much more detailed account in terms of actual content.

To model random disturbances of this kind, [Fudenberg and Harris \(1992\)](#) proposed a version of the replicator dynamics where the phenotype populations $Z_{i\alpha}$ are perturbed by “aggregate shocks”. More precisely, their starting point was that the growth rates of (1.1) are still driven by the phenotypes’ reproductive fitness $u_{i\alpha}$, but that they are also perturbed by independent “white noise” processes $W_{i\alpha}$, accounting for the fickle hand of nature. In other words, Fudenberg and Harris replaced the deterministic growth equation (1.1) with the stochastic differential equation (SDE):

$$dZ_{i\alpha} = Z_{i\alpha} (u_{i\alpha} dt + \sigma_{i\alpha} dW_{i\alpha}), \quad (1.5)$$

where the diffusion coefficients $\sigma_{i\alpha}$ measure the intensity of nature’s interference. In this way, by applying Itô’s lemma (the stochastic equivalent of the chain rule; see Chapters 3–5 for details), one obtains the *replicator dynamics with aggregate shocks*:

$$\begin{aligned} dX_{i\alpha} = & X_{i\alpha} \left[(u_{i\alpha}(X) - u_i(X)) - \left(\sigma_{i\alpha}^2 X_{i\alpha} - \sum_{\beta} \sigma_{i\beta}^2 X_{i\beta}^2 \right) \right] dt \\ & + X_{i\alpha} \left[\sigma_{i\alpha} dW_{i\alpha} - \sum_{\beta} \sigma_{i\beta} X_{i\beta} dW_{i\beta} \right]. \end{aligned} \quad (1.6)$$

Needless to say, these aggregate shocks complicate things by quite a bit: for starters, the noise coefficients $\sigma_{i\alpha}$ do not appear only in the diffusive term $X_{i\alpha} (\sigma_{i\alpha} dW_{i\alpha} - \sum_{\beta} \sigma_{i\beta} X_{i\beta} dW_{i\beta})$, but also in the deterministic drift of the dynamics (1.6). As a result, one expects that the corresponding rationality properties of the stochastic replicator dynamics will be similarly affected, mirroring the interference of nature with the game. To that effect, [Cabrales \(2000\)](#) showed that dominated phenotypes still die out, but only if the variance $\sigma_{i\alpha}$ of the shocks is low enough. More recently, the work of [Imhof \(2005\)](#) revealed that even equilibrational states emerge over time but, again, conditionally on the shocks being mild enough (see also [Hofbauer and Imhof, 2009](#)).

On the other hand, since the noise coefficients $\sigma_{i\alpha}$ owe their presence in the drift of (1.6) to the way that the aggregate shocks dynamics (1.6) are derived from the population growth equation (1.5), it is not a priori obvious whether the stochastic dynamics (1.6) remain appropriate from the point of view of learning. This is one of the central questions that this dissertation seeks to address and, indeed, it turns out that this is not the case: if the players adhere to an exponential learning scheme, the stochastic perturbations $W_{i\alpha}$ should instead be applied directly to the scores of (1.3), leading to the stochastic differential equation:

$$dU_{i\alpha} = u_{i\alpha} dt + \sigma_{i\alpha} dW_{i\alpha}. \quad (1.7)$$

So, if the probabilities with which players choose their actions are given by the Boltzmann distribution (1.4), the rules of Itô calculus yield the *stochastic replicator dynamics of exponential learning*:

$$\begin{aligned} dX_{i\alpha} = & X_{i\alpha} [u_{i\alpha}(X) - u_i(X)] dt \\ & + \frac{1}{2} X_{i\alpha} \left[\sigma_{i\alpha}^2 (1 - 2X_{i\alpha}) - \sum_{\beta} \sigma_{i\beta}^2 X_{i\beta} (1 - 2X_{i\beta}) \right] dt \\ & + X_{i\alpha} \left[\sigma_{i\alpha} dW_{i\alpha} - \sum_{\beta} \sigma_{i\beta} X_{i\beta} dW_{i\beta} \right]. \end{aligned} \quad (1.8)$$

(As before, we refer the reader to Chapters 4 and 5 for a more detailed account of these issues.)

Aggregate
shocks

and the stochastic
replicator dynamics.

Learning in the
presence of noise.

We thus see that in the presence of noise, evolution and learning lead to very different dynamics; hence, one would expect that the rationality properties of the replicator dynamics in these two contexts will be just as different. Indeed, in stark contrast to the results of [Cabrales](#) and [Imhof](#), we will see in [Chapter 4](#) that suboptimal choices always die out in the long run, *irrespective of the fluctuations' magnitude*; even more to the point, the game's (strict) equilibria are the only steady states of the stochastic dynamics (1.8) which remain attracting and stable with arbitrarily high probability. In other words, despite all the noise which interferes with the players' choices, exponential learning gives players a clearer picture of the underlying game than natural selection, allowing them to eventually settle down to a "rationally admissible" state.

Network applications.

This robustness quality is precisely what makes the exponential learning scheme (and learning methods in general) so appealing from the viewpoint of network theory (where stochastic perturbations are, quite literally, omnipresent). That said, noise plays a different part in each application of game theory to networks, so, although a general theory of stochastically fluctuating games is necessary for our work, it would be naïve of us to expect that it would also be the end of it. That being the case, the main objective of this dissertation will be twofold: first, to provide the foundations for such a general theory, and then to showcase its potential with a few concrete applications.

1.4 THESIS OUTLINE AND OVERVIEW OF RESULTS

This thesis consists of three parts. The first part ([Chapters 2 and 3](#)) is an introductory one where we give a brief overview of the game-theoretic material and stochastic calculus tools that we will use in our analysis. The other two parts roughly correspond to the two principal objectives mentioned above: in Part II ([Chapters 4 and 5](#)) we provide a theoretical analysis of stochastically perturbed games, while in Part III ([Chapters 6 and 7](#)), we analyze the applications of this theory to communication networks. A more detailed outline follows below – consult also the flow diagram at the beginning of this dissertation:

A brief excursion in game theory.

[CHAPTER 2](#) is mostly a reference chapter, consisting of an epigrammatic overview of game theory geared towards the specifics of our work. We begin with the static background of evolutionary game theory, giving the definitions for games in normal form (in their various incarnations) and their solution concepts (dominated strategies, Nash equilibria and Wardrop's principle, etc.). We then follow up with a discussion of evolutionary concepts (such as evolutionarily stable strategies) and present the basic properties of the replicator dynamics in deterministic environments.

Stochastic analysis.

[CHAPTER 3](#) is also expository in nature. In the interest of completeness, our goal is to give a bird's eye view of the tools from stochastic analysis that we will employ in our study of randomly fluctuating games: the calculus of (Itô) diffusions, the partial differential equations that are associated to their evolution (namely the Kolmogorov backward equation and its forward counterpart, the Fokker-Planck equation), and the stability properties of these diffusions.

Disturbances in Nash games.

[CHAPTER 4](#) is where our analysis proper begins. After deriving the stochastic replicator dynamics of exponential learning for Nash games with a finite number of players, we proceed to show that dominated strategies die out at an exponential rate, no matter how loud the noise becomes. In fact, by induction

on the rounds of elimination of dominated strategies, we show that this is true even for *iteratively* dominated strategies: despite all the random disturbances, only rationally admissible strategies can survive in the long run. Finally, with respect to equilibrial play, we also show that the strict Nash equilibria of N -person Nash games are stochastically stable in the dynamics of exponential learning, thus recovering the folk theorem of evolutionary game theory.

CHAPTER 5 extends our results to games with continuous player sets, accounting in this way for the large population requirements of evolutionary games. Other than an infinite player set, a big difference between these games and finite player Nash games is that payoffs here no longer need be multilinear (as is the case in the probabilistic interpretation of expected gains in Nash games). Despite this nonlinearity, much of our analysis carries through to games of this type, establishing a series of similar results. In addition to all that, if the game in question admits a *potential* (an important attribute which bears close ties to congestion mechanisms), then we are able to derive results of global convergence to the neighbourhood of evolutionarily stable strategies.

Disturbances in potential games.

CHAPTER 6 comprises our first concrete network application. In particular, we study the distribution of traffic in networks whose routers try to minimise their delays by adhering to a simple learning scheme inspired by the replicator dynamics of evolutionary game theory. The stable steady states of these dynamics coincide with the network's *Wardrop equilibria* and form a convex polytope whose dimension is determined by the network's *redundancy* (an important concept which measures the "linear dependence" of the users' paths). Despite this abundance of stationary points, the long-term behaviour of the replicator dynamics turns out to be remarkably simple: every solution orbit converges to a Wardrop equilibrium.

The routing problem.

On the other hand, a major challenge occurs when the users' delays fluctuate unpredictably due to random external factors. In that case, interior equilibria are no longer stationary, but *strict* equilibria remain stochastically stable irrespective of the fluctuations' magnitude. In fact, if the network has no redundancy and the users are patient enough, we show that the long-term average of the users' traffic flows converges to the vicinity of an equilibrium, and we also estimate the corresponding invariant measure.

CHAPTER 7 marks a turn towards wireless networks. The problem we consider is that of a large number of wireless users who are able to switch dynamically between multiple wireless access-points (possibly belonging to different wireless standards). Even if users start out completely uneducated and naïve, we show that, by using a fixed set of strategies to process a broadcasted training signal, they quickly evolve and converge to a *correlated* equilibrium. In order to measure efficiency in this stationary state, we adapt the notion of the price of anarchy to this wireless setting and we obtain an explicit analytic estimate for it by using methods from statistical physics (namely the theory of replicas). Surprisingly, we find that the price of anarchy does not depend on the specifics of the wireless nodes (e.g. their spectral efficiency) but only on the number of strategies per user and a particular combination of the number of nodes, the number of users and the size of the training signal.

Correlated anarchy in wireless networks.

CHAPTER 8 contains the conclusions of our work and a number of open research directions for the future.

To streamline our presentation, the most technical points of our proofs and calculations have been relegated to a series of appendices at the end.

NOTATIONAL CONVENTIONS

Before proceeding further, we should note here one fundamental underlying assumption which will help readers preserve their sanity:

“Things are as nice as possible.”

For example, we will rarely make a distinction between objects that are naturally isomorphic, functions will be assumed to be as smooth as needed on the fly, topological spaces will be taken to be as separable as desired, and, sometimes, we will even identify equivalence classes with their representatives or random variables with their laws.

This should not be taken to mean that rigor is not important to us. Quite the contrary: in most places, assumptions will be clearly stated because the extra degree of generality afforded (or denied) by them will be crucial for our purposes. However, there are also points where a meticulous statement of every assumption would simply confound the reader. When this dubious case arises, we will choose to err on the side of simplicity and readability, relying on the context and common sense to do away with technical nonsense.

So, without further ado, let $S = \{s_\alpha\}_{\alpha=0}^n$ be a finite set. We define the vector space spanned by S over \mathbb{R} (also known as the free vector space on S over \mathbb{R}) to be the set of all formal linear combinations of elements of S with real coefficients, i.e. the set of all functions $x : S \rightarrow \mathbb{R}$. In tune with standard set-theoretic notation, we will denote this space by $\mathbb{R}^S \equiv \text{Maps}(S, \mathbb{R})$.

In this way, \mathbb{R}^S admits a canonical basis $\{e_\alpha\}_{\alpha=0}^n$ consisting of the indicator functions $e_\alpha : S \rightarrow \mathbb{R}$ which take the value $e_\alpha(s_\alpha) = 1$ on s_α and vanish otherwise; in particular, if $x \in \mathbb{R}^S$ has $x(s_\alpha) = x_\alpha$, we will have $x = \sum_\alpha x_\alpha e_\alpha$, so, clearly, $\mathbb{R}^S \cong \mathbb{R}^{|S|}$ (where $|S|$ denotes the cardinality of S). Hence, under the natural identification $s_\alpha \mapsto e_\alpha$, we will make no distinction between the elements s_α of S and the corresponding basis vectors e_α of \mathbb{R}^S ; in fact, to avoid drowning in a morass of indices and ugly expressions such as $H_{e_{\alpha_i}}$, we will routinely use α to refer interchangeably to either the element s_α of S or the basis vector e_α of \mathbb{R}^S (writing e.g. “ $\alpha \in S$ ” instead of “ $s_\alpha \in S$ ”). In the same vein, we will also identify the set $\Delta(S)$ of probability measures on S with the standard n -dimensional simplex of \mathbb{R}^S : $\Delta(S) \equiv \{x \in \mathbb{R}^S : \sum_\alpha x_\alpha = 1 \text{ and } x_\alpha \geq 0\}$ (i.e. the convex hull of the basis vectors e_α).

Concerning players and their actions (pure strategies), we will follow the original convention of Nash (1951) and employ Latin indices (i, j, \dots) for players (or classes thereof), while reserving Greek ones (α, β, \dots) for their (pure) strategies. Also, when the need arises to differentiate between the strategies of an individual player, we will use α, β, \dots for strategy indices that start at 0 and μ, ν, \dots for those that start at 1.

In addition to all that, if the players’ action sets \mathcal{A}_i are disjoint (as is typically the case), we will identify their union $\cup_i \mathcal{A}_i$ with the *disjoint union* $\mathcal{A} \equiv \coprod_i \mathcal{A}_i = \cup_i \{(\alpha, i) : \alpha \in \mathcal{A}_i\}$ via the canonical isomorphism $\alpha \in \mathcal{A}_i \mapsto (\alpha, i) \in \mathcal{A}$. In this manner, if $\{e_{i\alpha}\}$ is the natural basis of $\mathbb{R}^{\mathcal{A}_i}$ and $\{e_\alpha\}$ denotes the corresponding basis of $\mathbb{R}^{\mathcal{A}} \cong \prod_i \mathbb{R}^{\mathcal{A}_i}$, we will occasionally drop the index i altogether and write $x = \sum_\alpha x_\alpha e_\alpha \in \mathbb{R}^{\mathcal{A}}$ instead of $x = \sum_{i,\alpha} x_{i\alpha} e_{i\alpha} \in \prod_i \mathbb{R}^{\mathcal{A}_i}$. Similarly, when it is clear from the context that we are summing over the pure strategies \mathcal{A}_i of player i , we will use the shorthand $\sum_\alpha^i \equiv \sum_{\alpha \in \mathcal{A}_i}$.

Finally, with regards to probability, if $X(t)$, $t \geq 0$, is some stochastic process in \mathbb{R}^n starting at $X(0) = x \in \mathbb{R}^n$, we will denote its law by $\mathbf{P}_{X;x}$. However, if there is no doubt as to the process which we are referring to, its law will

be denoted more simply by \mathbf{P}_x and, in that case, we will employ the term “almost surely” instead of the somewhat unwieldy “ \mathbf{P}_x -almost surely”. Also, in typical probabilistic fashion, we will drop the arguments of random variables whenever possible, writing e.g. $\mathbf{P}(X = x)$ instead of $\mathbf{P}\{\omega : X(\omega) = x\}$.

We should reiterate here that if spelling out things to an iota would add to the transparency of the presentation rather than subtract from it, we will not hesitate to break these conventions and carry around any additional notational baggage that the situation might call for. We do this in the hope that the reader will be generous enough to forgive the sacrifice of a small degree of consistency in the pursuit of clarity.

Part I

THEORETICAL PRELIMINARIES

2

A BRIEF EXCURSION IN GAME THEORY

EVEN THOUGH it has been over five decades since the foundational work of von Neumann and Morgenstern (1944) and the seminal papers of Nash (1950a,b, 1951), there appears to be no universal consensus for notation and nomenclature among game theorists, often leading to fairly confusing conflicts of terminology. To avoid such issues, and also for the sake of completeness, we will present here a brief overview of game-theoretic essentials, loosely based on the panoramic surveys of Fudenberg and Tirole (1991), Fudenberg and Levine (1998) and the excellent book by Weibull (1995).

To spare the reader the embarrassment of successive (re)definitions and extensive (re)adaptations of previous ones, we have chosen to be as general as possible right from the beginning. The price for this early generality is that it requires a degree of abstraction which some readers might find burdensome, especially those not already acquainted with some of the concepts presented.¹ To avoid this Bourbaki-type pitfall, we have endeavored to lace our presentation with concrete examples, and also to keep these examples as free as possible of technical details.

2.1 GAMES IN NORMAL FORM

Let us begin with an example of a true children’s game: Rock-Paper-Scissors.² Since Rock crushes Scissors, Scissors cut Paper, and Paper covers Rock, there is no “dominant” strategy and the outcome of the game is essentially based on chance, making it an excellent way to resolve children’s disputes.³ Of course, human players might attempt any number of psychological ruses to confound their opponents, such as shouting “Rock” before they play or attempting to second-guess their opponents’ pattern of play in a rationale similar to that of Harsányi’s “belief hierarchy” (Harsányi, 1967-1968). Ultimately however, no such trick can yield consistent gains against a truly random strategy which employs Rock, Paper or Scissors with equal probability – a consequence of the game’s perfect symmetry.⁴ For this reason, this “mixed” strategy marks the game’s “equilibrium”: a player who deviates from the probability distribution $(1/3, 1/3, 1/3)$ will, over time, lose if his opponent sticks to it with the faith of a true believer.

*Rock-Paper-Scissors
as a strategic game.*

-
- ¹ In defense of our choice, we should state that we really *do* need every bit of generality introduced.
 - ² Known as *jan-ken-pon* in Japan, where the game is equally widespread.
 - ³ Interestingly enough, these disputes may also involve adult children as well. In 2006, Federal Judge Gregory Presnell from the Middle District of Florida, exasperated by the endless legalistic debates of *Avista management v. Wausau Underwriters*, ordered opposing sides to settle a trivial point regarding the location of a witness deposition by playing a game of Rock-Paper-Scissors on the front steps of the federal courthouse in Tampa (*The Seattle Times*, June 10, 2006).
 - ⁴ In fact, a popular strategy in Rock-Paper-Scissors tournaments is to memorize the digits of some irrational number (such as π or e) mod 3, ignoring zeros (*The New York Times*, September 5, 2004). Thankfully, mathematicians have devoted more energy to the question of whether these distributions of digits are truly uniform rather than devising strategies of this kind.

Another example of a game is that which we inadvertently play whenever we connect to the Internet. Indeed, in its most abstract form, the Internet can be viewed as a network over which a user places some demands, e.g. download page “X” from server “Y”, split a video stream over channels “A” and “B”, etc. However, each time a user chooses (consciously or not) an Internet route, he will have to contend with all other users that share this path’s links, leading to congestion and great delays if many users make the same choice. In this way, Internet users can be seen as “players”, trying to choose their routes in such a way so as to minimize the delays that they experience.

Components of a
game in normal form:

We thus see that a game roughly consists of three components: 1) a set of *players*, representing the competing entities; 2) the possible *actions* that the players can take in the game; and 3) the *rewards* associated to the couplings of these actions. Once these components have been described, the resulting entity will be called a *game in normal* (or *strategic*) *form*.⁵

Players,

To make this more precise, our starting point will be a set of *players* \mathcal{N} , together with a finite measure ν on \mathcal{N} which “accounts” for all players $i \in \mathcal{N}$ (in the sense that the singletons $\{i\} \subseteq \mathcal{N}$ are all ν -measurable). This measure will allow us to “count” players: for instance, if the player set \mathcal{N} is itself finite, ν will always be taken to be the counting measure on \mathcal{N} ; alternatively, if we wish to study large, “continuous” populations of players, this measure will be nonatomic (such as Lebesgue measure on the interval $\mathcal{N} = [0, 1]$ for a population of “mass” 1).

their strategies,

The players’ possible actions in the game will be represented by their *strategy sets* Δ_i , $i \in \mathcal{N}$. For our purposes, we will assume that these sets are locally compact Hausdorff spaces and that the relative topologies induced on $\Delta_i \cap \Delta_j$ agree for all $i, j \in \mathcal{N}$. Thanks to this compatibility condition, $\Delta_0 \equiv \bigcup_j \Delta_j$ inherits a natural Borel structure arising from the union topology (the finest topology in which the inclusions $\Delta_i \hookrightarrow \Delta_0$ are continuous). In this way, an admissible *strategy profile* $x \in \prod_i \Delta_i$ will just be a measurable function $x : \mathcal{N} \rightarrow \Delta_0$ which maps $i \mapsto x_i \in \Delta_i$ for all players $i \in \mathcal{N}$ – that is to say, $x_i \in \Delta_i$ is just the action taken by player $i \in \mathcal{N}$, and we assume that this choice is made in a “measurable” way. For technical reasons, we will also require that the push-forward measure $x_*\nu$ induced on Δ_0 by x (given by $x_*\nu(U) = \nu(x^{-1}(U))$) for any Borel $U \subseteq \Delta_0$) be inner regular, and, hence, Radon (since ν is finite).

As is customary, we will identify two profiles which agree ν -almost everywhere, except when we need to focus on the strategy of a particular player $i \in \mathcal{N}$ against that of his *opponents* $\mathcal{N}_{-i} \equiv \mathcal{N} \setminus \{i\}$; in that case, we will use the shorthand $(x_{-i}; q_i)$ to denote the profile which agrees with x on \mathcal{N}_{-i} (ν -a.e.) and maps $i \mapsto q_i \in \Delta_i$. The set Δ of all such profiles $x \in \prod_i \Delta_i$ will then be referred to as the *strategy space* of the game – note that this space is itself Borel because it inherits the subspace topology from the product $\prod_i \Delta_i$.

and their payoffs.

Bearing all this in mind, the fitness of the players’ strategic choices will be determined by their *payoff functions* (or *utilities*) $u_i : \Delta \rightarrow \mathbb{R}$, $i \in \mathcal{N}$; in particular, $u_i(x) \equiv u_i(x_{-i}; x_i)$ will simply represent the reward that player $i \in \mathcal{N}$ receives in the strategy profile $x \equiv (x_{-i}; x_i) \in \Delta$, i.e. when he plays $x_i \in \Delta_i$ against his opponents’ strategy $x_{-i} \in \prod_{j \neq i} \Delta_j$. The only further assumptions that we will make is that these payoff functions be (Borel) measurable and that $u_i(x_{-i}; x_i) = u_i(x'_{-i}; x_i)$ whenever x and x' agree ν -a.e. on \mathcal{N}_{-i} .

This collection of *players* $i \in \mathcal{N}$, their *strategy sets* Δ_i , and their *payoff functions* $u_i : \Delta \rightarrow \mathbb{R}$ will be our working definition for a *game in normal form*, usually

⁵ The adjective “normal” is used here in contrast to games in “extensive form” (such as chess or Go) where one is interested in the particular move sequences and their ramifications.

- Players: $\mathcal{N} = \{1, 2\}$.
- Player actions: $\mathcal{A}_i = \{R, P, S\}, i = 1, 2$.
- Strategies: $\Delta_i = \Delta\{R, P, S\}, i = 1, 2$.
- Payoff matrix:

	R	P	S
R	(0, 0)	(-1, 1)	(1, -1)
P	(1, -1)	(0, 0)	(-1, 1)
S	(-1, 1)	(1, -1)	(0, 0)

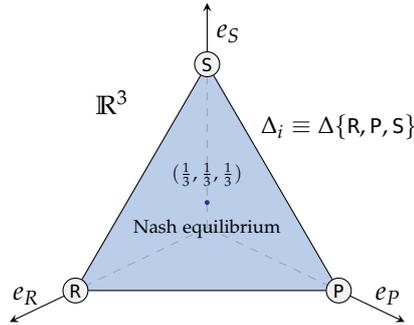


Figure 2.1: Rock-Paper-Scissors in strategic form.

denoted by $\mathfrak{G} \equiv \mathfrak{G}(\mathcal{N}, \Delta, u)$. Additionally, if the payoff functions $u_i : \Delta \rightarrow \mathbb{R}$ happen to be continuous, the game \mathfrak{G} will be called *continuous* as well.

Needless to say, this abstract definition is fairly elaborate, so we will immediately proceed with a few important examples to clarify the concept (and illustrate where each piece of generality is required).

2.1.1 N-person Games

As the name suggests, the players in N -person games are indexed by the finite set $\mathcal{N} = \{1, 2, \dots, N\}$, endowed with the usual counting measure. Since \mathcal{N} is finite, any strategy allocation $x : \mathcal{N} \rightarrow \bigcup_j \Delta_j$ will be measurable and, hence, the game's strategy space will simply be the finite product $\Delta \equiv \prod_j \Delta_j$ (thus doing away with most of the technical subtleties of the more general definition).

This is also the point where we recover the original scenario considered by Nash (1951). To see how, assume that every player $i \in \mathcal{N}$ comes with a finite set \mathcal{A}_i of actions (or pure strategies) which can be "mixed" according to some probability distribution $x_i \in \Delta(\mathcal{A}_i)$ (cf. our notational conventions in Chapter 1). In this interpretation, the players' strategy sets are just the simplices $\Delta_i \equiv \Delta(\mathcal{A}_i)$ (see also Fig. 2.1), and their payoff functions $u_i : \Delta \equiv \prod_i \Delta_i \rightarrow \mathbb{R}$ will be given by the multilinear expectations:

Nash games.

$$u_i(x) = u_i(x_1, \dots, x_N) = \sum_{\alpha_1 \in \mathcal{A}_1} \dots \sum_{\alpha_N \in \mathcal{A}_N} x_{1,\alpha_1} \dots x_{N,\alpha_N} u_i(\alpha_1, \dots, \alpha_N), \quad (2.1)$$

where $x_i = \sum_{\alpha} x_{i\alpha} e_{i\alpha}$ in the standard basis $\{e_{i\alpha}\}$ of $\mathbb{R}^{\mathcal{A}_i}$ and $u_i(\alpha_1, \dots, \alpha_N)$ is the reward that player i would obtain by choosing $\alpha_i \in \mathcal{A}_i$ against his opponents' action $\alpha_{-i} \in \mathcal{A}_{-i} \equiv \prod_{j \neq i} \mathcal{A}_j$.

Because of this (multi)linear structure, we will frequently refer to these Nash-type games as *linear games* to contrast them with more general N -person games where strategy spaces and payoff functions might fail to have any sort of linear structure whatsoever – for example, as is the case with the class of *concave games* considered by Rosen (1965).

2.1.2 Populations and Evolutionary Games

As we have already seen, the building blocks of evolutionary game theory are games which are played by uncountably many players. As such, these nonatomic *population games* require the full breadth afforded by our more

abstract definition – see also [Schmeidler \(1973\)](#) for what seems to be the first attempt to axiomatize nonatomic games.

The structure of
population games.

The first piece of additional structure encountered in these games is a measurable partition $\mathcal{N} = \bigcup_{k=1}^N \mathcal{N}_k$ of the player set \mathcal{N} into N disjoint *populations* (or *classes*) $\mathcal{N}_k \subseteq \mathcal{N}$, representing the “species” involved in the game. Every player $i \in \mathcal{N}$ belongs to a unique class \mathcal{N}_k which we denote by $\text{class}(i)$, and each of these classes is “measured” by the corresponding restriction ν_k of the measure ν on \mathcal{N}_k ($\nu_k(B) = \nu(B \cap \mathcal{N}_k)$ for any Borel $B \subseteq \mathcal{N}$). Usually, these classes consist of a “continuum” of players, in which case we assume that the measures ν_k are *nonatomic*: individual players have measure zero, reflecting the fact that they carry a negligible impact on the species to which they belong.

The second fundamental assumption is that this classification of players (their *taxonomy*) also determines how they interact with their environment and with each other. More precisely, this means that the strategy sets of two players that belong to the same population coincide: $\Delta_i = \Delta_j$ whenever $\text{class}(i) = \text{class}(j)$. Bearing this in mind, we will write \mathcal{A}_k for the common strategy set of the k -th population and \mathcal{A}_0 for the union of all these sets: $\mathcal{A}_0 = \bigcup_{k=1}^N \mathcal{A}_k = \bigcup_{i \in \mathcal{N}} \Delta_i$.

With all this said and done, note that every strategy profile $x : \mathcal{N} \rightarrow \mathcal{A}_0$ pushes forward a (by assumption, Radon) measure \hat{x}_k on \mathcal{A}_k in the usual way:

$$\hat{x}_k(U) \equiv (x_* \nu_k)(U) = \nu_k(x^{-1}(U)) = \nu\{i \in \mathcal{N}_k : x_i \in U\} \quad (2.2)$$

Strategy
distributions

for any Borel $U \subseteq \mathcal{A}_k$; in other words, $\hat{x}_k(U)$ is just the measure of players in population \mathcal{N}_k whose chosen strategy lies in $U \subseteq \mathcal{A}_k$. This collection of measures gives rise to the *strategy distribution* $\hat{x} = (\hat{x}_1, \dots, \hat{x}_N)$, which, by construction, is a Radon measure on the product space $\mathcal{A} \equiv \prod_k \mathcal{A}_k$.

Obviously, the relation between a strategy profile and its distribution is precisely that between a random variable and its law:⁶ whereas the profile x records the actual strategic choices of the players, the distribution \hat{x} measures their density. Accordingly, since players in population games are completely characterized by the species to which they belong, payoffs should depend solely on this strategy distribution and not on the individual strategic choices of each player.

and player
anonymity.

Specifically, if $P_0(\mathcal{A})$ is the space of all strategy distributions equipped with the topology of vague convergence (recall that we are working with Radon measures on locally compact Hausdorff spaces), we may assume that there exist continuous functions $\hat{u}_k : P_0(\mathcal{A}) \times \mathcal{A}_k \rightarrow \mathbb{R}$, $k = 1, \dots, N$, such that:

$$u_i(x) = \hat{u}_k(\hat{x}; x_i) \text{ for all } i \in \mathcal{N}_k. \quad (2.3)$$

Put differently, this means that players are assumed *anonymous*: as long as the overall strategy distribution \hat{x} stays the same, payoffs determined by (2.3) will remain unaffected even by positive-mass migrations of players from one strategy to another (and not only by migrations of measure zero).

Evolutionary
games.

EVOLUTIONARY GAMES Again, it would serve well to illustrate this abstract definition by means of a more concrete example. To wit, in evolutionary game theory, populations are usually represented by the intervals $\mathcal{N}_k = [0, m_k]$, where $m_k > 0$ denotes the “mass” of the population under Lebesgue measure.

⁶ After all, if we renormalize the (finite) population measure ν to unity, we see that the strategy profile $x : \mathcal{N} \rightarrow \Delta_0$ is just a random variable on the observation space Δ_0 .

The strategy spaces \mathcal{A}_k are also assumed to be finite,⁷ so a strategy distribution can be viewed as a point in the (finite-dimensional) product $\prod_k m_k \Delta(\mathcal{A}_k)$. Thus, if player $i \in \mathcal{N}_k$ chooses $\alpha \in \mathcal{A}_k$, his payoff will be given by:

$$u_{k\alpha}(x) \equiv u_k(x; \alpha), \quad (2.4)$$

where, in a slight abuse of notation, we removed the hats from \hat{u}_k and \hat{x} in order to stress that they are the fundamental quantities which describe the game (it will always be clear from the context whether we are referring to the strategy distribution in $P_0(\mathcal{A})$ or to the actual strategy profile $x : \mathcal{N} \rightarrow \mathcal{A}_0$).

This choice of notation is very suggestive for another reason as well: if we set $\Delta_k \equiv m_k \Delta(\mathcal{A}_k)$, then these simplices may be taken as the strategy sets of an associated N -person game whose players are indexed by $k = 1, \dots, N$ (that is, they correspond to the populations themselves). The only thing needed to complete this description is to define the payoff functions $u_k : \Delta \equiv \prod_k \Delta_k \rightarrow \mathbb{R}$, and a natural choice in that direction would be to take the *population averages*:

$$u_k(x) = m_k^{-1} \sum_{\alpha} x_{k\alpha} u_{k\alpha}(x), \quad (2.5)$$

where $x_{k\alpha}$ are the coordinates of x in the standard basis of $\prod_k \mathbb{R}^{\mathcal{A}_k}$.

These population averages play a very important role in the class of Nash-type evolutionary games (or *random matching* games) where payoffs are determined by pairing individuals in some Nash game.⁸ What this means is that when the game is played, one individual is chosen from each species \mathcal{N}_k (with probability determined by the population distribution $x \in \Delta$), and these individuals are matched to play a Nash-type game whose pure strategies coincide with the species' action sets \mathcal{A}_k . Assuming for simplicity that all populations have unit mass, the payoff $u_{k\alpha}(x)$ to the phenotype α in the distribution x will be given by $u_{k\alpha}(x) \equiv u_k(x_{-k}; \alpha)$, where $u_k(x)$ denotes the expectation:

$$u_k(x) = \sum_{\alpha_1 \in \mathcal{A}_1} \cdots \sum_{\alpha_N \in \mathcal{A}_N} x_{1,\alpha_1} \cdots x_{N,\alpha_N} u_i(\alpha_1, \dots, \alpha_N). \quad (2.6)$$

In this way, we see that the population averages $u_k(x) = \sum_{\alpha} x_{k\alpha} u_{k\alpha}(x)$ (recall that populations are assumed to have mass 1) essentially encapsulate the whole game: phenotype payoffs $u_{k\alpha}$ can be derived from the species average (2.6) by the substitution $x \rightarrow (x_{-k}; \alpha)$ or by the simple differentiation $u_{k\alpha} = \frac{\partial u_k}{\partial x_{k\alpha}}$ (a procedure which does not work in more general evolutionary games).

Interspecific scenaria like this are called *multi-population* (or *asymmetric*) random matching games, to differentiate them from the slightly different variation of *single-population* ones. As the name suggests, in single-population games there is a single species \mathcal{N} with finitely many phenotypes $\alpha \in \mathcal{A}$, and, at every instance of the game, a pair of individuals is randomly drawn to play a 2-person (symmetric) Nash game with strategies specified by the genetic programming of each phenotype $\alpha \in \mathcal{A}$.⁹ Thus, if the population share of α -type individuals is x_{α} , their expected payoff in such a match will be:

$$u_{\alpha}(x) = \sum_{\beta} u_{\alpha\beta} x_{\beta}, \quad (2.7)$$

⁷ See Cressman (2005) or Hofbauer et al. (2009) for an exception to this rule.

⁸ In fact, this term is often used synonymously to evolutionary games, much as N -person games are commonly taken to refer to Nash-type games.

⁹ One could also consider matching N -tuples in a N -person game, but single-population games of this type have not attracted much interest.

Random
matchings.

Multi- vs. single-
population games.

where $u_{\alpha\beta} \equiv u(\alpha, \beta)$ denotes the (symmetric) payoff matrix of the game. In this fashion, we obtain the *single-population average*:

$$u(x) = \sum_{\alpha} x_{\alpha} u_{\alpha}(x) = \sum_{\alpha, \beta} u_{\alpha\beta} x_{\alpha} x_{\beta}, \tag{2.8}$$

which, in contrast to the multi-population case (2.6), is *quadratic* in the distribution $x = \sum_{\alpha} x_{\alpha} e_{\alpha} \in \Delta(\mathcal{A})$ and not multilinear (or bilinear, as the case would have it). This disparity is obviously due to the self-interactions which are present in single-population games (and absent in multi-population ones), and it will have some very important consequences later on.

2.1.3 Potential Games

Another important subclass of evolutionary games arises when the payoffs $u_{k\alpha}$ of (2.4) satisfy the closedness condition:

$$\frac{\partial u_{k\alpha}}{\partial x_{\ell\beta}} = \frac{\partial u_{\ell\beta}}{\partial x_{k\alpha}} \text{ for all populations } k, \ell \text{ and all strategies } \alpha \in \mathcal{A}_k, \beta \in \mathcal{A}_{\ell}. \tag{2.9}$$

Nonatomic potential games.

This condition is commonly referred to as “externality symmetry” (Sandholm, 2001), and it describes games where a marginal increase of the players using strategy α has the same effect on the payoffs to players playing strategy β as the converse increase.¹⁰ However, the true origins of (2.9) do not lie in economics but in the geometry of physics: externality symmetry simply implies that the vector field $u_{k\alpha}$ (with k, α taken as a single index running through $\sum_k |\mathcal{A}_k|$) is irrotational and, hence, conservative (if interpreted as a force field).

Clearly then, since the strategy distributions of these games live in the simply connected polytope $\Delta = \prod_k \Delta_k$, condition (2.9) amounts to the existence of a *potential function* $F : \Delta \rightarrow \mathbb{R}$ such that:

$$u_{k\alpha}(x) = -\frac{\partial F}{\partial x_{k\alpha}}. \tag{2.10}$$

Thus, if a player $i \in \mathcal{N}_k$ makes the switch $\alpha \rightarrow \beta$, his payoff will change by:

$$u_{k\beta}(x) - u_{k\alpha}(x) = -\left(\frac{\partial F}{\partial x_{k\beta}} - \frac{\partial F}{\partial x_{k\alpha}} \right) = -dF_x(e_{k\beta} - e_{k\alpha}), \tag{2.11}$$

where $\{e_{k\alpha}\}$ denotes the standard basis of $\prod_k \mathbb{R}^{\mathcal{A}_k}$ and dF_x is the differential of F at x . In other words, *the strategy migration $\alpha \rightarrow \beta$ is profitable to a player iff the direction $e_{k\beta} - e_{k\alpha}$ descends the potential F* , a property of potential games whose ramifications will underlie a large part of our work.

Finite potential games.

Obviously, the differential nature of condition (2.9) depends crucially on the continuous (and nonatomic) structure of the player set \mathcal{N} . However, since evolutionary games can be seen as a limiting case of N -person games for very large numbers of players,¹¹ it would be of interest to see if the condition of externality symmetry also appears in the context of strategic games with a *finite* number of players.

The simplest way to do this would be to focus on the single-population case: N players sharing a common, finite set of strategies \mathcal{A} . So, let $n_{\alpha} \equiv n_{\alpha}(x), \alpha \in \mathcal{A}$

¹⁰ Strictly speaking, (2.9) implicitly requires the payoff functions $u_{r\alpha}$ to be defined over an open neighborhood of Δ , but this technicality will very rarely be an issue.

¹¹ The precise limiting procedure need not concern us here; instead, see Sandholm (2001).

denote the number of players choosing $\alpha \in \mathcal{A}$, in the strategy profile $x : \mathcal{N} \rightarrow \mathcal{A}$ – i.e. $n_\alpha(x)$ is just the cardinality of the set $x^{-1}(\alpha) = \{i \in \mathcal{N} : x_i = \alpha\}$. Based on our postulates for evolutionary games, we assume that the N (atomic) players are indistinguishable and anonymous, so that their payoffs depend only on the distribution (n_1, n_2, \dots) of their choices, i.e.:

$$u_i(x_{-i}; \alpha) \equiv u_\alpha(n_1, n_2, \dots). \quad (2.12)$$

In this way, to obtain the discrete equivalent of (2.10), we should posit the existence of a potential function $F \equiv F(n_1, n_2, \dots)$ such that:

$$u_i(x_{-i}; \alpha) = -(F(n_{-\alpha}; n_\alpha + 1) - F(n_{-\alpha}; n_\alpha)), \quad (2.13)$$

where, in obvious notation, $(n_{-\alpha}; n_\alpha)$ is just the standard shorthand for $(n_1, \dots, n_\alpha, \dots)$. If such a function exists, one can easily verify the discrete version of the externality symmetry condition (2.9):

$$u_\alpha(n_{-\beta}; n_\beta + 1) - u_\alpha(n_{-\beta}; n_\beta) = u_\beta(n_{-\alpha}; n_\alpha + 1) - u_\beta(n_{-\alpha}; n_\alpha), \quad (2.14)$$

which captures the marginal effect of a player switching strategies $\alpha \leftrightarrow \beta$.

The converse derivation of (2.13) from (2.14) is not much harder and can be found in the work of [Monderer and Shapley \(1996\)](#), where potential games were given their first thorough examination.¹² Therein, (2.13) is written in the equivalent (and, also, much more elegant) form:

$$u_i(x_{-i}; q_i) - u_i(x_{-i}; q'_i) = -(F(x_{-i}; q_i) - F(x_{-i}; q'_i)), \quad (2.15)$$

which has the added benefit of being immediately translatable to *any* N -person game (and not only to anonymous games with a finite set of common strategies).¹³ For this reason, we will use this last equation to define *finite* potential games, while retaining (2.10) for the continuous (nonatomic) case.

Remark. As an added consistency benefit, if the game is Nash, we will have $u_i(x) = \sum_\alpha x_{i\alpha} u_{i\alpha}(x)$ where $u_{i\alpha}(x) = u_i(x_{-i}; \alpha)$. On account of this multilinearity, it easily follows that the potential function F will be itself multilinear, and a simple differentiation of (2.15) for $q_i = e_{i\alpha}$ gives $u_{i\alpha}(x) = -\frac{\partial F}{\partial x_{i\alpha}}$. In other words, the definitions (2.10) and (2.15) are “equivalent” in Nash-type games, so we will often use (2.10) in the Nash setting without further discussion.

2.1.4 Congestion Games

The importance of the last class of examples that we will present lies in the fact that they can be used to model competitive situations where players must share a common set of facilities, such as the streets of a road network, water sources, etc. Quite fittingly, these games are called *congestion games*.

The first component of a congestion game is a set of *facilities* (or *resources*) \mathcal{S} , typically assumed to be finite. The strategy space Δ_i of each player consists of subsets of \mathcal{S} , corresponding to the resources that player i may employ: for

Atomic potential games.

The structure of congestion games.

¹² The first actual mention of potential games seems to date back to [Rosenthal \(1973\)](#).

¹³ The fact that this is the natural translation to N -person games can be seen by looking at Nash games. Indeed, let $\mathfrak{G} \equiv \mathfrak{G}(\mathcal{N}, \Delta, u)$ be a Nash game where the players’ pure strategies are drawn from the sets \mathcal{A}_i , $i \in \mathcal{N}$, and let $\mathfrak{G}^0 \equiv \mathfrak{G}^0(\mathcal{N}, \mathcal{A}, u)$ be the corresponding “pure” game where players are restricted to only play pure strategies. Then, \mathfrak{G} is a potential game in the sense of (2.13) if and only if \mathfrak{G}^0 is; furthermore, if F^0 is a potential for \mathfrak{G}^0 , then its multilinear extension F is a potential for \mathfrak{G} , and, conversely, the restriction of F to the vertices of Δ is a potential for \mathfrak{G}^0 .

NONATOMIC	There are no <i>atoms</i> (minimal sets of positive measure): every individual player has measure zero.
POPULATION	Games where the players' strategies and payoffs classify them in disjoint <i>populations</i> (or <i>classes</i>). Trivial if we allow populations with zero measure or consisting of a single atom.
ANONYMOUS	Payoffs depend only on the distribution of players among the various strategies and not on their individual strategic choices. Of interest only in games with non-trivial population partitions.
EVOLUTIONARY	Anonymous nonatomic games, typically with a finite number of populations and strategies per population.
POTENTIAL	Games where the payoff functions are derived from a potential.
CONGESTION	Anonymous population games played over an infrastructure of <i>facilities</i> . Equivalent to potential games.

Table 2.1: A short glossary of game-theoretic prefixes and adjectives.

example, if the set of facilities represents the various streets of a road network, a player's strategy set might consist of all subsets of streets that comprise a route from home to work. In this manner, players are segregated into disjoint classes \mathcal{N}_k , each class characterized by the common strategy space \mathcal{A}_k which is shared by all its players (in the road network analogy, a class would consist of all players with the same origin and destination).

The similarities to population games do not end here: players are again assumed to be anonymous, so that their payoffs only depend on their distribution among the various facilities. However, unlike population games, rewards and costs within a given class of players are unaffected by what happens in facilities not used by any member of the class: the cost of employing a set of resources is simply the sum of the individual facility costs, and each such cost only depends on the number of players utilizing it.

To put this in more rigorous terms, note that the union $\mathcal{A}_0 = \bigcup_k \mathcal{A}_k$ of the players' strategy spaces is itself a subset of the power set $2^{\mathcal{S}}$ which contains all possible combinations of facilities (recall that a choice of route entails a choice of a *set* of streets). In this way, a strategy profile $x : \mathcal{N} \rightarrow \mathcal{A}_0$ gives rise to a *load profile* $y(x) \in \mathbb{R}^{\mathcal{S}}$, defined in components by:

$$y_r(x) = \nu\{i : r \in x_i\}. \quad (2.16)$$

In plain terms, y reflects the load on each facility $r \in \mathcal{S}$: y_r is simply the measure of players that employ $r \in \mathcal{S}$. In a similar vein, we assume that the cost of employing $r \in \mathcal{S}$ is a continuous function $\phi_r(y_r)$ of the load on facility r , while the cost of employing a subset $\alpha \in 2^{\mathcal{S}}$ of facilities is the sum of the individual costs – after all, to go through a route in a road network, one must first traverse all the constituent streets. Hence, if player i uses the subset of resources $\alpha \subseteq \mathcal{S}$, his payoff will be given by:

$$u_{i\alpha}(x) \equiv u_i(x_{-i}; \alpha) = - \sum_{r \in \alpha} \phi_r(y_r(x)), \quad (2.17)$$

the “–” sign stemming from the inverse relation between costs and rewards.

So far, the most interesting feature of congestion games appears to be their applicational value and they do not seem to stand out in terms of theoret-

ical significance. That they actually do so is a consequence of a remarkable observation by [Monderer and Shapley \(1996\)](#) who showed that:

Theorem 2.1 ([Monderer and Shapley, 1996](#)). *Every congestion game is a potential game and every potential game is isomorphic to a congestion game.*

Equivalence with potential games.

“Isomorphic” here means that (after suitably relabelling strategies) there is a congestion game whose payoff structure exactly mirrors that of the original game. Also, we should note here that the original statement of the theorem concerns games with a finite number of players; however, it is not hard to extend the proof of Monderer and Shapley to account for continuous population games with a finite number of strategies (our working assumption).

At any rate, the converse direction of [Theorem 2.1](#) will not really be of interest us here. What will be far more important for our purposes is the actual potential which lurks beneath congestion games, and which comes in two flavors, continuous and discrete:

The Rosenthal potential.

$$F(x) = \sum_{r \in \mathcal{S}} \int_0^{y_r(x)} \phi_r(w) dw \quad (\text{continuous version}) \quad (2.18a)$$

$$F(x) = \sum_{r \in \mathcal{S}} \sum_{w=0}^{y_r(x)} \phi_r(w) \quad (\text{discrete version}) \quad (2.18b)$$

Owing to the work of [Rosenthal \(1973\)](#), this potential function is often called the *Rosenthal potential* of the game (conditions [\(2.10\)](#) and [\(2.15\)](#) are easily checked). Its importance will already become apparent in the next section and the concept will truly shine in [Chapter 6](#).

2.2 SOLUTION CONCEPTS IN GAME THEORY

Under the umbrella of rational behavior, selfish players will seek to play those strategies which deliver the best rewards against the choices of their opponents. This general postulate takes on many different guises, each leading to a different *solution concept* of a game. In what follows, we will give a brief account of these solution concepts, only pausing long enough to discuss their potential applications.

2.2.1 Dominated Strategies

Needless to say, rational players who seek to maximize their individual payoffs will avoid strategies that always lead to diminished payoffs against any play of their opponents. More precisely:

Definition 2.2 (Dominated Strategies). Let $\mathcal{G} \equiv \mathcal{G}(\mathcal{N}, \Delta, u)$ be a game in normal form. We will say that the strategy $x_i \in \Delta_i$ of player $i \in \mathcal{N}$ is *strictly dominated* by $x'_i \in \Delta_i$ and we will write $x_i \prec x'_i$ when:

Dominated strategies

$$u_i(x_{-i}; x_i) < u_i(x_{-i}; x'_i), \text{ for all strategies } x_{-i} \in \Delta_{-i} \text{ of } i\text{'s opponents.} \quad (2.19)$$

Remark. If the inequality [\(2.19\)](#) is not strict, we will say that x_i is *weakly dominated* by x'_i and we will write $x_i \preceq x'_i$ instead. Because our primary interest lies in strictly dominated strategies, “dominated” should always be taken to mean “strictly dominated” if it is not accompanied by another adjective.

An example of a dominated strategy occurs in the Prisoner’s Dilemma: as we mentioned in the introduction, “defect” always yields a lesser penalty

compared to “remaining silent”, suggesting the latter choice to be irrational (at the very least, from the point of view of game theory). With this in mind, dominated strategies can be effectively removed from the analysis of a game \mathcal{G} , because rational players will have little incentive to ever use them.

*and rationally
admissible ones.*

In the Prisoner’s Dilemma, this leaves “defecting” as the only rationally admissible strategy but, in games with a more elaborate strategic structure, this would simply lead to a reduced version of the original game, say \mathcal{G}_1 , where some strategies have been eliminated (incurring analogous restrictions on the domains of the payoff functions). However, by deleting a dominated strategy, a new strategy (perhaps of another player) might become dominated – for instance, if the deleted strategy was the only one that was performing worse than the new candidate for deletion. Proceeding ad infinitum, we thus expect to end up with a game \mathcal{G}_∞ with *no* dominated strategies: in this *resolved* game, all strategies are “rationally” admissible in the sense that they have survived every round of elimination.

Still, the endgame of this process might not be very interesting because the reduced strategy set might turn out to be empty. An easy (if a bit one-sided) example is a single-player game where the player chooses a real number $x \in \mathbb{R}$ and is awarded his choice as a payoff: $u(x) = x$. In this simple game, we obviously have $x \prec y$ if and only if $x < y$, so *all* strategies are dominated (and strictly so to boot)!

*Existence of
rational strategies.*

This example goes to show that there are very simple games with no rationally admissible strategies, just as there are conflicts which admit no rational resolution. What goes wrong here is that the the player’s payoff function does not assume a maximum value over his strategy space and, hence, no matter what the player chooses, there is always going to be a strictly better strategic choice. In a sense however, this is the only thing that can go wrong: if the payoff u_i to player i attains its maximum value at some $q_i \in \Delta_i$ for a fixed strategy x_{-i} of i ’s opponents, then the strategy q_i cannot be dominated by *any* other strategy in the full strategy space Δ_i , and will thus remain indomitable in any reduction of Δ_i as well.

This observation yields an important sufficient condition for the existence of rationally admissible strategies: if there exists some $x_{-i} \in \Delta_{-i}$ such that the “constrained” payoff function $u_i(x_{-i}; \cdot) : \Delta_i \rightarrow \mathbb{R}$ has a maximum value over Δ_i , then player i has a rationally admissible strategy. This condition is clearly satisfied in N -person games with compact strategy spaces and continuous payoff functions, but if the game is also Nash, we can take things a bit further: simply note that the maxima of the game’s multilinear payoffs are to be sought among the vertices of the strategy simplices $\Delta_i = \Delta(A_i)$. We thus get:

Proposition 2.3. *Every N -person game with compact strategy spaces and continuous payoffs possesses rationally admissible strategies. In addition, if the game is Nash, there exists a pure rationally admissible strategy.*

*Dominance
and rationality.*

Needless to say, even if rationally admissible strategies exist, they do not have to be unique, a fact well-known to any child who has played Rock-Paper-Scissors. In the fortuitous event that the set of these strategies really is a singleton (as in the Prisoner’s Dilemma), the game will be called *dominance solvable* and the sole surviving strategy will be the game’s *rational solution*.

2.2.2 Nash Equilibrium

On the other hand, not every game can be solved by deleting dominated strategies, so we are forced to explore other rational solution concepts instead.

The fundamental principle to keep in mind is this: if the strategy of a player's opponents is kept fixed, then that player would seek to play the strategy which maximizes his own payoff. This reasoning leads to the celebrated equilibrium concept put forth by Nash (1950a,b, 1951):

Definition 2.4 (Nash equilibrium). Let $\mathcal{G} \equiv \mathcal{G}(\mathcal{N}, \Delta, u)$ be a game in normal form. We will say that a strategy profile $q \in \Delta$ is at *Nash equilibrium* when:

Nash equilibrium.

$$u_i(q) \geq u_i(q_{-i}; q'_i) \quad \text{for almost every } i \in \mathcal{N} \text{ and all strategies } q'_i \in \Delta_i. \quad (2.20)$$

In addition, if the inequality (2.20) is strict for almost every player $i \in \mathcal{N}$, we will also say that q is a *strict* Nash equilibrium.

The set of Nash equilibria of \mathcal{G} will be denoted by $\Delta^* \equiv \Delta^*(\mathcal{G})$; similarly, we will write $\Delta_0^* \equiv \Delta_0^*(\mathcal{G})$ for the game's set of strict equilibria.

As a first consequence of this definition, we have:

Proposition 2.5. *A Nash equilibrium is rationally admissible for almost every player.*

The proof of this proposition is trivial: if it exists, a Nash equilibrium q will maximize almost every player's constrained payoff function $u_i(q_{-i}; \cdot)$, so it can never be strictly dominated. However, unlike rationally admissible strategies whose existence is relatively easy to establish, Nash equilibria are much more elusive. Indeed, the trouble with the principle of Nash is that every player will be trying to do the same thing at the same time and, in a competitive environment, these concurrent optimization efforts will hardly be compatible with one another. It is thus far from obvious that there exists a strategy profile which is *unilaterally stable* (in the sense that almost every player has no incentive to deviate if other players keep to their choices).

Rationality of Nash equilibria.

The first (and most influential) existence result in this direction was proven by Nash himself who, through an audacious use of Brouwer's fixed point theorem,¹⁴ showed that N -person linear games always possess an equilibrium. This result was subsequently extended by Rosen (1965) to the class of *concave games* (concave payoffs over convex strategy spaces), while Schmeidler (1973) essentially settled the issue for nonatomic population games with finite strategy sets – see also Ali Khan (1986) and Milchtaich (2000).

The issue of (in)existence.

Nevertheless, things can go sour even in very simple games: for example, if players do not mix their pure strategies in Rock-Paper-Scissors, there is no equilibrium – and, albeit discrete, this is even a continuous game. In the author's opinion, the question of whether equilibria exist in a given game or not is an intrinsic property of the game (or class of games) considered, much akin to the problem of deciding whether a topological space has the fixed point property or not. Fortunately however, we will not have to deal with this issue: the games that we will study will always have equilibria.

PROPERTIES OF EQUILIBRIA In Nash's original scenario (N -person linear games), there is an alternate characterization of Definition 2.4 which is very useful in practice. Since the payoff functions u_i are multilinear over the game's strategy sets $\Delta_i \equiv \Delta(\mathcal{A}_i)$, it follows that the equilibrium condition (2.20) will be satisfied if and only if:

Equilibria in Nash games.

$$u_{i\alpha}(q) \geq u_{i\beta}(q), \quad \text{for all } \alpha, \beta \in \mathcal{A}_i \text{ such that } q_{i\alpha} > 0, \quad (2.21)$$

¹⁴ Actually, his first proof (Nash, 1950a) employed Kakutani's fixed point theorem. The Kakutani proof lends itself better to subsequent generalizations of his result, but the elegance of the Brouwer proof is truly stunning – in the author's opinion, it was certainly taken from *The Book* of Erdős.

where $u_{i\alpha}(x) \equiv u_i(x_{-i}; \alpha)$ is the payoff that player i would have obtained by playing $\alpha \in \mathcal{A}_i$ against $x_{-i} \in \Delta_{-i} \equiv \prod_{j \neq i} \Delta_j$. In other words, a mixed profile q which is at Nash equilibrium will employ only those pure strategies $\alpha \in \mathcal{A}_i$ which yield the greatest payoffs, and these payoffs will necessarily be equal:

$$u_{i\alpha}(q) = u_{i\beta}(q) \text{ for all } \alpha, \beta \in \text{supp}(q_i), \quad (2.22)$$

where the *support* of q_i is defined as $\text{supp}(q_i) \equiv \{\alpha \in \mathcal{A}_i : q_{i\alpha} > 0\}$.

Wardrop's
principle.

In evolutionary games, Nash equilibria are captured by a similar condition known as *Wardrop's principle*:

$$u_{k\alpha}(q) \geq u_{k\beta}(q) \text{ for all } \alpha, \beta \in \mathcal{A}_k \text{ s.t. } q \text{ assigns positive mass to } \alpha \quad (2.23)$$

(here, $u_{k\alpha}(x)$ denotes the payoff to members of the class \mathcal{N}_k who chose $\alpha \in \mathcal{A}_k$). To see how this condition arises, note that if the strategy $\alpha \in \mathcal{A}_k$ has positive measure in the equilibrational profile q (that is, $v_k(q^{-1}(\alpha)) > 0$), then there exists a player $i \in \mathcal{N}_k$ (actually a positive mass of such players) with $q_i = \alpha$ and such that (2.20) holds.¹⁵ Hence, for every $\beta \in \mathcal{A}_k$, we readily get:

$$u_{k\alpha}(q) = u_i(q_{-i}; \alpha) \geq u_i(q_{-i}; \beta) = u_{k\beta}(q). \quad (2.24)$$

Wardrop's principle will be central in our analysis, especially in Chapters 5 and 6. En passant, we only note here that an analogous characterization can be laid down for population games with continuous strategy sets, but since we will not need this added generality, we will not press the issue further – see instead the recent work by [Cressman \(2005\)](#) or by [Hofbauer et al. \(2009\)](#).

Equilibria in
potential games.

Finally, if the game in question is also a potential one, we have seen that beneficial migrations descend the potential function, so the minima of the potential naturally correspond to strategy distributions where no unilateral improvement is possible. In fact, it can be shown that the Kuhn-Tucker constrained minimization conditions for the game's potential coincide precisely with the Wardrop characterisation (2.23):

Proposition 2.6 ([Sandholm, 2001](#)). *Let \mathcal{G} be a potential game with potential F . Then, the Nash set $\Delta^*(\mathcal{G})$ of the game coincides with the critical set of F .*

2.2.3 Justifying Equibrational Play: Correlation and Evolution

Justifications
of rationality.

From an *a posteriori* point of view, Nash equilibria represent a very sensible solution concept: in terms of individual rewards, every player is employing the best possible strategy against the background of his opponents' strategic choices. Still, the issue of why and how players might arrive to an equilibrational strategy profile in the first place remains an actively debated question. After all, the complexity of most games increases exponentially with the number of players and, hence, identifying a game's equilibria quickly becomes prohibitively difficult.¹⁶

Another important issue which was first pointed out by [Aumann \(1974\)](#) is that a player has no incentive to play his component of a Nash equilibrium unless he is convinced that all other players will play theirs. This argument

¹⁵ Since (2.20) holds for *almost* every player $i \in \mathcal{N}$, the positive measure requirement is a crucial one.

¹⁶ This complexity question essentially gave birth to the subdiscipline of algorithmic game theory. The issue was finally settled in Costis Daskalakis' widely (and justly) acclaimed doctoral dissertation, where he showed that the problem of computing Nash equilibria is PPAD-complete (in layman's terms, pretty complex; see [Daskalakis et al. \(2006\)](#) and [Daskalakis \(2008\)](#) for more details).

obviously gains additional momentum if the game in question has multiple Nash equilibria: in that case, even players with unbounded deductive capabilities will be hard-pressed to choose a strategy. From this point of view, rational individuals would appear to be more in tune with Aumann's notion of a *correlated equilibrium* where subjective beliefs are also taken into account, but this also brings to the table more of the unpredictable (and often irrational) "human" element that we sought to banish from our analysis.

Instead, as we have already mentioned, the seminal work of [Maynard Smith \(1974\)](#) on animal conflicts has cast Nash equilibria in a different light because it unearthed a profound connection between evolution and rationality: roughly speaking, one leads to the other. So, when different species contend for the limited resources of their habitat, evolution and natural selection are expected to steer the ensuing conflict to an equilibrial state which leaves no room for irrational behavior. As a consequence, instinctive "fight or flight" responses that are deeply ingrained in a species can be seen as a form of rationality, acquired over the species' evolutionary course.¹⁷

In what follows, we will give a brief account of these two justifications of equilibrial play, leading to a more refined concept (evolutionarily stable strategies) and a more general one (correlated equilibrium).

Correlated Equilibrium

The notion of a *correlated equilibrium* was first formulated by [Aumann \(1974, 1987\)](#) in an attempt to reconcile the inherently objective flavor of Nash equilibria with the players' subjective view of their world. In essence, we have already witnessed the gap between these two aspects of rationality: if a player employs his component of an equilibrial strategy only when he *believes* that all other players will employ theirs, then 1) a Nash equilibrium sounds much like a dog chasing its tail (consistent but hardly getting anywhere); and 2) in formulating an equilibrium concept, one should take into account players' beliefs about other players and the "world" in general.

Subjective beliefs and equilibrium.

A few years later, [Brian Arthur \(1994\)](#) put forth another argument that further strengthens this line of reasoning. While humans are only moderately strong in problems that can be solved by *deductive* reasoning (they do better than animals but much worse than computers), they excel in intuition and in solving problems by *inductive* reasoning. Since this "intuitive" approach rests heavily on what players believe is going on around them, an equilibrium will only be reachable if it also takes into account these beliefs. Sadly, Nash equilibria do not, so, to incorporate subjectivity, we will have to extend our game-theoretic framework somewhat.

For simplicity (we will not need more generality anyhow), let \mathcal{G} be a N -person Nash-type game, with actions (pure strategies) drawn from the sets \mathcal{A}_i , $i = 1, \dots, N$, and rewards $u_i : \prod_j \mathcal{A}_j \rightarrow \mathbb{R}$. Following [Aumann \(1987\)](#), let us also assume that the players' beliefs are formed by observing the "state of the world", i.e. the events that transpire in some (finite) probability space (Ω, \mathbf{P}) – ranging from rainfall in Beijing, to what another player had for breakfast, or to whatever other scrap of information a player might deem relevant.¹⁸ This data is observed, recorded and processed by the players, who then choose an

Correlation and subjective beliefs.

¹⁷ In his thesis ([Nash, 1950b](#), p. 21), Nash himself states that "[i]t is unnecessary to assume that the participants have full knowledge of the total structure of the game, or the ability and inclination to go through any complex reasoning processes. But the participants are supposed to accumulate empirical information on the relative advantages of the various pure strategies at their disposal."

¹⁸ Of course, this also brings up the issue of exactly what kind of information is actually observable by a player. To account for that, Aumann partitions Ω in player-specific σ -algebras which determine

action based on their subjective but *correlated* (since all players live in the same world) assessment of their surroundings.

Correlated strategies.

More formally, a *correlated strategy* will be a map $p : \Omega \rightarrow \mathcal{A} \equiv \prod_i \mathcal{A}_i$ whose components $p_i : \Omega \rightarrow \mathcal{A}_i$, $i = 1, \dots, N$, determine the players' responses $p_i(\omega) \in \mathcal{A}_i$ when the world is at the state $\omega \in \Omega$. Clearly, these strategies can be viewed as \mathcal{A} -valued random variables on Ω and, as such, we can "forget" about the underlying sample space Ω by looking at the probability law induced by p on \mathcal{A} :

$$P(\alpha_1, \dots, \alpha_N) = \mathbf{P} \{ \omega : p(\omega) = (\alpha_1, \dots, \alpha_N) \}. \quad (2.25)$$

Mimicking the notation in Nash games, when we want to focus on the strategic component $p_i : \Omega \rightarrow \mathcal{A}_i$ of player $i \in \mathcal{N}$ against the strategy of his opponents $p_{-i} : \Omega \rightarrow \mathcal{A}_{-i} \equiv \prod_{j \neq i} \mathcal{A}_j$, we will employ the shorthand $(p_{-i}; p_i) \equiv (p_1, \dots, p_i, \dots, p_N)$. Unlike Nash games however, these correlated strategies no longer live in the product of simplices $\Delta \equiv \prod_i \Delta(\mathcal{A}_i)$, but in the simplex $\Delta_c \equiv \Delta(\prod_i \mathcal{A}_i)$ which contains Δ as the subset of independent, *uncorrelated* strategies. In other words, if we set:

$$p_{i\alpha} \equiv \mathbf{P}(p_i = \alpha) = \sum_{\alpha_{-i} \in \mathcal{A}_{-i}} P(\alpha_{-i}; \alpha), \quad (2.26)$$

then the condition $P(\alpha_1, \dots, \alpha_N) = p_{1,\alpha_1} \cdots p_{N,\alpha_N} \in \prod_i \Delta(\mathcal{A}_i)$ which characterizes Nash games holds if and only if the individual strategies $p_i : \Omega \rightarrow \mathcal{A}_i$ are stochastically independent.

With all this machinery in place, the game is actually played as follows:

1. At the outset of the game, the players form their correlated strategies $p_i : \Omega \rightarrow \mathcal{A}_i$, $i = 1, \dots, N$ based on their prior beliefs;
2. The world is observed at some state $\omega \in \Omega$;
3. Players employ the actions $p_i(\omega) \in \mathcal{A}_i$ that are prescribed by their strategies in response to ω and receive the corresponding payoffs $u_i(p(\omega))$.

In this way, given a correlated strategy $p : \Omega \rightarrow \mathcal{A}$, payoffs can also be seen as random variables $u_i \circ p : \Omega \rightarrow \mathbb{R}$, an instance of which determines the payoffs $u_i(p(\omega))$ that players receive when the world is at state ω . Consequently, the expected payoff of a player will no longer be given by (2.1) (which presupposes that the players' strategies are independent of one another), but by the correlated expectation:

$$u_i(p) \equiv \sum_{\alpha_1 \in \mathcal{A}_1} \cdots \sum_{\alpha_N \in \mathcal{A}_N} P(\alpha_1, \dots, \alpha_N) u_i(\alpha_1, \dots, \alpha_N), \quad (2.27)$$

or, in more succinct form:

$$u_i(p) = \sum_{\alpha_{-i}}^{-i} \sum_{\alpha_i}^i P(\alpha_{-i}; \alpha_i) u_i(\alpha_{-i}; \alpha_i). \quad (2.28)$$

It is important to note that this notation meshes well with the one we used for the payoffs $u_i : \mathcal{A} \rightarrow \mathbb{R}$ of the original game. Indeed, if we stretch our notation a bit and also keep α_i for the "pure" strategy of player i which maps $\omega \mapsto \alpha_i$ for all $\omega \in \Omega$, then $u_i(\alpha_1, \dots, \alpha_N)$ defined as in (2.27) clearly coincides with the payoff to player i in the action profile $(\alpha_1, \dots, \alpha_N) \in \mathcal{A}$.

whether an event is "observable" (measurable) by a player or not. Since we want to keep our discussion as simple as possible, we will not concern ourselves with this issue here.

Having defined payoffs in this correlated context, we obtain:

Definition 2.7 (Correlated Equilibrium). Let \mathfrak{G} be a finite N -person game with action sets \mathcal{A}_i and payoffs $u_i : \mathcal{A} \equiv \prod_i \mathcal{A}_i \rightarrow \mathbb{R}$, and let (Ω, \mathbf{P}) be a finite probability space as above. A strategy $p = (p_1, \dots, p_N) : \Omega \rightarrow \mathcal{A}$ will be called a *correlated equilibrium* of \mathfrak{G} when:

$$u_i(p) \geq u_i(p_{-i}; p'_i) \quad (2.29)$$

for all strategies $p'_i : \Omega \rightarrow \mathcal{A}_i$ that factor through p_i (i.e. $p'_i = \sigma \circ p_i$ for some permutation $\sigma : \mathcal{A}_i \rightarrow \mathcal{A}_i$), and for all players $i = 1, \dots, N$.

The set of correlated equilibria of \mathfrak{G} will be denoted by $\Delta_c^* \equiv \Delta_c^*(\mathfrak{G})$.

At first sight, the requirement that the perturbation p'_i factor through p_i might appear artificial, but it is actually a vital ingredient of the way that correlated strategies work. Indeed, player i can be viewed more simply as a machine which receives the recommendation $p_i(\omega)$ (when the world is at state ω) from the “correlating device” $p_i : \Omega \rightarrow \mathcal{A}_i$ and then plays out its action. As a result, a player can either act based on the recommendation $p_i(\omega)$ of the strategy p_i , or disregard it altogether and play something totally different. However, since the only information that reaches the player in this picture is the suggestion $p_i(\omega)$ (and not the actual state ω), his action can only depend on $p_i(\omega)$, i.e. be of the form $\sigma(p_i(\omega))$ for some endomorphism $\sigma : \mathcal{A}_i \rightarrow \mathcal{A}_i$.

An alternative characterization of (2.29) which highlights precisely this feature of correlated equilibria is obtained by the simple rearrangement:

$$\begin{aligned} u_i(p_{-i}; p'_i) &= \sum_{\alpha_{-i}}^{-i} \sum_{\alpha'_i}^i \mathbf{P}(p_{-i} = \alpha_{-i}, p'_i = \alpha'_i) u_i(\alpha_{-i}; \alpha'_i) \\ &= \sum_{\alpha_{-i}}^{-i} \sum_{\alpha'_i}^i \left(\sum_{\alpha_i: \sigma(\alpha_i) = \alpha'_i} \mathbf{P}(p_{-i} = \alpha_{-i}, p_i = \alpha_i) \right) u_i(\alpha_{-i}; \alpha'_i) \\ &= \sum_{\alpha_{-i}}^{-i} \sum_{\alpha_i}^i P(\alpha_{-i}; \alpha_i) u_i(\alpha_{-i}; \sigma(\alpha_i)). \end{aligned} \quad (2.30)$$

Therefore, the correlated equilibrium condition (2.29) will be equivalent to:

$$\sum_{\alpha_{-i}}^{-i} \sum_{\alpha_i}^i P(\alpha_{-i}; \alpha_i) [u_i(\alpha_{-i}; \alpha_i) - u_i(\alpha_{-i}; \sigma(\alpha_i))] \geq 0 \quad (2.31)$$

for all players $i = 1, \dots, N$ and all maps $\sigma : \mathcal{A}_i \rightarrow \mathcal{A}_i$. In other words, a probability distribution P on \mathcal{A} will be at correlated equilibrium when, on average, a player cannot win more by unilaterally deviating from the suggestion $p_i(\omega)$ of an equilibrial strategy.

Except for this added “consistency of recommendations” requirement, (2.29) is quite similar to the Nash condition (2.20) and, sure enough, any Nash equilibrium of the game is also a correlated equilibrium. Conversely, correlated equilibria can be seen as Nash equilibria of a “correlated” version of \mathfrak{G} (Aumann, 1974, 1987; see also Chapter 7), but, in general, these two versions lead to different equilibrial sets. To better understand the nature of correlated equilibria, the following characterization will be very useful:

Proposition 2.8 (Aumann, 1987). Let \mathfrak{G} be a finite N -person game with action sets \mathcal{A}_i and payoffs $u_i : \mathcal{A} \equiv \prod_j \mathcal{A}_j \rightarrow \mathbb{R}$ as above. A probability distribution P on \mathcal{A} will be at correlated equilibrium if and only if:

$$\sum_{\alpha_{-i}}^{-i} P(\alpha_{-i}; \alpha) u_i(\alpha_{-i}; \alpha) \geq \sum_{\alpha_{-i}}^{-i} P(\alpha_{-i}; \alpha) u_i(\alpha_{-i}; \beta) \quad (2.32)$$

for all players $i = 1, \dots, N$ and for all $\alpha, \beta \in \mathcal{A}_i$.

*Correlated
Equilibrium.*

*Alternative
characterizations.*

*Correlated equilibrial
distributions.*

Proof. This classic characterization can be found in [Aumann \(1987\)](#) or [Fudenberg and Tirole \(1991, p. 74\)](#), and it is very easy to prove once one has (2.31) at one’s disposal; we give a proof here for completeness.

Obviously, the “if” direction is trivial: for any $\sigma : \mathcal{A}_i \rightarrow \mathcal{A}_i$, put $\beta = \sigma(\alpha)$ in (2.32) and sum over all $\alpha \in \mathcal{A}_i$. For the converse implication,¹⁹ fix $\alpha, \beta \in \mathcal{A}_i$ and define $\sigma : \mathcal{A}_i \rightarrow \mathcal{A}_i$ to be the map which takes $\alpha \mapsto \beta$ and leaves $\mathcal{A}_i \setminus \{\alpha\}$ untouched (i.e. player i follows the recommendation he receives unless he gets “ α ”, in which case he plays β). Then, the characterization (2.31) becomes:

$$\begin{aligned} 0 &\leq \sum_{\alpha_{-i}}^{-i} \sum_{\alpha_i \neq \alpha}^i P(\alpha_{-i}; \alpha_i) [u_i(\alpha_{-i}; \alpha_i) - u_i(\alpha_{-i}; \sigma(\alpha_i))] \\ &\quad + \sum_{\alpha_{-i}}^{-i} P(\alpha_{-i}; \alpha) [u_i(\alpha_{-i}; \alpha) - u_i(\alpha_{-i}; \sigma(\alpha))] \\ &= \sum_{\alpha_{-i}}^{-i} P(\alpha_{-i}; \alpha) [u_i(\alpha_{-i}; \alpha) - u_i(\alpha_{-i}; \beta)], \end{aligned} \tag{2.33}$$

which is just (2.32). □

Remark. We should also point out here that (2.32) admits yet another interpretation. If $p_{i\alpha} \equiv \mathbf{P}(p_i = \alpha) > 0$, we may divide both sides of (2.32) by $p_{i\alpha}$ to obtain the conditioned version:

$$\sum_{\alpha_{-i}}^{-i} P(\alpha_{-i} | \alpha) u_i(\alpha_{-i}; \alpha) \geq \sum_{\alpha_{-i}}^{-i} P(\alpha_{-i} | \alpha) u_i(\alpha_{-i}; \beta), \tag{2.34}$$

where, in obvious notation, $P(\alpha_{-i} | \alpha)$ denotes the conditional probability $P(\alpha_{-i} | \alpha) \equiv \mathbf{P}(p_{-i} = \alpha_{-i} | p_i = \alpha) = P(\alpha_{-i}; \alpha) / p_{i\alpha}$. Although this last form does not make (2.32) more useful in terms of calculations, it certainly elucidates its suggestion-deviation nature, so it will serve as our primary definition for correlated equilibria.

Structure of the set of correlated equilibria.

At any rate, Proposition 2.8 leads to a very sharp description of the set Δ_c^* of correlated equilibria of a game. Indeed, with (2.32) being linear in the coordinates $P(\alpha_1, \dots, \alpha_N)$ of the simplex $\Delta_c \equiv \Delta(\mathcal{A})$, it follows that these inequalities will cut out a (convex) polytope of Δ_c . Since equilibria in mixed strategies always exist in finite Nash games, and since Nash equilibria are also correlated equilibria, we easily obtain:

Proposition 2.9 ([Aumann, 1987](#)). *Let \mathfrak{G} be a finite game. Then, the set $\Delta_c^* \equiv \Delta_c^*(\mathfrak{G})$ of correlated equilibria of \mathfrak{G} is a nonempty convex polytope which contains the convex hull of the set $\Delta^* \equiv \Delta^*(\mathfrak{G})$ of Nash equilibria of \mathfrak{G} .*

Applicability and limitations.

Unfortunately, the notion of a correlated equilibrium is rife with assumptions concerning the behavior of players, their beliefs, and how players interpret their surroundings. Thus, even though it provides us with a set of formal tools to study the effects of correlation and subjectivity in game theory, it will inevitably be plagued by the same set of problems facing every mathematical model of human rationality: the human psyche is much more complex than the relatively simple structure of a correlated strategy can account for.

That said, this does not mean that this discussion has been in vain. Far from it actually: whenever there is a signalling device among players who are simpler than humans (thus rendering our analysis all the more applicable), correlated equilibrium becomes the de facto solution concept of the game. Accordingly, Aumann’s analysis will be very useful to us in Chapter 7, where we

¹⁹ The presentation of both [Aumann \(1987\)](#) and [Fudenberg and Tirole \(1991\)](#) is a bit unsatisfying (or, at least, imprecise) at this point. Specifically, [Aumann \(1987, p. 6\)](#) states that “the expectation [(2.29)] still obeys the inequality when conditioned on each possible value of $[p_i]$, i.e. on each possible “suggestion” to Player i ”, an argument which I find rather nebulous.

consider games between machines that are trained by monitoring a (randomly generated) common broadcast.

Evolutionarily Stable Strategies

As we have already mentioned, the applicational value of game theory chiefly stems from its close ties with entities whose behavior is regulated by their (genetic) programming. As such, we will devote the rest of this section to study the interplay of rationality and “natural” selection (or “artificial intelligence” in the case of machine evolution).

To begin with, consider an evolutionary game \mathfrak{G} played by N species \mathcal{N}_k , $k = 1, \dots, N$, each coming with a finite set of actions (phenotypes) \mathcal{A}_k and associated (continuous) payoff functions $u_{k\alpha}$, $\alpha \in \mathcal{A}_k$. Then, the population average of a species’ payoff in the population profile x will be:

$$u_k(x) = m_k^{-1} \sum_{\alpha} x_{k\alpha} u_{k\alpha}(x), \quad (2.35)$$

where $m_k = v(\mathcal{N}_k)$ denotes the “mass” of the k -th species and the notation \sum_{α}^k is shorthand for $\sum_{\alpha \in \mathcal{A}_k}$ – cf. the notational conventions of Chapter 1.

In keeping with the analogy between strategic choices and genetic programming, let us assume that there appears an influx of “mutated” individuals whose genetic traits (strategic choices) do not adhere to those of the incumbent population state $x \in \Delta \equiv \prod_k m_k \Delta(\mathcal{A}_k)$. Formally, this means that the species move to the perturbed state $x' = x + \varepsilon(y - x)$, where $y \in \Delta$ represents the mutant strain and $\varepsilon > 0$ measures the extent of the mutation. Then, the fundamental question that arises from an evolutionary standpoint is the following: *does the mutation y prevail over the incumbent population x ?*

Under the interpretation that payoffs represent reproductive fitness, the two factors that will decide this issue are the average fitness of the incumbent population and that of the mutants. This leads us to:

Definition 2.10 (Evolutionary Stability). A population state $x \in \Delta$ will be called *evolutionarily stable* when, for every sufficiently small $\varepsilon > 0$, and for every mutation x' of the form $x' = x + \varepsilon(y - x)$, $y \in \Delta \setminus \{x\}$, we have:

$$\sum_{k,\alpha} x_{k\alpha} u_{k\alpha}(x') > \sum_{k,\alpha} y_{k\alpha} u_{k\alpha}(x') \quad (2.36)$$

This stability criterion (originally due to Taylor, 1979) is quite natural from a biological perspective: it simply states that if the average reproductive fitness of the mutant strain is greater than that of the incumbent populations, then the mutation spreads, otherwise it becomes extinct. Nevertheless, there *is* some ambivalence between game theorists on how to define evolutionary stability in multiple-species environments. Some authors remove the averaging over the species in (2.36) and argue that it suffices if the mutation becomes extinct in *some* species (Cressman, 1992), while others take a stronger view and require that this unaveraged version of (2.36) hold for *all* species (Swinkels, 1992). All these definitions clearly agree in single-population environments and, even though they appear pretty different at first sight, they rarely clash with one another; we chose Taylor’s version because it represents the middle way.

At any rate, what is truly amazing with Definition (2.10) is that the simple natural selection criterion (2.36) also leads to rational behavior:

Proposition 2.11. *Evolutionarily stable strategies are at Nash equilibrium.*

Mutations in evolutionary games.

Evolutionary stability.

Rationality and evolution.

Proof. Let q be an evolutionarily stable strategy which is not a Nash equilibrium of the game. Then, by Wardrop's principle (2.23), there exists a population \mathcal{N}_k and an action $\alpha \in \mathcal{A}_k$ employed by a positive fraction $q_{k\alpha} > 0$ of \mathcal{N}_k which yields strictly worse rewards than some other action $\beta \in \mathcal{A}_k$: that is, $u_{k\alpha}(q) < u_{k\beta}(q)$ for some $\alpha \in \mathcal{A}_k$ with $q_{k\alpha} > 0$. In that case, if a small enough mass $\varepsilon > 0$ of individuals switch from α to β , they will obtain strictly better rewards as a result of the payoffs $u_{k\alpha}$ being continuous. More formally:

$$u_{k\alpha}(q_{-k}; q_k + \varepsilon(e_{k\beta} - e_{k\alpha})) < u_{k\beta}(q_{-k}; q_k + \varepsilon(e_{k\beta} - e_{k\alpha})) \text{ for some } \varepsilon > 0, \quad (2.37)$$

where $\{e_{k\alpha}\}$ is the canonical basis of $\mathbb{R}^{\mathcal{A}_k}$.²⁰ However, this also shows that the mutation $q' = q + \varepsilon(e_{k\beta} - e_{k\alpha})$ violates the evolutionary stability criterion (2.36), because the k -th species (which is the only one affected by it and, hence, the only one with a non-trivial presence in the species average) favors the transition $\beta \rightarrow \alpha$. Since q has been assumed evolutionarily stable, this yields a contradiction and completes our proof. \square

Equilibrium
selection.

This proposition was first proven by [Maynard Smith and Price \(1973\)](#) for single-population random matching games and is one of the cornerstones of evolutionary game theory (our more general version was motivated by our wish to venture beyond the narrow confines of Nash games). The reason for this is simple: it establishes the critical link between rationality and evolution.

Then again, it is important to note that not all equilibria are evolutionarily stable: for instance, if the members of a species are matched in a 2-person Nash game that returns equal rewards to all, then every strategy is at Nash equilibrium but there is no evolutionarily stable one. For a more interesting example, consider a single population $\mathcal{N} = [0, 1]$ playing a "coordination game" \mathfrak{G} with two strategies, "1" and "2", and payoff functions $u_1(x_1, x_2) = x_1$ and $u_2(x_1, x_2) = x_2$ respectively. It is then easy to show that this game has three equilibria, $(1, 0)$, $(0, 1)$ and $(1/2, 1/2)$, of which only the two pure ones are evolutionarily stable – if the population were evenly split, then a positive mass migration would yield better rewards to the mutants.

Criteria for
stability.

CHARACTERIZATIONS OF STABILITY The two examples above both show that evolution provides us with an important selection criterion between the various Nash equilibria of a game. In the following propositions we present two explicit characterizations of evolutionarily stable strategies which will be very important for us later on:

Proposition 2.12 ([Selten, 1980](#)). *Let \mathfrak{G} be a multi-population random matching game. Then, q is evolutionarily stable if and only if it is a strict equilibrium of \mathfrak{G} .*

Proposition 2.13. *Let \mathfrak{G} be a nonatomic potential game with potential function F . If q is a local minimum of F such that F is locally increasing along all rays emanating from q , then q is evolutionarily stable. In particular, if F is strictly convex, then the global minimum of F is the game's unique evolutionarily stable strategy.*

Proof. Proposition 2.12 is relatively well-known and its proof can be found e.g. in [Weibull \(1995, p. 167\)](#). As for Proposition 2.13, let q be a local minimum of F , and pick any $y \in \Delta \setminus \{x\}$. Then, if we set $z = y - q$, it suffices to show that:

$$\sum_{k,\alpha} z_{k\alpha} u_{k\alpha}(q + \varepsilon z) < 0 \text{ for all sufficiently small } \varepsilon > 0. \quad (2.38)$$

²⁰ Note that the requirement $q_{k\alpha} > 0$ is crucial: if $q_{k\alpha} = 0$ then ε will be de facto zero.

However, if we set $f(\theta) = F(q + \theta z)$, $0 \leq \theta \leq 1$, we readily obtain:

$$f'(\theta) = \sum_{k,\alpha} z_{k\alpha} \frac{\partial F}{\partial x_{k\alpha}} \Big|_{q+\theta z} = - \sum_{k,\alpha} z_{k\alpha} u_{k\alpha}(q + \theta z), \quad (2.39)$$

and since f is locally increasing by assumption ($q + \theta z$ is just a ray emanating from q after all), our assertion follows. \square

Proposition 2.13 will be of great importance to us because it ties us back to the physical origins of the potential function. In a sense, it tells us that although not all Nash equilibria minimize the potential, evolution selects precisely those that do so in an essential way, just as the stable points of a dynamical system correspond to the essential minima of its Hamiltonian. This observation will be central in our dynamical considerations and we will explore its ramifications in the chapters that follow.

2.3 EVOLUTION, LEARNING, AND THE REPLICATOR DYNAMICS

Let us recall the fundamental question that we posed in the beginning of Section 2.2.3: *how can competing individuals arrive to an equilibrial solution, and why would they feel inclined to do so in the first place?*

The notions of a correlated equilibrium and of evolutionary stability give a partial answer to the “why” in this question but, given that they are both *posterior*, static solution concepts, they hardly address the more dynamic “how”. Even the concept of an ESS (which has its roots in the theory of evolution, an inherently dynamic process) is little more than a stability requirement: the evolutionary trajectory that might lead players to an equilibrium is still veiled behind a shroud of deductive omnipotence.

Static vs. dynamic justifications.

As we mentioned in the introduction, there are two complementary approaches to this question: one concerns populations of mindless entities who evolve thanks to some natural selection mechanism, and the other revolves around slightly more intelligent automata which are able to “learn” and adapt to their surroundings. Quite fittingly, the dynamics stemming from the first approach have been dubbed *evolutionary*, reserving the designation *learning dynamics* for the latter.

2.3.1 The Replicator Dynamics

By far the most comprehensive account of evolutionary dynamics can be found in the review article by Hofbauer and Sigmund (2003) and the excellent book by Weibull (1995),²¹ where one encounters a plethora of natural selection mechanisms, based e.g. on “imitation of the fittest”, evolution driven by dissatisfaction, etc. The origins of each of these distinct selection mechanisms are quite disparate because they seek to address different aspects of the evolutionary process; however, most of these selection devices essentially boil down to the *replicator dynamics* of Taylor and Jonker (1978).

In what follows, we will focus our investigations on this replicator model of evolutionary growth, mostly because it can be derived from first principles with a minimal set of assumptions. The setup is as follows: we first consider N species \mathcal{N}_k , $k = 1, \dots, N$, each with an associated set \mathcal{A}_k of different phenotypes that all interact with each other in a common habitat (e.g. viral strains in

Evolution and growth.

²¹ The forthcoming book by Sandholm (2010) also promises to be a first-rate alternative.

bacterial environments). Then, if the population mass of the α -th phenotype of species \mathcal{N}_k is $z_{k\alpha}$, we assume that it follows the general growth equation:

$$dz_{k\alpha} = z_{k\alpha} u_{k\alpha}(x) dt, \quad (2.40)$$

where x denotes the profile of relative population shares $x_{k\alpha} = z_{k\alpha} / \sum_{\beta}^k z_{k\beta}$ and the (continuous) functions $u_{k\alpha}$ represent the per capita growth rate of the various species and phenotypes.²²

In a sense, (2.40) is the most general growth equation that we could write satisfying the conditions of perfect clonal reproduction (*replication*) and scale invariance. More precisely, the two assumptions that go into (2.40) are:

1. *Phenotypes Breed True*: phenotypes cannot arise spontaneously (or, conversely, become extinct) in finite time.²³
2. *Scale Invariance*: the growth rates $u_{k\alpha}$ are determined by the relative population shares x and not by the absolute populations z .

Needless to say, these assumptions are not to be taken lightly: spontaneous mutations are a key element in many evolutionary processes and, for example, it is also highly unlikely that the growth rate of a species depends only on the population distribution of its predators and not on the size of the preying species. Nevertheless, these limitations are not too restrictive: if the mutation mechanism is known, then it can be added in (2.40) appropriately, while predator-prey models (such as the Lotka-Volterra equation) can be rephrased in terms of (2.40) after a suitable change of variables. Truth be told, any tractable population model suffers from similar restrictions: the breadth of biological interactions that can be captured by (2.40) is not really diminished by the assumptions of replication and scale invariance.

The replicator dynamics.

At any rate, what is more important here is that a simple differentiation allows us to recast (2.40) solely in terms of the population shares $x_{k\alpha}$. Indeed, if we denote the population of the k -th species by $z_k = \sum_{\beta}^k z_{k\beta}$ and we also let $u_k(x) = \sum_{\alpha}^k x_{k\alpha} u_{k\alpha}(x)$ be its average (per capita) growth rate, we readily obtain the (*multi-population*) *replicator dynamics*:

$$\frac{dx_{k\alpha}}{dt} = \frac{z_k}{z_k^2} \frac{dz_{k\alpha}}{dt} - \frac{z_{k\alpha}}{z_k^2} \sum_{\beta}^k \frac{dz_{k\beta}}{dt} = x_{k\alpha} (u_{k\alpha}(x) - u_k(x)). \quad (2.41)$$

The replicator dynamics are one of the cornerstones of evolutionary game theory, but seeing as we have not yet made any mention of games in their derivation, one might be hard-pressed to see why. To recover this connection, one merely has to interpret the growth rates $u_{k\alpha}$ as payoffs to some underlying evolutionary game \mathfrak{G} whose action spaces correspond to the various phenotypes. As a result, we no longer have to worry about designing the game's payoffs so as to match the various interspecific interactions: they are simply the *observed* growth rates of a species' phenotypes and constitute measurable quantities of the evolutionary model (2.40).

From the point of view of learning theory, the situation is similar so, just as we focused on games with finitely many species and strategies in the evolutionary case, we will concentrate here on Nash games.²⁴ In particular,

²² The slight deviation from the notational scheme of Section 2.1.2 is in anticipation of the fact that the relative shares $x_{k\alpha}$ will be the key variables in this evolutionary context.

²³ This is hidden in the $z_{k\alpha}$ factor that appears in the RHS of (2.40).

²⁴ Unfortunately, a general theory of learning in non-linear games still eludes us; however, for an example, see Chapter 6.

let \mathfrak{G} be a Nash game played by N players with strategy sets $\Delta_i \equiv \Delta(\mathcal{A}_i)$ and payoffs $u_i : \Delta \equiv \prod_j \Delta_j \rightarrow \mathbb{R}$, extending the rewards $u_i(\alpha_1, \dots, \alpha_N)$ that correspond to pure strategy profiles $(\alpha_1, \dots, \alpha_N) \in \prod_j \mathcal{A}_j$. In this context, a *learning scheme* will be a dynamical system of the form:

Learning schemes

$$\frac{dx}{dt} = V(x), \text{ or, in coordinates: } \frac{dx_{i\alpha}}{dt} = V_{i\alpha}(x), \quad (2.42)$$

where $x = \sum_{i,\alpha} x_{i\alpha} e_{i\alpha} \in \Delta$ is the players' (mixed) strategy profile at time t and the vector field $V : \Delta \rightarrow \prod_j \mathbb{R}^{\mathcal{A}_j}$ plays the part of the "learning rule" in question – for simplicity, we will also take V to be smooth.²⁵

Of course, since the profile $x(t)$ evolves in Δ , $V(x)$ itself must lie on the tangent space $Z \equiv T_x \Delta$ of Δ ;²⁶ in coordinates, this amounts to:

$$\sum_{\alpha}^i V_{i\alpha}(x) = 0 \quad \text{for all players } i = 1, \dots, N. \quad (2.43)$$

Furthermore, let us also assume that V also leaves the *faces* of Δ invariant as well, in the sense that any individual trajectory $x_i(t)$ that begins at some face of Δ_i always remains in said face. The reason for this assumption is that a player who does not employ a particular action $\alpha \in \mathcal{A}_i$ with positive probability $x_{i\alpha} > 0$ has no way to gain information about the rewards that the action would yield. As such, there is no a priori reason that an adaptive learning rule would induce the user to sample it: in effect, such a learning rule would either fail to rely on readily observable information or would not necessarily be a very simple one.²⁷

This shows that the components $V_{i\alpha}(x)$ must vanish if $x_{i\alpha} = 0$, so if we set $V_{i\alpha}(x) = x_{i\alpha} \tilde{V}_{i\alpha}(x)$, we obtain the orthogonality condition $\sum_{\alpha}^i x_{i\alpha} \tilde{V}_{i\alpha}(x) = 0$. Thus, if we fix some function $v_i(x)$, any such vector field \tilde{V} may be written in the form $\tilde{V}_{i\alpha}(x) = v_{i\alpha}(x) - v_i(x)$ where the $v_{i\alpha}$ necessarily satisfy the condition $v_i(x) = \sum_{\alpha}^i x_{i\alpha} v_{i\alpha}(x)$. In this way, we obtain:

and learning dynamics.

$$\frac{dx_{i\alpha}}{dt} = x_{i\alpha} (v_{i\alpha}(x) - v_i(x)). \quad (2.44)$$

The similarity between (2.44) and the replicator equation (2.41) is obvious; all that remains is to link the functions $v_{i\alpha}(x)$ to the game's payoffs $u_{i\alpha}(x) \equiv u_i(x_{-i}; \alpha)$. Clearly, since (2.44) strengthens those actions $\alpha \in \mathcal{A}_i$ which have a larger than average $v_{i\alpha}$ component, the $v_{i\alpha}$ must factor through an increasing function of the payoffs $u_{i\alpha}$ of the game or, at the very least, to be "positively correlated" with them (Sandholm, 2001). So, the simplest choice is to just take $v_{i\alpha} = u_{i\alpha}$,²⁸ in which case we are again reduced to the replicator dynamics:

Replicators revisited.

$$\frac{dx_{i\alpha}}{dt} = x_{i\alpha} (u_{i\alpha}(x) - u_i(x)). \quad (2.41)$$

²⁵ Strictly speaking, this assumption requires V to be defined on an open neighborhood of Δ in \mathbb{R}^A but this inessential technicality need not concern us here.

²⁶ Since Δ is an affine subset of \mathbb{R}^n , all the subspaces $T_x \Delta$ will be isomorphic – for $x \in \text{Int}(\Delta)$ at least. As a result, we will not need to carry around the base point x in the notation for Z .

²⁷ In the probabilistic interpretation of Nash games, $x_{i\alpha}$ only represents the probability that a player will actually employ $\alpha \in \mathcal{A}_i$ in a given instance of the game, so a player does not really know the rewards to all actions that are present in x_i . The usual way to get around this obstacle is to consider a learning algorithm with two time scales: in the fast time scale, the users' probability of employing a strategy is roughly constant, so they are able sample all the strategies which they employ with positive probability in order to gain information about them and, then, in the slower time scale, the probabilities themselves evolve according to (2.42) – see also Leslie and Collins (2003) and Marsili et al. (2000).

²⁸ Or up to a proportionality constant: $v_{i\alpha} = \lambda_i u_{i\alpha}$; see Chapters 4, 5 and 6.

The *BNN* dynamics.

We thus see that the replicator dynamics emerge as a particularly simple model for evolution and learning. This does not mean that there are no other learning/evolutionary dynamics: a particularly well-behaved example from the theory of learning is the *Brown-von Neumann-Nash dynamics* (*BNN*):

$$\frac{dx_{i\alpha}}{dt} = \psi_{i\alpha}(x) - \psi_i(x) \quad (2.45)$$

where $\psi_{i\alpha}(x)$ denotes the “excess payoff”:

$$\psi_{i\alpha}(x) = [u_i(x) - u_{i\alpha}(x)]^+ = \max\{u_i(x) - u_{i\alpha}(x), 0\}. \quad (2.46)$$

These dynamics are very interesting in themselves (see e.g. [Brown and von Neumann, 1950](#); [Sandholm, 2001, 2005](#)) but they are somewhat less “elementary” because they incorporate more intricate assumptions than the replicator dynamics (for instance, the *BNN* dynamics assume that players have knowledge of the rewards to strategies that they never employed). For these reasons, the replicator dynamics will be our main model for the study of rationality in dynamically evolving environments.

2.3.2 Entropy and Rationality

Steady states of the replicator dynamics.

So, *what are the rationality properties of the replicator dynamics (2.41)*? An immediate observation concerning (2.41) is that Nash equilibria are stationary: if q is a Nash equilibrium, then Wardrop’s principle (2.21, 2.23) gives $u_{i\alpha}(q) = u_i(q)$ for all $\alpha \in \mathcal{A}_i$ with $q_{i\alpha} > 0$. Nonetheless, the same holds for *all* q' with equal payoffs along the actions that they employ with positive probability, and these profiles are not necessarily Nash (in the terminology of [Sandholm \(2001\)](#), this means that the replicator dynamics are “complacent”).²⁹ Consequently, the issue at hand is whether the replicator dynamics manage to single out Nash equilibria among other stationary states.

To phrase this issue in more precise terms, we will need some general definitions from the theory of dynamical systems. Recall first that in its most general form, a (continuous) dynamical system running over $[0, \infty)$ in some metric space M is defined in terms of an (continuous) *evolution function* $\theta : M \times [0, \infty) \rightarrow M$ which satisfies the consistency conditions:

$$\theta(x, 0) = x \quad (2.47a)$$

$$\theta(x, t + s) = \theta(\theta(x, t), s) \quad (2.47b)$$

for all $x \in M$ and for all $s, t \geq 0$.³⁰ We then have:

Dynamic stability.

Definition 2.14 (Notions of Stability). Let $\theta : M \times [0, \infty) \rightarrow M$ be a continuous dynamical system running over $[0, \infty)$ in the metric space M , and let $q \in M$. We will then say that:

1. q is an *ω -limit* of the trajectory $x(t) \equiv \theta(x, t)$ when $\lim x(t_k) = q$ for some increasing sequence of times $t_k \rightarrow \infty$.
2. q is *locally attracting* when there exists an open neighborhood U of q such that $\lim_{t \rightarrow \infty} \theta(x, t) = q$ for all $x \in U$.
3. q is *Lyapunov stable* when, for each neighborhood U of q , there exists a neighborhood V of q such that $\theta(x, t) \in U$ for all $t \geq 0$, whenever $x \in V$.

²⁹ However, it is easy to see that every *interior* stationary state of (2.41) must be at Nash equilibrium.

³⁰ In other words, the path $\theta^x \equiv \theta(x, \cdot)$ represents the solution orbit which starts at x at time $t = 0$

4. q is *locally asymptotically stable* when it is Lyapunov stable and attracting; if q attracts *all* initial conditions $x \in M$, then q will be called *globally asymptotically stable*.

In view of the above definition, we may now rephrase our original rationality question more precisely: which stationary states of the replicator dynamics (2.41) are attracting and/or stable? In particular, what kind of stability do Nash equilibria exhibit (if at all)?

Our most important tool in this direction will be the *relative entropy* (also known as the *Kullback-Leibler divergence*) between two strategies $q_i, x_i \in \Delta_i$:

The relative entropy function.

$$H_{q_i}(x_i) \equiv d_{\text{KL}}(q_i, x_i) = \sum_{\alpha \in \text{supp}(q_i)} q_{i\alpha} \log \frac{q_{i\alpha}}{x_{i\alpha}}, \quad (2.48)$$

where the sum is taken over the support of q_i : $\text{supp}(q_i) = \{\alpha : q_{i\alpha} > 0\}$. Of course, the entropy is finite only when x_i employs with positive probability all $\alpha \in \mathcal{A}_i$ that are present in q_i ; i.e. the domain of H_{q_i} is $\Delta_{q_i} \equiv \{x_i \in \Delta_i : q_i \ll x_i\}$, where “ \ll ” denotes absolute continuity of measures.³¹ All the same, it will matter little if we extend H_{q_i} continuously to all of Δ_i by setting $H_{q_i} = \infty$ outside Δ_{q_i} , so we will occasionally act as if H_{q_i} were defined over all of Δ_i .

Technicalities aside, the significance of the relative entropy lies in that it measures distance in probability space. Indeed, even though it is not a distance function per se (it fails to be symmetric and does not satisfy the triangle inequality), it is positive definite and strictly convex (Weibull, 1995, pp. 95–100). More importantly, if we fix a trajectory $x(t)$ of the replicator dynamics (2.41), a simple differentiation gives:

$$\begin{aligned} \frac{dH_{q_i}}{dt} &\equiv \sum_{j,\alpha} \frac{\partial H_{q_i}}{\partial x_{j\alpha}} \frac{dx_{j\alpha}}{dt} = \sum_{\alpha}^i \left(-\frac{q_{i\alpha}}{x_{i\alpha}} \right) x_{i\alpha} (u_{i\alpha}(x) - u_i(x)) \\ &= u_i(x) - \sum_{\alpha}^i q_{i\alpha} u_{i\alpha}(x), \end{aligned} \quad (2.49)$$

modulo the assumption that $x(t) \in \text{Int}(\Delta)$ to avoid inessential technicalities.

In Nash games, the convex combination $\sum_{\alpha}^i q_{i\alpha} u_{i\alpha}(x)$ is simply the payoff $u_i(x_{-i}; q_i)$ that the i -th player receives when playing $q_i \in \Delta_i$ against his opponents’ strategy $x_{-i} \in \Delta_{-i}$. This shows that the time derivative $\dot{H}_{q_i} \equiv dH_{q_i}/dt$ is controlled by the sign of the difference $u_i(x) - u_i(x_{-i}; q_i)$, so if we can estimate the latter, we will be able to get a good idea of the evolution of H_{q_i} (and, more importantly, the replicator trajectories $x(t)$). We can do this easily when the strategy q_i is a (strictly) dominated one:

Theorem 2.15 (Samuelson and Zhang, 1992). *Let $\mathfrak{G} \equiv \mathfrak{G}(\mathcal{N}, \Delta, u)$ be a Nash game, and let $x(t)$ be an interior solution orbit of the replicator dynamics (2.41). If the strategy $q_i \in \Delta_i$ is not rationally admissible for player i , then:*

Dominated strategies become extinct.

$$\lim_{t \rightarrow \infty} d_{\text{KL}}(q_i, x_i(t)) = \infty. \quad (2.50)$$

In particular, dominated strategies become extinct in the long run: if $\alpha \in \mathcal{A}_i$ is dominated, then $\lim_{t \rightarrow \infty} x_{i\alpha}(t) = 0$.

Proof. The proof of this proposition is relatively straightforward, but the technique will be important for our stochastic considerations in Chapter

³¹ Recall that a measure μ is called *absolutely continuous* with respect to ν ($\mu \ll \nu$) when every ν -null set is also μ -null. So, if q_i and x_i are interpreted as probability measures over \mathcal{A}_i , then $q_i \ll x_i$ means that $q_{i\alpha} = 0$ whenever $x_{i\alpha} = 0$ – equivalently, $\text{supp}(q_i) \subseteq \text{supp}(x_i)$.

4 (specifically, Proposition 4.1). Also, we will only show here that if q_i is strictly dominated, then $d_{\text{KL}}(q_i, x_i(t)) \rightarrow \infty$; the general case for rationally inadmissible strategies follows by induction on the rounds of elimination of dominated strategies (see also Theorem 4.3), while the special case for pure dominated strategies follows by noting that the relative entropy H_{q_i} reduces to $d_{\text{KL}}(\alpha, x_i) = -\log x_{i\alpha}$ when $q_i = \alpha \in A_i$ is pure.

So, assume that $q_i \prec q'_i$ and consider the difference $\dot{H}_{q_i}(x_i) - \dot{H}_{q'_i}(x_i)$. If $x(t)$ is an interior trajectory of the replicator dynamics, then (2.49) yields:

$$\dot{H}_{q_i}(t) - \dot{H}_{q'_i}(t) = \sum_{\alpha} (q'_{i\alpha} - q_{i\alpha}) u_{i\alpha}(x(t)) = u_i(x_{-i}(t); q'_i) - u_i(x_{-i}(t); q_i). \quad (2.51)$$

Since $q_i \prec q'_i$, it follows that the difference $u_i(x_{-i}; q'_i) - u_i(x_{-i}; q_i)$ will be positive for all $x_{-i} \in \Delta_{-i}$ and, since $\Delta_{-i} = \prod_{j \neq i} \Delta_j$ is compact, it will actually be bounded away from zero by some constant $m > 0$. We thus obtain:

$$\dot{H}_{q_i}(t) \geq \dot{H}_{q_i}(t) - \dot{H}_{q'_i}(t) \geq m > 0, \quad (2.52)$$

which implies that $\lim_{t \rightarrow \infty} H_{q_i}(x_i(t)) = \infty$. \square

The Kullback-Leibler entropy actually extends to an aggregate distance measure over all of Δ by summing over all the players of the game:

$$H_q(x) \equiv \sum_{i \in \mathcal{N}} H_{q_i}(x_i) = \sum_{\alpha \in \text{supp}(q)} q_{i\alpha} \log \frac{q_{i\alpha}}{x_{i\alpha}}, \quad (2.53)$$

where, as before, $\text{supp}(q) = \{\alpha : q_{i\alpha} > 0\}$. Unlike its player-specific summands H_{q_i} , the temporal evolution of H_q is controlled by the *aggregate* payoff difference $\sum_i (u_i(x) - \sum_{\alpha} q_{i\alpha} u_{i\alpha}(x))$, an expression which leads to more “global” results:

The “folk theorem”.

Theorem 2.16 (The Folk Theorem). *Let q be an ESS of an evolutionary game \mathfrak{G} .*

1. *If \mathfrak{G} is a single-population random matching game, then q is locally asymptotically stable in the replicator dynamics (2.41) and attracts all replicator trajectories which start at finite Kullback-Leibler distance from q ; in particular, if q is an interior ESS, then it attracts every interior solution orbit of (2.41).*
2. *If \mathfrak{G} is a multi-population random matching game, then q is locally asymptotically stable in the replicator dynamics (2.41); conversely, every locally asymptotically stable state of (2.41) is an ESS of \mathfrak{G} . In other words, the asymptotically stable states of the replicator dynamics (2.41) coincide precisely with the strict equilibria of \mathfrak{G} (Proposition 2.12).*

This theorem has been proven in a variety of different contexts and is affectionately known as the “folk theorem” of evolutionary game theory (for a more detailed account and proofs, see Weibull, 1995, pp.100–112).³² In a sense, it provides the justification that we have been looking for all along (at least, from an evolutionary perspective): *if species evolve according to the simple growth equation (2.40), then they converge to an ESS; alternatively, if the players of a Nash game follow the myopic learning scheme (2.41), then the only states to which they might be attracted to are the strict equilibria of the underlying game.*

This observation is the beating heart of this thesis: from this point on, our main goal will be to extend it to stochastic environments where the players’ evolutionary course is constantly subject to random perturbations.

³² Some authors append to the “folk theorem” an assortment of related results – for instance, that if a state is the ω -limit of an interior replicator trajectory, or if it is Lyapunov stable, then it is necessarily at Nash equilibrium (Bomze, 1986; Nachbar, 1990; Weibull, 1995).

3

ELEMENTS OF STOCHASTIC ANALYSIS

IF one takes the publication of Itô's seminal papers (Itô, 1944, 1946, 1951) as the starting point of stochastic analysis, one sees that it has been around for roughly the same amount of time as game theory. However, unlike game theory, stochastic calculus has grown as a branch of probability theory, inheriting in this way the robust structure of the latter and a universal conceptual consensus which is absent in the former. Thanks to this legacy, stochastic analysis has evaded the myriads of different (and often conflicting) definitions that have plagued game theory, and it has been blessed with a number of excellent textbooks, all consistent in terms of nomenclature.

As such, we will have no need to fix an extensive amount of notation or terminology, and the few fundamentals presented in this chapter only appear here for the sake of completeness. Our approach will be largely based on the brilliant books by Kuo (2006) and Øksendal (2007) to whom we refer the interested reader for more details – we are only scratching the surface here.

MOTIVATION In a nutshell, the object of stochastic analysis is to study differential equations with stochastic, “noise” terms (white or otherwise colored), such as the Langevin equation:

$$\frac{dx}{dt} = b(x, t) + \sigma(x, t)\eta(t) \quad (3.1)$$

where $\eta(t)$ represents the “noise” process.

What are the properties that such a stochastic process should have based on our intuitive interpretation of what constitutes (white) noise? At the very least, one must assume that 1) η is *unbiased*: $\langle \eta(t) \rangle = 0$ for all $t \geq 0$ (where $\langle \cdot \rangle$ denotes expectation); 2) that $\eta(t_1)$ is independent from $\eta(t_2)$ for $t_1 \neq t_2$ (i.e. $\langle \eta(t_1), \eta(t_2) \rangle = \delta(t_2 - t_1)$ in physics notation); and 3) that $\eta(t)$ is *stationary* in the sense that $\eta(t)$ and $\eta(t + \tau)$ are identically distributed for all $\tau > 0$.

As Øksendal (2007, p. 21) points out, the problem with this approach is that there is *no* “reasonable” stochastic process satisfying all these requirements, because *any* such process would not have continuous paths.¹ Itô's way around this obstacle was to rewrite (3.1) in discrete time as:

$$X(t_{k+1}) - X(t_k) = b(X(t_k), t_k)\delta_k + \sigma(X(t_k), t_k)[W(t_{k+1}) - W(t_k)], \quad (3.2)$$

where $\delta_k = t_{k+1} - t_k$ is the discrete time step and $W(t_k)$ is “some” stochastic process which somehow reflects our “white noise” postulates, but is also “regular” enough to yield reasonable results when we take the limit $\delta_k \rightarrow 0$.²

¹ A physicist would expect this: if one chooses a number in $(0, 1)$ randomly every τ units of time, then one could hardly expect a continuous path as $\tau \rightarrow 0$.

² It is important to note here that the difference between the last term of (3.2) and the alternative discretization $\sigma(X(t_k), t_k)\eta(t_k)\delta_k$ is crucial; the latter actually takes us nowhere (Kuo, 2006).

Of course, this program is much easier to state than to actually carry through. The first thing that needs to be done is to define the stochastic process $W(t)$ whose increments $\delta W_k = W(t_{k+1}) - W(t_k)$ will be used to represent white noise in continuous time. Unsurprisingly, physicists are intimately familiar with this process: it is none other than *Brownian motion* (known also as a *Wiener process*), and its construction (along with some of its properties) will be the subject of Section 3.1.

and the need for
stochastic integrals.

The second thing we need to do is to make rigorous sense of what “taking the limit” to continuous time really means (for example, in replacing the sum $\sum_k \sigma(X(t_k))\delta W_k$ with an integral). Essentially, this boils down to establishing an integration theory (in the Riemann-Stieltjes sense) for almost-nowhere differentiable integrators, i.e. to define integrals of the form:

$$I_g[f](t) \equiv \int_0^t f(s) dg(s), \text{ where } g \text{ is non-differentiable.} \quad (3.3)$$

This theory of *stochastic integration* was developed in a series of pioneering papers by Itô (1944, 1946, 1951), and was followed by Stratonovich (1966) who developed an “anticipative” version of stochastic integration, more useful in physics where noise is not always memoryless.³ We give a (very) rough sketch of this construction and the resulting calculus in Section 3.2, and then tie this material back to stochastic differential equations in Section 3.3.

3.1 BROWNIAN MOTION

Brownian motion is one of those concepts which is easy to treat informally but very hard to define rigorously. From a physical (and more intuitive) point of view, Brownian motion describes the movement of particles that wade through some medium and are constantly kicked around by random collisions with the molecules of the medium. For instance, the movement of pollen grains suspended in liquid (R. Brown’s original observation), smoke molecules in the air, or the diffusion of ink in water are all instances of Brownian motion.

Brownian motion as a
stochastic process

To capture this randomness on a formal level, recall first the definition of a stochastic process: given some underlying probability space (Ω, \mathbf{P}) , a stochastic process indexed by $t \in [0, \infty)$ is just a \mathbf{P} -measurable function $X : \Omega \times [0, \infty) \rightarrow \mathbb{R}$. Following standard notation and terminology, we will often drop the argument $\omega \in \Omega$ and write $X(t)$ for the random variable $X(\cdot, t) : \Omega \rightarrow \mathbb{R}$; we will only keep ω in reference to the *sample paths* $X^\omega \equiv X(\omega, \cdot) : [0, \infty) \rightarrow \mathbb{R}$ of X . In this way, Brownian motion in n dimensions may be described in terms of a stochastic process $W(t) = (W_1(t), \dots, W_n(t))$ that evolves in \mathbb{R}^n .

However, what should we assume about the properties of this process based on our empirical observations of Brownian motion? First of all, Brownian paths must be continuous: a particle cannot teleport from one place to another. Secondly, assuming a uniform medium, the *increments* $W(t + \tau) - W(t)$ should be independent of $W(t)$ for all $\tau > 0$: the position of the particle at time t should have no impact on what kind of stochastic jolts it will receive there. Finally, from Einstein’s work on molecular diffusions (Einstein, 1905, 1956) and the teachings of the Central Limit Theorem for random walks in discrete time (which are easier to analyze), we expect the increments $W(t + \tau) - W(t)$ to be normally distributed with variance proportional to $\sqrt{\tau}$.

³ For example, in climate models, a “slow” process is coupled to a “fast”, chaotic process which is usually modelled as noise. However, albeit chaotic, the faster process is still deterministic and, hence, its past retains a degree of correlation with its future evolution.

Bearing all this in mind, we are led to:

Definition 3.1 (Brownian motion). A stochastic process W in \mathbb{R}^n will be called an n -dimensional *Brownian motion* (or a *Wiener process*) starting at $x \in \mathbb{R}^n$ if:

Brownian motion.

1. $\mathbf{P}(W(0) = x) = 1$. By convention, if no initial condition $x \in \mathbb{R}^n$ is specified, $W(t)$ will be assumed to start at 0.
2. W has *independent increments*: for any collection of times t_1, t_2, \dots, t_k with $0 \leq t_1 < t_2 < \dots < t_k$, the random variables:

$$W(t_1), W(t_2) - W(t_1), \dots, W(t_k) - W(t_{k-1}) \text{ are independent.} \quad (3.4)$$

3. W has *stationary, normally distributed increments*: for any $t \geq 0$, $\tau > 0$, and for every Borel $U \subseteq \mathbb{R}^n$:

$$\mathbf{P}(W(t + \tau) - W(t) \in U) = \frac{1}{(2\pi\tau)^{n/2}} \int_U e^{-x^2/2\tau} dx. \quad (3.5)$$

4. Almost every sample path $W^\omega \equiv W(\cdot, \omega)$ of W is continuous, i.e.:

$$\mathbf{P}\{\omega : W^\omega \text{ is continuous}\} = 1. \quad (3.6)$$

Remark. There is an important point to note here. The assumption that the Brownian paths W^ω are continuous was the first one that we made and, from a physical point of view, it is the most fundamental one. However, it turns out that it is mathematically *redundant*: if a stochastic process \tilde{W} satisfies properties (2) and (3), then the paths of \tilde{W} will be almost surely continuous because of Kolmogorov's continuity theorem (see below). In fact, properties (1)–(3) characterize Brownian motion *completely* (Knight, 1981).

As an easy consequence of this definition, we obtain some fundamental properties of Brownian motion that will be used freely in the rest of the text (see also Kuo, 2006 or Øksendal, 2007):

Proposition 3.2. Let $W(t) = (W_1(t), \dots, W_n(t))$ be an n -dimensional Brownian motion starting at $x \in \mathbb{R}^n$. Then:

Properties of Brownian motion.

1. $W(t)$ is a Gaussian process.⁴ In particular, for all $s, t \geq 0$:

$$\mathbf{E}_x [W(t)] = x \quad (3.7a)$$

$$\mathbf{E}_x [(W(t) - x)(W(s) - x)] = n \min\{s, t\}. \quad (3.7b)$$

2. The components $W_j(t)$ of $W(t)$ are independent Brownian motions.
3. For any $t_0 \geq 0$, the shifted process $W(t) - W(t_0)$ is a Brownian motion.
4. For any $\lambda > 0$, the rescaled process $W(\lambda t) / \sqrt{\lambda}$ is a Brownian motion.

The proof of this proposition is trivial, so we prefer to omit it. However, what is not at all trivial is to show that Brownian motions really do exist: after all, it is far from obvious that there exists some stochastic process $W(t)$ satisfying the stringent postulates of Definition 3.1. There are three main ways to go about such a construction, each very deep and intuitive, but also involving a fair

Constructions of Brownian motion.

⁴ Recall that a process $X(t)$ is called Gaussian if, for all $0 \leq t_1 \leq \dots \leq t_k$, the random variable $Z(t) = (X(t_1), \dots, X(t_k))$ follows a (multi)normal distribution (Øksendal, 2007).

bit of technical craftsmanship. Since we will not need every detail involved, we will only present an outline of these constructions here, leaving the rest to dedicated sources such as [Hida \(1980\)](#), [Knight \(1981\)](#), [Revuz and Yor \(1999\)](#) or [Kuo \(2006\)](#).

Wiener's construction.

The first construction (Wiener's original approach) was to work backwards from the space $\Omega = C^0[0, 1]$ of continuous paths $\omega : [0, 1] \rightarrow \mathbb{R}$ and try to endow it with a probability measure that would "pick up" the Brownian paths (one could work directly with the entire domain $[0, \infty)$ but the compactness of $[0, 1]$ makes things easier). The starting point of Wiener's method was to establish a probability measure in this very large space by first assigning probabilities to the "cylinder sets":

$$C \equiv C(t_1, \dots, t_k; U) \equiv \{\omega \in C^\infty[0, 1] : (\omega(t_1), \dots, \omega(t_k)) \in U\}, \quad (3.8)$$

where $0 < t_1 < \dots < t_k \leq 1$ and U is a Borel subset of \mathbb{R}^k . So, in keeping with the Gaussian postulate of Definition 3.1, we let:

$$\mathbf{P}(C) = \int_U \prod_{j=1}^k \frac{1}{\sqrt{2\pi(t_j - t_{j-1})}} \exp \left[-\frac{(x_j - x_{j-1})^2}{2(t_j - t_{j-1})} \right] dx_1 \cdots dx_k, \quad (3.9)$$

which, if we take $U = \prod_j [a_j, b_j]$ for transparency, is just another way of saying:

$$\begin{aligned} \mathbf{P} \{ \omega(t_1) \in [a_1, b_1], \dots, \omega(t_k) \in [a_k, b_k] \} \\ = \int_{a_1}^{b_1} \cdots \int_{a_k}^{b_k} \prod_{j=1}^k p(x_{j-1}, x_j, t_j - t_{j-1}) dx_1 \cdots dx_k, \end{aligned} \quad (3.10)$$

where x_0 and t_0 are both taken to be 0 and p is the Gaussian transition kernel:

$$p(x, y, t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}}. \quad (3.11)$$

[Wiener \(1923\)](#) proved that the probability assignment $C \mapsto \mathbf{P}(C)$ is a well-defined σ -additive mapping on cylindrical sets and thus extends to the σ -algebra generated by all such sets.⁵ Since this algebra coincides with the Borel algebra $\Omega \equiv C^0[0, 1]$ induced by the topology of uniform convergence (see e.g. [Kuo, 2006](#), p. 24), it follows that \mathbf{P} gives rise to a probability measure on Ω . Hence, if we set $W(t, \omega) = \omega(t)$, $\omega \in C^0[0, 1]$, the construction of \mathbf{P} shows that $W(t)$ satisfies the postulates of Definition 3.1 ([Kuo, 2006](#), Theorem 3.1.2).

Kolmogorov's general method.

This method has the benefit of being very concise and elegant but it also runs "backwards", in the sense that it does not provide a way to construct Brownian paths; instead, it simply endows the space of *all* continuous paths with a probability measure which conforms to what Brownian paths should "look like". Another approach which is similar (at least in spirit) can be obtained by being less restrictive at the beginning and try to establish a Brownian measure on the space $\mathbb{R}^{[0, \infty)}$ of *all* paths $\omega : [0, \infty) \rightarrow \mathbb{R}$ instead (and not only on the continuous ones). What makes things simpler in this case is that the probability assignment (3.9) is *consistent* (that is, $\mathbf{P}(C)$ does not depend on how C is represented as a cylindrical set) and thus, by the Kolmogorov extension theorem ([Shiryayev, 1995](#), p. 167), we immediately recover a probability measure \mathbf{P} on $\mathbb{R}^{[0, \infty)}$ with all the desired Gaussian properties.

⁵ This means that $\mathbf{P} \left(\bigcup_{j=1}^{\infty} A_j \right) = \sum_{j=1}^{\infty} \mathbf{P}(A_j)$ for every countable family of disjoint sets A_j .

Indeed, if we set $W(t, \omega) = \omega(t)$, conditions (1)–(3) of Definition 3.1 are easy to verify; however, having admitted discontinuous paths, condition (4) takes a bit more work. To that end, note that $W(t)$ satisfies the regularity condition:

$$\mathbf{E} [(W(t) - W(s))^4] = 2(t - s)^2, \quad (3.12)$$

which tells us that, “on average”, the variations of $W(t)$ are Lipschitz. Therefore, by Kolmogorov’s continuity theorem (Kuo, 2006) (which is all about establishing the continuity of Lipschitz processes), we know that $W(t)$ does have a continuous realization, i.e. \mathbf{P} -almost every path $\omega(t)$ is continuous.

This procedure is a bit deeper than Wiener’s original approach because it relies on two fundamental theorems of probability theory. However, it is still not as intuitive as one would hope, mostly because it extols the Gaussian properties of Brownian motion and not the “random kicks” mechanism. The latter is captured remarkably well in Lévy’s interpolation method which produces Brownian paths by interpolating between Gaussian jumps.

The details of the method can be found in Hida (1980) or Kuo (2006), but its essence is easy to describe. Indeed, let ξ_n be a sequence of independent Gaussian random variables of mean 0 and variance 1, defined on some sample space Ω . Then, as a first step, we define the stochastic process $W_1(t)$:

$$W_1(t) = \xi_1 t, \quad 0 \leq t \leq 1, \quad (3.13)$$

that is, W_1 interpolates linearly between the points $(0, 0)$ and $(1, \xi_1)$. For W_2 , keep the endpoints 0, 1 fixed, perturb $W_1(1/2)$ by $2^{-1}\xi_2$ (to keep the correct variance) and define $W_2(t)$ by linearly interpolating between these values. So, by perturbing midpoints in this fashion, we get a sequence of stochastic processes W_n which, by construction, converges in variance for every dyadic rational t :

$$W_n(t) \xrightarrow{L^2(\Omega)} W(t) \text{ for every } t \text{ of the form } t = k/2^m, k = 1, \dots, 2^{m-1}. \quad (3.14)$$

From the definition of $W_n(t)$, we get $\mathbf{E}[W(t)] = 0$ and $\mathbf{E} [(W(t) - W(s))^2] = |t - s|$ for all dyadic rationals s, t (a straightforward, if somewhat tedious, calculation suffices to show that). This implies that the mapping $t \mapsto W(t)$ is continuous in $L^2(\Omega)$ so, with the dyadic rationals being dense in $[0, 1]$, we may extend $W(t)$ continuously over all of $[0, 1]$. The process $W(t)$, $t \in [0, 1]$ that we obtain in this way is called the Lévy limit and, with some more work, we can show that it satisfies all the postulates of Brownian motion (Hida, 1980).

Lévy’s construction is the one closest to our intuitive feel of how Brownian paths take shape as a series of random jolts. Given the astonishing richness of this phenomenon, we could remain on the topic of random walks and their properties nigh indefinitely, but since this discussion would take us too far afield, we refer the reader to any one of the excellent accounts on the subject listed in this chapter, Kuo (2006) being perhaps the most comprehensive one.

Lévy’s method
of interpolation.

3.2 STOCHASTIC INTEGRATION

Having established the existence and basic properties of Brownian motion, the second step in our program is to use it as a Riemann-Stieltjes integrator, i.e. to make sense of the “continuous-time” limit:

$$\sum_k f(t_k) \delta W_k \rightarrow \int f(s) dW(s), \quad (3.15)$$

where $f(s)$ is some stochastic process.

Problems in defining stochastic integrals.

A most natural question is the following: since we are essentially talking about a sum of random variables, can't we simply take the standard Riemann sums approach and prove that the resulting sum converges in some appropriate sense (probability, L^2 or whatnot)? Unfortunately, the short answer is "no", a slightly longer answer being that the limit depends crucially on how the Riemann sums are sampled.

Differences between left and right sums.

To understand this, consider the "stochastic integral" $\int_0^t W(s) dW(s)$. Indeed, if we take some partition $0 = t_0 < t_1 < \dots < t_k = t$ of $[0, t]$, the left and right Riemann sums of $W(s)$ with respect to $\{t_j\}$ will be:

$$L_k = \sum_{j=1}^k W(t_{j-1})(W(t_j) - W(t_{j-1})) \quad (3.16a)$$

$$R_k = \sum_{j=1}^k W(t_j)(W(t_j) - W(t_{j-1})). \quad (3.16b)$$

As a result, the difference between the two sums will be:

$$R_k - L_k = \sum_{j=1}^k (W(t_j) - W(t_{j-1}))^2, \quad (3.17)$$

and since $\mathbf{E}[(W(t_j) - W(t_{j-1}))^2] = t_j - t_{j-1}$, we readily get:

$$\mathbf{E}(R_k - L_k) = t, \quad (3.18)$$

which is not small, no matter how large k becomes.⁶ In other words, it makes a lot of difference (t to be exact) whether the stochastic integral is defined in terms of the left endpoints of the partition or of the right ones – or of some other arbitrary choice $t_j^* \in [t_{j-1}, t_j]$.

For reasons that we will explain later, the two most useful choices for sampling the integrand $f(s)$ of the (yet to be defined) stochastic integral $\int_0^t f(s) dW(s)$ are:

1. The left endpoint $t_j^* = t_{j-1}$ which leads to the *Itô stochastic integral*.
2. The midpoint $t_j^* = (t_{j-1} + t_j)/2$ which leads to the *Stratonovich integral*.

Adaptability and stochastic integrands.

We will first focus on the construction of the Itô integral, leaving the Stratonovich version for later. All the same, before proceeding any further, one must also ask: how general can the stochastic process $f(s)$ be? In both cases, a (mostly technical) requirement is that f must be square integrable on average, in the sense that $\mathbf{E} \left[\int_0^t f(s)^2 ds \right] < \infty$. More importantly however, it turns out that f must also be *adapted* to the Brownian motion (more properly to its *filtration*), i.e the value of $f(s)$ must be decidable by looking at the history of $W(s')$ for all times $s' \leq s$ (for example, $f(s) = s^2 + \max_{s' \leq s} \{W(s')\}$ is fine, whereas $f(s) = W(s+1)$ is not).

Itô integration step-by-step:

The precise definition of adaptability or of the class of processes that are admissible as stochastic integrands need not concern us here.⁷ Instead, we prefer to give a brief account of how the stochastic integral is actually brought to life. This is done in three steps: first, one defines the stochastic integral of (adapted) "step processes", then one approximates the integrand by a sequence of step processes, and, finally, one shows that this approximation behaves well with respect to the step integral.

⁶ Of course, this argument is a heuristic one: even though $\mathbf{E}(R_k - L_k) = t > 0$, we could still have $R_k - L_k \rightarrow 0$ almost surely. See [Kuo \(2006, Theorem 4.1.2\)](#) for a stronger result.

⁷ The interested reader will find an excellent account in [Kuo \(2006\)](#).

STEP 1: STEP INTEGRALS A step process is a process of the form:

$$f(s, \omega) = \sum_{j=1}^k \xi_{j-1}(\omega) \chi_{[t_{j-1}, t_j)}(s), \quad 0 \leq s \leq t, \omega \in \Omega, \quad (3.19)$$

where $0 = t_0 < t_1 < \dots < t_k = t$ is a partition of $[0, t]$, $\chi_{[a,b]}$ denotes the indicator function of the interval $[a, b]$, and the random variables ξ_j are square integrable ($\mathbf{E}(\xi_j^2) < \infty$) and only depend on the history of $W(s)$ up to time t_j . The stochastic integral of such a process is then defined in the natural way:

$$I(f) \equiv \int_0^t f(s) dW(s) = \sum_{j=1}^k \xi_{j-1} (W(t_j) - W(t_{j-1})), \quad (3.20)$$

and it can be shown that (Kuo, 2006; Øksendal, 2007):

$$\mathbf{E} [I(f)^2] = \int_0^t \mathbf{E} [f^2(s)] ds, \quad (3.21)$$

a result which is known as the *Itô isometry* for step processes.

STEP 2: APPROXIMATION BY STEP PROCESSES The next step consists of an approximation lemma to the effect that if $f(s)$ is an admissible integrand process, then there exists a sequence of adapted step processes $f_n(s)$ such that:

$$\int_0^t \mathbf{E} [(f(s) - f_n(s))^2] ds \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (3.22)$$

The construction of such a sequence is the most technical part of the whole procedure and involves three successive approximations. First, one approximates f in the sense of (3.22) by a sequence of bounded processes, then one approximates these bounded processes by continuous ones, and, finally, one shows that these continuous processes can be approximated by step processes.

STEP 3: THE ITÔ INTEGRAL With all this said and done, the Itô integral of f is defined in terms of its step process version (3.20) and the approximation (3.22) of f by step processes. More precisely, if f_n is a sequence of step processes which converges to f in the sense of (3.22), then we define:

$$I(f) \equiv \int_0^t f(s) dW_s \equiv \lim_{n \rightarrow \infty} \int_0^t f_n(s) dW(s). \quad (3.23)$$

Thanks to the isometry (3.21), it can be shown that this definition is independent of the choice of approximating step sequences f_n and we also obtain the following properties of the Itô integral by looking at step approximations:

Proposition 3.3. *Let f, g be square integrable adapted processes. Then, the Itô integrals $I(f), I(g)$ have the following properties:*

1. *Linearity:* $I(af + bg) = aI(f) + bI(g)$.
2. *Zero mean:* $\mathbf{E}[I(f)] = 0$.
3. *Itô Isometry:* $\mathbf{E} [I(f)I(g)] = \int_0^t \mathbf{E}[f(s)g(s)] ds$.

The proof of this proposition can be found in Øksendal (2007). With some more work that involves Doob's (sub)martingale inequality (Kuo, 2006), one can also show that Itô integrals are continuous:

1. *Step Processes.*

2. *Step process approximations.*

3. *Reconvergence.*

Properties of Itô integrals.

Proposition 3.4 (Continuity of Itô Integrals). *Let f be a square integrable process and let $X(t)$ denote the Itô integral:*

$$X(t) = \int_0^t f(s) dW(s). \quad (3.24)$$

Then, the paths $X^\omega \equiv X(\cdot, \omega)$ of X are continuous almost surely.

*The Stratonovich
Integral.*

THE STRATONOVICH INTEGRAL As we have already mentioned, the Stratonovich integral corresponds to the sampling choice $t_j^* = (t_{j-1} + t_j)/2$, which suggests replacing the step integral (3.20) with:

$$\sum_{j=1}^k f(t_j^*) (W(t_j) - W(t_{j-1})). \quad (3.25)$$

By taking the limit in probability of this sum as the width $\delta_k = \max_j |t_j - t_{j-1}|$ of the partition $0 = t_0 < t_1 < \dots < t_k = t$ goes to zero, it can be shown (Kuo, 2006, Theorem 8.3.7) that one obtains a partition-independent result which we denote by:

$$S(f) \equiv \int_0^t f(s) \partial W(s), \quad (3.26)$$

where we used ∂ to distinguish the “Stratonovich differential” $\partial W(s)$ from the corresponding Itô version $dW(s)$ that appears in (3.23).⁸

The main difference between these two definitions is that the Stratonovich choice is *anticipative*: even if f is adapted to Brownian motion, the sampled value $f(t_j^*)$ cannot be determined by looking at the history of the process up to time t_j . This “forward-looking” feature of the Stratonovich integral makes it inappropriate as an integration theory for processes which are completely unpredictable, but if the future evolution of a process is not completely uncorrelated with its history (that is, if it is not a *martingale*), then Stratonovich integration is frequently a better fit.

For this reason, Itô integrals are mostly used in biology and economics where random events are inherently nondeterministic (at least, in a lax interpretation of the word), while Stratonovich integrals are used in physics to model chaotic (but deterministic) processes that evolve in very fast time scales. This Itô-Stratonovich dilemma will be crucial to our work, and we will discuss it further in the next section and in the chapters that follow.

Itô vs. Stratonovich.

3.3 ITÔ CALCULUS AND STOCHASTIC DIFFERENTIAL EQUATIONS

*Stochastic differential
equations (SDE).*

Having established a theory of stochastic integration, we are now ready to go back to our program of writing a noisy version of the Langevin equation (3.1) in a meaningful way. Indeed, we may now take the continuous time limit of (3.2) and write:

$$\int_0^t X(s) ds = \int_0^t b(X(s), s) ds + \int_0^t \sigma(X(s), s) dW(s) \quad (3.27)$$

or, if we remove the integrals for notational convenience:

$$dX(t) = b(X(t), t) dt + \sigma(X(t), t) dW(t). \quad (3.28)$$

⁸ We follow the notation of Klebaner (2005) here; other authors (notably Kuo, 2006 and Øksendal, 2007) use the notation $\circ dW(s)$ instead of $\partial W(s)$.

As opposed to (3.1) which was formulated in terms of the non-existent “white noise” derivative $\eta(t) = \frac{dW}{dt}$, the stochastic differential equation (SDE) (3.28) is well-defined as the differential representation of the stochastic integral equation (3.27).⁹ So, the two main questions that arise are:

1. Are there existence and uniqueness results for (3.28)?
2. How can the SDE (3.28) be solved?

Without further ado (or proof), we state the answer to the first question:

Theorem 3.5 (Existence and Uniqueness of Solutions to SDE’s). *Consider the SDE (3.28), $0 \leq t \leq T$, with measurable coefficients b and σ , and assume that there exist constants C, D such that:*

Existence and uniqueness.

$$|b(x, t)| + |\sigma(x, t)| \leq C(1 + |x|) \quad (3.29a)$$

and:

$$|b(x, t) - b(y, t)| + |\sigma(x, t) - \sigma(y, t)| \leq D|x - y| \quad (3.29b)$$

for all $t \in [0, T]$ and for all $x, y \in \mathbb{R}$. Then, for every initial condition $x \in \mathbb{R}$, (3.28) admits a unique adapted solution $X(t)$ with continuous paths starting at x .

Remark. It is worth noting here that the theorem also holds for initial conditions x that are random variables with finite second moment; the only further assumption that needs to be made then is that x be suitably “independent” of the Brownian filtration – see Øksendal (2007) for details and proofs.

From a theoretical standpoint, this theorem is of paramount importance because it provides the necessary foundations for the theory of stochastic differential equations. However, from a more practical point of view, our second question on how to actually solve an SDE seems to be more relevant. Of course, this question is also meaningless in complete generality, much as there is no hope to provide a solution method for every ordinary differential equation. Hence, it is best to reformulate this question in terms of antiderivatives as:

- 2’. Is there a formula that allows us to compute Itô integrals?

In contrast to ordinary integrals where the fundamental theorem of calculus reduces the problem to finding an antiderivative of the integrand, the answer to this question is, in general, no. However, the answer is positive if we restrict ourselves to integrands which are (sufficiently smooth) functions of t and $W(t)$. This celebrated result is known as *Itô’s formula* (or lemma, reflecting the modesty of its finder) and it lies at the very heart of stochastic calculus:

Theorem 3.6 (Itô’s formula; Itô, 1944). *Let f be a C^2 function. Then, the process $Y(t) = f(W(t))$ satisfies:*

Itô’s formula.

$$dY(t) = f'(W(t)) dW(t) + \frac{1}{2} f''(W(t)) dt. \quad (3.30)$$

The proof of this stochastic variant of the chain rule can be found e.g. in Kuo (2006) or Øksendal (2007), and it leads to the following formula for calculation of Itô integrals:

$$\int_0^t f'(W(s)) dW(s) = f(W(t)) - f(W(0)) - \frac{1}{2} \int_0^t f''(W(s)) ds, \quad (3.31)$$

⁹ It should be stressed however that the Wiener differential $dW(s)$ is not a “true” differential in any sense of the word; (3.28) is just a more convenient way of expressing (3.27).

an expression which reduces the calculation of the stochastic integral $I(f)$ to an antiderivative problem and an ordinary Riemann-Stieltjes integral.

To get the most mileage out of (3.30), we will need a slightly more general version, adjusted to processes of the form $Y = f(X(t))$ where $X(t)$ is itself an Itô process, i.e. it satisfies:

$$dX(t) = \mu(t) dt + \sigma(t) dW(t), \quad (3.32)$$

with μ and σ being themselves adapted. Then, if $Y(t) = f(X(t))$, we have:

$$dY(t) = \left[f'(X(t))\mu(t) + \frac{1}{2}f''(X(t))\sigma^2(t) \right] dt + f'(X(t))\sigma(t) dW(t) \quad (3.33)$$

Formal multiplication
of Itô differentials.

Even though it reduces to the standard Itô formula for $\mu = 0$ and $\sigma = 1$, the expression (3.33) might appear a bit opaque, at least at first sight. Its proof is actually only slightly more convoluted than the proof of (3.30) (see e.g. Øksendal, 2007), but, at a purely formal level, it can be “derived” by taking the Taylor expansion of f and using the “differential multiplication table”:

\times	$dW(t)$	dt
$dW(t)$	dt	0
dt	0	0

We stress here that the equality $dW(t) \cdot dW(t) = dt$ is purely formal. It can be made precise by looking at the quadratic covariation of $W(t)$, but we will avoid the temptation of diving in this issue – instead, see Klebaner (2005).

To illustrate the use of this “multiplication table”, note that, at a (purely!) formal level, the “Taylor expansion” of $Y(t) = f(X(t))$ gives:

$$dY(t) = d(f(X(t))) = f'(X(t))dX(t) + \frac{1}{2}f''(X(t))(dX(t))^2 + \mathcal{O}((dX(t))^3), \quad (3.34)$$

and, by substituting $dX(t)$ from (3.32), we easily obtain (3.33). Though purely formal, calculations of this kind will be our bread and butter in the following chapters, especially for multidimensional Itô processes. In that case, we will follow exactly the same method as above, except that we will have to complete the differential multiplication table with the entries $dW_i(t) \cdot dW_j(t) = \delta_{ij} dt$ that correspond to the independent components W_i and W_j of multidimensional Wiener processes.

Stratonovich
Differentials.

STRATONOVICH DIFFERENTIALS Having established a set of calculus rules for Itô differentials, our last task in this section will be to establish a similar set of rules for Stratonovich differentials, defined in terms of (3.25) and (3.26). As it turns out, the Stratonovich integral is related to the Itô integral via the differential relation:

$$X \partial Y = X dY + \frac{1}{2}(dX) \cdot (dY), \quad (3.35)$$

where the “differential product” $dX \cdot dY$ is calculated according to the multiplication table that we presented above and corresponds to the *quadratic covariation* $[X, Y]$ of X and Y (for details, see Kuo, 2006, Section 8.3). As a result, if we set $X = 1$ and $Y(t) = f(W(t))$ in (3.35), Itô’s formula gives:

$$\partial Y = dY = f'(W(t)) dW(t) + \frac{1}{2}f''(W(t)) dt = f'(W(t)) \partial W(t), \quad (3.36)$$

where the last step follows from (3.35) and the formal derivation:

$$\begin{aligned} f'(W(t))\partial W(t) &= f'(W(t))dW(t) + \frac{1}{2}\left(f''(W(t))dW(t) + \frac{1}{2}f'''(W(t))dt\right)dW(t) \\ &= f'(W(t))dW(t) + \frac{1}{2}f''(W(t))dt. \end{aligned} \quad (3.37)$$

In this way, we see that Stratonovich differentials behave well with respect to the chain rule of ordinary calculus, a fact which makes them easier to deal with (calculationally at least). There are other reasons that support the use of Stratonovich differentials as well: the Wong-Zakai theorem (Wong and Zakai, 1965) states that if we take a smooth approximation to a Brownian path in the SDE (3.28) and consider the corresponding approximate solution of the ordinary differential equation that arises, then this solution approximates the stochastic solution of (3.28) only if it is interpreted in the Stratonovich sense – see the next section as well. On the other hand, the future-correlated nature of the Stratonovich interpretation renders it inappropriate for many applications so, in the end, the Itô-Stratonovich dilemma can only be attacked on a case-by-case basis.

3.4 DIFFUSIONS AND THEIR GENERATORS

We end this chapter by presenting a few facts concerning an important class of stochastic differential equations:

Definition 3.7 (Itô diffusion). A (time-homogeneous) Itô diffusion in \mathbb{R}^n is a stochastic process $X(t) = (X_1(t), \dots, X_n(t))$ which satisfies a SDE of the form:

Itô diffusions

$$dX_\alpha(t) = b_\alpha(X(t))dt + \sum_{\mu=1}^m \sigma_{\alpha\mu}(X(t))dW_\mu(t), \quad \alpha = 1, \dots, n, \quad (3.38)$$

where $W(t) = (W_1(t), \dots, W_m(t))$ is a Brownian motion in \mathbb{R}^m and the drift and diffusion coefficients b and σ are both Lipschitz continuous.

The reason that such processes are called diffusions is because they are closely related to heat-type differential equations and other diffusive phenomena. Physically, this is to be expected: after all, (3.38) simply represents the spread of Brownian particles in a medium where there is a “deterministic” drift b and a “diffusive mechanism” σ . From a mathematical point of view however, the connection between diffusion processes and parabolic (heat-type) differential equations is less clear. Thus, our first step in that direction will be to consider the (infinitesimal) generator of a diffusion (Øksendal, 2007):

Definition 3.8 (Generator of a Diffusion). Let $X(t)$ be an n -dimensional (time-homogeneous) Itô diffusion. Then, the generator \mathcal{L} of $X(t)$ is defined by:

and their generators.

$$\mathcal{L}f(x) = \lim_{t \rightarrow 0} \frac{\mathbf{E}_x[f(X(t))] - f(x)}{t}, \quad (3.39)$$

for all functions f for which the limit (3.39) exists for all $x \in \mathbb{R}^n$.

Remark. In accordance with our notational conventions, \mathbf{E}_x signifies here that the process $X(t)$ has the initial condition $X(0) = x \in \mathbb{R}^n$

An important result of stochastic analysis is that the class of functions for which the RHS of (3.39) exists contains the space $C_0^2(\mathbb{R}^n)$ of twice continuously

Generators as
differential operators.

differentiable functions on \mathbb{R}^n with compact support (Øksendal, 2007, Section 7.3). When restricted to this space, \mathcal{L} can actually be expressed in terms of the drift and diffusion coefficients of $X(t)$ in a particularly suggestive way:

$$\mathcal{L}f = \sum_{\alpha=1}^n b_{\alpha}(x) \frac{\partial f}{\partial x_{\alpha}} + \frac{1}{2} \sum_{\alpha,\beta=1}^n \sum_{\mu=1}^m \sigma_{\alpha\mu}(x) \sigma_{\beta\mu}(x) \frac{\partial^2 f}{\partial x_{\alpha} \partial x_{\beta}}, \quad (3.40)$$

or, in more succinct notation:

$$\mathcal{L} = \langle b, \nabla \rangle + \frac{1}{2} \langle \sigma \sigma^T, \nabla \otimes \nabla \rangle, \quad (3.41)$$

where ∇ represents the usual differential operator $\nabla_{\alpha} = \frac{\partial}{\partial x_{\alpha}}$, and $\langle \cdot, \cdot \rangle$ denotes the usual Euclidean metric pairing: $\langle x, y \rangle = \sum_{\alpha} x_{\alpha} y_{\alpha}$.

One cannot fail to notice here the resemblance of (3.40) to Itô's formula (3.33). Indeed, if we set $Y(t) = f(X(t))$ for some $f \in C_0^2(\mathbb{R}^n)$, then (3.33) readily yields:

$$dY(t) = \mathcal{L}f(X(t)) dt + \sum_{\alpha=1}^n \sum_{\mu=1}^m \frac{\partial f}{\partial x_{\alpha}} \Big|_{X(t)} \sigma_{\alpha\mu}(X(t)) dW_{\mu}(t). \quad (3.42)$$

In other words, $\mathcal{L}f$ merely captures the temporal drift of the process $Y(t) = f(X(t))$ and, in tune with Definition 3.8, \mathcal{L} can be seen as the stochastic analogue of the deterministic time derivative $\frac{d}{dt}$.

Dynkin's formula.

A consequence of this interpretation is Dynkin's formula (Øksendal, 2007):

$$\mathbf{E}_x[f(X(t))] = f(x) + \mathbf{E}_x \left[\int_0^t \mathcal{L}f(X(s)) ds \right]. \quad (3.43)$$

The key feature of this formula (which will become evident in Chapter 5) is that it can be used to calculate the evolution of functions of the original diffusion process. Essentially, this leads to Kolmogorov's backward equation (KBE):

Kolmogorov's
backward equation.

Theorem 3.9 (Kolmogorov's backward equation). *Let $f \in C_0^2(\mathbb{R}^n)$ be a twice continuously differentiable function with compact support, and let $X(t)$ be an n -dimensional diffusion process with generator \mathcal{L} . Then, if $u(x, t) = \mathbf{E}_x[f(X(t))]$, u will satisfy Kolmogorov's backward equation:*

$$\frac{\partial u}{\partial t} = \mathcal{L}u, \quad (3.44)$$

with initial condition $u(x, 0) = f(x)$.

The "backward" denomination of (3.44) is owed to the fact that it is an equation in the "backward" variables x which specify where the process $X(t)$ started. As the name implies, there is also a *forward* version of (3.44) which describes the forward evolution of the diffusion.

To be more precise, let $p(y, t) \equiv p_x(y, t)$ be the transition probability density with which a particle starting at x finds itself at y at time t . We will then have:

$$\mathbf{E}_x[f(X(t))] = \int_{\mathbb{R}^n} p(y, t) f(y) dy \quad \text{for all } f \in C_0^2(\mathbb{R}^n), \quad (3.45)$$

and, by employing Dynkin's formula, we easily get:

$$\int_{\mathbb{R}^n} f(y) p(y, t) dy = f(x) + \int_0^t \int_{\mathbb{R}^n} p(y, s) \mathcal{L}f(y) dy ds. \quad (3.46)$$

Thus, let \mathcal{L}^\dagger be the *adjoint generator* defined by:

$$\mathcal{L}^\dagger g(y) = - \sum_{\alpha=1}^n \frac{\partial}{\partial y_\alpha} \left[b_\alpha(y) g(y) \right] + \sum_{\alpha, \beta=1}^n \frac{\partial^2}{\partial y_\alpha \partial y_\beta} \left[\sum_{\mu} \sigma_{\alpha\mu}(y) \sigma_{\beta\mu}(y) g(y) \right]. \quad (3.47)$$

Then, if $\langle \cdot, \cdot \rangle$ denotes the L^2 inner product on \mathbb{R}^n , we will have:

$$\langle \mathcal{L}f(y), p(y, t) \rangle = \langle f(y), \mathcal{L}^\dagger p(y, t) \rangle \quad \text{for all } f \in C_0^2(\mathbb{R}^n). \quad (3.48)$$

In this way, differentiating (3.46) with respect to time finally yields:

Theorem 3.10 (Kolmogorov's forward equation). *Let $p(y, t) \equiv p_x(y, t)$ be the transition probability density of a diffusion process $X(t)$ with initial condition $X(0) = x$ and infinitesimal generator \mathcal{L} . Then, for all $y \in \mathbb{R}^n$ and all $t > 0$, $p(y, t)$ satisfies Kolmogorov's forward equation (KFE):*

Kolmogorov's forward equation.

$$\frac{\partial p}{\partial t} = \mathcal{L}^\dagger p(y, t), \quad (3.49)$$

\mathcal{L}^\dagger being the adjoint (3.47) of \mathcal{L} with respect to the L^2 inner product on \mathbb{R}^n .

Remark 1. This equation is deeply rooted in statistical physics where it is more commonly known *Fokker-Planck equation*. As with most of our treatment of stochastic analysis, we stress that we are only scratching the surface here; for an excellent account of this fascinating subject, see [Risken \(1989\)](#).

Remark 2. Kolmogorov's forward equation (3.49) finally justifies the name "diffusion process". If $X(t)$ is an n -dimensional Brownian motion, it is easy to see that its generator \mathcal{L} will be none other than the (self-adjoint) Laplacian $\frac{1}{2}\Delta = \frac{1}{2} \sum_{\alpha} \frac{\partial^2}{\partial x_\alpha^2}$. As a result, the transition probability density for particles that begin at the origin $0 \in \mathbb{R}^n$ will satisfy the heat equation:

$$\frac{\partial p}{\partial t} = \frac{1}{2}\Delta p = \frac{1}{2} \sum_{\alpha=1}^n \frac{\partial^2 p}{\partial x_\alpha^2}. \quad (3.50)$$

STABILITY AND ERGODICITY Having established the link between diffusion processes and the diffusive phenomena which they are used to model, there arises a key question: *what are the possible evolution patterns of a diffusion process?*

Types of trajectories.

In broad, qualitative terms, there are three possible types of trajectories that one might expect to observe:

1. *Stable trajectories:* the particle always stays in a neighborhood of where it started, possibly converging to a point (a *trap* of the diffusion).
2. *Recurrent trajectories:* the particle might stray far away from its starting point, but eventually returns again to its vicinity.
3. *Transient trajectories:* the particles escape to infinity.

There are many factors that affect the evolution of a diffusion, even within a class of "similar" processes: for example, Brownian motion is recurrent in dimensions 1 and 2, but it is transient for $n \geq 3$ ([Øksendal, 2007](#), pp. 125–126). To make these ideas precise, we offer the following definitions ([Arnold, 1974](#); [Bhattacharya, 1978](#); [Gikhman and Skorokhod, 1971](#)):

Definition 3.11 (Stochastic Stability). Let $X(t)$ be a diffusion process in \mathbb{R}^n . We will say that $q \in \mathbb{R}^n$ is *stochastically asymptotically stable* with respect to

Stochastic stability.

$X(t)$ when, for every neighbourhood U of q and every $\varepsilon > 0$, there exists a neighbourhood V of q such that:

$$\mathbf{P}_x \left(X(t) \in U \text{ for all } t \geq 0 \text{ and } \lim_{t \rightarrow \infty} X(t) = q \right) \geq 1 - \varepsilon \quad (3.51)$$

for every initial condition $X(0) = x \in V$.

*Recurrence
and transience.*

Definition 3.12 (Recurrence and Transience). Let $X(t)$ be a diffusion process in \mathbb{R}^n starting at $x \in \mathbb{R}^n$. We will say that $X(t)$ is *recurrent* when, for every $y \in \mathbb{R}^n$ and for every neighborhood U_y of y :

$$\mathbf{P}_x (X(t_k) \in U_y) = 1, \quad (3.52)$$

for some sequence of (random) times t_k that increases to infinity.

On the other hand, we will say that $X(t)$ is *transient* if:

$$\mathbf{P}_x \left(\lim_{t \rightarrow \infty} |X(t)| = \infty \right) = 1, \quad (3.53)$$

for every initial condition $X(0) = x \in \mathbb{R}^n$.

The important dichotomy between recurrent and transient processes has been explored, among others, by [Maruyama and Tanaka \(1959\)](#), [Khas'minskii \(1960\)](#) and [Bhattacharya \(1978\)](#). Under relatively mild conditions, it turns out that recurrence becomes synonymous with *ergodicity*: the transition probabilities of a recurrent diffusion converge in total variation to an invariant measure which represents the steady state distribution of the process.

This ergodicity property will be very important to us in Chapters 5 and 6 but, for now, we prefer to focus on the notion of stability. Much the same as in the deterministic case, stochastic asymptotic stability is often established by means of a (local) Lyapunov function:¹⁰

*The stochastic
Lyapunov method.*

Definition 3.13. Let $X(t)$ be a diffusion in \mathbb{R}^n with infinitesimal generator \mathcal{L} . We will say that f is a (local) *stochastic Lyapunov function* of the diffusion $X(t)$ with respect to the point $q \in \mathbb{R}^n$ when there exists some neighborhood U of q such that:

1. $f(x) \geq 0$ for all $x \in U$, with equality if and only if $x = q$;
2. there exists a constant $k > 0$ such that $\mathcal{L}f(x) \leq -kf(x)$ for all $x \in U$.

Whenever such a Lyapunov function exists, it is known that the point $q \in \mathbb{R}^n$ where f attains its minimum will be stochastically asymptotically stable ([Gikhman and Skorokhod, 1971](#), pp. 314–315). Thus, finding a Lyapunov function will be our principal method for establishing the stochastic stability of an equilibrium point.

¹⁰ It might be important here to recall that a Lyapunov function of a *deterministic* dynamical system of the form $\dot{x} = V(x)$, $x \in \mathbb{R}^n$, is a positive-definite function with $f(x) > 0$ unless $x = 0$ and such that $df(V(x)) < 0$ for all $x \neq 0$ – the origin was, of course, chosen for convenience. If such a function exists, then the origin is (globally) asymptotically stable (i.e. Lyapunov stable and globally attracting).

Part II

STOCHASTIC FLUCTUATIONS IN
NON-COOPERATIVE GAMES

4

STOCHASTIC PERTURBATIONS IN NASH GAMES

HAVING established the necessary background in game theory and stochastic analysis, we are finally free to tackle in more precise terms the question that we first phrased back in Chapter 2: *does rationality still emerge in the presence of stochastic perturbations that interfere with the game?*

In evolutionary games, these perturbations traditionally take the form of “aggregate shocks” that are applied directly to the population of each phenotype. This approach by Fudenberg and Harris (1992) has spurred quite a bit of interest, and there is a number of features that differentiate it from the theory of deterministic evolution that we presented in Section 2.3. To name but an example, if the variance of the shocks is low enough, Cabrales (2000) showed that dominated strategies indeed become extinct in the long run. More recently, the work of Imhof (2005) revealed that even equilibrial play arises over time but, again, conditionally on the shocks being mild enough relative to the payoff differences in the game.

The advent of stochasticity.

On the other hand, if one looks at games with a finite number of players, it is hardly relevant to consider shocks of this type, simply because there are no longer any populations to apply them to. Instead, any and all stochastic fluctuations should be reflected directly on the stimuli that incite players to play one strategy or another: their payoffs. This leads to a picture which is very different from the evolutionary one, and it is precisely the one that we will try to paint in this chapter.

OUTLINE In more precise terms, our aim will be to analyze the evolution of players in stochastically perturbed Nash games.¹ The particular stimulus-response model that we will consider is simple enough: players keep cumulative scores of their strategies’ performance and employ exponentially more often those that score better. In many ways, this “exponential learning” behavior is the simplest one that players could exhibit: as we will see, it reduces to the multi-population replicator dynamics, which are themselves the simplest scale-invariant growth model with no spontaneous mutations.

This approach is made precise in Section 4.1; then, in Section 4.2, we derive the stochastic replicator equation that governs the behavior of players when their learning curves are subject to random disturbances. The replicator equation that we get is different from the “aggregate shocks” approach of Fudenberg and Harris (1992) and, as a result, it exhibits markedly different rationality properties as well. In stark contrast to the results of Cabrales (2000) and Imhof (2005), we show in Section 4.3 that dominated strategies become extinct *irrespective of the noise level* (Proposition 4.1). In fact, by induction on the rounds of elimination of dominated strategies, we show that this is true even for *iteratively* dominated strategies: despite the noise, *only rationally admissible*

¹ See also Mertikopoulos and Moustakas (2009, 2010b).

strategies can survive the long run (Theorem 4.3). Then, as an easy corollary of the above, we infer that players converge to a strict equilibrium (Corollary 4.4) whenever the underlying game is a dominance-solvable one.

The similarities to the deterministic case do not stop here. In Section 4.4, we make a suggestive detour in the land of congestion games, and we show that if the noise is relatively mild with respect to the learning rate of the players, then the game's strict equilibria are stochastically asymptotically stable; in fact, if the game is *dyadic* (i.e. players only have two choices), the tameness condition can be dropped altogether.

Encouraged by these results, we attack the general case in Section 4.5: as it turns out, strict equilibria are *always* stochastically stable in the replicator dynamics of stochastically perturbed exponential learning (Theorem 4.8). This is a direct port of the deterministic (multi-population) folk theorem 2.16 and, as such, it begs to be compared with the results of Imhof (2005) and Hofbauer and Imhof (2009), where it is the equilibria of a suitably modified game that are stable, and not those of the actual game being considered. In this way, we see that stochasticity separates the processes of learning and of natural selection: exponential learning seems to give a clearer picture of the underlying game and obviates the need for evolution-type modifications.

4.1 EXPONENTIAL LEARNING

Let $\mathfrak{G} \equiv \mathfrak{G}(\mathcal{N}, \Delta, u)$ be a N -player Nash-type game, and let \mathcal{A}_i be the set of possible actions (pure strategies) that player $i \in \mathcal{N}$ can take in the game. If we recall the form of the replicator equation (2.41), we immediately get:

$$\frac{d}{dt} \log(x_{i\alpha}) = u_{i\alpha}(x) - u_i(x), \quad (4.1)$$

Revisiting the replicator equation.

where, as before, $x = \sum_{i,\alpha} x_{i\alpha} e_{i\alpha} \in \Delta$ represents the players' mixed strategy profile and $u_{i\alpha}(x) = u_i(x_{-i}; \alpha)$ is the "pure" payoff that player i would receive by playing $\alpha \in \mathcal{A}_i$ against his opponents' mixed strategy $x_{-i} \in \Delta_{-i}$ (as always, $\{e_{i\alpha} : \alpha \in \mathcal{A}_i\}$ represents the canonical basis of $\mathbb{R}^{\mathcal{A}_i}$; cf. Chapter 1). Then, by integrating (4.1) and taking the ratio $x_{i\alpha}/x_{i\beta}$, we readily obtain:

$$\frac{x_{i\alpha}(t)}{x_{i\beta}(t)} = \frac{e^{U_{i\alpha}(t)}}{e^{U_{i\beta}(t)}}, \quad (4.2)$$

where $U_{i\alpha}(t) = \int_0^t u_{i\alpha}(x(s)) ds$. However, since $\sum_{\beta} x_{i\beta} = 1$ (recall that x_i is a probability distribution in $\Delta_i \equiv \Delta(\mathcal{A}_i)$), a simple summation of this last equation leads to the *Gibbs distribution*:

$$x_{i\alpha}(t) = \frac{e^{U_{i\alpha}(t)}}{\sum_{\beta} e^{U_{i\beta}(t)}}. \quad (4.3)$$

In the probabilistic interpretation of mixed strategies (where x_i represents the probability distribution of player i over his action set \mathcal{A}_i), this last equation and the independence of the x_i 's imply that the probability $P(\alpha_1, \dots, \alpha_N)$ of observing the players at the state $(\alpha_1, \dots, \alpha_N) \in \mathcal{A} \equiv \prod_i \mathcal{A}_i$ will be given by:

$$P(\alpha_1, \dots, \alpha_N) = \frac{\exp(\sum_i U_{i\alpha_i}(t))}{\sum_{\beta_1} \dots \sum_{\beta_N} \exp(\sum_i U_{i\beta_i}(t))}. \quad (4.4)$$

In other words, if players evolve according to the replicator equation (2.41), the $U_{i\alpha}$ can be seen as (the negatives of) player-specific Hamiltonians which the players are inherently trying to maximize over time.

This optimization effort is implicit in the replicator dynamics (2.41), but it is much easier to understand if we read this line of reasoning backwards. Indeed, the integral $U_{i\alpha}(t) = \int_0^t u_{i\alpha}(x(s)) ds$ in (4.3) simply represents the *total* payoff that the strategy $\alpha \in \mathcal{A}_i$ would have garnered for player i up to time t (of course, paired against the evolving mixed strategy $x_{-i}(t)$ of i 's opponents). In its turn, this suggests an alternative interpretation of the replicator equation (2.41), which is even more intuitive than our derivation in Section 2.1.2 (and which will be the focal point of our stochastic considerations as well).

So, before introducing any rules of dynamic evolution, let us assume that players keep track of their strategies' performance by means of the *scores* $U_{i\alpha}$:

$$dU_{i\alpha}(t) = u_{i\alpha}(x(t)) dt, \quad (4.5)$$

where $x(t)$ represents the players' strategy profile at time t and, in the absence of initial bias, we assume that $U_{i\alpha}(0) = 0$ for all players $i \in \mathcal{N}$ and all actions $\alpha \in \mathcal{A}_i$.² These scores obviously reinforce the perceived success of each strategy as measured by the average payoff it yields and, hence, it stands to reason that players will lean towards the strategy with the highest score.

How do the players do that? A first choice might be to choose a strategy with probability proportional to its score, that is, to set $x_{i\alpha} = U_{i\alpha} / \sum_{\beta} U_{i\beta}$; however, this choice is problematic because scores might well be negative. In a small leap of faith, we assume (or, more fittingly, *propose*) that players update their strategies according to the namesake law of *exponential learning*:

$$x_{i\alpha}(t) = \frac{e^{\lambda_i U_{i\alpha}(t)}}{\sum_{\beta} e^{\lambda_i U_{i\beta}(t)}}, \quad (4.6)$$

which, except for the *learning rates* $\lambda_i > 0$ is just a reiteration of (4.3).

At first sight, this choice might seem arbitrary and, indeed, if our only criterion were to specify the variables $x_{i\alpha}$ in terms of $U_{i\alpha}$ in a consistent way, we might as well have gone with squaring instead of exponentiating (or any other positive smooth function). The reason that makes the updating rule (4.6) a natural choice by itself is its relation to the *logit model* and the learning scheme of *logistic fictitious play* (Fudenberg and Levine, 1998).³

The idea behind logistic learning is quite simple: given the strategy of one's opponents, one would like to play the best possible response to it, i.e. the strategy which maximizes one's gains in this constrained context. However, it might well happen that this best-response is not uniquely determined: for instance, if there exist two actions $\alpha, \beta \in \mathcal{A}_i$ of player i which yield the same maximum reward $u_i(x_{-i}; \alpha) = u_i(x_{-i}; \beta)$, then any convex combination of these two strategies will also be a best response to x_{-i} .

To resolve this lack of uniqueness, one resorts to a "smoothing" trick and perturbs the payoff functions $u_i(x)$ by a strictly concave function $h_i(x_i)$ whose

Scoring
strategies.

Logistic
learning.

² In discrete time, equation (4.5) would take the form: $U_{i\alpha}(t + \delta) = U_{i\alpha}(x(t)) + u_{i\alpha}(x(t))\delta$. We will concentrate here on continuous time; for the discrete-time case, see e.g. Benaïm (1999).

³ In the literature, the term "logit learning" is often used to describe (4.6). I find this choice somewhat unfortunate: the logit function is defined as $\text{logit } p \equiv \log(p/(1-p))$, and it is its *inverse*, the *logistic* function which essentially appears in (4.6). In this respect, the designation "logistic" appears far superior to "logit".

derivatives blow up at the boundary of Δ_i . More precisely, one considers a “perturbed” version $\mathfrak{G}(\tau)$ of \mathfrak{G} with payoffs given by:

$$\tilde{u}_i(x) = u_i(x) + \tau_i h_i(x_i), \quad (4.7)$$

where τ_i is a player-specific temperature-like parameter which “freezes” the game $\mathfrak{G}(\tau)$ back to $\mathfrak{G}(0) \equiv \mathfrak{G}$ when it becomes zero.

Connections with thermodynamics.

Geometrically, if one pictures the constrained payoff $u_i(x_{-i}; \cdot) : \Delta_i \rightarrow \mathbb{R}$ in terms of its graph over Δ_i , then the perturbation h_i amounts to attaching a strictly concave “cap” over this graph, itself a simplex over \mathcal{A}_i (owing to the multilinearity of u_i). Clearly then, the strategy $\beta_i(x) \in \Delta_i$ which corresponds to the summit of this cap will be the *unique* best reply of player i to x_{-i} in the perturbed game, and, as it turns out, the mapping $x \mapsto \beta_i(x)$ is actually smooth (Fudenberg and Levine, 1998).

Obviously, this *smooth best-reply correspondence* $\beta_i : \Delta \rightarrow \Delta_i$ depends on the perturbation function h_i . A most natural choice in this regard would be to take the *entropy* of x_i itself and set:

$$h_i(x_i) = - \sum_{\alpha \in \text{supp}(x_i)} x_{i\alpha} \log x_{i\alpha}. \quad (4.8)$$

With this choice, one is essentially asked to maximize:

$$\tilde{u}_i(x_{-i}; x_i) = u_i(x_{-i}; x_i) - \tau_i \sum_{\alpha}^i x_{i\alpha} \log x_{i\alpha} \quad (4.9)$$

for fixed $x_{-i} \in \Delta_{-i}$, and subject to the condition $\sum_{\alpha}^i x_{i\alpha} = 1$. Thus, by solving this (convex) optimization problem, we obtain the *logistic best-reply*:

$$\beta_{i\alpha}(x) = \frac{\exp\left(\frac{1}{\tau_i} u_{i\alpha}(x)\right)}{\sum_{\beta}^i \exp\left(\frac{1}{\tau_i} u_{i\beta}(x)\right)}, \quad (4.10)$$

which freezes to a myopic (and usually ill-defined) best-reply scheme when the temperature parameters τ_i go to zero.

Learning rates and temperature.

Under the light of the above, the exponential learning rule (4.6) becomes a very reasonable choice: the players are maximizing their expected scores, while simultaneously keeping the entropy of their strategies as high as possible, in an effort to sample them as uniformly as the disparity of their payoffs allows. This score-entropy balance is controlled by the players’ learning rates λ_i which, as can be seen by comparing (4.6) to (4.10), act as “inverse temperatures” of sorts: in high temperatures (small λ_i), the players’ learning curves are “soft” and payoff differences are toned down; on the contrary, if $\lambda_i \rightarrow \infty$, exponential learning freezes down to a hard best-reply process.

To justify the term “learning rate” – and also to recover the connection to the replicator dynamics – simply note that if we differentiate (4.6) to decouple it from (4.5), we obtain:

$$\begin{aligned} \frac{dx_{i\alpha}}{dt} &= \frac{\lambda_i u_{i\alpha}(x) \sum_{\beta} e^{\lambda_i U_{i\beta}} - \lambda_i e^{\lambda_i U_{i\alpha}} \sum_{\beta} u_{i\beta}(x) e^{\lambda_i U_{i\beta}}}{\left(\sum_{\beta} e^{\lambda_i U_{i\beta}}\right)^2} \\ &= \lambda_i x_{i\alpha} (u_{i\alpha}(x) - u_i(x)). \end{aligned} \quad (4.11)$$

This *rate-adjusted* replicator equation shows that the rates λ_i also control the timescale at which players evolve: the greater a player’s learning rate, the

faster his strategies will be updated – clearly, when players all learn at the “standard” rate $\lambda_i = 1$, we recover the uniform replicator dynamics (2.41).

4.2 LEARNING IN THE PRESENCE OF NOISE

The rate-adjusted dynamics (4.11) obviously inherit the rationality properties of the replicator dynamics in multi-population random matching games, so there is really not much left to say here: dominated strategies become extinct (Theorem 2.15) and the folk theorem (Theorem 2.16) applies. However, one could argue that these properties are a direct consequence of the players’ receiving accurate information about the game when they update their scores, and this is a requirement that cannot always be met. The interference of nature with the game or imperfect readings of one’s rewards invariably introduce stochastic fluctuations in (4.5), which, in turn, will propagate all the way down to the replicator dynamics (4.11).

Learning and uncertainties.

To account for these random perturbations, we will assume that the players’ scores are governed instead by the *stochastic* differential equation:

$$dU_{i\alpha}(t) = u_{i\alpha}(X(t)) dt + \sigma_{i\alpha}(X(t)) dW_{i\alpha}(t), \quad (4.12)$$

where, as before, the strategy profile $X(t) \in \Delta$ is given by the logistic law:

$$X_{i\alpha}(t) = \frac{e^{\lambda_i U_{i\alpha}(t)}}{\sum_{\beta}^i e^{\lambda_i U_{i\beta}(t)}}, \quad (4.13)$$

$W(t) = \sum_{i,\alpha} W_{i\alpha}(t) e_{i\alpha}$ is a Wiener process in $\mathbb{R}^{\mathcal{A}} \cong \prod_i \mathbb{R}^{\mathcal{A}_i}$, and the noise coefficients $\sigma_{i\alpha} : \Delta \rightarrow \mathbb{R}$ are essentially bounded functions $\sigma_{i\alpha} : \Delta \rightarrow \mathbb{R}$ which measure the impact of the noise on the players’ scoring system.⁴

Remark. A very important instance of the dependence of the noise coefficients on the state of the game can be seen if $\sigma_{i\alpha}(x_{-i}; \alpha) = 0$ for all players $i \in \mathcal{N}$ and all $\alpha \in \mathcal{A}_i$, $x_{-i} \in \Delta_{-i}$. In that case, equation (4.12) becomes a convincing model for the case of imperfect information: when a player actually employs a strategy, his payoff observations are accurate, but with regards to strategies that he rarely employs, his readings could be arbitrarily off the mark.

Bearing all this in mind, we may decouple (4.12) and (4.13) by applying Itô’s formula (3.30). To that end, recall that $W(t)$ has stochastically independent components across both players and their strategies, so we will have $dW_{j\beta} \cdot dW_{k\gamma} = \delta_{jk} \delta_{\beta\gamma}$. In this way, we get:

The propagation of stochasticity.

$$\begin{aligned} dX_{i\alpha} &= \sum_{j,\beta} \frac{\partial X_{i\alpha}}{\partial U_{j\beta}} dU_{j\beta} + \frac{1}{2} \sum_{j,\beta} \sum_{k,\gamma} \frac{\partial^2 X_{i\alpha}}{\partial U_{j\beta} \partial U_{k\gamma}} dU_{j\beta} \cdot dU_{k\gamma} \\ &= \sum_{\beta}^i \left(u_{i\beta}(X) \frac{\partial X_{i\alpha}}{\partial U_{i\beta}} + \frac{1}{2} \sigma_{i\beta}^2(X) \frac{\partial^2 X_{i\alpha}}{\partial U_{i\beta}^2} \right) dt \\ &\quad + \sum_{\beta}^i \sigma_{i\beta}(X) \frac{\partial X_{i\alpha}}{\partial U_{i\beta}} dW_{i\beta}. \end{aligned} \quad (4.14)$$

Remark. Though unremarkable in itself, the stochastic differentiation (4.14) is actually the most general form of the replicator dynamics because it encodes

⁴ Recall that a function f is called essentially bounded when $\text{ess sup } |f| < \infty$, i.e. when it is bounded modulo a Lebesgue-null set.

the propagation of the noise that is inherent to the processes $U_{i\alpha}$ down to the observable variables $X_{i\alpha}$.

At any rate, an ordinary differentiation of (4.6) immediately yields:

$$\frac{\partial X_{i\alpha}}{\partial U_{i\beta}} = \lambda_i X_{i\alpha} (\delta_{\alpha\beta} - X_{i\beta}) \quad (4.15a)$$

$$\frac{\partial^2 X_{i\alpha}}{\partial U_{i\beta}^2} = \lambda_i^2 X_{i\alpha} (\delta_{\alpha\beta} - X_{i\beta}) (1 - 2X_{i\beta}) \quad (4.15b)$$

and, hence, by plugging these expressions back into (4.14), we get:

$$\begin{aligned} dX_{i\alpha} &= \lambda_i X_{i\alpha} (u_{i\alpha}(X) - u_i(X)) dt \\ &\quad + \frac{\lambda_i^2}{2} X_{i\alpha} \left(\sigma_{i\alpha}^2(X) (1 - 2X_{i\alpha}) - \sum_{\beta}^i \sigma_{i\beta}^2(X) X_{i\beta} (1 - 2X_{i\beta}) \right) dt \\ &\quad + \lambda_i X_{i\alpha} \left(\sigma_{i\alpha}(X) dW_{i\alpha} - \sum_{\beta}^i \sigma_{i\beta}(X) X_{i\beta} dW_{i\beta} \right), \end{aligned} \quad (4.16)$$

or, in the “uniform” case $\lambda_i = 1$:

$$\begin{aligned} dX_{i\alpha} &= X_{i\alpha} (u_{i\alpha}(X) - u_i(X)) dt \\ &\quad + \frac{1}{2} X_{i\alpha} \left(\sigma_{i\alpha}^2(X) (1 - 2X_{i\alpha}) - \sum_{\beta}^i \sigma_{i\beta}^2(X) X_{i\beta} (1 - 2X_{i\beta}) \right) dt \\ &\quad + X_{i\alpha} \left(\sigma_{i\alpha}(X) dW_{i\alpha} - \sum_{\beta}^i \sigma_{i\beta}(X) X_{i\beta} dW_{i\beta} \right). \end{aligned} \quad (4.17)$$

The stochastic replicator dynamics.

The rate-adjusted equation (4.16) and its uniform sibling (4.17) will constitute our stochastic version of the replicator dynamics and thus merit some discussion in and by themselves. First, note that these dynamics admit a (unique) strong solution for any initial state $X(0) = x \in \Delta$, even though they do not satisfy the linear growth condition $|b(x)| + |\sigma(x)| \leq C(1 + |x|)$ that is required for the existence and uniqueness theorem for SDE’s (Theorem 3.5). Nevertheless, an addition over $\alpha \in \mathcal{A}_i$ reveals that every simplex $\Delta_i \subseteq \Delta$ remains invariant under (4.16): if $X_i(0) = x_i \in \Delta_i$, then $d\left(\sum_{\alpha}^i X_{i\alpha}\right) = 0$ and, hence, $X_i(t)$ will stay in Δ_i for all $t \geq 0$.⁵

Strong solutions.

So, let ϕ be a smooth bump function that is equal to 1 on some open neighborhood $U \supseteq \Delta$ and which vanishes outside some compact set $K \supseteq U$. Then, if we denote the drift and diffusion coefficients of (4.16) by b and c respectively, the SDE

$$dX_{i\alpha} = \phi(X) \left(b_{i\alpha}(X) dt + \sum_{\beta} c_{i,\alpha\beta}(X) dW_{i\beta} \right) \quad (4.18)$$

will have bounded diffusion and drift coefficients and will thus admit a unique strong solution. But since this last equation agrees with (4.16) on Δ and any solution of (4.16) always stays in Δ , we can easily conclude that our perturbed replicator dynamics admit a unique strong solution for any initial $X(0) = x \in \Delta$.

Evolution vs. learning.

It is also important to compare the dynamics (4.16), (4.17) to the “aggregate shocks” approach of Fudenberg and Harris (1992) that has become the principal incarnation of the replicator dynamics in stochastic environments. So, let us first recall how aggregate shocks enter the replicator dynamics in the first place. The main idea is that the reproductive fitness of an individual is not

⁵ Actually, it is not harder to see that every face of Δ is a *trap* for $X(t)$ in the sense that any trajectory which starts at a face of Δ will remain in the face with probability 1.

only affected by deterministic factors but is also subject to stochastic shocks due to the “weather” and the interference of nature with the game.

More precisely, if $Z_{i\alpha}(t)$ denotes the population size of phenotype $\alpha \in \mathcal{A}_i$ of the species $i \in \mathcal{N}$ in some multi-population random matching game \mathfrak{G} , its growth will be determined by:

$$dZ_{i\alpha}(t) = Z_{i\alpha}(t)(u_{i\alpha}(X(t)) dt + \sigma_{i\alpha} dW_{i\alpha}(t)), \quad (4.19)$$

where, as in (2.40), $X(t) \in \Delta$ denotes the population shares $X_{i\alpha} = Z_{i\alpha} / \sum_{\beta}^i Z_{i\beta}$. In this way, an Itô differentiation yields the (*multi-population*) replicator dynamics with aggregate shocks:

$$\begin{aligned} dX_{i\alpha} = & X_{i\alpha} \left[(u_{i\alpha}(X) - u_i(X)) - \left(\sigma_{i\alpha}^2 X_{i\alpha} - \sum_{\beta} \sigma_{i\beta}^2 X_{i\beta} \right) \right] dt \\ & + X_{i\alpha} \left[\sigma_{i\alpha} dW_{i\alpha} - \sum_{\beta} \sigma_{i\beta} X_{i\beta} dW_{i\beta} \right]. \end{aligned} \quad (4.20)$$

We thus see that the effects of noise propagate differently in the case of exponential learning and in the case of evolution. Indeed, if we compare equations (4.16) (or, more fittingly, the uniform version (4.17)) with (4.20) term by term, we see that their drifts are not quite the same. Even though the payoff adjustment $u_{i\alpha} - u_i$ ties both equations back together in the deterministic limit ($\sigma = 0$), the two expressions differ by the term:

$$\frac{1}{2} X_{i\alpha} \left(\sigma_{i\alpha}^2 - \sum_{\beta} \sigma_{i\beta}^2 X_{i\beta} \right) dt. \quad (4.21)$$

Innocuous as this term might seem, it is actually crucial for the rationality properties of exponential learning in games with randomly perturbed payoffs. As we shall see in the next sections, it leads to some miraculous cancellations that allow rationality to emerge in all noise levels.

This difference further suggests that we can pass from (4.16) to (4.20) simply by modifying the game’s payoffs to $\tilde{u}_{i\alpha} = u_{i\alpha} + \frac{1}{2}\sigma_{i\alpha}^2$. Of course, this presumes that the noise coefficients $\sigma_{i\alpha}$ be constant – the general case would require us to allow for games whose payoffs may not be multilinear. This apparent lack of generality does not really change things but we prefer to keep things simple, so, for the time being, it suffices to point out that this modified game was precisely the one that came up in the analysis of [Imhof \(2005\)](#) and [Hofbauer and Imhof \(2009\)](#). As a result, this modification appears to play a pivotal role in setting apart learning and evolution in a stochastic setting: whereas the modified game is deeply ingrained in the process of natural selection, exponential learning seems to give players a clearer picture of the actual underlying game.

*Adjusting
for noise.*

While on the point of evolution versus learning, there is yet another important observation to be made. More precisely, note that if we compare the deterministic scoring and growth equations (4.5) and (2.40) respectively, we readily obtain:

$$dU_{i\alpha} = u_{i\alpha}(x) dt = \frac{dz_{i\alpha}}{z_{i\alpha}} = d(\log z_{i\alpha}), \quad (4.22)$$

so that $z_{i\alpha}(t) = e^{U_{i\alpha}(t)}$. In the stochastic case however, Itô’s lemma gives:

$$d(\log Z_{i\alpha}) = \frac{dZ_{i\alpha}}{Z_{i\alpha}} - \frac{1}{2} \frac{1}{Z_{i\alpha}^2} (dZ_{i\alpha})^2 = \frac{dZ_{i\alpha}}{Z_{i\alpha}} - \frac{1}{2} \sigma_{i\alpha}^2 dt. \quad (4.23)$$

Therefore, a comparison of (4.12) and (4.19) yields:

$$dU_{i\alpha}(t) = d(\log Z_{i\alpha}) + \frac{1}{2}\sigma_{i\alpha}^2 dt, \quad (4.24)$$

or, equivalently:

$$Z_{i\alpha}(t) = \exp\left(\int_0^t u_{i\alpha}(X(s)) - \frac{1}{2}\sigma_{i\alpha}^2(X(s)) ds\right). \quad (4.25)$$

The Stratonovich interpretation.

In a certain sense, this last equation illustrates the role of the modified game $\tilde{u}_{i\alpha} = u_{i\alpha} + \frac{1}{2}\sigma_{i\alpha}^2$ even more clearly than the extra term (4.21). More importantly, it also serves to explain the observation of Hofbauer and Imhof (2009) regarding the behavior of Stratonovich-type perturbations (see also Khas'minskii and Potsepun, 2006): since $\partial(\log Z_{i\alpha}) = \frac{\partial Z_{i\alpha}}{Z_{i\alpha}}$, we see that, essentially, Itô perturbations in exponential learning correspond to Stratonovich perturbations in evolution.

4.3 EXTINCTION OF DOMINATED STRATEGIES

Armed with the stochastic replicator equations (4.16) and (4.17) to model exponential learning in noisy environments, the logical next step is to see if the rationality properties of the deterministic dynamics carry over to this stochastic setting.

The cross entropy.

In this direction, we will first show that dominated strategies always become extinct in the long run and that only the rationally admissible ones survive. As in Cabrales (2000) (implicitly) and Imhof (2005) (explicitly), the key ingredient of our approach will be the *cross entropy* between two mixed strategies $q_i, x_i \in \Delta_i$ of player $i \in \mathcal{N}$:

$$H(q_i, x_i) \equiv - \sum_{\alpha: q_{i\alpha} > 0} q_{i\alpha} \log(x_{i\alpha}) = H(q_i) + d_{\text{KL}}(q_i, x_i) \quad (4.26)$$

where $H(q_i) = -\sum_{\alpha} q_{i\alpha} \log q_{i\alpha}$ is the *entropy* of q_i and d_{KL} is the Kullback-Leibler distance of x_i from q_i .

As we pointed out in the proof of Theorem 2.15 and the discussion immediately preceding it, $d_{\text{KL}}(q_i, x_i) < \infty$ if and only if x_i employs with positive probability all pure strategies $\alpha \in \mathcal{A}_i$ that are present in q_i (i.e. iff q_i is absolutely continuous w.r.t. x_i). Therefore, if $d_{\text{KL}}(q_i, x_i) = \infty$ for all dominated strategies q_i of player i , it immediately follows that x_i cannot be dominated itself. Following this line of reasoning, we obtain:

Dominated strategies always become extinct.

Proposition 4.1. *Let $X(t)$ be a solution trajectory of the stochastic replicator dynamics (4.17) for some interior initial condition $X(0) = x \in \text{Int}(\Delta)$. Then, if $q_i \in \Delta_i$ is (strictly) dominated, we will have:*

$$\lim_{t \rightarrow \infty} d_{\text{KL}}(q_i, X_i(t)) = \infty \quad \text{almost surely.} \quad (4.27)$$

In particular, if $q_i = \alpha \in \mathcal{A}_i$ is pure, we will have $\lim_{t \rightarrow \infty} X_{i\alpha}(t) = 0$ (a.s.): strictly dominated strategies do not survive in the long run.

Proof. Note first that $X(0) = x \in \text{Int}(\Delta)$ and, hence, $X_i(t)$ will almost surely stay in $\text{Int}(\Delta_i)$ for all $t \geq 0$ – a simple consequence of the uniqueness of strong solutions and the invariance of the faces of Δ_i under the dynamics (4.17).

Let us now consider the cross entropy $G_{q_i}(t)$ between q_i and $X_i(t)$:

$$G_{q_i}(t) \equiv H(q_i, X_i(t)) = - \sum_{\alpha} q_{i\alpha} \log X_{i\alpha}(t). \quad (4.28)$$

As a result of $X_i(t)$ being an interior path, $G_{q_i}(t)$ will remain finite for all $t \geq 0$ (a.s.). So, by applying Itô's formula, we get:

$$\begin{aligned} dG_{q_i} &= \sum_{\beta}^i \frac{\partial G_{q_i}}{\partial X_{i\beta}} dX_{i\beta} + \frac{1}{2} \sum_{\beta, \gamma}^i \frac{\partial^2 G_{q_i}}{\partial X_{i\gamma} \partial X_{i\beta}} dX_{i\beta} \cdot dX_{i\gamma} \\ &= - \sum_{\beta}^i \frac{q_{i\beta}}{X_{i\beta}} dX_{i\beta} + \frac{1}{2} \sum_{\beta}^i \frac{q_{i\beta}}{X_{i\beta}^2} (dX_{i\beta})^2, \end{aligned} \quad (4.29)$$

and, after substituting $dX_{i\beta}$ from the replicator dynamics (4.17), this last equation becomes:

$$\begin{aligned} dG_{q_i} &= \sum_{\beta}^i q_{i\beta} \left[u_i(X) - u_{i\beta}(X) + \frac{1}{2} \sum_{\gamma}^i \sigma_{i\gamma}^2(X) X_{i\gamma} (1 - X_{i\gamma}) \right] dt \\ &\quad + \sum_{\beta}^i q_{i\beta} \sum_{\gamma}^i (X_{i\gamma} - \delta_{\beta\gamma}) \sigma_{i\gamma}(X) dW_{i\gamma}. \end{aligned} \quad (4.30)$$

Accordingly, if $q'_i \in \Delta_i$ is another mixed strategy of player i , we obtain:

$$\begin{aligned} dG_{q_i} - dG_{q'_i} &= (u_i(X_{-i}; q'_i) - u_i(X_{-i}; q_i)) dt \\ &\quad + \sum_{\beta}^i (q'_{i\beta} - q_{i\beta}) \sigma_{i\beta}(X) dW_{i\beta}, \end{aligned} \quad (4.31)$$

so, after integrating, we obtain:

$$\begin{aligned} G_{q_i - q'_i}(t) &= H(q_i - q'_i, x) + \int_0^t u_i(X_{-i}(s); q'_i - q_i) ds \\ &\quad + \sum_{\beta}^i (q'_{i\beta} - q_{i\beta}) \int_0^t \sigma_{i\beta}(X(s)) dW_{i\beta}(s), \end{aligned} \quad (4.32)$$

where, in obvious notation, $G_{q_i - q'_i} \equiv G_{q_i} - G_{q'_i}$.

Suppose then that $q_i \prec q'_i$ and let $v_i = \inf\{u_i(x_{-i}; q'_i - q_i) : x_{-i} \in \Delta_{-i}\}$. With Δ_{-i} compact, it easily follows that $v_i > 0$ and the first term of (4.32) will be bounded from below by $v_i t$. However, since monotonicity fails for Itô integrals,⁶ the second term must be handled with more care. To that end, let $\xi_i(s) = \sum_{\beta}^i (q'_{i\beta} - q_{i\beta}) \sigma_{i\beta}(X(s))$ and use the Cauchy-Schwarz inequality to get:

$$\xi_i^2(s) \leq A_i \sum_{\beta}^i (q'_{i\beta} - q_{i\beta})^2 \sigma_{i\beta}^2(X(s)) \leq A_i \sigma_i^2 \sum_{\beta}^i (q'_{i\beta} - q_{i\beta})^2 \leq 2A_i \sigma_i^2, \quad (4.33)$$

where $A_i = |\mathcal{A}_i|$ is the number of actions (pure strategies) available to player i and $\sigma_i = \text{ess sup}\{|\sigma_{i\beta}(x)| : x \in \Delta, \beta \in \mathcal{A}_i\}$ – recall also that $q_i, q'_i \in \Delta_i$ for the last step. Therefore, if $\psi_i(t) = \sum_{\beta}^i (q'_{i\beta} - q_{i\beta}) \int_0^t \sigma_{i\beta}(X(s)) dW_{i\beta}(s)$ denotes the martingale part of (4.31) and $\rho_i(t)$ is its quadratic variation, (4.33) yields:

$$\rho_i(t) = [\psi_i, \psi_i](t) = \int_0^t \xi_i^2(s) ds \leq 2S_i \sigma_i^2 t. \quad (4.34)$$

Now, if $\lim_{t \rightarrow \infty} \rho_i(t) = \infty$, it follows from the time-change theorem for martingales (see e.g. Karatzas and Shreve, 1998, Theorem 3.4.6) that there

The bound of dominance.

⁶ For instance, $\int_0^1 1 dW(s) = W(1)$, and the latter is positive or negative with equal probability.

exists a (time-rescaled) Wiener process \tilde{W}_i with $\psi_i(t) = \tilde{W}_i(\rho_i(t))$. Hence, the law of the iterated logarithm for Brownian motion (Lamperti, 1977) yields:

$$\begin{aligned} \liminf_{t \rightarrow \infty} G_{q_i - q'_i}(t) &\geq H(q_i - q'_i, x) + \liminf_{t \rightarrow \infty} \left(v_i t + \tilde{W}_i(\rho_i(t)) \right) \\ &\geq H(q_i - q'_i, x) + \liminf_{t \rightarrow \infty} \left(v_i t - \sqrt{2\rho_i(t) \log \log \rho_i(t)} \right) \\ &\geq H(q_i - q'_i, x) + \liminf_{t \rightarrow \infty} \left(v_i t - 2\sigma_i \sqrt{A_i t \log \log (2A_i \sigma_i^2 t)} \right) \\ &= \infty \quad (\text{almost surely}). \end{aligned} \quad (4.35)$$

On the other hand, if $\lim_{t \rightarrow \infty} \rho_i(t) < \infty$, it is trivial to obtain $G_{q_i - q'_i}(t) \rightarrow \infty$ by letting $t \rightarrow \infty$ in (4.32). Therefore, with $G_{q_i}(t) \geq G_{q_i}(t) - G_{q'_i}(t) \rightarrow \infty$, we readily get $\lim_{t \rightarrow \infty} d_{\text{KL}}(q_i, X_i(t)) = \infty$ (a.s.); and since $G_\alpha(t) = -\log X_{i\alpha}(t)$ for all pure strategies $\alpha \in \mathcal{A}_i$, our proof is complete. \square

As in Imhof (2005), we now obtain the following estimate for the lifespan of pure dominated strategies:

Rate of extinction. **Proposition 4.2.** *Let $X(t)$ be a solution path of (4.17) with initial condition $X(0) = x \in \text{Int}(\Delta)$; assume further that the pure strategy $\alpha \in \mathcal{A}_i$ is dominated. Then, for any $M > 0$ and for t large enough, we will have:*

$$\mathbf{P}_x \left(X_{i\alpha}(t) < e^{-M} \right) \geq \frac{1}{2} \operatorname{erfc} \left(\frac{M - h_i(x_i) - v_i t}{2\sigma_i \sqrt{A_i t}} \right) \quad (4.36)$$

where $A_i = |\mathcal{A}_i|$ is the number of strategies available to player i , $\sigma_i = \operatorname{ess\,sup} \{ |\sigma_{i\beta}(y)| : y \in \Delta, \beta \in \mathcal{A}_i \}$ and the constants $v_i > 0$ and $h_i(x_i)$ do not depend on t .

Proof. The proof is pretty straightforward and for the most part follows Imhof (2005). Sure enough, if $\alpha \prec p_i \in \Delta_i$, we will have (in the notation of the previous proposition):

$$\begin{aligned} -\log X_{i\alpha}(t) &= G_\alpha(t) \geq G_\alpha(t) - G_{p_i}(t) \\ &\geq H(e_{i\alpha}, x) - H(p_i, x) + v_i t + \tilde{W}_i(\rho_i(t)) \\ &= h_i(x_i) + v_i t + \tilde{W}_i(\rho_i(t)) \end{aligned} \quad (4.37)$$

where $v_i \equiv \min_{x_{-i}} \{ u_i(x_{-i}; p_i) - u_i(x_{-i}; \alpha) \} > 0$ and we have set $h_i(x_i) \equiv \log x_{i\alpha} - \sum_{\beta} p_{i\beta} \log x_{i\beta}$. We thus obtain:

$$\begin{aligned} \mathbf{P}_x \left(X_{i\alpha}(t) < e^{-M} \right) &\geq \mathbf{P}_x \left(\tilde{W}_i(\rho_i(t)) > M - h_i(x_i) - v_i t \right) \\ &= \frac{1}{2} \operatorname{erfc} \left(\frac{M - h_i(x_i) - v_i t}{\sqrt{2\rho_i(t)}} \right), \end{aligned} \quad (4.38)$$

and since (4.34) bounds the quadratic variation $\rho_i(t)$ by $2A_i\sigma_i^2 t$ from above, the estimate (4.36) will hold for all sufficiently large t . \square

Remark 1. This result should be contrasted to those of Cabrales (2000) and Imhof (2005) where dominated strategies die out only if the noise coefficients (shocks) $\sigma_{i\alpha}$ satisfy certain tameness conditions. The origin of this notable difference is the form of the replicator equation (4.17) and, in particular, the extra terms (4.21) that are propagated by exponential learning – and which are absent from the aggregate shocks dynamics (4.20). As can be seen from

the derivations in Proposition 4.1, these terms are precisely the ones that allow players to pick up on the true payoffs $u_{i\alpha}$ instead of the modified ones $\tilde{u}_{i\alpha} = u_{i\alpha} + \frac{1}{2}\sigma_{i\alpha}^2$ that come up in the work of Hofbauer and Imhof (2009).

Remark 2. It turns out that the way that the noise coefficients $\sigma_{i\beta}$ depend on the profile $x \in \Delta$ is not really crucial: as long as $\sigma_{i\beta}(x)$ is continuous (or essentially bounded), our arguments are not affected. The only way in which a specific dependence influences the extinction of dominated strategies is seen in Proposition 4.2: a sharper estimate of the quadratic variation $\rho_i(t)$ could conceivably yield a more accurate estimate for the CDF (4.36).

In light of the above, it is only natural to ask if proposition 4.1 can be extended to strategies that are only *iteratively* dominated. This is indeed so:

Theorem 4.3. *Let $X(t)$ be a solution path of (4.17) starting at $X(0) = x \in \text{Int}(\Delta)$. If $q_i \in \Delta_i$ is iteratively dominated:*

$$\lim_{t \rightarrow \infty} d_{\text{KL}}(q_i, X_i(t)) = \infty \quad \text{almost surely,} \quad (4.39)$$

Only rational strategies survive.

i.e. only rationally admissible strategies survive in the long run.

Proof. As in the deterministic case studied by Samuelson and Zhang (1992), the main idea is that $X(t)$ gets progressively closer to the faces of Δ that are spanned by the pure strategies which have not been eliminated. Following Cabrales (2000), we will prove this by induction on the rounds of elimination of dominated strategies; Proposition 4.1 is simply the case $n = 1$.

To proceed, let $M_i \subseteq \Delta_i$, $M_{-i} \subseteq \Delta_{-i}$, and denote by $\text{Adm}(M_i, M_{-i})$ the set of strategies $q_i \in M_i$ that are admissible (i.e. not dominated) with respect to any strategy $q_{-i} \in M_{-i}$. So, if we start with $\Delta_i^0 = \Delta_i$ and $\Delta_{-i}^0 = \prod_{j \neq i} \Delta_j^0$, we may define inductively the set of strategies that remain admissible after n elimination rounds as $\Delta_i^n \equiv \text{Adm}(\Delta_i^{n-1}, \Delta_{-i}^{n-1})$, where $\Delta_i^{n-1} = \prod_{j \neq i} \Delta_j^{n-1}$; similarly, the pure strategies that have survived after n such rounds will be denoted by $\mathcal{A}_i^n \equiv \mathcal{A}_i \cap \Delta_i^n$. Clearly, this sequence forms a descending chain $\Delta_i^0 \supseteq \Delta_i^1 \supseteq \dots \supseteq \Delta_i^n \supseteq \dots$ and the (nonempty by Proposition 2.3) set $\Delta_i^\infty \equiv \bigcap_{n=0}^\infty \Delta_i^n$ will consist precisely of the strategies of player i that are rationally admissible.

Iterated elimination.

Assume then that the cross entropy $G_{q_i}(t) = -\sum_{\alpha} q_{i\alpha} \log X_{i\alpha}(t)$ diverges as $t \rightarrow \infty$ for all strategies $q_i \notin \Delta_i^k$ that die out within the first k rounds; in particular, if $\alpha \notin \mathcal{A}_i^k$, this implies that $X_{i\alpha}(t) \rightarrow 0$ as $t \rightarrow \infty$. We will show that the same holds if q_i survives for k rounds but is eliminated in the next one.

Indeed, if $q_i \in \Delta_i^k$ but $q_i \notin \Delta_i^{k+1}$, there will exist some $q'_i \in \Delta_i^{k+1}$ such that:

$$u_i(x_{-i}; q'_i) > u_i(x_{-i}; q_i) \quad \text{for all } x_{-i} \in \Delta_{-i}^k. \quad (4.40)$$

Note now that any $x_{-i} \in \Delta_{-i}$ can be decomposed as $x_{-i} = x_{-i}^{\text{adm}} + x_{-i}^{\text{dom}}$, where x_{-i}^{adm} is the ‘‘admissible’’ part of x_{-i} , i.e. the projection of x_{-i} on the subspace spanned by the surviving vertices $\mathcal{A}_{-i}^k = \prod_{j \neq i} \mathcal{A}_j^k$. Thus, if $v_i = \min\{u_i(\alpha_{-i}; q'_i) - u_i(\alpha_{-i}; q_i) : \alpha_{-i} \in \mathcal{A}_{-i}^k\}$, we will have $v_i > 0$, and hence:

The induction mechanism.

$$u_i(x_{-i}^{\text{adm}}; q'_i) - u_i(x_{-i}^{\text{adm}}; q_i) \geq v_i > 0 \quad \text{for all } x_{-i} \in \Delta_{-i} \quad (4.41)$$

by linearity. Moreover, by the induction hypothesis, we also have $X_{-i}^{\text{dom}}(t) \rightarrow 0$ as $t \rightarrow \infty$. Thus, there exists some t_0 such that:

$$\left| u_i(X_{-i}^{\text{dom}}(t), q'_i) - u_i(X_{-i}^{\text{dom}}(t), q_i) \right| < v_i/2 \quad (4.42)$$

for all $t \geq t_0$ (recall that $X_{-i}^{\text{dom}}(t)$ is spanned by already eliminated strategies). Therefore, as in the proof of Proposition 4.1, we obtain for $t \geq t_0$:

$$G_{q_i}(t) - G_{q'_i}(t) \geq M + \frac{1}{2}v_it + \sum_{\beta}^i (q'_{i\beta} - q_{i\beta}) \int_0^t \sigma_{i\beta}(X(s)) dW_{i\beta}(s) \quad (4.43)$$

where M is a constant depending only on t_0 . In this way, the same reasoning as before gives $\lim_{t \rightarrow \infty} G_{q_i}(t) = \infty$ and our assertion follows. \square

As a result, if there exists only one rationally admissible strategy, we get:

Solvable games.

Corollary 4.4. *Let $X(t)$ be an interior solution path of the replicator equation (4.17) for some dominance-solvable game \mathfrak{G} , and let $x_0 \in \mathcal{A}$ be the (unique) strict equilibrium of \mathfrak{G} . Then:*

$$\lim_{t \rightarrow \infty} X(t) = x_0 \quad \text{almost surely,} \quad (4.44)$$

that is, players converge to the game's strict equilibrium almost surely.

The effects of temperature.

In concluding this section, it is important to note that all our results on the extinction of dominated strategies remain true in the rate-adjusted dynamics (4.16) as well (it is essentially a matter of rescaling, but see also Chapters 5 and 6). The only difference that comes about when players use different learning rates λ_i is in Proposition 4.2 where the CDF estimate (4.36) becomes:

$$\mathbf{P}_x \left(X_{i\alpha}(t) < e^{-M} \right) \geq \frac{1}{2} \operatorname{erfc} \left(\frac{M - h_i(x_i) - \lambda_i v_i t}{2\lambda_i \sigma_i \sqrt{S_i t}} \right). \quad (4.45)$$

As it stands, this is not a significant difference in itself because the two estimates are asymptotically equal for large times. Nonetheless, it is this very lack of contrast that clashes with the deterministic setting where faster learning rates accelerate the emergence of rationality. The reason for this gap is that an increased learning rate λ_i also carries a commensurate increase in the noise coefficients σ_i , and thus deflates the benefits of sharpening payoff differences. In fact, as we shall see in the next sections, the learning rates do not really allow players to learn any faster as much as they help diminish their shortsightedness: *slow learners are better at averaging out the noise.*

4.4 CONGESTION GAMES: A SUGGESTIVE DIGRESSION

Having established that irrational choices die out in the long run, we turn now to the question of whether equilibrial play is stable in the stochastic replicator dynamics of exponential learning. Before tackling this issue in complete generality, it will be quite illustrative to pay a visit to the class of (finite) congestion games where the presence of a potential simplifies things considerably. In this way, the results we obtain here should be considered as a motivating precursor to the general case analysed in Section 4.5.

Descending the potential.

To begin with, it is easy to see that the potential function F of (2.10) is Lyapunov for the deterministic replicator dynamics (4.11). Indeed, assume that player $i \in \mathcal{N}$ is learning at a rate $\lambda_i > 0$, and let $x(t)$ be a solution path of the rate-adjusted dynamics (4.11). Then, a differentiation of $V(t) \equiv F(x(t))$ gives:

$$\begin{aligned} \frac{dV}{dt} &= \sum_{i,\alpha} \frac{\partial F}{\partial x_{i\alpha}} \frac{dx_{i\alpha}}{dt} = - \sum_{i,\alpha} u_{i\alpha}(x) \lambda_i x_{i\alpha} (u_{i\alpha}(x) - u_i(x)) \\ &= - \sum_i \lambda_i \left(\sum_{\alpha}^i x_{i\alpha} u_{i\alpha}^2(x) - u_i^2(x) \right) \leq 0, \end{aligned} \quad (4.46)$$

the last step following from Jensen's inequality – recall that $u_i(x) = \sum_{\alpha}^i x_{i\alpha} u_{i\alpha}(x)$ and also that $\frac{\partial F}{\partial x_{i\alpha}} = -u_{i\alpha}(x)$ on account of (2.10). In particular, this implies that the trajectories $x(t)$ will be attracted to the local minima of F , and since these minima coincide with the strict equilibria of the game, we painlessly infer that strict equilibrial play is asymptotically stable in (2.41).⁷

It is therefore reasonable to ask whether similar conclusions can be drawn in the noisy setting of (4.16). Mirroring the deterministic case, a promising way to go about this question would be to take the potential function F of the game and try to show that it is stochastically Lyapunov in the sense of definition 3.13. Indeed, if $q_0 = (e_{1,0}, \dots, e_{N,0}) \in \Delta$ is a local minimum of F (and hence, a strict equilibrium of the underlying game), we may assume without loss of generality that $F(q_0) = 0$ so that $F(x) > 0$ in a neighbourhood of q_0 . We are thus left to examine the negativity condition of Definition 3.13, i.e. whether there exists some $k > 0$ such that $\mathcal{L}F(x) \leq -kF(x)$ for all x close to q_0 .

To that end, recall that $\frac{\partial F}{\partial x_{i\alpha}} = -u_{i\alpha}$ and that $\frac{\partial^2 F}{\partial x_{i\alpha}^2} = 0$ (by multilinearity). Then, an application of the infinitesimal generator \mathcal{L} of (4.16) to F produces:

$$\begin{aligned} \mathcal{L}F(x) &= - \sum_{i,\alpha} \lambda_i x_{i\alpha} u_{i\alpha}(x) (u_{i\alpha}(x) - u_i(x)) \\ &\quad - \sum_{i,\alpha} \frac{\lambda_i^2}{2} x_{i\alpha} u_{i\alpha}(x) \left[\sigma_{i\alpha}^2 (1 - 2x_{i\alpha}) - \sum_{\beta}^i \sigma_{i\beta}^2 x_{i\beta} (1 - 2x_{i\beta}) \right], \end{aligned} \quad (4.47)$$

where, for simplicity, we assumed that the noise coefficients $\sigma_{i\alpha}$ are constant. This suggests the following plan: study (4.47) term-by-term in a neighborhood of q_0 by considering the perturbed strategies $x_i = (1 - \varepsilon_i)e_{i,0} + \varepsilon_i y_i$ where y_i belongs to the face of Δ_i that lies opposite to $e_{i,0}$ (i.e. $y_{i\mu} \geq 0$, $\mu = 1, 2, \dots$ and $\sum_{\mu} y_{i\mu} = 1$), and $\varepsilon_i > 0$ measures the L^1 distance of x_i from $e_{i,0}$.

Proceeding in this way, we obtain:

$$\begin{aligned} u_i(x) &= \sum_{\alpha}^i x_{i\alpha} u_{i\alpha}(x) = (1 - \varepsilon_i)u_{i,0}(x) + \varepsilon_i \sum_{\mu}^i y_{i\mu} u_{i\mu}(x) \\ &= u_{i,0}(x) + \varepsilon_i \sum_{\mu}^i y_{i\mu} [u_{i\mu}(x) - u_{i,0}(x)] \\ &= u_{i,0}(x) - \varepsilon_i \sum_{\mu}^i y_{i\mu} \Delta u_{i\mu} + \mathcal{O}(\varepsilon_i^2), \end{aligned} \quad (4.48)$$

where $\Delta u_{i\mu} = u_{i,0}(q_0) - u_{i\mu}(q_0) > 0$. Then, by going back to (4.47), we get:

$$\begin{aligned} &\sum_{\alpha}^i x_{i\alpha} u_{i\alpha}(x) [u_{i\alpha}(x) - u_i(x)] \\ &= (1 - \varepsilon_i)u_{i,0}(x) [u_{i,0}(x) - u_i(x)] + \varepsilon_i \sum_{\mu}^i y_{i\mu} u_{i\mu}(x) [u_{i\mu}(x) - u_i(x)] \\ &= \varepsilon_i \sum_{\mu}^i y_{i\mu} u_{i,0}(q_0) \Delta u_{i\mu} - \varepsilon_i \sum_{\mu}^i y_{i\mu} u_{i\mu}(q_0) \Delta u_{i\mu} + \mathcal{O}(\varepsilon_i^2) \\ &= \varepsilon_i \sum_{\mu}^i y_{i\mu} (\Delta u_{i\mu})^2 + \mathcal{O}(\varepsilon_i^2). \end{aligned} \quad (4.49)$$

As for the second term of (4.47), some easy algebra reveals that:

$$\begin{aligned} &\sigma_{i,0}^2 (1 - 2x_{i,0}) - \sum_{\beta}^i \sigma_{i\beta}^2 x_{i\beta} (1 - 2x_{i\beta}) \\ &= -\varepsilon_i \sigma_{i,0}^2 (1 - 2\varepsilon_i) - \varepsilon_i \sum_{\mu}^i \sigma_{i\mu}^2 y_{i\mu} + 2\varepsilon_i^2 \sum_{\mu}^i \sigma_{i\mu}^2 y_{i\mu}^2 \\ &= -\varepsilon_i \left(\sigma_{i,0}^2 + \sum_{\mu}^i y_{i\mu} \sigma_{i\mu}^2 \right) + \mathcal{O}(\varepsilon_i^2). \end{aligned} \quad (4.50)$$

⁷ As we already mentioned in Section 2.2.2, we plead guilty to a slight abuse of terminology in assuming that equilibria in pure strategies are also strict.

Thus, after a (somewhat painful) series of calculations, we finally get:

$$\begin{aligned}
& \sum_{\alpha}^i x_{i\alpha} u_{i\alpha}(x) \left(\sigma_{i\alpha}^2 (1 - 2x_{i\alpha}) - \sum_{\beta}^i \sigma_{i\beta}^2 x_{i\beta} (1 - 2x_{i\beta}) \right) \\
&= (1 - \varepsilon_i) u_{i,0}(x) \left(\sigma_{i,0}^2 (1 - 2x_{i,0}) - \sum_{\beta}^i \sigma_{i\beta}^2 x_{i\beta} (1 - 2x_{i\beta}) \right) \\
&+ \varepsilon_i \sum_{\mu}^i y_{i\mu} \left(\sigma_{i\mu}^2 (1 - 2x_{i\mu}) - \sum_{\beta}^i \sigma_{i\beta}^2 x_{i\beta} (1 - 2x_{i\beta}) \right) \\
&= -\varepsilon_i u_{i,0}(q_0) \left(\sigma_{i,0}^2 + \sum_{\mu}^i y_{i\mu} \sigma_{i\mu}^2 \right) \\
&+ \varepsilon_i \sum_{\mu}^i y_{i\mu} u_{i\mu}(q_0) (\sigma_{i\mu}^2 + \sigma_{i,0}^2) + \mathcal{O}(\varepsilon_i^2) \\
&= -\varepsilon_i \sum_{\mu}^i y_{i\mu} \Delta u_{i\mu} \left(\sigma_{i\mu}^2 + \sigma_{i,0}^2 \right) + \mathcal{O}(\varepsilon_i^2). \tag{4.51}
\end{aligned}$$

Hence, if we assume without loss of generality that $F(q_0) = 0$ and set $\xi = x - q_0$ (i.e. $\xi_{i,0} = -\varepsilon_i$ and $\xi_{i\mu} = \varepsilon_i y_{i\mu}$), we readily get:

$$\begin{aligned}
F(x) &= \sum_{i,\alpha} \frac{\partial F}{\partial x_{i\alpha}} \xi_{i\alpha} + \mathcal{O}(\xi^2) = - \sum_{i,\alpha} \frac{\partial u_i}{\partial x_{i\alpha}} \Big|_{q_0} \xi_{i\alpha} + \mathcal{O}(\varepsilon^2) \\
&= - \sum_{i,\alpha} u_{i\alpha}(q_0) \xi_{i\alpha} + \mathcal{O}(\varepsilon^2) = \sum_i \varepsilon_i \sum_{\mu}^i y_{i\mu} \Delta u_{i\mu} + \mathcal{O}(\varepsilon^2), \tag{4.52}
\end{aligned}$$

where $\varepsilon^2 = \sum_i \varepsilon_i^2$. Therefore, by combining equations (4.49), (4.51) and (4.52), the negativity condition $\mathcal{L}F(x) \leq -kF(x)$ becomes:

$$\sum_i \lambda_i \varepsilon_i \sum_{\mu}^i y_{i\mu} \Delta u_{i\mu} \left[\Delta u_{i\mu} - \frac{\lambda_i}{2} (\sigma_{i\mu}^2 + \sigma_{i,0}^2) \right] \geq k \sum_i \varepsilon_i \sum_{\mu}^i y_{i\mu} \Delta u_{i\mu} + \mathcal{O}(\varepsilon^2). \tag{4.53}$$

So, if $\Delta u_{i\mu} > \frac{\lambda_i}{2} (\sigma_{i\mu}^2 + \sigma_{i,0}^2)$ for all $\mu \in \mathcal{A}_i^*$, this last inequality will be satisfied for some $k > 0$ whenever ε is small enough. Essentially, this proves:

Proposition 4.5. *Let $q = (\alpha_1 \dots \alpha_N)$ be a strict equilibrium of a congestion game \mathcal{G} with potential function F and assume for convenience that $F(q) = 0$. Assume further that the learning rates λ_i are sufficiently small, so that:*

$$F(q_{-i}, e_{i\mu}) > \frac{\lambda_i}{2} (\sigma_{i\mu}^2 + \sigma_{i,0}^2) \quad \text{for all players } i \in \mathcal{N} \text{ and for all } \mu \in \mathcal{A}_i \setminus \{\alpha_i\}. \tag{4.54}$$

Then q is stochastically asymptotically stable in the rate-adjusted dynamics (4.16).

We thus see that no matter how loud the noise σ_i might be, stochastic stability is always guaranteed if the players choose a learning rate that is slow enough as to allow them to average out the noise. Of course, it can be argued here that it is highly unrealistic to expect players to be able to estimate the amount of nature's interference and choose a suitably small rate λ_i . On top of that, the very form of the condition (4.54) is strongly reminiscent of the "modified" game of Hofbauer and Imhof (2009), a similarity which seems to contradict our statement that exponential learning favours rational reactions in the *original* game. The catch here is that condition (4.54) is only *sufficient* and proposition (4.5) merely highlights the role of a potential function in a stochastic environment. As we shall see in the following section, nothing stands in the way of choosing a different Lyapunov candidate and dropping the requirement (4.54) altogether.

THE DYADIC CASE To gain some further intuition into why the condition (4.54) is redundant, it will be particularly helpful to examine the case where

players compete for the resources of only two facilities (i.e. $\mathcal{A}_i = \{0, 1\}$ for all $i \in \mathcal{N}$) and try to learn the game with the help of the uniform replicator equation (4.17). This is the natural setting for the El Farol bar problem posed by Brian Arthur (1994) and the ensuing minority game (Marsili et al., 2000) where players choose to “buy” or “sell” and are rewarded when they are in the minority – buyers in a sellers’ market or sellers in an abundance of buyers.

As has been shown by Milchtaich (1996), such games always possess strict equilibria, even when players have different payoff functions. So, by relabelling indices if necessary, let us assume that $q_0 = (e_{1,0}, \dots, e_{N,0})$ is such a strict equilibrium and let us set $x_i \equiv x_{i,0}$. Then, the generator of the replicator equation (4.16) takes the form:

$$\begin{aligned} \mathcal{L} = & \sum_i x_i(1-x_i) \left[\Delta u_i(x) + \frac{1}{2}(1-2x_i)\sigma_i^2(x) \right] \frac{\partial}{\partial x_i} \\ & + \frac{1}{2} \sum_i x_i^2(1-x_i)^2 \sigma_i^2(x) \frac{\partial^2}{\partial x_i^2}, \end{aligned} \quad (4.55)$$

where now $\Delta u_i \equiv u_{i,0} - u_{i,1}$ and $\sigma_i^2 = \sigma_{i,0}^2 + \sigma_{i,1}^2$.

It thus appears particularly appealing to introduce a new set of variables y_i such that $\frac{\partial}{\partial y_i} = x_i(1-x_i) \frac{\partial}{\partial x_i}$. This is just the logit transformation $y_i = \text{logit } x_i \equiv \log \frac{x_i}{1-x_i}$, and in these variables, (4.55) assumes the form:

$$\mathcal{L} = \sum_i \left(\Delta u_i \frac{\partial}{\partial y_i} + \frac{1}{2} \sigma_i^2 \frac{\partial^2}{\partial y_i^2} \right) \quad (4.56)$$

which reveals that the noise coefficients can be effectively decoupled from the payoffs (cf. (4.55) where the noise coefficients also appear in the $\frac{\partial}{\partial x}$ terms).

To take advantage of this decoupling, we will act with \mathcal{L} on the Lyapunov candidate $f(y) = \sum_i e^{-a_i y_i}$ ($a_i > 0$):

$$\mathcal{L}f(y) = - \sum_i a_i \left(\Delta u_i - \frac{1}{2} a_i \sigma_i^2 \right) e^{-a_i y_i}. \quad (4.57)$$

So, if a_i is chosen small enough so that $\Delta u_i - \frac{1}{2} a_i \sigma_i^2 \geq m_i > 0$ for all sufficiently large y_i (recall that $\Delta u_i(q_0) > 0$ since q_0 is a strict equilibrium), we get:

$$\mathcal{L}f(y) \leq - \sum_i a_i m_i e^{-a_i y_i} \leq -k f(y) \quad (4.58)$$

where $k = \min_i \{a_i m_i\} > 0$. Therefore, since f is strictly positive for $y_{i,0} > 0$ and only vanishes as $y \rightarrow \infty$ (i.e. at the equilibrium q_0), a trivial modification of the Lyapunov method (Gikhman and Skorokhod, 1971, pp. 314–315) yields:

Proposition 4.6. *The strict equilibria of minority games are stochastically asymptotically stable in the uniform replicator equation (4.17).*

Unconditional stability.

Remark 1. It is trivial to see that strict equilibria of minority games will also be stable in the rate-adjusted dynamics (4.16) for any choice of learning rates: simply choose a_i such that $\Delta u_i - \frac{1}{2} a_i \lambda_i \sigma_i^2 \geq m_i > 0$.

Remark 2. A closer inspection of the calculations leading to Proposition 4.6 reveals that nothing hinges on the congestion mechanism per se: it is (4.56) that is crucial to our analysis and \mathcal{L} takes this form whenever the underlying game is a dyadic one (i.e. $|\mathcal{A}_i| = 2$ for all $i \in \mathcal{N}$). In other words, we have:

Proposition 4.7. *Strict equilibria of dyadic games are stochastically asymptotically stable in the stochastic replicator dynamics (4.16), (4.17) of exponential learning.*

4.5 STABILITY OF EQUILIBRIAL PLAY

As we have already noted, in deterministic environments, the “folk theorem” of evolutionary game theory provides some pretty strong ties between equilibrial play and stability: strict equilibria are asymptotically stable in the multi-population replicator dynamics (2.41) (Weibull, 1995). In our stochastic setting, we have already seen that this is always true in two important classes of games: those that can be solved by iterated elimination of dominated strategies (Corollary 4.4) and dyadic ones (Proposition 4.7).

Towards a stochastic folk theorem.

Although interesting in themselves, these results clearly fall short of adding up to a decent analogue of the folk theorem for stochastically perturbed games. Nevertheless, they are quite strong omens in that direction, an expectation which is vindicated in the following:

Theorem 4.8. *The strict equilibria of a Nash game \mathfrak{G} are stochastically asymptotically stable in the replicator dynamics (4.16), (4.17) of exponential learning.*

The adjusted learning scores.

Before proving Theorem 4.8, we should first take a slight detour in order to properly highlight some of the issues at hand. On that account, assume again that the profile $q_0 = (e_{1,0}, \dots, e_{N,0})$ is a strict equilibrium of \mathfrak{G} . Then, if q_0 is to be stochastically stable, say in the uniform dynamics (4.17), one would expect the strategy scores $U_{i,0}$ of player i to grow much faster than the scores $U_{i\mu}$, $\mu \in \mathcal{A}_i^* \equiv \mathcal{A}_i \setminus \{0\}$, of his other strategies. This is captured remarkably well by the “adjusted” scores:

$$Z_{i,0} = \lambda_i U_{i,0} - \log \left(\sum_{\mu}^i e^{\lambda_i U_{i\mu}} \right), \quad (4.59a)$$

$$Z_{i\mu} = \lambda_i (U_{i\mu} - U_{i,0}) \quad (4.59b)$$

where the *sensitivity parameters* $\lambda_i > 0$ are akin (but not identical) to the rates of equation (4.16) (the choice of notation is fairly premeditated though).

Clearly, whenever $Z_{i,0}$ is large, $U_{i,0}$ will be much greater than any other score $U_{i\mu}$ and, hence, the strategy $e_{i,0} \in \mathcal{A}_i$ will be employed by player i far more often. To see this in more detail, it is convenient to introduce the variables:

$$Y_{i,0} \equiv e^{Z_{i,0}} = \frac{e^{\lambda_i U_{i,0}}}{\sum_{\nu}^i e^{\lambda_i U_{i\nu}}}, \quad (4.60a)$$

$$Y_{i\mu} \equiv \frac{e^{Z_{i\mu}}}{\sum_{\nu}^i e^{Z_{i\nu}}} = \frac{e^{\lambda_i U_{i\mu}}}{\sum_{\nu}^i e^{\lambda_i U_{i\nu}}}, \quad (4.60b)$$

The transformed replicator variables.

where $Y_{i,0}$ is a measure of how close X_i is to $e_{i,0} \in \mathcal{A}_i$ while $\sum_{\mu}^i Y_{i\mu} e_{i\mu} \in \Delta(\mathcal{A}_i^*)$ is a direction indicator. The two sets of coordinates are then related by the transformation $Y_{i\alpha} = X_{i\alpha}^{\lambda_i} / \sum_{\mu} X_{i\mu}^{\lambda_i}$, $\alpha \in \mathcal{A}_i$, $\mu \in \mathcal{A}_i^*$, and thus, to show that the strict equilibrium $q_0 = (e_{1,0}, \dots, e_{N,0})$ is stochastically asymptotically stable in the replicator equation (4.17), it will suffice to show that $Y_{i,0}$ diverges to infinity as $t \rightarrow \infty$ with arbitrarily high probability.

Our first step in this direction will be to derive an SDE for the evolution of the $Y_{i\alpha}$ processes. To that end, Itô's lemma gives:

$$\begin{aligned} dY_{i\alpha} &= \sum_{j,\beta} \frac{\partial Y_{i\alpha}}{\partial U_{j\beta}} dU_{j\beta} + \frac{1}{2} \sum_{j,\beta} \sum_{k,\gamma} \frac{\partial^2 Y_{i\alpha}}{\partial U_{j\beta} \partial U_{k\gamma}} dU_{j\beta} \cdot dU_{k\gamma} \\ &= \sum_{\beta}^i \left(u_{i\beta} \frac{\partial Y_{i\alpha}}{\partial U_{i\beta}} + \frac{1}{2} \sigma_{i\beta}^2 \frac{\partial^2 Y_{i\alpha}}{\partial U_{i\beta}^2} \right) dt + \sum_{\beta}^i \sigma_{i\beta} \frac{\partial Y_{i\alpha}}{\partial U_{i\beta}} dW_{i\beta}. \end{aligned} \quad (4.61)$$

On the other hand, a simple differentiation of (4.60a) yields:

$$\frac{\partial Y_{i,0}}{\partial U_{i,0}} = \lambda_i Y_{i,0} \quad \frac{\partial^2 Y_{i,0}}{\partial U_{i,0}^2} = \lambda_i^2 Y_{i,0} \quad (4.62a)$$

$$\frac{\partial Y_{i,0}}{\partial U_{iv}} = -\lambda_i Y_{i,0} Y_{iv} \quad \frac{\partial^2 Y_{i,0}}{\partial U_{iv}^2} = -\lambda_i^2 Y_{i,0} Y_{iv} (1 - 2Y_{iv}), \quad (4.62a')$$

and, similarly, from (4.60b):

$$\frac{\partial Y_{i\mu}}{\partial U_{i,0}} = 0 \quad \frac{\partial^2 Y_{i\mu}}{\partial U_{i,0}^2} = 0 \quad (4.62b)$$

$$\frac{\partial Y_{i\mu}}{\partial U_{iv}} = \lambda_i Y_{i\mu} (\delta_{\mu v} - Y_{iv}) \quad \frac{\partial^2 Y_{i\mu}}{\partial U_{iv}^2} = \lambda_i^2 Y_{i\mu} (\delta_{\mu v} - Y_{iv}) (1 - 2Y_{iv}). \quad (4.62b')$$

In this way, by plugging everything back into (4.61), we finally obtain:

$$\begin{aligned} dY_{i,0} &= \lambda_i Y_{i,0} \left[u_{i,0} - \sum_{\mu}^i Y_{i\mu} u_{i\mu} + \frac{\lambda_i}{2} \sigma_{i,0}^2 - \frac{\lambda_i}{2} \sum_{\mu}^i Y_{i\mu} (1 - 2Y_{i\mu}) \sigma_{i\mu}^2 \right] dt \\ &\quad + \lambda_i Y_{i,0} \left[\sigma_{i,0} dW_{i,0} - \sum_{\mu}^i \sigma_{i\mu} Y_{i\mu} dW_{i\mu} \right], \end{aligned} \quad (4.63a)$$

$$\begin{aligned} dY_{i\mu} &= \lambda_i Y_{i\mu} \left[u_{i\mu} - \sum_{\nu}^i u_{i\nu} Y_{i\nu} \right] dt \\ &\quad + \frac{\lambda_i^2}{2} Y_{i\mu} \left[\sigma_{i\mu}^2 (1 - 2Y_{i\mu}) - \sum_{\nu}^i \sigma_{i\nu}^2 Y_{i\nu} (1 - 2Y_{i\nu}) \right] dt \\ &\quad + \lambda_i Y_{i\mu} \left[\sigma_{i\mu} dW_{i\mu} - \sum_{\nu}^i \sigma_{i\nu} Y_{i\nu} dW_{i\nu} \right], \end{aligned} \quad (4.63b)$$

where we have suppressed the arguments of u_i and σ_i in order to reduce notational clutter.

This last SDE is particularly revealing: roughly speaking, we see that if λ_i is chosen small enough, the deterministic term $u_{i,0} - \sum_{\mu} Y_{i\mu} u_{i\mu}$ will dominate the rest (compare this to the “soft” learning rates of proposition 4.5). Then, since we know that strict equilibria are asymptotically stable in the deterministic case, it is plausible to expect the SDE (4.63) to behave in a similar fashion.

Proof of theorem 4.8. Tying in with our previous discussion, we will establish stochastic asymptotic stability of strict equilibria in the dynamics (4.17) by looking at the processes $Y_i = (Y_{i,0}, Y_{i,1}, \dots) \in \mathbb{R} \times \Delta^{S_i-1}$ of equation (4.60). In the spirit of the previous section, we will accomplish this with the help of the stochastic Lyapunov method that we outlined in Chapter 3 (Definition 3.13).

Unsurprisingly, our first task will be to calculate the generator of the diffusion $Y = (Y_1, \dots, Y_N)$, i.e. the second order differential operator:

$$\mathcal{L} = \sum_{i,\alpha} b_{i\alpha}(y) \frac{\partial}{\partial y_{i\alpha}} + \frac{1}{2} \sum_i \sum_{\alpha,\beta}^i \left(c_i(y) c_i^T(y) \right)_{\alpha\beta} \frac{\partial^2}{\partial y_{i\alpha} \partial y_{i\beta}}, \quad (4.64)$$

where b_i and c_i are the drift and diffusion coefficients of (4.63) respectively. In particular, if we restrict our attention to sufficiently smooth functions of the form $f(y) = \sum_i f_i(y_{i,0})$, the application of \mathcal{L} yields:

$$\begin{aligned} \mathcal{L} f(y) &= \sum_i \lambda_i y_{i,0} \left[u_{i,0} + \frac{\lambda_i}{2} \sigma_{i,0}^2 - \sum_{\mu}^i y_{i\mu} \left(u_{i\mu} - \frac{\lambda_i}{2} (1 - 2y_{i\mu}) \sigma_{i\mu}^2 \right) \right] \frac{\partial f_i}{\partial y_{i,0}} \\ &\quad + \frac{1}{2} \sum_i \lambda_i^2 y_{i,0}^2 \left[\sigma_{i,0}^2 + \sum_{\mu}^i \sigma_{i\mu}^2 y_{i\mu}^2 \right] \frac{\partial^2 f_i}{\partial y_{i,0}^2}. \end{aligned} \quad (4.65)$$

*Bounding
the noise.*

Therefore, let us consider the function $f(y) = \sum_i 1/y_{i,0}$ for $y_{i,0} > 0$. With $\frac{\partial f}{\partial y_{i,0}} = -1/y_{i,0}^2$ and $\frac{\partial^2 f}{\partial y_{i,0}^2} = 2/y_{i,0}^3$, equation (4.65) becomes:

$$\mathcal{L}f(y) = -\sum_i \frac{\lambda_i}{y_{i,0}} \left[u_{i,0} - \sum_\mu^i u_{i\mu} y_{i\mu} - \frac{\lambda_i}{2} \sigma_{i,0}^2 - \frac{\lambda_i}{2} \sum_\mu^i y_{i\mu} (1 - y_{i\mu}) \sigma_{i\mu}^2 \right]. \quad (4.66)$$

However, since $q_0 = (e_{1,0}, \dots, e_{N,0})$ has been assumed to be a strict Nash equilibrium of \mathfrak{G} , we will have $u_{i,0}(q_0) > u_{i\mu}(q_0)$ for all $\mu \in \mathcal{A}_i^*$. Thus, by continuity, there exists some positive constant $v_i > 0$ with $u_{i,0} - \sum_\mu u_{i\mu} y_{i\mu} \geq v_i > 0$ whenever $y_{i,0}$ is large enough (recall that $\sum_\mu y_{i\mu} = 1$). So, if we pick positive $\lambda_i < v_i/\sigma_i^2$ where $\sigma_i = \text{ess sup}\{|\sigma_{i\beta}(x)| : x \in \Delta, \beta \in \mathcal{A}_i\}$, we will get:

$$\mathcal{L}f(y) \leq -\sum_i \frac{\lambda_i v_i}{2} \frac{1}{y_{i,0}} \leq -\frac{1}{2} \min_i \{\lambda_i v_i\} f(y) \quad (4.67)$$

for all sufficiently large $y_{i,0}$. Moreover, f is strictly positive for $y_{i,0} > 0$ and vanishes only as $y_{i,0} \rightarrow \infty$. Hence, as in the proof of Proposition 4.6, our claim follows on account of f being a (local) stochastic Lyapunov function, and, to obtain the general case for the rate-adjusted replicator dynamics (4.16), we only need to rescale the parameters λ_i accordingly. \square

At the other end of the spectrum, equilibria which are not pure are not even stationary states of the replicator dynamics (4.16), thus leading to the following incarnation of the folk theorem of evolutionary game theory:

The stochastic folk theorem.

Corollary 4.9. *Let \mathfrak{G} be a Nash game. Then, the stochastically asymptotically stable states of the replicator dynamics (4.16) coincide with the strict equilibria of \mathfrak{G} .*

Remark 1. If we trace our steps back to the coordinates $X_{i\alpha}$, our Lyapunov candidate takes the form $f(x) = \sum_i x_{i,0}^{-\lambda_i} \sum_\mu^i x_{i\mu}^{\lambda_i}$. It thus begs to be compared to the Lyapunov function $\sum_\mu x_\mu^\lambda$ employed by Hofbauer and Imhof (2009) to derive a conditional version of Theorem 4.8 in the evolutionary setting. As it turns out, the obvious extension $f(x) = \sum_i \sum_\mu^i x_{i\mu}^{\lambda_i}$ works in our case as well, but the calculations are much more cumbersome and they are also shorn of their ties to the adjusted scores (4.59).

Remark 2. We should not neglect to highlight here the dual role that the learning rates λ_i play in our analysis. In the logistic learning model (4.6) they measure the players' convictions and how strongly they react to a given stimulus (the scores $U_{i\alpha}$); in this role, they are fixed at the outset of the game and form an intrinsic part of the rate-adjusted dynamics (4.16). On the other hand, they also make a virtual appearance as free temperature parameters in the adjusted scores (4.59), to be softened until we get the desired result. For this reason, even though Theorem 4.8 remains true for any choice of learning rates, the function $f(x) = \sum_i x_{i,0}^{-\lambda_i} \sum_\mu^i x_{i\mu}^{\lambda_i}$ is Lyapunov only if the sensitivity parameters λ_i are small enough. It might thus seem unfortunate that we chose the same notation in both cases, but we feel that our decision is justified by the intimate relation of the two parameters.

The contrast with aggregate shocks.

To contrast things with single-population evolutionary models (the subject of Imhof, 2005), recall that an evolutionarily stable strategy (ESS) is a strategy which is robust against invasion by mutant phenotypes (Definition 2.10). As we have already noted, these strategies are precisely the ones which attract evolutionary processes (Theorem 2.16), so one might well ask why we have not included them in our stochastic considerations.

The reason for this omission is pretty simple: even the weakest evolutionary criteria in multi-population models tend to reject all strategies which are not strict Nash equilibria (Proposition 2.12). Therefore, since the learning model (4.11) essentially corresponds to the multi-population random matching dynamics (2.41), we lose nothing by focusing our analysis on the strict equilibria of the game. If anything, this equivalence between ESS and strict equilibria in multi-population settings further highlights the importance of the latter.

However, this also brings out the gulf between the single-population setting and our own, even if we restrict ourselves to 2-player matchings (which are the norm in single-population models). Indeed, the single-population version of the dynamics (4.20) is:

$$\begin{aligned} dX_\alpha &= X_\alpha \left[(u_\alpha(X) - u(X, X)) - \left(\sigma_\alpha^2 X_\alpha - \sum_\beta \sigma_\beta^2 X_\beta^2 \right) \right] dt \\ &+ X_\alpha \left[\sigma_\alpha dW_\alpha - \sum_\beta \sigma_\beta X_\beta dW_\beta \right], \end{aligned} \quad (4.68)$$

and, because of the term $u(X, X) = \sum_{\alpha, \beta} u_{\alpha\beta} X_\alpha X_\beta$, these dynamics are *cubic* in X – and not quadratic, as is the case with (4.16). Because of this self-interacting structure (which is not shared by the multilinear payoff functions of (4.16)), it turns out that in the presence of an interior ESS and mild enough shocks, the solution paths of (4.68) are recurrent and the transition probabilities of the processes $X(t)$ converge in total variation to an invariant distribution (Imhof, 2005, Theorem 2.1).

Theorem 4.8 rules out such behaviour in the case of strict equilibria (the multi-population analogue of an ESS), but does not answer the following question: if the underlying game only has mixed equilibria, will the solution $X(t)$ of the dynamics (4.16) be recurrent? In single-population matchings, the comparison of the payoff matrix to the (diagonal) matrix of noise coefficients leads to an array of global estimates which enabled Imhof (2005) to derive a set of conditions for recurrence. In multi-population environments however, this comparison is uneven and we need more (global) information in order to attack this question; when such information is available (e.g. in the form of a potential function), studying the long-term behavior of the replicator dynamics near interior points becomes again a feasible undertaking.

Evolutionary stability.

APPLICATIONS TO NETWORK DESIGN In concluding this chapter, it is worth pointing out noise and uncertainty in networks have two general sources. The first of these has to do with the time variability of the connections which may be due to the fluctuations of the link quality because of mobility in the wireless case or because of external factors (e.g. load conditions) in wireline networks. This variability is usually dependent on the state of the network and was actually our original motivation in considering noise coefficients $\sigma_{i\alpha}$ that are functions of the players' strategy profile – see Chapter 6 for more on this topic.⁸

Potential applications.

The second source stems from errors in the measurement of the payoffs themselves (e.g. the throughput obtained in a particular link) and also from the lack of information on the payoff of strategies that were not employed. The variability of the noise coefficients $\sigma_{i\alpha}$ again allows for a reasonable approximation to this problem. Indeed, if $\sigma_{i\alpha} : \Delta \rightarrow \mathbb{R}$ is continuous and satisfies $\sigma_{i\alpha}(x_{-i}; \alpha) = 0$ for all $i \in \mathcal{N}, \alpha \in \mathcal{A}_i$, this means that there are only

⁸ Incidentally, this was also our original motivation for considering randomly fluctuating payoffs in the first place: travel times and delays in traffic models are not determined solely by the players' choices, but also by the fickle interference of nature.

errors in estimating the payoffs of strategies that were not employed (or small errors for pure strategies that are employed with high probability). Of course, this does not yet give the full picture: in many cases, one should consider a discrete-time dynamical system (and games where players are not allowed to mix their strategies), but we conjecture that our results will remain essentially unaltered – see also Chapters 6 and 7.

5

STOCHASTIC PERTURBATIONS IN POPULATION GAMES

IN terms of structure, Nash-type games are relatively simple to analyze: there is a finite number of players, each with a finite number of available actions, and the players' (expected) payoffs are multilinear functions of the players' (mixed) strategies. As a result of this straightforward configuration, it is also relatively easy to write down equations of evolutionary growth, for which the folk theorem (Theorem 2.16) provides a definitive set of conditions that guarantee convergence to equilibrium (or, alternatively, lock players in perpetual cycles such as the ones observed in symmetric games of the Rock-Paper-Scissors variety).

In light of all this, it was only natural to pick Nash games as the starting point of our stochastic considerations, just like Fudenberg and Harris (1992), Cabrales (2000) and Imhof (2005) used random Nash matchings to study evolution in the presence of aggregate shocks. Our aim in this chapter will be to take things one step further and investigate what happens in nonatomic population games played by a continuum of players who employ a Gibbs-type learning scheme similar to (4.13). As such, we will still be working with the replicator dynamics of exponential learning (4.16) (and not with the biologically oriented version (4.20)) but, now, the payoff functions will no longer adhere to the multilinear form (2.1).

Despite this added degree of generality, we will see that most of our results in Chapter 4 carry over to this nonlinear environment without major modifications. At the same time however, this generality also gives rise to a (potentially) much richer landscape of equilibria in the interior of the strategy space, and these equilibria cannot be explored by the techniques of the last chapter. To probe the stability properties of these interior steady states, we will need information on the players' aggregate behavior and, for this reason, a large part of this chapter's analysis will be devoted to (continuous) potential games à la Sandholm (2001, 2005).

OUTLINE We begin by exploring the rationality properties of the replicator dynamics in deterministic environments. Just as in random matching scenarios, dominated strategies become extinct in any population game, and, with a little more work, it is not hard to see that the strict equilibria of any such game are asymptotically stable. As it turns out, this last result is a consequence of a much deeper property of the replicator dynamics in general population games: *any* population state which is evolutionarily stable (in the sense of Definition 2.10) will also be asymptotically stable (Proposition 5.3).

In view of the folk theorem, one would hope that the converse of this statement also holds, namely that *only* evolutionarily stable strategies can be attracting (at least in multi-population games). This, however, is not true, and we provide a counterexample of a game (an adjusted version of the

*A return to
population games.*

Matching Pennies game) which has an interior equilibrium that attracts *all* interior replicator trajectories, but which is not evolutionarily stable.

A key notion in all these considerations is the *evolutionary index* of a population distribution, a function whose sign determines whether a mutation is destructive to the state of the population or not. If the underlying game admits a potential function that satisfies a mild growth condition (e.g. if it is convex), then the study of this index reveals that the global minimum of the potential will be evolutionarily stable and will attract all interior replicator orbits, thus leading to a partial equivalence between dynamical and evolutionary stability. Even more importantly however, by estimating the growth of this index, we are able to show that the replicator dynamics actually converge to equilibrium at an *exponential* rate (Theorems 5.5 and 5.6).

Section 5.2 is devoted to the effect of stochastic fluctuations in this nonatomic setting. Working with the stochastic replicator dynamics that are derived from the scheme of exponential learning, an extension of the methodology that we developed in Chapter 4 for Nash games shows that dominated strategies become extinct in any population game irrespective of the noise level (Theorem 5.8), and that strict equilibria remain (stochastically) asymptotically stable (Theorem 5.9). In fact, if the game admits a (strictly) convex potential function and the players are learning at sufficiently high “learning temperatures”, we are even able to estimate the average time it takes them to hit a neighborhood of the strict equilibrium in question (Theorem 5.10).

In conjunction with the stochastic stability of strict equilibria, this last result allows us to conclude that when a strict equilibrium exists, the stochastic replicator dynamics of exponential learning will converge to it almost surely (Corollary 5.12). On the other hand, given that strict equilibria do not always exist, we must also determine what happens in the case of *interior* evolutionarily stable states – which are no longer stationary in the stochastic replicator dynamics, and, hence, cannot be stable either. Nevertheless, if the populations are “patient enough”, we prove that the replicator trajectories are *recurrent* in (strictly) convex potential games, and we use this fact to prove that the long-term average of the population distribution concentrates its mass in the vicinity of the game’s unique ESS (Theorem 5.13).

5.1 DETERMINISTIC EVOLUTION IN POPULATION GAMES

Population
games,

To begin with, let us briefly recall the general structure of (nonatomic) population games as we presented them in Chapter 2 (where we refer the reader for more details). As the name suggests, the core ingredient of a population game \mathfrak{G} is a collection of N continuous populations (species) of players represented by the intervals $\mathcal{N}_k = [0, m_k]$, $k = 1, \dots, N$, where $m_k > 0$ denotes the “mass” of the k -th population under Lebesgue measure. As is the norm in games of this type, we will assume that the players of each population share a common (and finite) set of possible actions $\mathcal{A}_k = \{\alpha_0, \alpha_1, \dots\}$, and also that their payoffs only depend (anonymously) on the (continuous) distribution of the various players among the action sets \mathcal{A}_k .¹

More precisely, such a distribution (also referred to as a *population state*) will be represented as a point in the strategy space $\Delta \equiv \prod_k \Delta_k$ where Δ_k is the (weighted) simplex $\Delta_k \equiv m_k \Delta(\mathcal{A}_k) = \{x_k \in \mathbb{R}^{\mathcal{A}_k} : x_{k\alpha} \geq 0 \text{ and } \sum_{\alpha} x_{k\alpha} = m_k\}$. In this way, our anonymity assumption above simply means that the payoff to those members of population \mathcal{N}_k who employed $\alpha \in \mathcal{A}_k$ will be given by some

¹ For the distinction between strategy profiles and distributions, recall our discussion in Chapter 2.

(continuous) function $u_{k\alpha} : \Delta \rightarrow \mathbb{R}$. As a result of this setup, the equilibrial set $\Delta^* \equiv \Delta^*(\mathcal{G})$ of \mathcal{G} will be characterized by *Wardrop's principle* (2.23):

$$u_{k\alpha}(q) \geq u_{k\beta}(q) \text{ for all } \alpha, \beta \in \mathcal{A}_k \text{ such that } q_{k\alpha} > 0, \quad (5.1a)$$

which, in its strictest form for pure equilibria (vertices of Δ) becomes:

$$u_{k,0}(q) > u_{k\mu}(q) \text{ for all } \mu \in \mathcal{A}_k^* \equiv \mathcal{A}_k \setminus \{0\}, \quad (5.1b)$$

where $q = \sum_k m_k e_{k,0} \in \prod_k \mathbb{R}^{\mathcal{A}_k}$ is the pure equilibrium in question (of course, the choice of the vertex $\sum_k m_k e_{k,0}$ in Δ was arbitrary).

The parting point of this definition from the more traditional “random matching” scenario considered in evolutionary game theory is that these payoff functions are not assumed multilinear over the coordinates of the population state $x \in \Delta$. This complicates things considerably when trying to calculate the equilibria of a game, so, to keep things tractable, we will restrict much of our analysis to the class of *potential games*, whose payoffs are derived from a *potential function*:

$$u_{k\alpha}(x) = -\frac{\partial F}{\partial x_{k\alpha}} \quad (5.2)$$

(see Section 2.1.3 for more details). If such a potential function exists, then the game's equilibrial set coincides with the critical set of F (Sandholm, 2001).

5.1.1 Exponential Learning and Rational Behavior

Now, as in Nash-type games, we assume (or, rather, propose) that each player tries to adapt to his opponents by employing the “exponential learning” scheme that we introduced in Section 4.1. In particular, this means that at each instance of the game, a player looks at his payoff and at the rewards of the other “phenotypes” (actions) in his “species”, and then updates the cumulative scores:

$$dU_{k\alpha}(t) = u_{k\alpha}(x(t)) dt, \quad (5.3)$$

where $x(t)$ denotes the population state at time t .²

In this way, if players choose their actions based on the Gibbs probability distribution $p_{k\alpha} = \exp(\lambda_k U_{k\alpha}) / \sum_{\beta}^k \exp(\lambda_k U_{k\beta})$ with learning rates (population-specific inverse temperatures) $\lambda_k > 0$, $k = 1, \dots, N$, the population state $x(t)$ at time t will be given by:

$$x_{k\alpha}(t) = \frac{e^{\lambda_k U_{k\alpha}(t)}}{\sum_{\beta}^k e^{\lambda_k U_{k\beta}(t)}} m_k, \quad (5.4)$$

since, clearly, the population shares $x_{k\alpha}$ will have to sum up to the total population mass of the k -th species: $\sum_{\alpha}^k x_{k\alpha} = m_k$.

² Appearances would dictate otherwise, but there is a fundamental difference between (5.3) and its Nash counterpart (4.5): players in population games do not need to know the full strategy profile of their opponents. Indeed, thanks to the way that payoffs are determined in population games, a player only needs to look at the rewards obtained by those neighbors in his population who employed different actions. Because this observation is empirical and not fictitious (there are no probabilities involved, only actual, observable quantities), exponential learning becomes even more relevant in population games than it is in Nash games.

Wardrop's principle,

and potential functions.

Exponential learning in population games.

Thus, if we differentiate (5.4) to decouple it from (5.3), we again obtain the rate-adjusted replicator equation:

$$\frac{dx_{k\alpha}}{dt} = \lambda_k x_{k\alpha} (u_{k\alpha}(x) - u_k(x)), \quad (5.5)$$

where $u_k(x)$ denotes the *population average*:

$$u_{k\alpha}(x) = \frac{1}{m_k} \sum_{\beta}^k x_{k\beta} u_{k\beta}(x) \quad (5.6)$$

(contrast this with the multilinear expectation (2.1) for Nash games).

*Dominance
and evolution.*

Before proceeding to the stochastic version of (5.5), it would be a good idea to see which rationality properties of the replicator dynamics continue to hold in this nonlinear context. The first solution concept that we will attack is that of a dominated strategy, which, in contrast to the Nash case, can only be *pure* in population games: $\alpha \in \mathcal{A}_k$ is dominated by $\beta \in \mathcal{A}_k$ ($\alpha \prec \beta$) iff $u_{k\alpha}(x) < u_{k\beta}(x)$ for all $x \in \Delta$ – after all, it makes little sense for a population state $x_k \in \Delta_k$ to be “dominated” in our setting. Despite this structural difference, the proof of Theorem 2.15 applied to pure strategies remains unchanged and yields:

Proposition 5.1. *Let \mathfrak{G} be a population game and let $x(t)$ be an interior solution orbit of the replicator dynamics (5.5). If the strategy $\alpha \in \mathcal{A}_k$ is dominated, then:*

$$\lim_{t \rightarrow \infty} x_{k\alpha}(t) = 0. \quad (5.7)$$

More generally, the same is true if $\alpha \in \mathcal{A}_k$ is not rationally admissible.

In fact, by adapting the same proof to strict equilibria, we also obtain:

Proposition 5.2. *The strict equilibria of population games are always asymptotically stable in the replicator dynamics (5.5).*

Proof. Though straightforward, the proof of this proposition will have remarkable ramifications for our stochastic considerations later on (and has a deep connection to our approach in Chapter 4), so it is worth examining it in detail.

On that account, let us assume without loss of generality that $q = \sum_k m_k e_{k,0}$ is a strict equilibrium of the game, so that $u_{k,0}(q) > u_{k\mu}(q)$ for all $\mu \in \mathcal{A}_k^* \equiv \mathcal{A}_k \setminus \{0\}$, and for all populations $k = 1, \dots, N$. Then, for any solution trajectory $x(t)$ which starts at finite Kullback-Leibler distance from q , the replicator equation (5.5) readily yields:

$$\begin{aligned} \frac{d}{dt} H_{q_k}(x_k(t)) &= -\frac{d}{dt} \log x_{k,0}(t) = -\lambda_k (u_{k,0}(x) - u_k(x)) \\ &= -\frac{\lambda_k}{m_k} \left(u_{k,0}(x)(m_k - x_{k,0}) - \sum_{\mu}^k x_{k\mu} u_{k\mu}(x) \right). \end{aligned} \quad (5.8)$$

Consider now the perturbed population state $x_k = q_k + \varepsilon_k m_k z_k$ where $\varepsilon_k > 0$ and the components of $z_k = \sum_{\alpha} z_{k\alpha} e_{k\alpha}$ satisfy:

$$\sum_{\mu}^k z_{k\mu} = 1 = -z_{k,0} \text{ and } z_{k\mu} \geq 0 \text{ for all } \mu \in \mathcal{A}_k^* \text{ and for all } k = 1, \dots, N \quad (5.9)$$

(i.e. z_k is tangent to Δ_k at q_k and points towards its interior). We will then have $m_k - x_{k,0} = \varepsilon_k m_k$, and (5.8) becomes:

$$\frac{dH_{q_k}}{dt} = -\lambda_k \varepsilon_k \left(u_{k,0}(x) - \sum_{\mu}^k z_{k\mu} u_{k\mu}(x) \right) = -\lambda_k \varepsilon_k \sum_{\mu}^k z_{k\mu} \Delta u_{k\mu}(x), \quad (5.10)$$

where $\Delta u_{k\mu}(x) = u_{k,0}(x) - u_{k\mu}(x)$ (note here the similarities with the proof of Proposition 4.5).

However, since $u_{k,0}(q) > u_{k\mu}(q)$ for all $\mu \in \mathcal{A}_k^*$, the convex combination $\sum_{\mu} z_{k\mu} \Delta u_{k\mu}(q)$ will also be positive, and, by continuity, the same will hold for all x sufficiently close to q (that is, for small enough $\varepsilon_k > 0$). Consequently, if we sum (5.9) over all $k = 1, \dots, N$, we see that the Kullback-Leibler entropy $H_q = \sum_k H_{qk}$ is a local Lyapunov function for the replicator dynamics (5.5), and our assertion follows. \square

So far, the rationality properties of the replicator dynamics in population games appear identical to their properties in (multi-population) random matching Nash games: dominated strategies become extinct and strict equilibria are asymptotically stable. One would thus hope to complete this picture by supplying a converse to Proposition 5.2, to the effect that *only* strict equilibria can be stable under the replicator dynamics (5.5). In many ways, such a result would be the perfect complement to Theorem 2.16 and would provide a remarkably robust (and simple) equilibrium selection mechanism. However, it turns out that this is a “bridge too far”: as we shall see, there are many (classes of) games which possess interior equilibria that attract *every* interior solution of the replicator dynamics.³

So, instead of looking only at the strict equilibria of the game (when they even exist), we should expand our search to include more diverse equilibria as well. To that end, let $q \in \text{Int}(\Delta)$ be an interior equilibrium of some population game \mathcal{G} which evolves according to the rate-adjusted replicator dynamics (5.5) with learning rates $\lambda_k = 1$ (for simplicity). Then, the time derivative of the relative entropy function $H_q(x) = \sum_{k,\alpha} q_{k\alpha} \log(q_{k\alpha}/x_{k\alpha})$ will be:

$$\frac{dH_q}{dt} = \sum_{k,\alpha} \frac{\partial H_q}{\partial x_{k\alpha}} \frac{dx_{k\alpha}}{dt} = - \sum_{k,\alpha} q_{k\alpha} (u_{k\alpha}(x) - u_k(x)) = -L_q(x), \quad (5.11)$$

where, after a simple rearrangement, we have set:

$$L_q(x) = - \sum_k \sum_{\alpha} (x_{k\alpha} - q_{k\alpha}) u_{k\alpha}(x). \quad (5.12)$$

Obviously, if $L_q(x)$ is positive (resp. negative) in a neighborhood of q , then q will be locally attracting and stable (resp. repelling and unstable); on the other hand, if the sign of L_q changes depending on the direction that we take away from q , then more complicated phenomena might arise.

This last observation already makes the function L_q crucial for our purposes, but its true essence lies in its connection to the concept of evolutionary stability. Indeed, if we turn back to Definition 2.10, we cannot fail to notice the similarity between the evolutionary stability condition (2.36) and the expression (5.12); to boot, if q is evolutionarily stable, (2.36) is equivalent to the requirement:

$$\sum_{k,\alpha} z_{k\alpha} u_{k\alpha}(q + \varepsilon z) < 0, \quad (5.13)$$

for all sufficiently small $\varepsilon > 0$ and for every tangent vector $z \in T_q\Delta$ with Euclidean norm $\|z\| = 1$.

³ This failure is also encountered in the ingenious approach of Ritzberger and Vogelsberger (1990) – see also Weibull (1995, pp. 227–228). Working with multi-population random matching games, they showed that the replicator vector field of (5.5) has the same orbits as a divergence-free vector field; from Liouville’s theorem, it then follows that only pure states can be attracting. However, if the game’s payoff functions do not satisfy the multilinearity condition $\frac{\partial u_{k\alpha}}{\partial x_{k\alpha}} = 0$, the divergence of this “Nash field” does not necessarily vanish and this approach breaks down.

Domains of stability.

In other words, we see that the mutated state $x = q + \varepsilon z$ will be “evolutionarily destructive” for q if and only if $L_q(x) < 0$. Prompted by this observation, $L_q(x)$ will be called the *evolutionary index* of x (with respect to q) and, accordingly, we will say that x is:

- *congruous* to q (or *preserving*) when $L_q(x) > 0$;
- *incongruous* to q (or *destructive*) when $L_q(x) < 0$;
- *neutral* with respect to q when $L_q(x) = 0$.

In a similar fashion, the *stability domain* of q will be the (necessarily open) set $\Delta_q^+ = \{x \in \Delta : L_q(x) > 0\}$ of points that are congruous to q – obviously, the *unstable* and *neutral* domains Δ_q^- and Δ_q^0 of q are defined analogously.

Evolutionary stability vs. asymptotic stability.

The rationale behind this terminology should be clear enough: q is evolutionarily stable if and only if every state x in some neighborhood U of q preserves q . Obviously, this last condition amounts to $L_q(x)$ being positive in a neighborhood of q , so, by (5.11) and Lyapunov’s theorem, we get:

Proposition 5.3. *Evolutionarily stable states are asymptotically stable in the replicator dynamics (5.5).*

Remark. Strictly speaking, the proof of Proposition 5.3 only works in the “uniform” replicator dynamics with learning rates $\lambda_k = 1$. In the general case, (5.11) does not hold and we should consider instead the *rate-adjusted* entropy:

$$H_q(x; \lambda) \equiv \sum_k \lambda_k^{-1} H_{q_k}(x_k) = \sum_k \lambda_k^{-1} \sum_{\alpha: q_{k\alpha} > 0} q_{k\alpha} \log \frac{q_{k\alpha}}{x_{k\alpha}}. \quad (5.14)$$

The rate-adjusted entropy $H_q(x; \lambda)$ has the same properties as its uniform counterpart $H_q(x)$ (which is recovered when we set $\lambda_k = 1$), and the same calculations as above show that $\dot{H}_q(x; \lambda) = -L_q(x)$ in the replicator dynamics (5.5). For this reason, we will not make any distinction between the two versions and, for simplicity, we will use the notation $H_q(x)$ for both (however, we *will* make explicit mention of the learning rates λ_k when needed).

Now, a closer inspection of Theorem 2.16 reveals that Proposition 5.3 is, essentially, the “if” part of the folk theorem in general population games. As a result, the “correct” characterization of asymptotically stable states should be phrased as a converse to Proposition 5.3, and not Proposition 5.2 as we tried before – after all, the equivalence between strict equilibria and evolutionarily stable strategies in multi-population Nash games is just an artifact of multilinearity.

Matching Pennies as a counterexample.

This statement appears even more appealing than our previous attempt to characterize the asymptotically stable points of the replicator dynamics, but, unfortunately, our expectations are not vindicated this time either. A counterexample in multi-population games is provided by the following adaptation of the “Matching Pennies” game, inspired by Weibull (1995, pp. 198–199) – a single-population example was first given by Zeeman (1981).⁴ Two species, U and V , both of mass 1, choose between “heads” (0) or “tails” (1), gaining respectively the “raw” rewards:

$$\begin{aligned} u_0(x, y) &= 2y + c & v_0(x, y) &= 2 + c - 2x \\ u_1(x, y) &= 2 + c - 2y & v_1(x, y) &= 2x + c \end{aligned} \quad (5.15)$$

⁴ In the original Matching Pennies game, two players, A and B , each hold up a penny: if the faces of the pennies match, player A takes both, otherwise the pennies go to B .

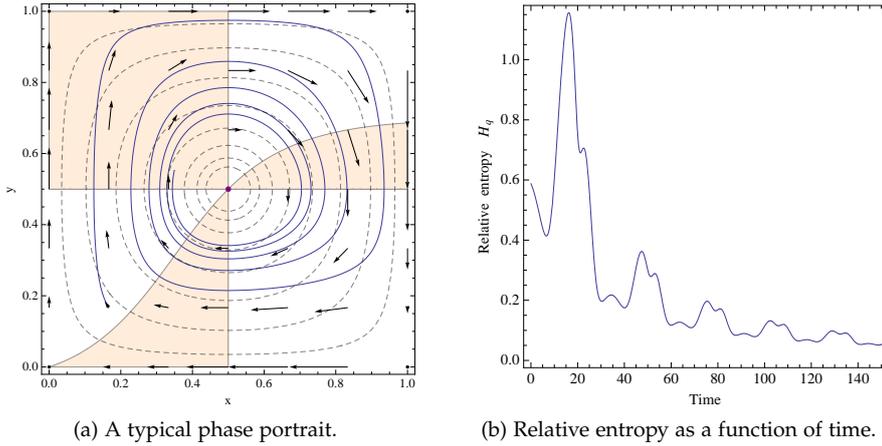


Figure 5.1: An attracting equilibrium which is not evolutionarily stable in the payoff-adjusted Matching Pennies game ($c = 0.01$). The grey contours in Fig. 5.1a are the level sets of the relative entropy $H_q(x)$, while the orange region represents the evolutionarily unstable domain $\Delta_q^- = \{x : L_q(x) < 0\}$. As can be seen in Fig. 5.1b, the K-L divergence increases when the replicator trajectories enter the domain of evolutionary instability.

where x (resp. y) is the share of species U (resp. V) that picked 0, and $c > 0$ is a control parameter. These raw rewards are then averaged over each species (giving $u(x, y) = xu_0(x, y) + (1 - x)u_1(x, y) = c + 2[xy + (1 - x)(1 - y)]$ and a similar expression for v), and the actual payoff functions to the two species are the “adjusted” versions:

$$\begin{aligned} \bar{u}_0(x, y) &= u_0(x, y)/u(x, y) & \bar{v}_0(x, y) &= v_0(x, y)/v(x, y) \\ \bar{u}_1(x, y) &= u_1(x, y)/u(x, y) & \bar{v}_1(x, y) &= v_1(x, y)/v(x, y). \end{aligned} \quad (5.16)$$

In this context, the role of the control parameter c is to artificially increase the raw rewards u and v by the same amount, thus ensuring that the adjusted payoffs \bar{u} and \bar{v} remain well-defined – the standard version of Matching Pennies has $c = -1$. As a result, given that this adjustment is uniform across populations, the equilibrium of the game will not change, and the adjusted version will still have a unique equilibrium at $q = (1/2, 1/2)$. Nevertheless, Fig. 5.1 shows that this equilibrium is *not* evolutionarily stable (every neighborhood of q intersects the instability domain Δ_q^-), even though it is attracting: interior trajectories converge to equilibrium, but the relative entropy of a solution fluctuates quite a bit before eventually settling down.

5.1.2 Evolution in Potential Games

So, does all this amount to a clear account of the stability properties of the replicator dynamics in population games? On the one hand, yes, and in a definitive manner: Proposition 5.3 shows that evolutionarily stable states are also asymptotically stable, and the Matching Pennies counterexample shows that we cannot hope to obtain a converse to this statement. On the other hand, this answer is, in a certain sense, not much more than a restatement of the problem (an illustrative one, to be sure, but a restatement nonetheless): we are still left to know whether a population game actually has any evolutionarily stable states or not.

The role of the evolutionary index.

This last question boils down to showing that the evolutionary index $L_q(x)$ is positive in a neighborhood of a candidate equilibrium $q \in \Delta$. As such, the answer will depend on the game (or class thereof) considered and, essentially, it will be an ad hoc one. Nevertheless, we can obtain pretty general results if we restrict ourselves to the class of continuous potential games, where the existence of a potential function allows us to express the evolutionary index L_q in a particularly suggestive way.

Indeed, if \mathfrak{G} is a potential game with potential $F : \Delta \rightarrow \mathbb{R}$, we immediately obtain:

$$L_q(x) = - \sum_k \sum_\alpha^k (x_{k\alpha} - q_{k\alpha}) u_{k\alpha}(x) = \sum_k \sum_\alpha^k (x_{k\alpha} - q_{k\alpha}) \frac{\partial F}{\partial x_{k\alpha}}, \quad (5.17)$$

an expression which suggests that the evolutionary index L_q may be interpreted as a sort of directional derivative of F . To make this idea precise, note first that if q_k is an interior point of Δ_k , the tangent space $T_{q_k} \Delta_k$ to Δ_k at q_k will be naturally isomorphic to the subspace Z_k of $\mathbb{R}^{\mathcal{A}_k}$ which is parallel to Δ_k :

$$T_{q_k} \Delta_k \cong Z_k = \{z_k \in \mathbb{R}^{\mathcal{A}_k} : \sum_\alpha^k z_{k\alpha} = 0\}. \quad (5.18)$$

Of course, an analogous statement will also hold for $T_q \Delta$: if we set $\mathcal{A} \equiv \coprod_k \mathcal{A}_k$, then $T_q \Delta$ will be isomorphic to the subspace of $\mathbb{R}^{\mathcal{A}}$ which is parallel to Δ :

$$T_q \Delta \cong Z = \{z \in \mathbb{R}^{\mathcal{A}} : \sum_\alpha^k z_{k\alpha} = 0 \text{ for all } k = 1, \dots, N\}. \quad (5.19)$$

(Obviously $Z \cong \prod_k Z_k$ under the identification $\mathbb{R}^{\mathcal{A}} \cong \prod_k \mathbb{R}^{\mathcal{A}_k}$.)

Tangent cones.

However, for points on the boundary of Δ , we will need to take a slightly different approach than the differential geometric one where $T_{q_k} \Delta_k$ has dimension equal to that of the lowest-dimensional face of Δ_k containing q_k . Instead, we will regularly need to consider all “rays” that emanate from q and which are (at least partially) contained in Δ , i.e. those vectors $z \in Z$ such that $q + \theta z \in \Delta$ for some $\theta > 0$. We will thus denote this (solid) tangent cone by

$$T_q^c \Delta = \{z \in Z : z_{k\alpha} \geq 0 \text{ for all } \alpha \in \mathcal{A}_k \text{ with } q_{k\alpha} = 0\}, \quad (5.20)$$

so that $T_q^c \Delta$ always contains $T_q \Delta$ and coincides with it whenever $q \in \text{Int}(\Delta)$.

In this manner, if we set $f(\theta) = F(q + \theta z)$ and $h(\theta) = L_q(q + \theta z)$, $z \in T_q^c \Delta$, (5.17) may be rewritten as:

$$f'(\theta) = \sum_k \sum_\alpha^k \frac{\partial F}{\partial x_{k\alpha}} z_{k\alpha} = \theta^{-1} h(\theta), \quad (5.21)$$

The growth of the potential function.

for all sufficiently small $\theta > 0$ (that is, all $\theta > 0$ with $q + \theta z \in \Delta$). We have already encountered this expression in Proposition 2.13: if $f'(\theta) > 0$ for every $z \in T_q^c \Delta$ and small enough $\theta > 0$, then the evolutionary index $L_q(x)$ will be positive in a neighborhood of q , and, hence, q will be an ESS. If, however, this condition does not hold, (i.e. there exist rays $z \in T_q^c \Delta$ with $f'(\theta) = 0$ arbitrarily close to q) then q cannot be evolutionarily stable.⁵

This expression also allows us to obtain a series of growth estimates for the evolutionary index L_q based on estimates for the potential F . To obtain the first estimate of this kind, we will make the relatively mild assumption that F

⁵ This also shows that even *strict* minima of the potential function may fail to be evolutionarily stable – for a cute example, look at the global minimum of $F(x) = x^2(1 + \sin^2(1/x))$.

grows *superlinearly* along every ray that emanates from some point $q \in \Delta$, i.e. that there exists a point $q \in \Delta$ such that the function:

$$\frac{F(q + \theta z) - F(q)}{\theta}, \theta > 0, \text{ is increasing for every } z \in T_q^c \Delta. \quad (5.22)$$

Clearly, if the potential function F satisfies (5.22), the point q will be the unique minimum of F , just as in the case of (strictly) convex functions.⁶ Furthermore, on account of (5.22), we immediately get $\theta f'(\theta) \geq f(\theta) - f(0)$, or, for $x = q + \theta z$:

$$L_q(x) = \theta f'(\theta) \geq f(\theta) - f(0) = F(x) - F(q) \geq 0. \quad (5.23)$$

Therefore, since q is the unique minimum of F , we immediately see that q attracts every interior solution trajectory of the replicator dynamics (5.5).

In a similar vein, we also get the following growth estimate for L_q :

Lemma 5.4. *Let \mathfrak{G} be a potential game whose potential function $F : \Delta \rightarrow \mathbb{R}$ satisfies the growth condition (5.22). Then, if $q = \sum_k m_k e_{k,0}$ is a strict equilibrium of \mathfrak{G} and $\Delta u_k = \min_{\mu \neq 0} \{u_{k,0}(q) - u_{k\mu}(q)\}$, we will have:*

A linear estimate for the evolutionary index.

$$L_q(q + \theta z) \geq \frac{1}{2} \theta \sum_k \|z_k\|_1 \Delta u_k \quad \text{for all } z \in T_q^c \Delta \text{ and } \theta \geq 0 \text{ s.t. } q + \theta z \in \Delta. \quad (5.24)$$

Proof. Clearly, since $z \in T_q^c \Delta$ and $q = \sum_k m_k e_{k,0}$, the individual components $z_k \in T_{q_k}^c \Delta_k$ of z will be of the form:

$$z_k = \sum_{\mu}^k z_{k\mu} (e_{k\mu} - e_{k,0}), \text{ where } z_{k\mu} \geq 0 \text{ for all } \mu \in \mathcal{A}_k^* \equiv \mathcal{A}_k \setminus \{0\}. \quad (5.25)$$

So, if $f(\theta) = F(q + \theta z)$, the growth condition (5.22) easily implies:

$$\frac{f(\theta) - f(0)}{\theta} \geq \lim_{\theta \rightarrow 0} \frac{f(\theta) - f(0)}{\theta} = f'(0), \quad (5.26)$$

or, equivalently, $f(\theta) \geq f(0) + \theta f'(0)$ for all $\theta \geq 0$ such that $q + \theta z \in \Delta$. However, by (5.21), we will also have $f'(0) = -\sum_{k,\alpha} z_{k\alpha} u_{k\alpha}(q)$, and, thanks to (5.25), we may rewrite this last sum as:

$$\begin{aligned} f'(0) &= -\sum_{k,\alpha} z_{k\alpha} u_{k\alpha}(q) = \sum_k \sum_{\mu}^k z_{k\mu} [u_{k,0}(q) - u_{k\mu}(q)] \\ &\geq \sum_k \sum_{\mu}^k z_{k\mu} \Delta u_k = \frac{1}{2} \sum_k \Delta u_k \|z_k\|_1, \end{aligned} \quad (5.27)$$

the last equality following from the definition of the L^1 norm of z_k : $\|z_k\|_1 = \sum_{\alpha}^k |z_{k\alpha}| = |z_{k,0}| + \sum_{\mu}^k |z_{k\mu}| = 2 \sum_{\mu}^k z_{k\mu}$ (recall that $\sum_{\alpha}^k z_{k\alpha} = 0$ and, hence, $z_{k,0} = -\sum_{\mu}^k z_{k\mu}$). In this way, (5.23) finally gives:

$$L_q(q + \theta z) \geq f(\theta) - f(0) \geq \frac{1}{2} \theta \sum_k \Delta u_k \|z_k\|_1. \quad \square$$

This linear growth estimate for the evolutionary index L_q that we just established will play a key role in our stochastic considerations in the next section. For now however, Lemma 5.4 provides us with a (global!) estimate of the rate at which the solution trajectories of the replicator dynamics converge to a strict equilibrium:

⁶ Any strictly convex potential satisfies (5.22), but the converse is not true; in fact, (5.22) does not even guarantee (weak) convexity along rays.

Exponential stability
of strict equilibria.

Theorem 5.5. *Let \mathfrak{G} be a potential game satisfying the growth condition (5.22), and let q be a strict equilibrium of \mathfrak{G} . Any solution orbit of the replicator dynamics (5.5) which starts at a finite Kullback-Leibler distance from q satisfies:*

$$H_q(x(t)) \leq h_0 e^{-ct}, \quad (5.28)$$

where $h_0 = H_q(x(0))$ and $c = h_0^{-1} \min_k \left\{ m_k (1 - e^{-\lambda_k h_0 / m_k}) \Delta u_k \right\}$.

Proof. To begin with, assume without loss of generality that $q = \sum_k m_k e_{k,0}$ and note that any $x_k \in \Delta_k \setminus \{q_k\}$ can be uniquely expressed in the form:

$$x_k = q_k + \theta_k z_k \text{ for some } \theta_k \geq 0 \text{ and } z_k \in T_{q_k}^c \Delta \text{ with } z_{k,0} = -m_k. \quad (5.29)$$

So, with $x_{k,0} = m_k(1 - \theta_k)$, we readily obtain:

$$H_q(x) = \sum_k \frac{m_k}{\lambda_k} \log \frac{m_k}{x_{k,0}} = - \sum_k \frac{m_k}{\lambda_k} \log(1 - \theta_k), \quad (5.30)$$

where $\lambda_k, k = 1, \dots, N$, are the learning rates of (5.5).

Now, let θ_k^* be defined by the equation $h_0 = \lambda_k^{-1} H_{q_k}(q_k + \theta_k z_k)$ (that is, $\theta_k^* = 1 - e^{-\lambda_k h_0 / m_k}$), implying that $-\frac{m_k}{\lambda_k} \log(1 - \theta_k) \leq h_0 \theta_k / \theta_k^*$ if and only if $0 \leq \theta_k \leq \theta_k^*$ (recall that h is convex). We then claim that:

$$H_q(x(t)) = - \sum_k \frac{m_k}{\lambda_k} \log(1 - \theta_k(t)) \leq \sum_k h_0 \theta_k(t) / \theta_k^* \quad \text{for all } t \geq 0, \quad (5.31)$$

Estimating the decay
of the relative entropy.

where $\theta_k(t)$ is defined by (5.29). To see that this is so, it clearly suffices to show that $\theta_k(t) \leq \theta_k^*$ for all $t \geq 0$. However, if $\theta_k(t) > \theta_k^*$ for some $t \geq 0$, then we would also have $H_{q_k}(x_k(t)) > h_0$, and, hence, $H_q(x(t)) > H_q(x(0))$ as well, a clear contradiction because $H_q(x(t))$ is decreasing – recall that $\dot{H}_q(x) = -L_q(x) \leq 0$ on account of (5.23).

On the other hand, an application of Lemma 5.4 immediately yields:

$$\frac{d}{dt} H_q(x(t)) = -L_q(x(t)) \leq - \sum_k m_k \theta_k(t) \Delta u_k. \quad (5.32)$$

As a result, if we set $c = h_0^{-1} \min_k \left\{ m_k (1 - e^{-\lambda_k h_0 / m_k}) \Delta u_k \right\}$ (so that $m_k \Delta u_k \geq c h_0 / \theta_k^*$ for all $k = 1, \dots, N$), this last inequality becomes:

$$\frac{d}{dt} H_q(x(t)) \leq -c \sum_k h_0 \theta_k(t) / \theta_k^* \leq -c H_q(x(t)) \quad \text{for all } t \geq 0. \quad (5.33)$$

and the estimate (5.28) follows trivially from Gronwall's lemma. \square

Remark. It is interesting to note here that the exponential rate of convergence c of (5.28) is essentially $\mathcal{O}(\lambda)$ – a consequence of the rate-adjusted relative entropy $H_q(x) = \sum_k \lambda_k^{-1} \sum_\alpha q_{k\alpha} \log(q_{k\alpha} / x_{k\alpha})$ being $\mathcal{O}(1/\lambda)$. This sits very well with our interpretation of the learning parameters λ_k : the players' convergence rate is proportional to their learning rate.

The projective
distance function.

To extend this result to interior evolutionarily stable strategies, we will first need to study the decomposition (5.29) in a little more detail. To that end, let $q \in \text{Int}(\Delta)$ be an interior point of Δ and set $S_q = \{z \in T_q \Delta : q + z \in \text{bd}(\Delta)\}$. Then, as before, any $x \in \Delta \setminus \{q\}$ may be uniquely expressed in the form $x = q + \theta z$ for some $z \in S_q$ and $\theta \in [0, 1]$. We thus define the *projective distance* $\Theta_q(x)$ of x from q to be:

$$\Theta_q(x) = \theta \Leftrightarrow x = q + \theta z \text{ for some } z \in S_q \text{ and } 0 \leq \theta \leq 1. \quad (5.34)$$

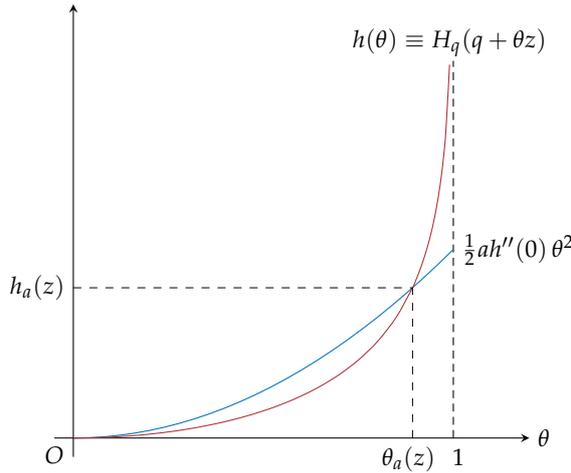


Figure 5.2: Illustrating the quadratic estimate of Theorem 5.6 and Lemma 5.7.

Much like the relative entropy H_q , Θ_q is not a bona fide distance function by itself, so our choice of terminology might appear somewhat unfortunate. It is justified however by the fact that Θ_q closely resembles the L^1 norm: the “projective balls” $B_\theta = \{x \in \Delta : \Theta_q(x) \leq \theta\}$ are rescaled copies of Δ (S_q is the “unit sphere” in this picture), and the graph $\text{gr}(\Theta_q) \equiv \{(x, \theta) \in \Delta \times \mathbb{R} : \Theta_q(x) = \theta\}$ of Θ_q is simply a cone over the polytope Δ .⁷

The second thing of note in the case of interior equilibria is that any first order growth estimate for the evolutionary index $L_q(x)$ (such as (5.24)) will be useless for our purposes. Indeed, if $q \in \text{Int}(\Delta)$ is evolutionarily stable, it will also be an interior local minimum of F , and we will thus have:

$$\left. \frac{d}{d\theta} \right|_{\theta=0} F(q + \theta z) = 0 \quad \text{for all } z \in T_q \Delta. \tag{5.35}$$

This means that we will need to strengthen our superlinearity condition (5.22) to provide a meaningful second-order bound for the growth of $f(\theta)$. The easiest (and most economical) way to do this is to (finally) assume that the potential F is (strictly) convex:

Theorem 5.6. *Let \mathfrak{G} be a strictly convex potential game whose potential function $F : \Delta \rightarrow \mathbb{R}$ attains its (global) minimum at $q \in \text{Int}(\Delta)$. Then, q is evolutionarily stable and attracts every interior solution $x(t)$ of the replicator dynamics (5.5). Furthermore, we will also have:*

Exponential stability of interior equilibria.

$$H_q(x(t)) \leq H_q(x(0)) \exp\left(-\frac{1}{2} \frac{r \zeta_0^2}{h_c} t\right), \tag{5.36}$$

where r is the minimum eigenvalue of the Hessian of F (taken over all $x \in \Delta$), $h_c \geq H_q(x(0))$ is a positive constant which is equal to $H_q(x(0))$ if the latter is small enough, and $\zeta_0^2 = \min_{x \in \Delta} \{\|x - q\|^2 : H_q(x) = h_c\}$.

The proof of this theorem is not particularly difficult, but it is rife with technical details. These are simplified considerably by the following lemma which provides a quadratic upper bound for the growth of the relative entropy $H_q(x)$ near an interior point $q \in \text{Int}(\Delta)$:

⁷ The definition of the projective distance Θ_q may actually be extended to points q in the boundary of Δ by a suitable redefinition of the “unit sphere” S_q so that the decomposition (5.34) remain unique. If q is a vertex of Δ , it can be easily seen that $\Theta_q(x)$ is just a scaled variant of $\|x - q\|_1$.

A quadratic growth estimate.

Lemma 5.7. *Let $q \in \text{Int}(\Delta)$ and set $h(\theta) = H_q(q + \theta z)$, $0 \leq \theta < 1$, $z \in S_q$. Then, for every $a > 1$, the equation:*

$$H_q(q + \theta z) = \frac{a}{2} \sum_k \frac{1}{\lambda_k} \sum_\alpha \frac{z_{k\alpha}^2}{q_{k\alpha}} \theta^2 \quad (5.37)$$

has a unique positive root $\theta_a \equiv \theta_a(z)$. Accordingly, if $h_a \equiv H_q(q + \theta_a z)$, then:

$$H_q(q + \theta z) \leq h_a \theta^2 / \theta_a^2 \quad \text{for all } \theta \leq \theta_a. \quad (5.38)$$

Proof. The proof is essentially a study in the properties of the function $h(\theta) = H_q(q + \theta z)$, $0 \leq \theta < 1$, $z \in S_q$. Indeed, an easy differentiation yields:

$$h^{(n)}(\theta) = (-1)^n (n-1)! \sum_k \frac{1}{\lambda_k} \sum_\alpha \frac{q_{k\alpha} z_{k\alpha}^n}{(q_{k\alpha} + \theta z_{k\alpha})^n}, \quad (5.39)$$

showing that $h^{(n)}(\theta)$ is strictly positive for all even $n > 0$. So, if we consider the quadratic estimate $g(\theta) = \frac{a}{2} \sum_k \lambda_k^{-1} \sum_\alpha \frac{z_{k\alpha}^2}{q_{k\alpha}} = \frac{a}{2} h''(0) \theta^2$, we are left to show that the equation $h(\theta) = g(\theta)$ has a unique positive root.

The actual existence of a solution is obvious: to begin with, note that $h'(0) = -\sum_k \lambda_k^{-1} \sum_\alpha z_{k\alpha} = 0$ on account of z being tangent to Δ at q (i.e. $\sum_\alpha z_{k\alpha} = 0$ for all $k = 1, \dots, N$). So, if we introduce the difference $w(\theta) = h(\theta) - g(\theta)$, we will have $w(0) = w'(0) = 0$, whereas $w''(0) = (1-a)h''(0) < 0$. This means that w will initially decrease to negative values, and since $\lim_{\theta \rightarrow 1} w(\theta) = \infty$, the existence of a positive root is guaranteed.

The local growth of the relative entropy.

On the other hand, to establish uniqueness, it is necessary to study higher order derivatives of h . In particular, we have $h^{(4)}(\theta) > 0$, so the third derivative $h'''(\theta)$ will be strictly increasing for all θ . In its turn, this means that $h''(\theta)$ is strictly convex, so the equation $h''(\theta) = ah''(0)$ – or, equivalently, the equation $w''(\theta) = 0$ – will have a unique positive solution (recall that $a > 1$).

Now, assume in absurdum that w has two distinct positive roots, say θ_1, θ_2 with $0 < \theta_1 < \theta_2$. Given that 0 is also a root of w , the mean value theorem implies that there exist ξ_1, ξ_2 with $0 < \xi_1 < \theta_1 < \xi_2 < \theta_2$ and $w'(\xi_1) = w'(\xi_2) = 0$. However, we also have $w'(0) = 0$ as well, so a second application of the mean value theorem provides ξ'_1, ξ'_2 with $0 < \xi'_1 < \xi_1 < \xi'_2 < \xi_2$ and $w''(\xi'_1) = w''(\xi'_2) = 0$. This contradicts our previous result that w'' has a unique positive root and the last part of the lemma follows by rewriting the inequality $w(\theta) \leq 0$, $\theta \leq \theta_a$, in terms of h and g . \square

Bounding the decay of the relative entropy.

Proof of Theorem 5.6. Since F is convex, it will clearly satisfy the growth condition (5.22), so the first part of the theorem follows by the estimate (5.23) for $L_q(x)$ (which shows that L_q is Lyapunov for the replicator dynamics).

Now, to obtain the exponential bound (5.36), write $x \in \Delta$ in the projective form $x = q + \theta z$ where $z \in S_q$ and $\theta = \Theta_q(x)$. Then, if we set $f(\theta) = F(q + \theta z)$ as before, we will have:

$$f''(\theta) = \sum_{k,\alpha} \sum_{\ell,\beta} z_{k\alpha} z_{\ell\beta} \frac{\partial^2 F}{\partial x_{k\alpha} \partial x_{\ell\beta}}. \quad (5.40)$$

However, on account of the (strict) convexity of F , the restriction of the Hessian matrix $\frac{\partial^2 F}{\partial x_{k\alpha} \partial x_{\ell\beta}}$ of F on $T\Delta$ will be positive definite for all $x \in \Delta$. Therefore, its smallest eigenvalue will always be positive, and with Δ compact, it will be

bounded away from 0 by some positive constant $r > 0$.⁸ In this way, a first order Taylor expansion with Lagrange remainder yields $f(\theta) \geq f(0) + \frac{1}{2}r\theta^2\|z\|^2$, or, on account of the inequality (5.23):

$$L_q(x) \geq F(x) - F(q) \geq \frac{1}{2}r\|x - q\|^2, \quad (5.41)$$

where $\|\cdot\|$ denotes the ordinary Euclidean norm.

Mimicking the proof of Theorem 5.5, we are left to obtain a similar (quadratic) inequality for $H_q(x)$. This was the whole purpose behind Lemma 5.7, so fix some $a > 1$ and pick a tangent vector $z \in S_q$. Then, if $\theta_a \equiv \theta_a(z)$ and $h_a \equiv H_q(q + \theta_a(z)z)$ are defined as in the proof of Lemma 5.7, we will have $H_q(q + \theta z) \leq h_a\theta^2/\theta_a^2$ for all $\theta \leq \theta_a$.

On the other hand, if we write $x(t)$ in the projective form $x(t) = q + \theta(t)z(t)$ where $z(t) \in S_q$ and $\theta(t) \equiv \Theta_q(x(t))$ is the projective distance of $x(t)$ from q , the quadratic estimate (5.38) that we just alluded to will only hold if $\theta(t) \leq \theta_a(z(t))$ for all $t \geq 0$. Unfortunately, since $x(0)$ might be arbitrarily away (in terms of K-L distance) from q this need not be the case, so we will need to tighten our quadratic bound even further.

To that end, let $h_c = H_q(x(0)) \vee \max\{h_a(z) : z \in S_q\}$. By following the same strategy as in Lemma 5.7, it is easy to see that the equation $H_q(q + \theta z) = h_c$ has a unique positive solution $\theta_0 \equiv \theta_0(z)$ for all $z \in S_q$. Then, since $h_c \geq h_a(z)$, we will also have $\theta_0(z) \geq \theta_a(z)$, and, hence, we obtain:

$$H_q(q + \theta z) \leq h_c\theta^2/\theta_0^2(z) \quad \text{for all } \theta \leq \theta_0. \quad (5.42)$$

Moreover, we claim that $\theta(t) \leq \theta_0(z(t))$ for all $t \geq 0$: indeed, if $\theta(t)$ ever exceeded $\theta_0(z(t))$, we would also have $H_q(x(t)) > h_c \geq H_q(x(0))$, a contradiction because $H_q(x(t))$ is decreasing.

As a result of all the above, we will have for all $t \geq 0$:

$$H_q(x(t)) \leq h_c \frac{\theta^2(t)}{\theta_0^2(z(t))} = h_c \frac{\|\theta(t)z(t)\|^2}{\|\theta_0(z(t))z(t)\|^2} \leq \frac{h_c}{\zeta_0^2} \|x(t) - q\|^2, \quad (5.43)$$

where $\zeta_0^2 = \min\{\|x - q\|^2 : x \in \Delta, H_q(x) = h_c\}$. Then, in conjunction with the estimate (5.41), we obtain:

$$\frac{d}{dt}H_q(x(t)) \leq -L_q(x(t)) \leq -\frac{1}{2}r\|x(t) - q\|^2 \leq -\frac{1}{2}\frac{r\zeta_0^2}{h_c}H_q(x(t)) \quad (5.44)$$

and the exponential bound (5.36) follows again from Gronwall's lemma. \square

Remark 1. As written, the exponent of the exponential bound (5.36) implicitly depends on the "internal" constant $a > 1$ that we introduced in its proof. Obviously, h_a is an increasing function of a , so, to sharpen our estimate, one simply has to let $a \rightarrow 1^+$.

In a similar vein, there are many points where we could tighten our bounds further, but, as in Theorem 5.5, these refinements would not yield significant

⁸ This actually shows that F is *strongly* convex over Δ but our lack of argumentation might not sound very convincing – after all, the smallest root of a real polynomial *does not* depend continuously on the polynomial's coefficients. Nevertheless, if we identify a complex n -th degree polynomial with a point in \mathbb{C}^n and let $\mathbb{C}_{sym}^n \equiv \mathbb{C}^n/S_n$ be the space of its roots modulo the symmetric group S_n (i.e. we ignore ordering), then the map $\rho : \mathbb{C}^n \rightarrow \mathbb{C}_{sym}^n$ will be continuous (a consequence of Rouché's theorem). Thus, if $H \subseteq \mathbb{C}^n$ is the set of all characteristic polynomials of positive-definite matrices, the restriction of ρ on H will also be continuous and its image will be contained in (the positive orthant of) \mathbb{R}^n/S_n . Therefore, given that the minimum of an n -tuple of real numbers is a continuous function on \mathbb{R}^n which descends to the quotient \mathbb{R}^n/S_n , our claim follows.

gains. What is more important for our purposes is to note that h_c is still of order $\mathcal{O}(1/\lambda)$, which means that the equilibration time of the replicator dynamics is itself of order $\mathcal{O}(\lambda)$, just as in the strict equilibrium case.

Remark 2. The requirement that the potential function be strictly convex might appear somewhat restrictive but, in fact, it can be weakened considerably. Indeed, the proofs of Theorems 5.5 and 5.6 both carry through if the potential function F is only *locally* (strictly) convex near q (that is, as long as q is a “deep” minimum of the potential) and the starting point of the dynamics lies within the domain of convexity of F .

5.2 THE EFFECT OF STOCHASTIC FLUCTUATIONS

Without further ado, let us now see what happens in the presence of stochastic fluctuations that interfere with the game. As in Chapter 4, we will assume that the players’ scores $U_{k\alpha}$ evolve according to the SDE:

$$dU_{k\alpha}(t) = u_{k\alpha}(X(t)) dt + \sigma_{k\alpha}(X(t)) dW_{k\alpha}(t), \quad (5.45)$$

where the population state $X(t)$ is given by the Gibbs distribution:

$$X_{k\alpha}(t) = m_k \frac{\exp(\lambda_k U_{k\alpha}(t))}{\sum_{\beta} \exp(\lambda_k U_{k\beta}(t))} \quad (5.46)$$

with learning rates (population-specific inverse temperatures) $\lambda_k > 0$ – as before, the noise coefficients $\sigma_{k\alpha} : \Delta \rightarrow \mathbb{R}$ are assumed essentially bounded over Δ and $W(t)$ is a Wiener process in $\mathbb{R}^{\mathcal{A}} \cong \prod_k \mathbb{R}^{\mathcal{A}^k}$. Therefore, by applying Itô’s formula to (5.46), we obtain the stochastic replicator equation:

The stochastic replicator dynamics.

$$\begin{aligned} dX_{k\alpha} &= \lambda_k X_{k\alpha} (u_{k\alpha}(X) - u_k(X)) dt \\ &+ \frac{\lambda_k^2 X_{k\alpha}}{2 m_k} \left(\sigma_{k\alpha}^2(X) (m_k - 2X_{k\alpha}) - \frac{1}{m_k} \sum_{\beta} \sigma_{k\beta}^2(X) X_{k\beta} (m_k - 2X_{k\beta}) \right) dt \\ &+ \lambda_k X_{k\alpha} \left(\sigma_{k\alpha}(X) dW_{k\alpha} - \frac{1}{m_k} \sum_{\beta} \sigma_{k\beta}(X) X_{k\beta} dW_{k\beta} \right) \end{aligned} \quad (5.47)$$

(see Section 4.2 for the essentials of the derivation).

The difference between this equation and its Nash sibling (4.16) is that *a*) the masses of the various populations introduce the scaling factors m_k in (5.47) (obviously, (4.16) is recovered in the uniform case $m_k = 1$); and *b*) the payoffs $u_{k\alpha}(x)$ are not multilinear over Δ (even though $u_k(x)$ is still given by the convex combination $u_k(x) = m_k^{-1} \sum_{\alpha} x_{k\alpha} u_{k\alpha}(x)$). As a result, we will proceed in much the same fashion as we did in Section 4.2 and, for the sake of simplicity, we will drop the dependence of the noise coefficients $\sigma_{k\alpha}$ on the population state $x \in \Delta$, replacing them instead with their essential supremum: $\sigma_{k\alpha} \leftrightarrow \text{ess sup}_{x \in \Delta} |\sigma_{k\alpha}(x)|$. We also note that (5.47) admits a unique strong solution for every initial condition $X(0) \in \Delta$ by the same reasoning as in Section 4.2, so the two equations (4.16) and (5.47) appear indistinguishable.

These similarities are strengthened by the following result (which is proved exactly as its counterparts in Chapter 4):

Dominated strategies become extinct.

Theorem 5.8. *Let \mathfrak{G} be a population game and let $X(t)$ be an interior solution path of the stochastic replicator dynamics (5.47) with initial condition $X(0) = x \in \text{Int}(\Delta)$. Then, if $\alpha \in A_k$ is strictly dominated, we will have:*

$$\lim_{t \rightarrow \infty} X_{k\alpha}(t) = 0 \quad \text{almost surely.} \quad (5.48)$$

In fact, $X_{k\alpha}(t)$ also converges to 0 in probability as well:

$$\mathbf{P}_x \left(X_{k\alpha}(t) < e^{-M} \right) \geq \frac{1}{2} \operatorname{erfc} \left(\frac{M - h_k(x_k) - \lambda_k v_k t}{2\lambda_k \sigma_k \sqrt{A_k t}} \right), \quad \text{for all } M > 0, \quad (5.49)$$

where $h_k(x_k)$ and $v_k > 0$ are constants, $A_k = |A_k|$ is the number of actions available to population A_k , and $\sigma_k = \max_{\beta} \{|\sigma_{k\beta}|\}$.

In fact, (5.48) still holds if $\alpha \in A_k$ is only iteratively dominated: in the long run, the only strategies that survive are the rationally admissible ones.

Proof. As we already mentioned, the proof of the theorem essentially follows that of Propositions 4.1 and 4.2, so we will only give a brief sketch here.

Since α is pure, we will have $d_{\text{KL}}(e_{k\alpha}, x) = \frac{m_k}{\lambda_k} \log(m_k/x_{k\alpha})$ – recall that we are employing the rate-adjusted version of the relative entropy which is appropriate to the rate-adjusted dynamics (5.47). So, for any $\beta \in A_k$, the same series of (Itô) differentiations that led to (4.31) now give:

$$d \left(\log(X_{k\alpha}/X_{k\beta}) \right) = \left(u_{k\beta}(X) - u_{k\alpha}(X) \right) dt + \sigma_{k\beta} dW_{k\beta} - \sigma_{k\alpha} dW_{k\alpha} \quad (5.50)$$

By the definition of dominated strategies, if $\alpha \prec \beta$, we will have $u_{k\alpha}(x) < u_{k\beta}(x)$ for all $x \in \Delta$, and, given that Δ is compact, this also gives $v_k = \inf\{u_{k\beta}(x) - u_{k\alpha}(x) : x \in \Delta\} > 0$. Therefore, since the proof of Propositions 4.1/4.2 and Theorem 4.3 only depend on this bound being positive, our assertion follows. \square

Exactly the same technique also yields:

Theorem 5.9. *The strict equilibria of a population game \mathfrak{G} are stochastically asymptotically stable in the replicator dynamics (5.47) of exponential learning.*

Stochastic stability of strict equilibria.

Proof. By defining the “adjusted scores” $Z_{k\alpha}$ as in (4.59) and scaling the “adjusted distributions” $Y_{k\alpha}$ (4.60) by m_k , we see that the proof of Theorem 4.8 essentially hinges on the condition $u_{k,0}(q) > u_{k,\mu}(q)$, where $q = \sum_k m_k e_{k,0}$ is the strict equilibrium in question and $\mu \in A_k^*$. Since (by Wardrop’s principle) this condition characterizes strict equilibria in population games as well, the theorem follows. \square

So, what new insights are there to be gained by looking at population games? In the general case, we have nothing to offer, much as we could not hope to go further with ordinary Nash games with no kind of “global” structure present (such as a congestion mechanism). However, the landscape changes dramatically if we focus our efforts on potential games; in that case, just as in the deterministic setting of the previous section, we are able to estimate (average) rates of convergence and even some global convergence results.

The role of the potential function.

Our first result is an estimate of the time it takes a replicator path to get near a strict equilibrium. To put this in precise terms, we will measure distances in Δ with the L^1 -norm: $\|z\|_1 = \sum_{k,\alpha} |z_{k,\alpha}|$. In this norm, it is not too hard to see that Δ has a diameter of $2 \sum_k m_k$, so pick some positive $\delta < 2 \sum_k m_k$ and let $K_\delta = \{x \in \Delta : \|x - q\|_1 \leq \delta\}$ be the corresponding compact δ -neighborhood of q . Then, to see if $X(t)$ ever hits K_δ , we will examine the hitting time τ_δ :

$$\tau_\delta \equiv \tau_{K_\delta} = \inf\{t > 0 : X(t) \in K_\delta\} = \inf\{t > 0 : \|X(t) - q\|_1 \leq \delta\}. \quad (5.51)$$

Thereby, our chief concern will be this: is the hitting time τ_δ finite with high probability? And if it is, is its expected value also finite? If the players are

“patient enough” (i.e. their learning temperature is not too low), it turns out that this indeed the case:

Strict equilibrium
hitting times.

Theorem 5.10. *Let $q = \sum_k m_k e_{k,0}$ be a strict Wardrop equilibrium of a potential game \mathfrak{G} whose potential function satisfies the growth condition (5.22); assume further that the players’ learning rates λ_k satisfy the equation:*

$$\lambda_k \sigma_{k,max}^2 < \Delta u_k \quad \text{for all } k = 1, \dots, N, \quad (5.52)$$

where $\sigma_{k,max}^2 = \max_{\alpha \in \mathcal{A}_k} \{\sigma_{k\alpha}^2\}$ and $\Delta u_k = \min_{\mu \neq 0} \{u_{k,0}(q) - u_{k\mu}(q)\}$. Then, for all $\delta > 0$ and for every initial condition $X(0) = x \in \Delta$ with finite relative entropy $H_q(x) < \infty$, the hitting time τ_δ will be finite (on average and almost surely):

$$\mathbf{E}_x[\tau_\delta] \leq \frac{2H_q(x)}{a\delta}, \quad \text{where } a = \min_k \{\Delta u_k - \lambda_k \sigma_{k,max}^2\}. \quad (5.53)$$

Before proving this theorem, we will first need to pave the way with a preparatory calculation:

Stochastic evolution
of the relative entropy.

Lemma 5.11. *Let \mathcal{L} be the infinitesimal generator of the replicator dynamics (5.47). Then, for any $q \in \Delta$, we will have:*

$$\begin{aligned} \mathcal{L}H_q(x) &= -L_q(x) - \frac{1}{2} \sum_k \frac{\lambda_k}{m_k} \sum_\beta \sigma_{k\beta}^2 (x_{k\beta} - q_{k\beta})^2 \\ &\quad + \frac{1}{2} \sum_k \frac{\lambda_k}{m_k} \sum_\beta \sigma_{k\beta}^2 (m_k - 2q_{k\beta})(x_{k\beta} - q_{k\beta}) \\ &\quad + \frac{1}{2} \sum_k \frac{\lambda_k}{m_k} \sum_\beta \sigma_{k\beta}^2 q_{k\beta} (m_k - q_{k\beta}), \end{aligned} \quad (5.54)$$

where $L_q(x) = -\sum_{k,\alpha} (x_{k\alpha} - q_{k\alpha}) u_{k\alpha}(x)$ is the evolutionary index of (5.12).

Proof. The proof of this lemma is a straightforward exercise in stochastic calculus. Indeed, exactly as in the proof of Proposition 4.1, Itô’s formula gives:

$$\begin{aligned} dH_q(X(t)) &= -L_q(X(t))dt + \frac{1}{2} \sum_k \frac{\lambda_k}{m_k} \sum_\beta \sigma_{k\beta}^2 X_{k\beta}(t)(m_k - X_{k\beta}(t))dt \\ &\quad - \sum_{k,\beta} q_{k\beta} \left(\sigma_{k\beta} dW_{k\beta} - \frac{1}{m_k} \sum_\gamma \sigma_{k\gamma} X_{k\gamma}(t) dW_{k\gamma} \right), \end{aligned} \quad (5.55)$$

the term in the brackets corresponding to $\mathcal{L}H_q$. Then, by setting $z = x - q$, we readily obtain:

$$\mathcal{L}H_q(x) = -L_q(x) + \frac{1}{2} \sum_k \frac{\lambda_k}{m_k} \sum_\beta \sigma_{k\beta}^2 (q_{k\beta} + z_{k\beta})(m_k - q_{k\beta} - z_{k\beta}), \quad (5.56)$$

and (5.54) follows by simply carrying out the algebra. \square

Proof of Theorem 5.10. As in the case of Lemma 5.4, any $x_k \in \Delta_k$ may be written as $x_k = q_k + \theta_k z_k$, where $0 \leq \theta_k \leq 1$ and $z_k \in T_{q_k}^c \Delta_k$ is of the form:

$$z_{k,0} = -m_k \text{ and } z_{k\mu} = m_k \zeta_{k\mu}, \mu \in \mathcal{A}_k^*, \text{ with } \zeta_{k\mu} \geq 0 \text{ and } \sum_\mu \zeta_{k\mu} = 1, \quad (5.57)$$

that is, $z_k = -m_k e_{k,0} + m_k \zeta_k$ for some $\zeta_k \in \Delta(\mathcal{A}_k^*)$. Thus, by substituting in (5.54), we get the expression:

$$\begin{aligned} \mathcal{L}H_q(x) &= -L_q(x) - \frac{1}{2} \sum_k \lambda_k m_k \theta_k^2 \left(\sigma_{k,0}^2 + \sum_\mu \sigma_{k\mu}^2 \zeta_{k\mu}^2 \right) \\ &\quad + \frac{1}{2} \sum_k \lambda_k m_k \theta_k \left(\sigma_{k,0}^2 + \sum_\mu \sigma_{k\mu}^2 \zeta_{k\mu} \right), \end{aligned} \quad (5.58)$$

which, if we use (5.23) and drop the quadratic term, becomes:

$$\mathcal{L}H_q(x) \leq -\sum_k m_k \Delta u_k \theta_k + \sum_k \lambda_k m_k \sigma_{k,max}^2 \theta_k, \quad (5.59)$$

since, clearly, $\sigma_{k,max}^2 \equiv \max_{\alpha \in \mathcal{A}_k} \{\sigma_{k\alpha}^2\} \geq \frac{1}{2} \left(\sigma_{k,0}^2 + \sum_{\mu} \sigma_{k\mu}^2 \bar{\xi}_{k\mu} \right)$ for every direction vector $\bar{\xi} \in \Delta(\mathcal{A}_k^*)$.

Assume now that $\Delta u_k > \lambda_k \sigma_{k,max}^2$ for all $k = 1, \dots, N$, and let $K_\delta = \{x \in \Delta : \|x - q\|_1 \leq \delta\}$ be a compact L^1 neighborhood of q in Δ . We will then have $x \in K_\delta$ if and only if $2\sum_k m_k \theta_k \leq \delta$ and, hence, for all $x \notin K_\delta$:

$$\mathcal{L}H_q(x) \leq -\sum_k m_k \left(\Delta u_k - \lambda_k \sigma_{k,max}^2 \right) \theta_k \leq -a \sum_k m_k \theta_k \leq -\frac{1}{2} a \delta, \quad (5.60)$$

where $a = \min_k \{\Delta u_k - \lambda_k \sigma_{k,max}^2\}$. Therefore, by a simple (but very useful!) estimate of Durrett (1996, Theorem 5.3 in page 268), we finally get:

$$\mathbf{E}_x[\tau_\delta] \leq \frac{2H_q(x)}{a\delta} \quad \square$$

In essence, Theorem 5.10 tells us that if the players' learning rates are slow enough, then every solution trajectory $X(t)$ will come arbitrarily close to the game's (necessarily unique) strict equilibrium in finite time. Thus, since Theorem 5.9 ensures that q attracts all nearby trajectories with arbitrarily high probability, applying these two theorems in tandem yields:

Corollary 5.12. *Let \mathfrak{G} be a potential game \mathfrak{G} whose potential function satisfies the growth condition (5.22), and assume that the players' learning rates satisfy (5.52). Then, every solution path of the replicator dynamics (5.47) which starts at finite K-L distance from a strict equilibrium q will converge to it almost surely.*

Convergence to strict equilibria.

Remark 1 (Tightness). We should state here that the estimate (5.53) for the hitting time τ_δ is not very tight: the quadratic term of (5.54) has the same sign as the evolutionary index $L_q(x)$, and we could also take a direction dependent bound in (5.58) instead of the "uniform" bound $\sigma_{k,max}^2$ – of course, if the distribution of the noise coefficients across \mathcal{A}_k is uniform, this last point is moot. This would lead to a sharper bound for $\mathbf{E}_x(\tau_\delta)$, but the extra complication is not really worth the effort.⁹

Remark 2 (Temperance and Temperature). The "slow-learning" condition (5.52) shows that the replicator dynamics reward *patience*: players who take their time in learning the game manage to weed out the noise and eventually converge to equilibrium. This begs to be compared with the (inverse) temperature analogy for the learning rates: if the "learning temperatures" $T_k = 1/\lambda_k$ are too low, the players' learning schemes become very rigid and this intemperance amplifies any random variations in the experienced delays. On the other hand, when the temperature rises above the critical threshold $T_c = \sigma_{k,max}^2 / \Delta u_k$, the stochastic fluctuations are toned down and the deterministic drift draws users to equilibrium.

In any event, since strict equilibria do not always exist, we should return to the generic case of *interior* equilibria $q \in \text{Int}(\Delta)$. We have already seen that these equilibria are not very well-behaved in stochastic environments: they are not stationary and (5.54) shows that $\mathcal{L}H_q$ is actually *positive* in their vicinity. Despite all that, if players learn at sufficiently high temperatures, we have:

⁹ On the other hand, it would be pretty important if we could drop condition (5.52) altogether, but, unfortunately, it has resisted all our attempts to do so.

The invariant distribution.

Theorem 5.13. *Let $q \in \text{Int}(\Delta)$ be an interior equilibrium of a (strictly) convex potential game \mathfrak{G} , and assume that the players' learning rates satisfy the condition:*

$$D_k^2 > d_k^2(\lambda_k) \equiv \lambda_k m_k r^{-1} \sigma_{k,\max}^2 \left(1 - \frac{1}{A_k}\right), \quad (5.61)$$

where $D_k = \text{dist}(q_k, \text{bd}(\Delta_k))$, $A_k = |\mathcal{A}_k|$, $\sigma_{k,\max}^2 = \max_{\alpha \in \mathcal{A}_k} \{\sigma_{k\alpha}^2\}$, and $r > 0$ is the minimum eigenvalue of the Hessian of the game's potential over Δ .

Then, for any interior initial condition $X(0) = x \in \text{Int}(\Delta)$, the trajectories $X(t)$ are recurrent (a.s.) and their time averages are concentrated in a neighborhood of q . More precisely, if $\delta_k > d_k(\lambda_k)$ and $\tau_\delta = \inf\{t > 0 : \|X_k(t) - q_k\| \leq \delta_k \text{ for all } k\}$, we will have:

$$a) \quad \mathbf{E}_x[\tau_\delta] \leq \frac{2H_q(x)}{r(\delta^2 - d_\lambda^2)} \quad (5.62a)$$

$$b) \quad \mathbf{E}_x \left[\frac{1}{t} \int_0^t \|X(s) - q\|^2 ds \right] \leq d_\lambda^2 + \frac{2H_q(x)}{rt}, \quad (5.62b)$$

where $\delta^2 = \sum_k \delta_k^2$ and $d_\lambda^2 \equiv \sum_k d_k^2(\lambda_k)$.

Moreover, the transition probabilities of $X(t)$ converge in total variation to an invariant probability measure π on Δ which concentrates mass around q ; in particular, if $B_\delta = \{x \in \Delta : \|x - q\| \leq \delta\}$ is a δ -ball centered at q and contained in Δ , then:

$$\pi(B_\delta) \geq 1 - d_\lambda^2/\delta^2. \quad (5.63)$$

Proof. By Definition 3.12, recurrence here means that for every $y \in \text{Int}(\Delta)$ and every neighbourhood U_y of y , the diffusion $X(t)$ has the property:

$$\mathbf{P}_x(X(t_k) \in U_y) = 1, \quad (5.64)$$

The role of recurrence.

for some sequence of (random) times t_k that increases to infinity. Hence, using the recurrence criteria of [Bhattacharya \(1978\)](#), we will prove our claim by showing that $X(t)$ hits a compact neighbourhood of q in finite time, and that the generator of a suitably transformed stochastic process is elliptic.¹⁰

To prove the first part of this claim, recall that Lemma 5.11 gives:

$$\mathcal{L}H_q(x) = -L_q(x) + \frac{1}{2} \sum_k \frac{\lambda_k}{m_k} \sum_\beta \sigma_{k\beta}^2 x_{k\beta} (m_k - x_{k\beta}). \quad (5.65)$$

As in the proof of Theorem 5.6, the first term of (5.65) will be bounded from above by $-\frac{1}{2}r^2\|x - q\|^2$ where $r > 0$ is the minimum eigenvalue of the Hessian of the game's potential. Then, to bound the second term, a simple constrained optimization argument gives (recall that $\sum_\beta x_k = m_k$):

$$\sum_\beta \sigma_{k\beta}^2 x_{k\beta} (m_k - x_{k\beta}) \leq \sigma_{k,\max}^2 \sum_\beta x_{k\beta} (m_k - x_{k\beta}) \leq \sigma_{k,\max}^2 m_k^2 \left(1 - \frac{1}{A_k}\right), \quad (5.66)$$

where $\sigma_{k,\max}^2 = \max_{\beta \in \mathcal{A}_k} \{\sigma_{k\beta}^2\}$ and $A_k = |\mathcal{A}_k|$ is the number of choices at the disposal of population k . We thus obtain:

$$\begin{aligned} \mathcal{L}H_q(x) &\leq -\frac{1}{2}r\|x - q\|^2 + \frac{1}{2} \sum_k \lambda_k m_k \sigma_{k,\max}^2 \left(1 - \frac{1}{A_k}\right) \\ &= -\frac{1}{2}r \sum_k \left(\|x_k - q_k\|^2 - d_k^2(\lambda_k)\right), \end{aligned} \quad (5.67)$$

¹⁰ This approach was pioneered by [Imhof \(2005\)](#), so, to allow for easy comparison, we have tried to keep the formulation of our theorem as close as possible to Imhof's original results.

where $d_k^2(\lambda_k) \equiv \lambda_k m_k r^{-1} \sigma_{k,max}^2 \left(1 - \frac{1}{A_k}\right)$.

So, pick $\delta_k > 0$ with $d_k(\lambda_k) < \delta_k < D_k$, and let $K_\delta = \{x \in \Delta : \|x_k - q_k\| \leq \delta_k, k = 1, \dots, N\}$ be a product of balls that is wholly contained in Δ – recall that $d_k(\lambda_k) < \text{dist}(q_k, \text{bd}(\Delta_k)) \equiv D_k$ by (5.76). Then, for $x \notin K_\delta$, (5.67) becomes:

$$\mathcal{L}H_q(x) \leq -\frac{1}{2}r \sum_k \left(\delta_k^2 - d_k^2(\lambda_k)\right) = -\frac{1}{2}r \left(\delta^2 - d_\lambda^2\right) < 0, \quad (5.68)$$

and the bound (5.62a) follows from Theorem 5.3 in (Durrett, 1996, p. 268).

Furthermore, Dynkin's formula (see e.g. Øksendal, 2007, Theorem 7.4.1) applied to (5.67) gives:

$$\mathbf{E}_x [H_q(X(t))] = H_q(x) + \mathbf{E}_x \left[\int_0^t \mathcal{L}H_q(X(s)) ds \right] \quad (5.69)$$

$$\leq H_q(x) - \frac{1}{2}r \mathbf{E}_x \left[\int_0^t \|X(s) - q\|^2 ds \right] + \frac{1}{2}r d_\lambda^2 t, \quad (5.70)$$

and, with $\mathbf{E}_x[H_q(X(t))] \geq 0$, we easily get the estimate (5.62b):

$$\mathbf{E}_x \left[\frac{1}{t} \int_0^t \|X(s) - q\|^2 ds \right] \leq d_\lambda^2 + \frac{2H_q(x)}{rt}. \quad (5.62b)$$

Having established these crucial bounds, the rest of the proof essentially works just as in Imhof (2005). With this in mind, consider the transformed process $Y(t) = \Psi(X(t))$ given by $\Psi_{i\mu}(x) = \log x_{i\mu} / x_{i,0}$, $\mu \in \mathcal{A}_i^* \equiv \mathcal{A}_i \setminus \{\alpha_{i,0}\}$. Since $\frac{\partial \Psi_{i\mu}}{\partial x_{i\mu}} = 1/x_{i\mu}$ and $\frac{\partial \Psi_{i\mu}}{\partial x_{i,0}} = -1/x_{i,0}$, Itô's formula readily yields:

$$dY_{i\mu} = \mathcal{L}\Psi_{i\mu}(X(t)) dt + \sigma_{k\mu} dW_{k\mu} - \sigma_{k,0} dW_{k,0}, \quad (5.71)$$

and the form of the martingale part of this process shows that the infinitesimal generator of Y is elliptic. Additionally, (5.62a) shows that $Y(t)$ will hit a compact neighborhood of the origin in finite time, and, hence, by the criteria of Bhattacharya (1978, Lemma 3.4), it follows that $Y(t)$ is recurrent. However, given that Ψ is invertible in $\text{Int}(\Delta)$, the same will also hold for $X(t)$ as well, thus showing that the transition probabilities of the diffusion $X(t)$ will converge in total variation to an invariant probability measure π on Δ .

We are thus left to establish the bound $\pi(B_\delta) \geq 1 - d_\lambda^2/\delta^2$. To that end, we will use the ergodic property of $X(t)$, namely that:

$$\pi(B_\delta) = \lim_{t \rightarrow \infty} \mathbf{E}_x \left[\frac{1}{t} \int_0^t \chi_{B_\delta}(X(s)) ds \right], \quad (5.72)$$

where χ_{B_δ} denotes the indicator function of B_δ . Indeed, since $\|x - q\| \geq \delta$ outside B_δ , it easily follows that:

$$\mathbf{E}_x \left[\frac{1}{t} \int_0^t \chi_{B_\delta}(X(s)) ds \right] \geq \mathbf{E}_x \left[\frac{1}{t} \int_0^t \left(1 - \|X(s) - q\|^2/\delta^2\right) ds \right], \quad (5.73)$$

so the estimate (5.63) follows by letting $t \rightarrow \infty$ in (5.62b). \square

We conclude this section with a few remarks:

Remark 1 (Learning vs. Noise). The nature of our bounds reveals a most interesting feature of the replicator equation (5.47). On the one hand, as $\lambda_k \rightarrow 0$, we also get $d_k(\lambda_k) \rightarrow 0$, and the invariant measure π converges vaguely to a point

Hitting time estimates.

The criterion of Bhattacharya.

mass at q . Hence, if the learning rates λ_k are slow enough (or, equivalently, if the noise level σ is low enough), we recover the convergence part of Theorem 5.6 (as we should!). On the other hand, there is a clear downside to using very slow learning rates (i.e. very high learning temperatures): our bound for the expected time to hit a neighbourhood of an equilibrium increases with λ_k and actually diverges when they approach the critical value determined by the “slow-learning” condition (5.76):

$$\lambda_k^c = \frac{rD_k^2}{m_k\sigma_{k,max}^2} \left(1 - \frac{1}{A_k}\right)^{-1} \quad (5.74)$$

As a result, choosing learning rates is a delicate process and users will have to balance the rate versus the desired sharpness of their convergence.

Remark 2 (Sharpness). As in the case of Theorem 5.10, the “slow-learning” condition (5.76) is not particularly sharp and can be tightened in a number of ways. For example, the bound (5.66) can be improved by noting that the maximum of the expression $\sum_{\beta} \sigma_{k\beta}^2 x_{k\beta} (m_k - x_{k\beta})$ over $\Delta_k \equiv m_k \Delta(\mathcal{A}_k)$ is actually:

$$\frac{1}{4} A_k m_k^2 \left[\sigma_{k,a}^2 - \left(1 - \frac{2}{A_k}\right)^2 \sigma_{k,h}^2 \right] \leq m_k^2 \sigma_{k,max}^2 \left(1 - \frac{1}{A_k}\right), \quad (5.75)$$

where $\sigma_{k,a}^2$ and $\sigma_{k,h}^2$ are the arithmetic and harmonic means of the noise coefficients $\sigma_{k\beta}^2$ respectively.

Similar improvements can be made in many other parts of the proof of Theorem 5.13, but as the reader might suspect, most of them are not really worth the trouble. Nevertheless, one notable exception occurs if we observe that it suffices for (5.65) to be positive when $x \in \text{bd}(\Delta)$. In that case, the same constrained maximization arguments (which now take place over the various faces of Δ) lead to the alternative “slow-learning” condition:

$$D_k^2 > \lambda_k m_k r^{-1} \sigma_{k,max}^2 \left(1 - \frac{1}{A_k - 1}\right). \quad (5.76)$$

On the face of it, this condition is similar in appearance to (5.61), but there is a huge difference when $A_k = 2$ (in which case the k -th summand in the RHS of (5.76) vanishes):

Proposition 5.14. *If the members of the k -th population only have two choices at their disposal, then, assuming (5.76) holds for all populations with more than two choices, the trajectories $X(t)$ will be recurrent for any choice of $\lambda_k > 0$.*

In particular, if the game is dyadic, the solution paths of the replicator dynamics will be recurrent for all $\lambda_k > 0$.

By following the same line of reasoning as in the proof of Theorem 5.13, we can also obtain sharper bounds for the time averages in dyadic games of this sort (cf. the results in Section 4.4). However, since the end expressions are quite complicated, we prefer to leave things as they stand – at least until the next chapter, where we will explore a similar direction.

*Dyadic games:
an important case.*

Part III

APPLICATIONS TO NETWORKS

6

THE ROUTING PROBLEM

THE underlying problem of managing the flow of traffic in a large-scale network is as simple to state as it is challenging to resolve: given the rates of traffic generated by the users of the network, one is asked to identify and realize the most “satisfactory” distribution of traffic among the network’s routes.

Of course, given that this notion of “satisfaction” depends on the users’ optimization criteria, it would serve well to keep a concrete example in mind. Perhaps the most illustrative one is that of the Internet itself, where the primary concern of its users is to minimise the travel times of their data flows. However, since the time needed to traverse a link in the network increases (nonlinearly even) as the link becomes more congested, the users’ concurrent minimization efforts invariably lead to game-like interactions whose complexity precludes even the most rudimentary attempts at coordination. In this way, a traffic distribution will be considered “satisfactory” by a user when there is no unilateral move that he could make in order to further decrease the delays (or *latencies*) that he experiences.

*Balancing
network traffic.*

As we have noted earlier (Chapter 2), this Nash-type condition is aptly captured by *Wardrop’s principle* (Wardrop, 1952): given the level of congestion caused by other users, every user will seek to employ the minimum-latency path available to him. In fact, this principle was first stated as a “traffic equilibrium” condition (J. G. Wardrop himself was a transportation engineer) and, as might be expected, it has attracted a great deal of interest in the networks community.

One of the earliest (and most important) results regarding these *Wardrop equilibria* was the discovery that they can be calculated by solving a convex optimization problem (Beckmann et al., 1956; Dafermos and Sparrow, 1969). Among others, this characterization enabled Roughgarden and Tardos (2002, 2004) to quantify the efficiency of these equilibrium states by estimating their “price of anarchy”, i.e. the ratio between the aggregate delay of a flow at Wardrop equilibrium and the minimum achievable (aggregate) latency (Koutsoupas and Papadimitriou, 1999; Papadimitriou, 2001). Still, the size of large-scale networks makes computing these equilibria a task of considerable difficulty, clearly beyond the users’ individual deductive capabilities. Moreover, a user has no incentive to actually play out his component of an equilibrium traffic allocation unless he is convinced that his opponents will also employ theirs (an argument which gains additional momentum if there are multiple equilibria). It is thus more reasonable to take a less centralized approach and instead ask: *is there a simple learning procedure which leads users to Wardrop equilibrium?*

*The characterization
of Wardrop equilibria.*

Even though the static properties of Wardrop equilibria have been studied quite extensively, this question has been left relatively unexplored. In fact, it was only recently that the work of Sandholm (2001) showed that a good

candidate for such a learning scheme would be the replicator dynamics of evolutionary game theory. In our congestion setting, the populations modeled by these dynamics correspond to the users' traffic flows, so the convex optimization formulation of Beckmann, McGuire, and Winsten allows us to recast our problem in terms of a (nonatomic) *potential game*, like the ones studied in the previous chapter. Indeed, Wardrop equilibria can be located by looking at the minimum of the *Rosenthal potential* (Rosenthal, 1973) and, hence, Sandholm's analysis shows that they are Lyapunov stable rest points of the replicator dynamics. This fact was also recognized independently by Fischer and Vöcking (2004) who additionally showed that the (interior) solution orbits of the replicator dynamics converge to the set of Wardrop equilibria.¹

Uniqueness
and redundancy.

Rather surprisingly, when there is not a unique equilibrium, the structure of the Wardrop set itself seems to have been overlooked in the above considerations. Specifically, it has been widely assumed that, if the network's delay functions are strictly increasing, then there exists a unique Wardrop equilibrium (for instance, see Sandholm, 2001, Corollary 5.6). As a matter of fact, this uniqueness property is only true in *irreducible* networks, i.e. those networks whose paths are "independent" of one another (in a sense made precise by Definition 6.1). In general, the Wardrop set of a network is a convex polytope whose dimension is determined by the network's *redundancy*, a notion which quantifies precisely this "linear dependence". Nonetheless, despite this added structure, we show that the expectations of Fischer and Vöcking are vindicated in that the long-term behavior of the replicator dynamics remains disarmingly simple: (almost) every replicator orbit converges to a Wardrop flow and not merely to the *set* of such flows (Theorem 6.12).

Nature's
interference.

Having said that, the imitation procedure inherent in the replicator dynamics implicitly presumes itself that users have perfectly accurate information at their disposal. Unfortunately however, this assumption is not very realistic in networks which exhibit wild delay fluctuations as the result of interference by random exogenous factors (commonly gathered under the collective moniker "nature"). In population biology, these disturbances are usually modelled by introducing "aggregate shocks" to the replicator dynamics (Fudenberg and Harris, 1992) and, as one would expect, these shocks complicate the situation considerably. For instance, to reiterate what we have already seen, if the variance of the shocks is mild enough compared to the payoffs of the game, Cabrales (2000) proved that dominated strategies become extinct in (multi-population) random matching games. This restriction is also featured in the work of Imhof (2005) who showed that even equilibrial play arises over time but, again, conditionally on the noise processes not being too loud (see also Benaïm et al., 2008; Hofbauer and Imhof, 2009). On the other hand, the "exponential learning" approach that we put forth in Chapters 4 and 5 (where the replicator dynamics are not perturbed as an evolutionary birth-death process, but, instead, as a learning one), shows that similar rationality properties continue to hold no matter how loud the noise becomes (see also Mertikopoulos and Moustakas, 2009, 2010b).

Differences from
Nash environments.

All the same, with the exception of our work in the previous chapter, these approaches have chiefly focused on Nash-type games where payoffs are multilinear functions over a product of simplices; for example, payoffs in single-population random matching games are determined by the bilinear form which is associated to the matrix of the game. This linear structure

¹ Actually, the authors suggest that these orbits converge to a point, but their analysis only holds when there is a unique equilibrium.

simplifies things considerably but, unfortunately, congestion models rarely adhere to it; additionally, the notions of Nash and Wardrop equilibrium are at variance in many occasions, a disparity which also calls for a different approach; and, finally, the way that stochastic fluctuations propagate to the users' choices in a network leads to a new stochastic version of the replicator dynamics where the noise processes are no longer independent across users (different paths might share a common subset of links over which disturbances are strongly correlated). On that account, the effect of stochastic fluctuations in congestion models cannot be understood by simply translating previous work on the stochastic replicator dynamics.

OUTLINE In this chapter, we study the distribution of traffic in networks whose links are subject to constant stochastic perturbations that randomly affect the delays experienced by individual traffic elements. This model is presented in detail in Section 6.1, where we also adapt our game-theoretic machinery to our network needs: specifically, we introduce the notion of a network's *redundancy* in Section 6.1.1 and we examine its connection to Wardrop equilibria in Section 6.1.2. We then derive the rationality properties of the deterministic replicator dynamics in Section 6.2, where we show that (almost) every solution trajectory converges to a Wardrop equilibrium (the content of Theorem 6.12).

Section 6.3 is devoted to the stochastic considerations which constitute the core of this chapter. Our first result is that *strict* Wardrop equilibria remain stochastically asymptotically stable *irrespective* of the fluctuations' magnitude (Theorem 6.14); in fact, if the users are "patient enough", we are able to estimate the average time it takes them to hit a neighborhood of the equilibrium in question (Theorem 6.15). In conjunction with stochastic stability, this allows us to conclude that when a strict equilibrium exists, users converge to it almost surely (Corollary 6.16). On the other hand, given that such equilibria do not always exist, we also prove that the replicator dynamics in irreducible networks are recurrent (again under the assumption that the users are patient enough), and we use this fact to show that the long-term average of their traffic distributions concentrates mass in the neighborhood of an interior Wardrop equilibrium (Theorem 6.17).

Remark. Before proceeding any further, we should note that this chapter is, in many ways, an application of the techniques of Chapter 5. Indeed, the similarities between the congestion models that we consider here and the potential games of the previous chapter are many: after all, every congestion game *does* admit a potential function (the Rosenthal potential).

There are two main differences which merit the length of an entire chapter to be expounded upon. The first is that the Rosenthal potential of a congestion model does not always adhere to the (strict version of the) growth conditions that we introduced in the previous chapter (superlinearity or strict convexity), and, as a result, our methods require some (slight) tweaking. More importantly however, the very nature of the underlying network problem leads to a new version of the replicator dynamics which is different both from the aggregate shocks approach of Fudenberg and Harris (1992) and also from the "exponential learning" version of Chapters 4 and 5. Accordingly, the effect of any stochastic perturbations will be equally dissimilar in this network setting (different paths might exhibit correlated fluctuations), a fact which is mirrored in the rationality properties of the replicator dynamics.

Connections with potential games.

6.1 NETWORKS AND CONGESTION MODELS

Stated somewhat informally, our chief interest lies in networks whose nodes produce traffic that seeks to reach its destination as quickly as possible. However, since the time taken to traverse a path in a network increases as the network becomes congested, it is hardly an easy task to pick the “path of least resistance” – especially given that users compete against each other in their endeavors. As a result, the game-theoretic setup of Chapter 2 turns out to be remarkably appropriate for the analysis of these traffic flows.

6.1.1 Networks and Flows

Modelling
a network.

Following Roughgarden and Tardos (2002, 2004), let $\mathcal{G} \equiv \mathcal{G}(\mathcal{V}, \mathcal{E})$ be a (finite) directed graph with node set \mathcal{V} and edge set \mathcal{E} , and let $\sigma = (v, w)$ be an *origin-destination* pair in \mathcal{G} (that is, an ordered pair of nodes $v, w \in \mathcal{V}$ that can be joined by a path in \mathcal{G}). Suppose further that the origin v of σ outputs traffic towards the destination node w at some rate $\rho > 0$; then, the pair σ together with the rate ρ will be referred to as a *user* of \mathcal{G} . In this way, a *network* $\mathcal{Q} \equiv \mathcal{Q}(\mathcal{N}, \mathcal{A})$ in \mathcal{G} will consist of a set of *users* \mathcal{N} (indexed by $i = 1, \dots, N$), together with an associated collection $\mathcal{A} \equiv \prod_i \mathcal{A}_i$ of sets of paths (or *routes*) $\mathcal{A}_i = \{\alpha_{i,0}, \alpha_{i,1}, \dots\}$ joining v_i to w_i (where $\sigma_i = (v_i, w_i)$ is the origin-destination pair of user $i \in \mathcal{N}$).

Before we proceed, some remarks of a mostly book-keeping nature are in order. First off, since we will only be interested in users with at least a modicum of choice on how to route their traffic, we will take $|\mathcal{A}_i| \geq 2$ for all $i \in \mathcal{N}$. Secondly, we will be assuming that the origin-destination pairs of distinct users are themselves distinct. Fortunately, neither assumption is crucial: if there is only one route available to user i , the traffic rate ρ_i can be considered as a constant load on the route; and if two users $i, j \in \mathcal{N}$ with rates ρ_i, ρ_j share the same origin-destination pair, we will replace them by a single user with rate $\rho_i + \rho_j$ (see also Section 6.1.2). This means that the sets \mathcal{A}_i can be assumed disjoint and, as a pleasant byproduct, the path index $\alpha \in \mathcal{A}_i$ fully characterizes the user i to whom it belongs – recall also our notational conventions in Chapter 1.

Traffic flows

So, if $x_{i\alpha} \equiv x_\alpha$ denotes the amount of traffic that user i routes via the path $\alpha \in \mathcal{A}_i$, the corresponding traffic flow of user i may be represented by the point $x_i = \sum_\alpha x_{i\alpha} e_{i\alpha}$, where $\{e_{i\alpha}\}$ is the standard basis of the free vector space $V_i \equiv \mathbb{R}^{\mathcal{A}_i}$. However, for such a flow to be admissible, we must also have $x_{i\alpha} \geq 0$ and $\sum_\alpha x_{i\alpha} = \rho_i$; hence, the set of admissible flows for user i will be the simplex $\Delta_i \equiv \rho_i \Delta(\mathcal{A}_i) = \{x_i \in V_i : x_{i\alpha} \geq 0 \text{ and } \sum_\alpha x_{i\alpha} = \rho_i\}$. Then, by collecting all these individual flows in a single profile, a *flow in the network* \mathcal{Q} will simply be a point $x = \sum_i x_i \in \Delta \equiv \prod_i \Delta_i$.

and load
profiles.

An alternative (and very useful!) description of a flow $x \in \Delta$ can be obtained by looking at the traffic *load* that the flow induces on the edges of the network, i.e. at the amount of traffic y_r that circulates in each edge $r \in \mathcal{E}$ of \mathcal{G} . In particular, let us set:

$$y_r = \sum_i y_{ir} = \sum_i \sum_{\alpha \ni r} x_{i\alpha} \quad (6.1)$$

where $y_{ir} = \sum_{\alpha \ni r} x_{i\alpha}$ denotes the load that is induced on $r \in \mathcal{E}$ by the individual flow $x_i \in \Delta$. Then, a very important question that arises is whether these two descriptions are equivalent; put differently, can one recover the flow distribution $x \in \Delta$ from the loads y_r on the edges of the network?

To answer this question, let $\{\varepsilon_r\}$ be the standard basis of the space $W \equiv \mathbb{R}^{\mathcal{E}}$ which is spanned by the edges \mathcal{E} of \mathcal{G} , and consider the *indicator map* $P^i : V_i \rightarrow W$ which sends a path $\alpha \in \mathcal{A}_i$ to the sum of its constituent edges: $P^i(e_{i\alpha}) = \sum_{r \in \alpha} \varepsilon_r$.² Obviously, if we set $P^i(e_{i\alpha}) = \sum_r P_{r\alpha}^i \varepsilon_r$, we see that the entries of the matrix P^i will be:

The network's
indicator matrix.

$$P_{r\alpha}^i = \begin{cases} 1, & \text{if } r \in \alpha, \\ 0, & \text{otherwise.} \end{cases} \quad (6.2)$$

We can then aggregate this construction over all $i \in \mathcal{N}$ by considering the product space $V \equiv \mathbb{R}^{\mathcal{A}} \cong \prod_i V_i$ and the corresponding indicator matrix $P = P^1 \oplus \dots \oplus P^N$ whose entries take the value $P_{r\alpha} = 1$ if the path $\alpha \in \mathcal{A}$ employs the edge r and vanish otherwise. By doing just that, (6.1) takes the simpler form $y_r = \sum_{\alpha} P_{r\alpha} x_{\alpha}$ – or, even more succinctly, $y = P(x)$ – and, therefore, the question of whether a flow can be recovered from a load profile can be answered in the positive if the indicator map $P : V \rightarrow W$ is injective.³

This, however, is not the end of the matter because the individual flows $x_i \in \Delta_i$ are actually restricted to live in the affine subspaces $p_i + Z_i$ where $p_i = \rho_i A_i^{-1} \sum_{\alpha} e_{i\alpha}$ is the barycentre of Δ_i and $Z_i \equiv T_{p_i} \Delta_i = \{z_i \in V_i : \sum_{\alpha} z_{i\alpha} = 0\}$ is the tangent space to Δ_i at p_i .⁴ As a result, what is actually of essence here is the action of P on the subspaces $Z_i \leq V_i$, i.e. the restriction $Q \equiv P|_Z : Z \rightarrow W$ of P on the subspace $Z \equiv \prod_i Z_i$. In this way, any two flows $x, x' \in \Delta$ will have $z = x' - x \in Z$, and the respective loads $y, y' \in W$ will satisfy:

The network's
redundancy matrix.

$$y' - y = P(x') - P(x) = P(z) = Q(z), \quad (6.3)$$

so that $y' = y$ iff $x' - x \in \ker Q$. Under this light, it becomes clear that a flow $x \in \Delta$ can be recovered from the corresponding load profile $y \in W$ if and only if Q is injective. For this reason, the map $Q : Z \rightarrow W$ will be called the *redundancy matrix of the network* Q , giving rise to:

Definition 6.1. Let \mathcal{Q} be a network in a graph \mathcal{G} and let Q be the redundancy matrix of \mathcal{Q} . The *redundancy red*(\mathcal{Q}) of \mathcal{Q} is defined to be:

Redundancy

$$\text{red}(\mathcal{Q}) \equiv \dim(\ker Q). \quad (6.4)$$

If $\text{red}(\mathcal{Q}) = 0$, the network \mathcal{Q} will be called *irreducible*; otherwise, \mathcal{Q} will be called *reducible*.

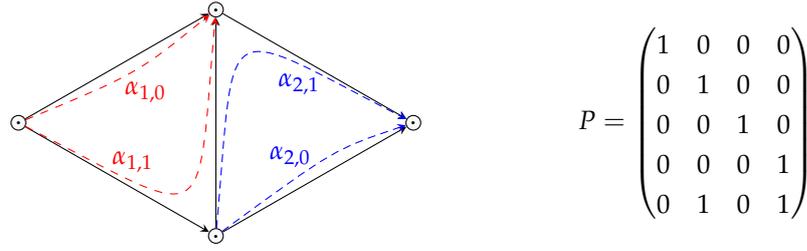
The rationale behind this terminology should be clear enough: when a network \mathcal{Q} is reducible, some of its routes are “linearly dependent” and the respective directions in $\ker Q$ will be “redundant” (in the sense that they are not reflected on the edge loads). By comparison, the degrees of freedom of irreducible networks are all “active” and any statement concerning the network’s edges may be translated to one concerning its routes.

This dichotomy between reducible and irreducible networks will be quite significant for our purposes, so it is worth dwelling on Definition 6.1 for a bit longer; specifically, it will be important to have a simple recipe with which to compute the redundancy matrix Q of a network \mathcal{Q} . To that end, let $Q^i \equiv P^i|_{Z_i}$ be the restriction of P^i on Z_i and, as before, let $\{e_{i,0}, e_{i,1}, \dots\}$ be the standard

² The relation $r \in \alpha$ means that the edge r is contained in the path α , viewed as a subset of \mathcal{E} .

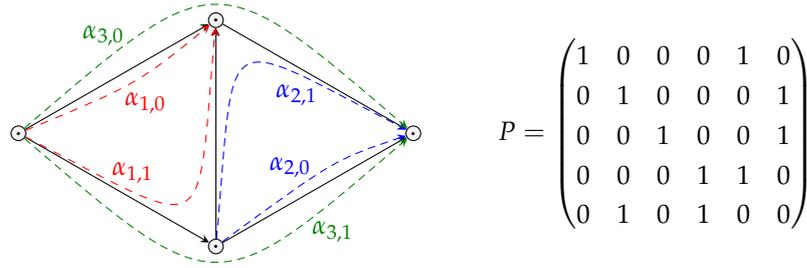
³ For obvious reasons, this map is sometimes called the *arc-path incidence matrix* of the network (for instance, see Braess et al., 2005).

⁴ As before, it is also worth keeping in mind that if we set $\mathcal{A}_i^* = \mathcal{A}_i \setminus \{\alpha_{i,0}\}$, then $Z_i \cong \mathbb{R}^{\mathcal{A}_i^*}$.



No linearly dependent paths.

(a) An irreducible network: $\text{red}(\mathcal{Q}) = 0$.



$$\alpha_{1,0} + \alpha_{2,1} + \alpha_{3,1} = \alpha_{1,1} + \alpha_{2,0} + \alpha_{3,0}$$

(b) A reducible network: $\text{red}(\mathcal{Q}) = 1$.

Figure 6.1: The addition of a user may increase the redundancy of a network.

basis of $V_i = \mathbb{R}^{\mathcal{A}_i}$. Then, the vectors $\tilde{e}_{i\mu} = e_{i\mu} - e_{i,0}$, $\mu \in \mathcal{A}_i^* \equiv \mathcal{A}_i \setminus \{0\}$, constitute a basis for Z_i , and it is easy to see that the matrix elements of Q^i in this basis will be given by:

$$Q_{r\mu}^i = P_{r\mu}^i - P_{r,0}^i. \tag{6.5}$$

The above suggests that if there are too many users in a network, then it is highly unlikely that the network will be irreducible. Indeed, we have:

A bound on the redundancy.

Proposition 6.2. *Let $\mathcal{Q}(\mathcal{N}, \mathcal{A})$ be a network in the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ and let $\mathcal{E}' \subseteq \mathcal{E}$ be the set of edges that are present in \mathcal{Q} . Then:*

$$\text{red}(\mathcal{Q}) \geq |\mathcal{N}| - |\mathcal{E}'|. \tag{6.6}$$

Hence, a network will always be reducible if the number of users exceeds the number of available links.

Proof. From the definition of $Q : Z \rightarrow W$ we can easily see that $\text{im } Q$ is contained in the subspace of W that is spanned by \mathcal{E}' ; furthermore, since every user has $|\mathcal{A}_i| \geq 2$ routes to choose from, it follows that $\dim Z = \sum_i (|\mathcal{A}_i| - 1) \geq |\mathcal{N}|$. Therefore: $\text{red}(\mathcal{Q}) = \dim(\ker Q) = \dim Z - \dim(\text{im } Q) \geq |\mathcal{N}| - |\mathcal{E}'|$. \square

6.1.2 Congestion Models and Equilibrium

The time spent by a traffic element on an edge $r \in \mathcal{E}$ of the graph \mathcal{G} will be a function $\phi_r(y_r)$ of the traffic load y_r on the edge in question – for example, if the edge represents an M/M/1 queue with capacity μ_r , then $\phi_r(y_r) = 1/(\mu_r - y_r)$. In tune with tradition, we will assume that these *latency* (or *delay*) functions

are strictly increasing, and also, to keep things simple, that they are at least C^1 with $\phi'_r > 0$.

On that account, the time needed to traverse an entire route $\alpha \in \mathcal{A}_i$ will be:

$$\omega_{i\alpha}(x) = \sum_{r \in \alpha} \phi_r(y_r) = \sum_r P_{r\alpha}^i \phi_r(y_r), \quad (6.7)$$

where, as before, $y_r = \sum_{\beta} P_{r\beta} x_{\beta}$. In summary, we then have:

Definition 6.3. A congestion model $\mathfrak{C} \equiv \mathfrak{C}(\mathcal{Q}, \phi)$ in a graph $\mathfrak{G}(\mathcal{V}, \mathcal{E})$ is a network $\mathcal{Q}(\mathcal{N}, \mathcal{A})$ of \mathfrak{G} , equipped with a family of increasing latency functions $\phi_r, r \in \mathcal{E}$.

Congestion models.

The similarities between this definition and that of a game in normal form should be evident: all that is needed to turn Definition 6.3 into a N -person game is to specify its payoff functions. One way to go about this would be to consider the *user averages*:

$$\omega_i(x) = \frac{1}{\rho_i} \sum_{\alpha}^i x_{i\alpha} \omega_{i\alpha}(x) = \frac{1}{\rho_i} \sum_r y_{ir} \phi_r(y_r), \quad (6.8)$$

where the last equality follows from (6.7) and the definition of $y_{ir} = \sum_{\alpha}^i x_{i\alpha}$. Thus, in keeping with the equilibrium condition (2.20), a flow q will be at Nash equilibrium in the game $\mathfrak{G}_1 \equiv \mathfrak{G}_1(\mathcal{N}, \Delta, -\omega)$ when:

$$\omega_i(q) \leq \omega_i(q_{-i}; q'_i) \text{ for every user } i \in \mathcal{N} \text{ and all flows } q'_i \in \Delta_i. \quad (\text{NE1})$$

For many classes of latency functions ϕ_r (e.g. the expected delays in M/M/1 queues), the average delays ω_i turn out to be convex and the existence of an equilibrium is assured by the results of Rosen (1965). However, not only is this not always the case but, more importantly, the user averages (6.8) do not necessarily reflect the users' actual optimization objectives either.

Notions of equilibrium.

Indeed, another equally justified choice of payoffs would be to take the *worst delay* that users experience:

$$\tilde{\omega}_i(x) = \max_{\alpha: x_{i\alpha} > 0} \{\omega_{i\alpha}(x)\}, \quad (6.9)$$

i.e. the time at which a user's last traffic packet reaches its destination. In that case, a flow q will be at equilibrium for the game $\mathfrak{G}_2 \equiv \mathfrak{G}_2(\mathcal{N}, \Delta, -\tilde{\omega}_i)$ when:

$$\tilde{\omega}_i(q) \leq \tilde{\omega}_i(q_{-i}; q'_i) \text{ for every user } i \in \mathcal{N} \text{ and all flows } q'_i \in \Delta_i. \quad (\text{NE2})$$

Unfortunately, the payoff functions $\tilde{\omega}_i$ may be discontinuous along intersections of faces of Δ_i because the support $\text{supp}(x_i) = \{\alpha \in \mathcal{A}_i : x_{i\alpha} > 0\}$ of x_i changes there as well. Consequently, the existence of equilibrial flows cannot be inferred from the general theory in this instance either.

On the other hand, if we go back to our original motivation (the Internet), we see that our notion of a "user" more accurately portrays the network's *routers* and not its "real-life" users (humans, applications, etc.). However, since routers are not selfish in themselves, conditions (NE1) and (NE2) do not necessarily point to the right direction either. Instead, the routers' selfless task is to ensure that the *nonatomic* traffic elements circulating in the network (the actual selfish entities of the game) remain satisfied. It is thus more reasonable to go back to Wardrop's principle (2.23) and define:

Definition 6.4. A flow $q \in \Delta$ will be at *Wardrop equilibrium* when

$$\omega_{i\alpha}(q) \leq \omega_{i\beta}(q) \text{ for all } i \in \mathcal{N} \text{ and for all routes } \alpha, \beta \in \mathcal{A}_i \text{ with } q_{i\alpha} > 0, \quad (6.10)$$

i.e. when every traffic element employs the fastest path available to it.

Condition (6.10) holds as an equality for all routes $\alpha, \beta \in \mathcal{A}_i$ that are employed in a Wardrop profile q . This gives $\omega_i(q) = \omega_{i\alpha}(q)$ for all $\alpha \in \text{supp}(q_i)$, and leads to the following alternative characterization of Wardrop flows:

$$\omega_i(q) \leq \omega_{i\beta}(q) \text{ for all } i \in \mathcal{N} \text{ and for all } \beta \in \mathcal{A}_i. \quad (6.10')$$

Even more importantly, Wardrop equilibria can also be harvested from the (global) minimum of the *Rosenthal potential* (Rosenthal, 1973):

$$\Phi(y) = \sum_r \Phi_r(y_r) = \sum_r \int_0^{y_r} \phi_r(w) dw. \quad (6.11)$$

Rosenthal's potential.

As we noted back in Section 2.1.4, the reason for calling this function a potential is twofold: firstly, it is the nonatomic generalization of the potential function introduced by Monderer and Shapley (1996) to describe finite congestion games; secondly, the payoff functions $\omega_{i\alpha}$ can be obtained from Φ by a simple differentiation.

To be sure, if we set $F(x) = \Phi(y)$ where $y = P(x)$, we readily obtain:

$$\frac{\partial F}{\partial x_{i\alpha}} = \sum_r \frac{\partial \Phi}{\partial y_r} \frac{\partial y_r}{\partial x_{i\alpha}} = \sum_r \phi_r(y_r) P_{r\alpha}^i = \sum_{r \in \alpha} \phi_r(y_r) = \omega_{i\alpha}(x), \quad (6.12)$$

which is exactly the definition of a potential function in the sense of (2.10) – note also that Sandholm's "externality symmetry" condition (2.9) can be verified independently:

$$\frac{\partial \omega_{i\alpha}}{\partial x_{j\beta}} = \sum_r P_{r\alpha}^i \phi_r'(y_r) P_{r\beta}^j = \sum_{r \in \alpha \cap \beta} \phi_r'(y_r) = \frac{\partial \omega_{j\beta}}{\partial x_{i\alpha}}. \quad (6.13)$$

Characterizing Wardrop equilibria.

To describe the exact relation between Wardrop flows and the minima of Φ , consider the (convex) set $P(\Delta)$ of all load profiles y that result from admissible flows $x \in \Delta$. Since the latency functions ϕ_r are increasing, Φ will be strictly convex over $P(\Delta)$ and it will thus have a unique (global) minimum $y^* \in P(\Delta)$. Amazingly enough, the Kuhn-Tucker conditions that characterise this minimum coincide with the Wardrop condition (2.23) (Beckmann et al., 1956; Dafermos and Sparrow, 1969; Roughgarden and Tardos, 2002), so the Wardrop set of the congestion model \mathfrak{C} will be given by:

$$\Delta^*(\mathfrak{C}) \equiv \Delta^* = \{x \in \Delta : P(x) = y^*\} = P^{-1}(y^*) \cap \Delta. \quad (6.14)$$

Proposition 6.5. *Let $\mathfrak{C} \equiv \mathfrak{C}(\mathcal{Q}, \phi_r)$ be a congestion model with strictly increasing latencies ϕ_r and let Δ^* be its set of Wardrop equilibria. Then:*

1. any two Wardrop flows exhibit equal loads and delays.
2. Δ^* is a nonempty convex polytope with $\dim(\Delta^*) \leq \text{red}(\mathcal{Q})$; moreover, if there exists an interior equilibrium $q \in \text{Int}(\Delta)$, then $\dim(\Delta^*) = \text{red}(\mathcal{Q})$.

Since $P^{-1}(y^*)$ is an affine subspace of $\mathbb{R}^{\mathcal{A}} \cong \prod_i \mathbb{R}^{\mathcal{A}_i}$ and Δ is a product of simplices, there is really nothing left to prove (simply observe that if q is an interior Wardrop flow, then $P^{-1}(y^*)$ intersects the full-dimensional interior of Δ). The only surprise here is that this result seems to have been overlooked in much of the literature concerning congestion models: for instance, both Sandholm (2001, Corollary 5.6) and Fischer and Vöcking (2004, Propositions 2 and 3) presume that Wardrop equilibria are unique in networks with increasing

latencies. However, if there are two distinct flows x, x' leading to the same load profile y (e.g. as in the simple network of Fig. 6.1b), then the potential function $F(x) \equiv \Phi(P(x))$ is no longer strictly convex: it is in fact *constant* along every null direction of the redundancy matrix Q .

We thus see that a Wardrop equilibrium is unique if and only if *a*) the network \mathcal{Q} is *irreducible*, or *b*) $P^{-1}(y^*)$ only intersects Δ at a vertex. This last condition suggests that the vertices of Δ play a special role so, in analogy with Nash games, we define:

Definition 6.6. A Wardrop equilibrium q will be called *strict* if *a*) q is *pure*: $q = \sum_i \rho_i e_{i,\alpha_i}$, $\alpha_i \in \mathcal{A}_i$; and *b*) $\omega_{i\alpha_i}(q) < \omega_{i\beta}(q)$ for all paths $\beta \in \mathcal{A}_i \setminus \{\alpha_i\}$.

In Nash games, a pure equilibrium occasionally fails to be strict, but only by a hair: an arbitrarily small perturbation of a player's pure payoffs $u_i(\alpha_1, \dots, \alpha_N)$ resolves a pure equilibrium into a strict one without affecting the payoffs of the other players. In congestion models however, there is no such guarantee because, whenever two users' paths overlap, one cannot perturb the delays of one user independently of the other's. As a matter of fact, the existence of a strict Wardrop equilibrium actually *precludes* the existence of other equilibria:

Proposition 6.7. *If q is a strict Wardrop equilibrium of a congestion model \mathcal{C} , then q is the unique Wardrop equilibrium of \mathcal{C} .*

Strict equilibria and uniqueness.

Proof. Without loss of generality, let $q = \sum_i \rho_i e_{i,0}$ be a strict Wardrop equilibrium of \mathcal{C} , and suppose ad absurdum that $q' \neq q$ is another Wardrop flow. If we set $z = q' - q \in \ker Q$, it follows that the convex combination $q + \theta z$ will also be Wardrop for all $\theta \in [0, 1]$; moreover, for small enough $\theta > 0$, $q + \theta z$ employs at least one path $\mu \in \mathcal{A}_i^* \equiv \mathcal{A}_i \setminus \{0\}$ that is not present in q (recall that q is pure). As a result, we get $\omega_{i\mu}(q + \theta z) = \omega_{i,0}(q + \theta z)$ for all sufficiently small $\theta > 0$, and because the latency functions $\omega_{i\alpha}$ are continuous, this yields $\omega_{i,0}(q) = \omega_{i\mu}(q)$. On the other hand, since q is a *strict* equilibrium which does not employ μ , we must also have $\omega_{i,0}(q) < \omega_{i\mu}(q)$, a contradiction. \square

In other words, even if q is a strict equilibrium of a reducible network, then the redundant rays in $q + \ker Q$ will only intersect Δ at q . On the other hand, if q is merely a pure equilibrium, $q + \ker Q$ might well intersect the open interior of Δ ; in that case, there is no arbitrarily small perturbation of the delay functions that could make q into a strict equilibrium.

EQUILIBRIA AND OBJECTIVES On account of the above, we will focus our investigations on Wardrop equilibria. However, we should mention here that this equilibrated notion can also be reconciled (to some extent at least) with the optimization objectives represented by the payoffs (6.8) and (6.9) as well.

First, with respect to the average delays $\omega_i(x) = \rho_i^{-1} \sum_{\alpha} x_{i\alpha} \omega_{i\alpha}(x)$, the *optimal* traffic distributions which minimise the aggregate delay $\omega(x) = \sum_i \rho_i \omega_i(x)$ actually coincide with the Wardrop equilibria of a suitably modified game. This was first noted by Beckmann et al. (1956), who observed the inherent duality in Wardrop's principle: just as Wardrop equilibria occur at the minimum of the Rosenthal potential, so can one obtain the minimum of the aggregate latency ω by looking at the Wardrop equilibria of an associated congestion model. More precisely, the only change that needs to be made is to consider the "marginal" latency functions $\phi_r^*(y_r) = \phi_r(y_r) + y_r \phi_r'(y_r)$ (see also Roughgarden and Tardos, 2004). Then, to study these "socially optimal" flows, we simply have to redress our analysis to fit these "marginal latencies" instead (see Section 6.4 for more details).

Braess's paradox.

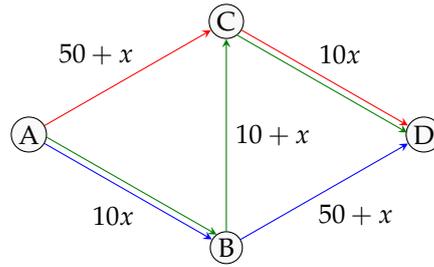


Figure 6.2: Braess's paradox and the disparity between Wardrop and Nash equilibria.

Secondly, Wardrop equilibria also have close ties with the Nash condition (NE2) which corresponds to the “worst-delays” (6.9). Specifically, one can easily see that the Nash condition (NE2) is equivalent to the Wardrop condition (2.23) when every user only has 2 possible paths to choose from (every amount of traffic diverted from one path increases the delay at the user's other path). However, if a user has 3 or more paths at his disposal, then the situation can change dramatically because of *Braess's paradox* (Braess, 1968).

The essence of this paradox is that there exist networks which perform better if one removes their fastest link. An example of such a network is given in Fig. 6.2, where it is assumed that a user seeks to route 6 units of traffic from A to D using the three paths $A \rightarrow B \rightarrow D$ (blue), $A \rightarrow C \rightarrow D$ (red) and $A \rightarrow B \rightarrow C \rightarrow D$ (green). In that case, the Wardrop condition (6.10) calls for equidistribution: 2 units are routed via each path, leading to a delay of 92 time units along all paths. Alternatively, if the user sends 3 traffic units via the red and blue paths and ignores the green one, all traffic will experience a delay of 83. Paradoxically, even though the green path has a latency of only 70, the Nash condition (NE2) is satisfied: if traffic is diverted from, say, the red path to the faster green one, then the latency of the blue path will also increase, thus increasing the worst delay $\tilde{\omega}_i$ as well.

*The price
of anarchy.*

This paradox is what led to the original investigations in the efficiency of selfish routing (Koutsoupias and Papadimitriou, 1999; Roughgarden and Tardos, 2002), and it seems that it is also what causes this disparity between Wardrop and Nash equilibria. A thorough investigation of this matter is a worthy project but, since it would take us too far afield, we will not pursue it here. Henceforward, we will focus almost exclusively on Wardrop flows which represent the most relevant equilibrium concept for our purposes.

6.2 LEARNING, EVOLUTION AND RATIONAL BEHAVIOR

Unfortunately, locating the Wardrop equilibria of a network is a rather arduous process which entails a good deal of global calculations (namely the minimization of a nonlinear convex functional with exponentially many variables over a convex polytope). Since such calculations exceed the deductive capabilities of individual users (especially if they do not have access to global information), it is of great interest to see whether there are learning schemes which allow users to reach an equilibrium without having to rely on centralized computations.

6.2.1 Learning and the Replicator Dynamics

For our purposes, a learning scheme will be a rule which trains users to route their traffic in an efficient way by processing information that is readily

available. On the other hand, since this information must be “local” in nature, any such learning scheme should be similarly “distributed” across users: for example, the play of one’s opponents or the exact form of the network’s latency functions are not easily accessible pieces of information. Furthermore, we should also be looking for a learning scheme which is simple enough for users to apply in real-time, without having to perform a huge number of calculations at each instant.

*Distributed
learning schemes.*

Recalling the discussion in Chapter 2, such a learning scheme may be cast as a dynamical system in continuous time:

$$\frac{dx}{dt} = v(x) \text{ or, in coordinates: } \frac{dx_{i\alpha}}{dt} = v_{i\alpha}(x), \quad (6.15)$$

where $x(t) \in \Delta$ denotes the flow at time t and the vector field $v : \Delta \rightarrow \mathbb{R}^A$ plays the part of the “learning rule” in question – for simplicity, we will also take v to be smooth. Of course, since the flow $x(t)$ evolves in Δ , v itself must lie on the tangent space Z of Δ ; we thus require that $\sum_{\alpha}^i v_{i\alpha}(x) = 0$ for all $i \in \mathcal{N}$.

Furthermore, v should also leave the *faces* of Δ invariant : any trajectory $x_i(t)$ that begins at some face of Δ_i must always remain in said face. This is actually an essential consequence of our postulates: if a user does not employ a route $\alpha \in \mathcal{A}_i$, then he has *no* information on it and, as such, there is no a priori reason that an adaptive learning rule would induce the user to sample it. In effect, such a learning rule would either fail to rely solely on readily observable information or would not necessarily be a very simple one.

This shows that $v_{i\alpha}(x)$ must vanish if $x_{i\alpha} = 0$, so if we set $v_{i\alpha}(x) = x_{i\alpha} \tilde{v}_{i\alpha}(x)$, we obtain the orthogonality condition $\sum_{\alpha}^i x_{i\alpha} \tilde{v}_{i\alpha}(x) = 0$. Accordingly, $\tilde{v}_{i\alpha}$ may be written in the form:

$$\tilde{v}_{i\alpha}(x) = u_{i\alpha}(x) - u_i(x), \quad (6.16)$$

where the $u_{i\alpha}$ satisfy no further constraints, and, as can be shown by a simple summation, the function $u_i(x)$ is just the *user average*: $u_i(x) = \rho_i^{-1} \sum_{\beta}^i x_{i\beta} u_{i\beta}(x)$ (recall that $\sum_{\beta}^i x_{i\beta} = \rho_i$). This shows that *any* learning rule which leaves the faces of Δ invariant must necessarily adhere to the replicator form:

$$\frac{dx_{i\alpha}}{dt} = x_{i\alpha} (u_{i\alpha}(x) - u_i(x)). \quad (6.17)$$

In our case, the most natural choice for the per capita growth rates $u_{i\alpha}$ of each “phenotype” $\alpha \in \mathcal{A}_i$ is to use the delay functions $\omega_{i\alpha}(x)$ and set $u_{i\alpha} = -\omega_{i\alpha}$. In so doing, we obtain:

$$\frac{dx_{i\alpha}}{dt} = x_{i\alpha} (\omega_i(x) - \omega_{i\alpha}(x)). \quad (6.18)$$

In keeping with our “local information” mantra, users do not need to know the delays along paths that they do not employ because the replicator field vanishes when $x_{i\alpha} = 0$. Thus, users that evolve according to (6.18) are oblivious to their surroundings, even to the existence of other users: they simply use (6.18) to respond to the stimuli $\omega_{i\alpha}(x)$ in the hope of minimizing their delays.

Alternatively, if players learn at different rates $\lambda_i > 0$ as a result of varied stimulus-response characteristics, we obtain the *rate-adjusted* dynamics:

$$\frac{dx_{i\alpha}}{dt} = \lambda_i x_{i\alpha} (\omega_i(x) - \omega_{i\alpha}(x)) \quad (6.19)$$

(naturally, the uniform case (6.18) is recovered when all players learn at the “standard” rate $\lambda_i = 1$). As with the “exponential learning” case, these learning rates can also be viewed as (player-specific) inverse temperatures: in high temperatures (small λ_i), the differences between routes are toned down and players evolve along the slow time-scales $\lambda_i t$; at the other end of the spectrum, if $\lambda_i \rightarrow \infty$, equation (6.19) “freezes” to a rigid (and myopic) best-reply process (see also [Börger and Sarin, 1997](#)).

6.2.2 Entropy and Rationality

From the discussion in Chapters 2 and 5, we know that all Wardrop flows are stationary in (6.17), but the converse does not hold: every flow that exhibits equal latencies along the paths in its support will be a rest point of (6.17), and such flows are not always Wardrop (at least if they are not interior points of Δ). Consequently, the issue at hand remains whether or not the replicator dynamics manage to single out Wardrop equilibria among other steady states.

At first glance, the existence of a potential function combined with our results in Chapter 5 would seem to render this a trivial question: the Wardrop equilibria are at the minimum of the potential, and, hence, interior trajectories will converge to equilibrium. Nonetheless, the underlying assumption for most of our results in Chapter 5 is that the game’s potential be strictly convex (at least locally), and this condition need not hold in congestion settings: if the network is reducible, then the potential $F(x) \equiv \Phi(y(x))$ fails to be strictly convex because of the redundant directions along which F is constant.

To examine this issue in more detail, note that if y^* is the minimum of the Rosenthal potential $\Phi(y)$, then the function $F_0(x) = \Phi(P(x)) - \Phi(y^*)$ is a *semi-definite* Lyapunov function for the replicator dynamics (6.19). Indeed, F_0 vanishes on the Wardrop set Δ^* , it is positive otherwise, and it satisfies:

$$\frac{dF_0}{dt} = \sum_{i,\alpha} \frac{\partial F_{i\alpha}}{\partial x_{i\alpha}} \frac{dx_{i\alpha}}{dt} = \sum_i \lambda_i \rho_i \left[\omega_i^2(x) - \rho_i^{-1} \left(\sum_\alpha x_{i\alpha} \omega_{i\alpha}^2(x) \right) \right] \leq 0, \quad (6.20)$$

the last step following from Jensen’s inequality – equality only holds when $\omega_{i\alpha}(x) = \omega_i(x)$ for all $\alpha \in \text{supp}(x)$. Thus, by standard results in the theory of dynamical systems, it follows that the solution orbits of (6.19) descend the potential F_0 and eventually converge to a connected subset of rest points – see also [Sandholm \(2001\)](#), where the property (6.20) is referred to as “positive correlation”.

Nevertheless, since not all stationary points of (6.19) are Wardrop equilibria, this result tells us little about the rationality properties of the replicator dynamics in reducible networks. To sharpen these results, we will fall back to the study of the *relative entropy* $H_q(x) = \sum_k \lambda_i^{-1} \sum_\alpha q_{i\alpha} \log(q_{i\alpha}/x_{i\alpha})$:

Lemma 6.8. *Let $\mathcal{C}(\mathcal{Q}, \{\phi_r\})$ be a congestion model with increasing latencies, and let $q \in \Delta$ be a Wardrop flow of \mathcal{C} . Then, the time derivative $\dot{H}_q(x)$ of the relative entropy $H_q(x)$ vanishes on the Wardrop set $\Delta^*(\mathcal{C})$ and is negative otherwise.*

In particular, if the network \mathcal{Q} is irreducible ($\text{red}(\mathcal{Q}) = 0$) Δ^ consists of a single point and H_q is Lyapunov for the replicator dynamics (6.18).*

To prove this lemma, note that, in our case, (5.11) reads:

$$\dot{H}_q(x) = \sum_{i,\alpha} \frac{\partial H_q}{\partial x_{i\alpha}} \dot{x}_{i\alpha} = - \sum_{i,\alpha} q_{i\alpha} (\omega_i(x) - \omega_{i\alpha}(x)) \equiv -L_q(x), \quad (6.21)$$

The potential as a Lyapunov function.

The relative entropy as a potential function.

where $L_q(x) = \sum_{i\alpha}^i q_{i\alpha}(\omega_i(x) - \omega_{i\alpha}(x))$ is the evolutionary index of x with respect to q (see Section 5.1.1 for a more in-depth discussion). The interesting twist here is that some linear algebra yields:

$$L_q(x) \equiv \sum_{i,\alpha} (x_{i\alpha} - q_{i\alpha})\omega_{i\alpha}(x) = \sum_r (y_r - y_r^*)\phi_r(y_r) \equiv \Lambda(y), \quad (6.22)$$

where $y = P(x)$ and $y^* = P(q)$ are the loads which correspond to the flows x and q respectively.⁵ As a result, the equation $L_q(x) \equiv \Lambda(y)$ is consistent with any choice of $q \in \Delta^*$, and we see that the evolutionary index in congestion models does not so much depend on the actual flow x as it does on the load profile $y(x)$.

Now, since the game's potential $F(x) = \Phi(y(x))$ in irreducible networks does not necessarily satisfy the growth condition (5.22) (which is the starting point for most of the theory surrounding the evolutionary index), we will need to look at the specific form of the Rosenthal potential before applying the techniques of Chapter 5. To that end, given that the latency functions ϕ_r , are increasing, we obtain the growth estimate:

$$\Lambda(y) = \sum_r (y_r - y_r^*)\phi_r(y_r) \geq \sum_r \int_{y_r^*}^{y_r} \phi_r(w) dw = \Phi(y) - \Phi(y^*). \quad (6.23)$$

Thus, going back to the proof of Lemma 6.8, we finally get:

$$\dot{H}_q(x) = -L_q(x) = -\Lambda(y) = -(\Phi(y) - \Phi(y^*)) \leq 0, \quad (6.24)$$

with equality holding if and only if $y = y^*$, that is, if and only if $x \in \Delta^*$.

Needless to say, there is a great degree of similarity between our results so far and our analysis for strictly convex games in Chapter 5. In fact, if we also assume that q is strict (say, $q = \sum_i \rho_i e_{i,0}$ for convenience), we get the same bound for the evolutionary index as in Lemma 5.4:

Lemma 6.9. *Let $q = \sum_i \rho_i e_{i,0}$ be a strict Wardrop equilibrium of some congestion model \mathfrak{C} with increasing latency functions. Then, if $\Delta\omega_i = \min_{\mu \neq 0} \{\omega_{i\mu}(q) - \omega_{i,0}(q)\}$, we will have:*

$$L_q(q + \theta z) \geq \frac{1}{2}\theta \sum_i \Delta\omega_i \|z_i\|_1, \quad (6.25)$$

for all $z \in T_q^c \Delta$ and $\theta \geq 0$ such that $q + \theta z \in \Delta$.

This lemma shows that L_q increases at least linearly along all “inward” rays $q + \theta z$.⁶ This is not so if q is an interior equilibrium:

Lemma 6.10. *Let $q \in \text{Int}(\Delta)$ be an interior Wardrop equilibrium of some congestion model \mathfrak{C} with increasing latency functions, and let $z \in T_q \Delta$. Then, if $m = \inf\{\phi_r'(y_r) : r \in \mathcal{E}, y \in P(\Delta)\}$, we will have:*

$$L_q(q + \theta z) \geq \frac{1}{2}m \|P(z)\|^2 \theta^2 \quad \text{for all } \theta \geq 0 \text{ such that } q + \theta z \in \Delta. \quad (6.26)$$

⁵ In more detail, this all stems from the invariance identity:

$$\sum_{\alpha} z_{\alpha} \omega_{\alpha} = \sum_{\alpha} \sum_r z_{\alpha} P_{r\alpha} \phi_r = \sum_r w_r \phi_r, \quad z \in V,$$

where $P : V \rightarrow W$ is the indicator matrix of the network \mathcal{Q} and $w = P(z)$. It is then easy to verify that $L_q(x) = L_q(x')$ whenever $x' - x \in \ker Q$ and also that $L_q = L_{q'}$ iff $q' - q \in \ker Q$ – thus justifying the notation of (6.22).

⁶ It is interesting to note here the relation with Proposition 6.7: if the ray $q + \theta z$ is inward-pointing, then z cannot be “redundant”, i.e. we cannot have $z \in \ker P$.

Remark. If γ_0^2 denotes the minimum of the Rayleigh quotient $R(z) = \langle z, P^T Pz \rangle$, $z \in T_q \Delta$, $\|z\| = 1$, then the growth estimate (6.26) can be written in the more familiar form $L_q(x) \geq \frac{1}{2} m \gamma_0^2 \|x - q\|^2$ – cf. equation (5.41).

Proof. Following the proof of Lemma 5.4, let $f(\theta) = F(q + \theta z)$ so that:

$$f'(0) = \sum_i \sum_\alpha^i z_{i\alpha} \omega_{i\alpha}(q) = \sum_i \sum_\alpha^i z_{i\alpha} \omega_i(q) = 0, \tag{6.27}$$

the second equality following from the fact that q is an interior equilibrium (that is, $\omega_{i\alpha}(q) = \omega_i(q)$ for all paths $\alpha \in \mathcal{A}_i$), and the last one being a consequence of z being tangent to Δ (so that $\sum_\alpha^i z_{i\alpha} = 0$). Following up on this, a second differentiation gives:

$$f''(\theta) = \frac{d^2}{d\theta^2} \sum_r \Phi_r(y_r^* + \theta w_r) = \sum_r w_r^2 \phi_r'(y_r^* + \theta w_r), \tag{6.28}$$

where $w = P(y)$. Clearly, since the set $P(\Delta)$ of load profiles y is compact and the (continuous) functions ϕ_r' are positive, we will also have $m = \inf\{\phi_r'(y_r) : r \in \mathcal{E}, y \in P(\Delta)\} > 0$. We thus get $f(\theta) \geq \frac{1}{2} m \|w\|^2 \theta^2$, and a first order Taylor expansion with Lagrange remainder yields:

$$L_q(q + \theta z) = \Lambda(y^* + \theta w) \geq f(\theta) - f(0) \geq \frac{1}{2} m \|P(z)\|^2 \theta^2. \quad \square$$

In view of the above bounds for the evolutionary index L_q , we obtain:

Exponential convergence to Wardrop equilibrium.

Theorem 6.11. *Let $\mathfrak{C} \equiv \mathfrak{C}(\mathcal{Q}, \{\phi_r\})$ be a congestion model with increasing latency functions ϕ_r , and let q be a Wardrop equilibrium of \mathfrak{C} .*

1. *If \mathfrak{C} is irreducible, or if q is strict, then q is evolutionarily stable, and we have:*

$$H_q(x(t)) \leq h_0 e^{-ct}, \tag{6.29}$$

where $h_0 = H_q(x(0))$, $c = h_0^{-1} \min_i \left\{ \rho_i (1 - e^{-\lambda_i h_0 / \rho_i}) \Delta \omega_i \right\}$ and $\Delta \omega_i = \min_{\mu \neq 0} \{ \omega_{i\mu}(q) - \omega_{i,0}(q) \}$.

2. *If \mathfrak{C} is reducible and q is not strict, then q is neutrally stable and the evolutionarily neutral directions coincide with the redundant ones; moreover:*

$$H_q(x(t)) \leq h_0 \exp\left(-\frac{1}{2} \frac{m \gamma_0^2}{h_c} t\right), \tag{6.30}$$

where $m = \inf\{\phi_r'(y_r) : r \in \mathcal{E}, y \in P(\Delta)\}$, $h_c \geq h_0$ is a constant, and γ_0^2 is the minimum value of the Rayleigh quotient $R(z) = \langle z, P^T Pz \rangle$, $z \in T_q^c \Delta$, $\|z\| = 1$.

The first part of this theorem is an immediate application of Theorem 5.5 and the second one can be proved in exactly the same way as Theorem 5.6, so the proof itself holds little intrinsic value. What is more important here is to note that Theorem 6.11 seems to deepen the dichotomy between reducible and irreducible networks even further: although convergence to equilibrium is guaranteed in irreducible networks, γ_0^2 vanishes in irreducible ones, so convergence is not guaranteed there.

The trouble with semi-definiteness.

The reason for this is that the relative entropy is only a *semi-definite* Lyapunov function, and such a function is not enough to establish convergence by itself. For instance, if we consider the homogeneous system:

$$\dot{x} = yz, \quad \dot{y} = -xz, \quad \dot{z} = -z^2, \tag{6.31}$$

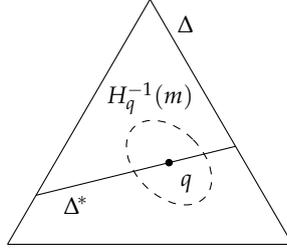


Figure 6.3: The various sets in the proof of Theorem 6.12.

with $z \geq 0$, we see that it admits the semi-definite Lyapunov function $H(x, y, z) = x^2 + y^2 + z^2$ whose time derivative only vanishes on the x - y plane. However, the general solution of (6.31) in cylindrical coordinates (ρ, ϕ, z) is:

$$\rho(t) = \rho_0, \quad \phi(t) = \phi_0 - \log(1 + z_0 t), \quad z(t) = \frac{z_0}{1 + z_0 t}, \quad (6.32)$$

and this represents a helix of constant radius whose coils become topologically dense as the solution orbits approach the x - y plane. We thus see that the solutions of (6.31) approach a set of stationary points, but do not converge to a specific one.

That said, there is much more at work in the replicator dynamics (6.19) than a single semi-definite Lyapunov function: there exists a whole *family* of such functions, one for each Wardrop flow $q \in \Delta$. So, undeterred by potential degeneracies, we see that the replicator dynamics actually *do* converge to equilibrium:

Theorem 6.12. *Let $\mathcal{C} \equiv \mathcal{C}(\Omega, \{\phi_r\})$ be a congestion model. Then, every interior solution trajectory of the replicator dynamics (6.19) converges to a Wardrop equilibrium of \mathcal{C} ; in particular, if the network Ω is irreducible, $x(t)$ converges to the unique Wardrop equilibrium of \mathcal{C} .*

Unconditional convergence.

Proof. It will be useful to shift our point of view to the *evolution function* $\theta(x, t)$ of the dynamics (6.19) which describes the solution trajectory that starts at x at time $t = 0$ – see also the discussion preceding Definition 2.14.

With this in mind, fix the initial condition $x \in \text{Int}(\Delta)$ and let $x(t) = \theta(x, t)$ be the corresponding solution orbit. If $q \in \Delta^*$ is a Wardrop equilibrium of \mathcal{C} , then, in view of Lemma 6.8, the function $V_q(t) \equiv H_q(\theta(x, t))$ will be decreasing and will converge to some $m \geq 0$ as $t \rightarrow \infty$. It thus follows that $x(t)$ converges itself to the level set $H_q^{-1}(m)$.

Suppose now that there exists some increasing sequence of times $t_n \rightarrow \infty$ such that $x_n \equiv x(t_n)$ does not converge to Δ^* . By compactness of Δ (and by descending to a subsequence if necessary), we may assume that $x_n = \theta(x, t_n)$ converges to some $x^* \notin \Delta^*$ – but necessarily in $H_q^{-1}(m)$. Hence, for any $t > 0$, we will have:

$$H_q(\theta(x, t_n + t)) = H_q(\theta(\theta(x, t_n), t)) \rightarrow H_q(\theta(x^*, t)) < H_q(x^*) = m \quad (6.33)$$

where the (strict) inequality stems from the fact that $\dot{H}_q < 0$ outside Δ^* . On the other hand, $H_q(\theta(x, t_n + t)) = V_q(t_n + t) \rightarrow m$, a contradiction.

Since the sequence t_n was arbitrary, this shows that $x(t)$ converges to the set Δ^* . Consequently, if q' is an ω -limit point of $x(t)$, then we will also have $V_{q'}(t'_n) = H_{q'}(x(t'_n)) \rightarrow 0$ for some increasing sequence of times $t'_n \rightarrow \infty$.

However, given that $V_{q'}(t)$ is decreasing, this means that $\lim_{t \rightarrow \infty} V_{q'}(t) = 0$ as well, and, seeing as $H_{q'}$ only vanishes at q' , we conclude that $x(t) \rightarrow q'$. \square

Remark 1 (Previous Work). In the context of potential games, Sandholm (2001) examined dynamics $v_{i\alpha}$ which are “positively correlated” to the game’s payoff functions $u_{i\alpha} = -\omega_{i\alpha}$, in the sense that $\sum_{i,\alpha} v_{i\alpha}(x)\omega_{i\alpha}(x) \geq 0$. It was then shown that if the rest points of these dynamics coincide with the game’s Wardrop equilibria (the “non-complacency” condition), then all solution orbits converge to set of Wardrop equilibria. Unfortunately, as we have already pointed out, the replicator dynamics are “complacent” and, in that case, Sandholm’s results only ensure that Wardrop equilibria are Lyapunov stable.

A common misconception.

To the best of our knowledge, the stronger convergence properties of Theorem 6.12 were first suggested by Fischer and Vöcking (2004) who identified the link between Wardrop equilibrium and evolutionary stability. In particular, the authors showed that Wardrop equilibria are robust against “mutations” that lead to greater delays; however, in networks with more than one users (the “multi-commodity” case as they call it), their approach rests heavily on the (implicit) assumption of irreducibility. If this is not the case, the evolutionary index L_q is only *semi-definite* and the approach of Fischer and Vöcking breaks down because Wardrop equilibria are only *neutrally* stable – this is also the problem with Corrolary 5.1 in Sandholm (2001). In fact, we have:

Absence of limit cycles.

Remark 2 (Evolution and Friction). Since the replicator trajectories converge to a Wardrop equilibrium, it follows that there can be no limit cycles. On the other hand, limit cycles are a common occurrence in evolutionary games: for example, Matching Pennies and Rock-Paper-Scissors (in their unadjusted versions) both exhibit limit cycles in the standard replicator dynamics (Weibull, 1995). So, while the evolutionary energy of large populations may remain undiminished over time, Theorem 6.12 shows that congestion models are *dissipative* and traffic flows invariably settle down to a steady state.

Remark 3 (Non-interior Trajectories). Clearly, if $x(0)$ does not employ all routes that are present in a Wardrop flow q , $x(t)$ cannot have q as a limit point – a consequence of the fact that the replicator dynamics leave the faces of Δ invariant. All the same, one can simply quotient out the routes that are not initially present until $x(0)$ becomes an interior point in the reduced strategy space Δ_{eff} that ensues. In that case, Theorem 6.12 can be applied to the (similarly reduced) congestion model $\mathfrak{C}_{\text{eff}}$ to show that $x(t)$ converges to Wardrop equilibrium in $\mathfrak{C}_{\text{eff}}$ (cf. the “restricted equilibria” of Fischer and Vöcking, 2004).

6.3 STOCHASTICALLY PERTURBED CONGESTION MODELS

Going back to our original discussion of learning schemes, we see that the users’ evolution hinges on the feedback that they receive about their choices, namely the delays $\omega_{i\alpha}(x)$ that they observe. We have already noted that this information is based on actual observations, but this does not necessarily mean that it is also *accurate* as well. For instance, the interference of nature with the game or imperfect readings of one’s payoffs might perturb this information considerably; additionally, if the users’ traffic flows are not continuous in time but consist of discrete segments instead (e.g. datagrams in communication networks), the queueing latencies $\omega_{i\alpha}$ only represent the users’ expected delays. Hence, the delays that users actually record might only be a randomly fluctuating estimate of the underlying payoffs, and this could negatively affect the rationality properties of the replicator dynamics.

6.3.1 The Stochastic Replicator Dynamics

Our goal here will be to determine the behavior of the replicator dynamics under stochastic perturbations of the kind outlined above. To that end, write the delay that users experience along the edge $r \in \mathcal{E}$ as $\hat{\phi}_r = \phi_r + \eta_r$ where η_r denotes the perturbation process. Then, the latency $\hat{\omega}_{i\alpha}$ along $\alpha \in \mathcal{A}_i$ will just be $\hat{\omega}_{i\alpha} = \omega_{i\alpha} + \eta_{i\alpha}$, where, in obvious notation, $\eta_{i\alpha} = \sum_{r \in \alpha} P_{r\alpha}^i \eta_r$. In this way, the replicator dynamics (6.18) become:

$$\frac{dx_{i\alpha}}{dt} = x_{i\alpha} (\hat{\omega}_i - \hat{\omega}_{i\alpha}) = x_{i\alpha} (\omega_i - \omega_{i\alpha}) + x_{i\alpha} (\eta_i - \eta_{i\alpha}), \quad (6.34)$$

where $\hat{\omega}_i = \rho_i^{-1} \sum_{\beta}^i x_{i\beta} \hat{\omega}_{i\beta}$ and $\eta_i = \rho_i^{-1} \sum_{\beta}^i x_{i\beta} \eta_{i\beta}$.

The exact form of the perturbations η_r clearly depends on the particular situation at hand. Still, since we are chiefly interested in stochastic fluctuations around the underlying delays $\omega_{i\alpha}$, it is reasonable to take these perturbations to be some sort of white noise that does not bias users towards one direction or another. In that case, we should rewrite (6.34) as a stochastic differential equation of the form:

$$dX_{i\alpha} = X_{i\alpha} [\omega_i(X) - \omega_{i\alpha}(X)] dt + X_{i\alpha} \left[dU_{i\alpha} - \rho_i^{-1} \sum_{\beta}^i X_{i\beta} dU_{i\beta} \right], \quad (6.35)$$

where $dU_{i\alpha}$ describes the total noise along the path $\alpha \in \mathcal{A}_i$:

$$dU_{i\alpha} = \sum_{r \in \alpha} \sigma_r dW_r = \sum_r P_{r\alpha}^i \sigma_r dW_r, \quad (6.36)$$

and $W(t) = \sum_r W_r(t) \varepsilon_r$ is a Wiener process in $\mathbb{R}^{\mathcal{E}}$, the space spanned by the edges \mathcal{E} of the network.

In a similar vein, if players learn at different rates λ_i , we obtain the rate-adjusted dynamics:

$$\begin{aligned} dX_{i\alpha} &= \lambda_i X_{i\alpha} [\omega_i(X) - \omega_{i\alpha}(X)] \\ &\quad + \lambda_i X_{i\alpha} \left[dU_{i\alpha} - \rho_i^{-1} \sum_{\beta}^i X_{i\beta} dU_{i\beta} \right]. \end{aligned} \quad (6.37)$$

Just as in the previous chapter, this last equation will constitute our stochastic version of the replicator dynamics, so we will devote the rest of this section to understand some of its general properties, and also to see how it compares to the other stochastic incarnations of the replicator dynamics that we have already encountered.

A first remark to be made concerns the noise coefficients σ_r : even though we have written them in a form that suggests they are constant, they need not be so – after all, the intensity of the noise on an edge might well depend on the edge loads $Y_r = \sum_{\alpha} P_{r\alpha} X_{\alpha}$. On that account, we will only assume that these coefficients are essentially bounded functions of the loads y . Nonetheless, in an effort to reduce notational clutter, we will not indicate this dependence explicitly; instead, we simply remark here that our results continue to hold if we replace σ_r with the worst-case scenario $\sigma_r \leftrightarrow \text{ess sup}_y \sigma_r(y)$.

Secondly, it is also important to compare (6.37) to the “aggregate shocks” version of Fudenberg and Harris (1992) and the “exponential learning” approach that we introduced in Chapters 4 and 5 (see also Mertikopoulos and Moustakas, 2009, 2010b). As we have already seen in the case of the former, one perturbs the replicator equation (6.17) by accounting for the (stochastic) interference of nature with the reproduction rates of the various species (Fu-

Stochastic perturbations in congestion models.

Comparison with other dynamics.

denberg and Harris (1992), Cabrales (2000), Imhof (2005), and Hofbauer and Imhof (2009)):

$$\begin{aligned} dX_{i\alpha} &= X_{i\alpha} [u_{i\alpha}(X) - u_i(X)] dt \\ &\quad - X_{i\alpha} \left[\sigma_{i\alpha}^2 X_{i\alpha} - \sum_{\beta}^i \sigma_{i\beta}^2 X_{i\beta} \right] dt \\ &\quad + X_{i\alpha} \left[\sigma_{i\alpha} dW_{i\alpha} - \sum_{\beta}^i \sigma_{i\beta} X_{i\beta} dW_{i\beta} \right], \end{aligned} \quad (6.38)$$

where $W = \sum_{i,\alpha} W_{i\alpha} e_{i\alpha}$ is a Wiener process that evolves in $\mathbb{R}^A \cong \prod_i \mathbb{R}^{A_i}$.

By comparison, in the exponential learning case it is assumed that the players of a Nash game adhere to the Boltzmann-type learning scheme (5.4) and employ a strategy with exponential sensitivity on the strategy's score. However, if the information that players receive is imperfect, the errors propagate to their learning curves and instead lead to the stochastic dynamics (5.47):

$$\begin{aligned} dX_{i\alpha} &= \lambda_i X_{i\alpha} [u_{i\alpha}(X) - u_i(X)] dt \\ &\quad + \frac{\lambda_i^2}{2} X_{i\alpha} \left[\sigma_{i\alpha}^2 (1 - 2X_{i\alpha}) - \sum_{\beta}^i \sigma_{i\beta}^2 X_{i\beta} (1 - 2X_{i\beta}) \right] dt \\ &\quad + \lambda_i X_{i\alpha} \left[\sigma_{i\alpha} dW_{i\alpha} - \sum_{\beta}^i \sigma_{i\beta} X_{i\beta} dW_{i\beta} \right]. \end{aligned} \quad (5.47)$$

In light of the above, it is easy to see that (6.37) admits a (unique) strong solution for any initial condition $X(0) = x \in \Delta$ by following the same reasoning as in Chapter 4. Nevertheless, there are two notable traits of (6.37) that set it apart from its other stochastic relatives. First off, the drift of (6.37) coincides with the deterministic replicator dynamics (6.19), whereas the drift coefficients of (6.38) and (5.47) do not. Secondly, the martingale processes U that appear in (6.37) are not uncorrelated components of some Wiener process (as is the case for both (6.38) and (5.47)): instead, depending on whether the paths $\alpha, \beta \in \mathcal{A}$ have edges in common or not, the processes U_α, U_β might be highly correlated or not at all.

Correlated noise processes.

To make this last observation more precise, recall that the Wiener differentials dW_r are orthogonal across the edges $r \in \mathcal{E}$ of the network: $dW_r \cdot dW_s = d[W_r, W_s] = \delta_{rs} dt$. In its turn, this implies that the stochastic differentials dU_α, dU_β satisfy:

$$\begin{aligned} dU_\alpha \cdot dU_\beta &= \left(\sum_r P_{r\alpha} \sigma_r dW_r \right) \cdot \left(\sum_s P_{s\beta} \sigma_s dW_s \right) \\ &= \sum_{r,s} P_{r\alpha} P_{s\beta} \sigma_r \sigma_s \delta_{rs} dt = \sum_{r \in \alpha\beta} \sigma_r^2 dt = \sigma_{\alpha\beta}^2 dt, \end{aligned} \quad (6.39)$$

where $\sigma_{\alpha\beta}^2 = \sum_r P_{r\alpha} P_{r\beta} \sigma_r^2$ gives the variance of the noise along the intersection $\alpha\beta \equiv \alpha \cap \beta$ of the paths $\alpha, \beta \in \mathcal{A}$ (note also that we used our notational conventions to avoid cumbersome expressions such as $\sigma_{i\alpha, j\beta}^2$). We thus see that the processes U_α and U_β are uncorrelated if and only if the paths $\alpha, \beta \in \mathcal{A}$ have no common edges; at the other extreme, we have:

$$(dU_\alpha)^2 = \sum_{r \in \alpha} \sigma_r^2 dt = \sigma_\alpha^2 dt \quad (6.40)$$

where $\sigma_\alpha^2 \equiv \sigma_{\alpha\alpha}^2 = \sum_r P_{r\alpha} \sigma_r^2$ measures the intensity of the noise on the route $\alpha \in \mathcal{A}$.⁷ These expressions will be key to our analysis, so we will make liberal use of them in the rest of this chapter.

⁷ This notation is also consistent with the common intersection notation: $\alpha\alpha \equiv \alpha \cap \alpha = \alpha$.

6.3.2 Stochastic Fluctuations and Rationality

As in the deterministic setting, our main tool will be the (rate-adjusted) relative entropy $H_q(x) = \sum_i \lambda_i^{-1} \sum_\alpha q_{i\alpha} \log(q_{i\alpha}/x_{i\alpha})$ which we will study with the help of the generator \mathcal{L} of the diffusion (6.37). To that end, recall that the infinitesimal generator \mathcal{L} of the Itô diffusion:

$$dX_\alpha(t) = \mu_\alpha(X(t)) dt + \sum_\beta \sigma_{\alpha\beta}(X(t)) dW_\beta(t), \quad (6.41)$$

is defined to be the second order differential operator:

$$\mathcal{L} = \sum_\alpha \mu_\alpha(x) \frac{\partial}{\partial x_\alpha} + \frac{1}{2} \sum_{\alpha,\beta} \left(\sigma(x) \sigma^T(x) \right)_{\alpha\beta} \frac{\partial^2}{\partial x_\alpha \partial x_\beta} \quad (6.42)$$

In this manner, if f is a sufficiently smooth function (namely, if f is C^2), we have seen that $\mathcal{L}f$ captures the drift of the process $f(X(t))$:

$$df(X(t)) = \mathcal{L}f(X(t)) dt + \sum_{\alpha,\beta} \frac{\partial f}{\partial x_\alpha} \Big|_{X(t)} \sigma_{\alpha\beta}(X(t)) dW_\beta(t). \quad (6.43)$$

Of course, in the case of the diffusion (6.37), the martingales U are not the components of a Wiener process, so (6.43) cannot be applied right off the shelf. However, a straightforward (if a bit cumbersome) application of Itô's lemma yields the following analogue of Lemma 5.11:

Some preliminary Itô calculations.

Lemma 6.13. *Let \mathcal{L} be the generator of (6.37). Then, for any $q \in \Delta$, we will have:*

$$\begin{aligned} \mathcal{L}H_q(x) &= -L_q(x) + \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \sum_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 (x_{i\beta} - q_{i\beta})(x_{i\gamma} - q_{i\gamma}) \\ &\quad + \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \sum_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}), \end{aligned} \quad (6.44)$$

where $L_q(x) = \sum_{i\alpha} (x_{i\alpha} - q_{i\alpha}) \omega_{i\alpha}(x)$ is the evolutionary index (5.12).

Proof. Let $V_q(t) = H_q(X(t))$. We then have:

$$\begin{aligned} dV_q &= \sum_{i,\alpha} \frac{\partial H_q}{\partial x_{i\alpha}} dX_{i\alpha} + \frac{1}{2} \sum_{i,\alpha} \sum_{j,\beta} \frac{\partial^2 H_q}{\partial x_{i\alpha} \partial x_{j\beta}} (dX_{i\alpha}) \cdot (dX_{j\beta}) \\ &= - \sum_{i,\alpha} \frac{1}{\lambda_i} \frac{q_{i\alpha}}{x_{i\alpha}} dX_{i\alpha} + \frac{1}{2} \sum_{i,\alpha} \frac{1}{\lambda_i} \frac{q_{i\alpha}}{x_{i\alpha}^2} (dX_{i\alpha})^2. \end{aligned} \quad (6.45)$$

However, with $X(t)$ being as in (6.37), we readily obtain:

$$\begin{aligned} (dX_{i\alpha})^2 &= \lambda_i^2 X_{i\alpha}^2 \left(dU_{i\alpha} - \rho_i^{-1} \sum_\beta^i X_{i\beta} dU_{i\beta} \right)^2 \\ &= \lambda_i^2 X_{i\alpha}^2 \left[(dU_{i\alpha})^2 - \frac{2}{\rho_i} \sum_\beta^i X_{i\beta} dU_{i\alpha} \cdot dU_{i\beta} + \frac{1}{\rho_i^2} \sum_{\beta,\gamma}^i X_{i\beta} X_{i\gamma} dU_{i\beta} \cdot dU_{i\gamma} \right] \\ &= \lambda_i^2 X_{i\alpha}^2 \left[\sigma_{i\alpha}^2 - \frac{2}{\rho_i} \sum_\beta^i \sigma_{\alpha\beta}^2 X_{i\beta} + \frac{1}{\rho_i^2} \sum_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 X_{i\beta} X_{i\gamma} \right] dt. \end{aligned} \quad (6.46)$$

As a result, we may combine the two equations (6.45) and (6.46) to obtain:

$$\begin{aligned} dV_q &= - \sum_{i,\alpha} q_{i\alpha} [\omega_i(X) - \omega_{i\alpha}(X)] dt - \sum_{i,\alpha} q_{i\alpha} \left[dU_{i\alpha} - \rho_i^{-1} \sum_\beta^i X_{i\beta} dU_{i\beta} \right] \\ &\quad + \frac{1}{2} \sum_{i,\alpha} \lambda_i q_{i\alpha} \left[\sigma_{i\alpha}^2 - \frac{2}{\rho_i} \sum_\beta^i \sigma_{\alpha\beta}^2 X_{i\beta} + \frac{1}{\rho_i^2} \sum_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 X_{i\beta} X_{i\gamma} \right] dt. \end{aligned} \quad (6.47)$$

Therefore, if we focus at a particular user $i \in \mathcal{N}$, the last term of (6.47) gives:

$$\begin{aligned}
& \sum_{\alpha} q_{i\alpha} \left[\sigma_{i\alpha}^2 - \frac{2}{\rho_i} \sum_{\beta} \sigma_{\alpha\beta}^2 X_{i\beta} + \frac{1}{\rho_i^2} \sum_{\beta,\gamma} \sigma_{\beta\gamma}^2 X_{i\beta} X_{i\gamma} \right] \\
&= \sum_{\alpha} q_{i\alpha} \sigma_{i\alpha}^2 - \frac{2}{\rho_i} \sum_{\alpha,\beta} \sigma_{\alpha\beta}^2 q_{i\alpha} X_{i\beta} + \frac{1}{\rho_i} \sum_{\beta,\gamma} \sigma_{\beta\gamma}^2 X_{i\beta} X_{i\gamma} \\
&= \sum_{\alpha} q_{i\alpha} \sigma_{i\alpha}^2 - \frac{1}{\rho_i} \sum_{\beta,\gamma} q_{i\beta} q_{i\gamma} \sigma_{\beta\gamma}^2 \\
&+ \frac{1}{\rho_i} \left[\sum_{\beta,\gamma} q_{i\beta} q_{i\gamma} \sigma_{\beta\gamma}^2 - 2 \sum_{\beta,\gamma} X_{i\beta} q_{i\gamma} \sigma_{\beta\gamma}^2 + \sum_{\beta,\gamma} X_{i\beta} X_{i\gamma} \sigma_{\beta\gamma}^2 \right] \\
&= \frac{1}{\rho_i} \left[\sum_{\beta,\gamma} q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}) \sigma_{\beta\gamma}^2 + \sum_{\beta,\gamma} \sigma_{\beta\gamma}^2 (X_{i\beta} - q_{i\beta})(X_{i\gamma} - q_{i\gamma}) \right] \quad (6.48)
\end{aligned}$$

and the lemma follows by substituting (6.48) into (6.47) and keeping only the resulting drift – that is, the first and third terms of (6.47). \square

In a certain sense, this lemma can be viewed as the stochastic analogue of Lemma 6.8 (which is recovered immediately if we set $\sigma = 0$). However, it also shows that the stochastic situation is much more intricate than the deterministic one. For example, if q is a Wardrop equilibrium, (6.44) gives:

$$\mathcal{L}H_q(q) = \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \sum_{\beta,\gamma} q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}) \sigma_{\beta\gamma}^2, \quad (6.49)$$

and if we focus on user $i \in \mathcal{N}$, we readily obtain:

$$\begin{aligned}
\sum_{\beta,\gamma} q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}) \sigma_{\beta\gamma}^2 &= \sum_{\beta,\gamma} q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}) \sum_r P_{r\beta}^i P_{r\gamma}^i \sigma_r^2 \\
&= \sum_r \sigma_r^2 \rho_i y_{ir} - \sum_r \sigma_r^2 y_{ir}^2 = \sum_r \sigma_r^2 y_{ir} (\rho_i - y_{ir}), \quad (6.50)
\end{aligned}$$

where $y_{ir} = \sum_{\alpha} P_{r\alpha}^i q_{i\alpha} \leq \rho_i$ is the load induced on edge $r \in \mathcal{E}$ by the i -th user. This shows that (6.49) is *positive* if at least one user mixes his routes, thus ruling out negative definiteness (even semi-definiteness) for $\mathcal{L}H_q$.

In view of the above, unconditional convergence to Wardrop equilibrium appears to be a “bridge too far” in our stochastic environment, especially when the equilibrium in question is not pure – after all, mixed equilibria are not even traps of (6.37). Instead, just as in the previous chapter, we obtain:

*Stability of
strict equilibria.*

Theorem 6.14. *Strict Wardrop equilibria are stochastically asymptotically stable in the replicator dynamics (6.37).*

Proof. By relabeling indices if necessary, assume that $q = \sum_i \rho_i e_{i,0}$ is the strict Wardrop equilibrium of \mathcal{C} . We will then show that H_q is a local stochastic Lyapunov function, i.e. that $\mathcal{L}H_q(x) \leq -kH_q(x)$ for some $k > 0$ and for all x sufficiently close to q ; then, our assertion will follow from Theorem 4 in [Gikhman and Skorokhod \(1971, pp. 314–315\)](#).

With this in mind, consider a perturbed flow $x = \sum_{i,\alpha} x_{i\alpha} e_{i\alpha}$ of the form:

$$x_{i,0} = \rho_i(1 - \varepsilon_i), \quad x_{i\mu} = \varepsilon_i \rho_i \xi_{i\mu} \text{ for } \mu = 1, 2, \dots \in \mathcal{A}_i^*, \quad (6.51)$$

where $\varepsilon_i > 0$ controls the L^1 distance between x_i and q_i , and $\xi_i \in \Delta(\mathcal{A}_i^*)$ is a point in the face of Δ_i which lies opposite to q (i.e. $\xi_{i\mu} \geq 0$ and $\sum_{\mu} \xi_{i\mu} = \rho_i$). Then, in view of Lemma 6.9, $L_q(x)$ will be bounded below by:

$$L_q(x) \geq \sum_i \rho_i \varepsilon_i \Delta\omega_i, \quad (6.52)$$

where $\Delta\omega_i = \min_{\mu} \{\omega_{i\mu}(q) - \omega_{i,0}(q)\} > 0$ (recall that q is strict).

Therefore, since the second term of $\mathcal{L}H_q(x)$ in (6.44) is clearly of order $\mathcal{O}(\varepsilon^2)$ (where $\varepsilon^2 = \sum_i \varepsilon_i^2$), we obtain:

$$\mathcal{L}H_q(x) \leq -\sum_i \rho_i \varepsilon_i \Delta \omega_i + \mathcal{O}(\varepsilon^2). \quad (6.53)$$

On the other hand, we also have:

$$H_q(x) = \sum_i \frac{\rho_i}{\lambda_i} \log \frac{\rho_i}{x_{i,0}} = -\sum_i \frac{\rho_i}{\lambda_i} \log(1 - \varepsilon_i) = \sum_i \frac{\rho_i}{\lambda_i} \varepsilon_i + \mathcal{O}(\varepsilon^2). \quad (6.54)$$

Thus, if we pick some positive $k < \min_i \{\lambda_i \Delta \omega_i\}$, a bit of algebra gives:

$$\mathcal{L}H_q(x) \leq -k \sum_i \frac{\rho_i}{\lambda_i} \varepsilon_i + \mathcal{O}(\varepsilon^2) = -kH_q(x) + \mathcal{O}(\varepsilon^2), \quad (6.55)$$

which shows that H_q is locally Lyapunov (in the sense of Definition 3.13). \square

In other words, Theorem 6.14 implies that trajectories which start sufficiently close to a strict equilibrium will remain in the vicinity of the equilibrium and will eventually converge to it with arbitrarily high probability. Nonetheless, this is a *local* result: if the users' initial traffic distribution is not close to a strict equilibrium itself, Theorem 6.14 does not apply; specifically, if $X(0)$ is an arbitrary initial condition in Δ , we cannot even tell if the trajectory $X(t)$ will ever come close to q .

To put this in more precise terms, we will follow the same approach as in Chapter 5 and we will measure distances in Δ with the L^1 -norm: $\|\sum_\alpha z_\alpha e_\alpha\|_1 = \sum_\alpha |z_\alpha|$. In this norm, Δ has a diameter of $2\sum_i \rho_i$, so let $K_\delta = \{x \in \Delta : \|x - q\|_1 \leq \delta\}$ be the corresponding compact neighbourhood of q ($\delta < 2\sum_i \rho_i$). Then, to see if $X(t)$ ever hits K_δ , we will examine the hitting time τ_δ :

$$\tau_\delta \equiv \tau_{K_\delta} = \inf\{t > 0 : X(t) \in K_\delta\} = \inf\{t > 0 : \|X(t) - q\|_1 \leq \delta\}. \quad (6.56)$$

Thereby, our chief concern is this: is the hitting time τ_δ finite with high probability? And if it is, is its expected value also finite?

To make our lives easier, let us consider the collective expressions:

$$\begin{aligned} \rho &= \sum_i \rho_i, & \Delta \omega &= \rho^{-1} \sum_i \rho_i \Delta \omega_i \\ \sigma^2 &= \sum_r \sigma_r^2, & \lambda &= \rho^{-1} \sum_i \rho_i \lambda_i. \end{aligned} \quad (6.57)$$

where $\Delta \omega_i = \min_{\mu \neq \alpha_i} \{\omega_{i\mu}(q) - \omega_{i\alpha_i}(q)\} > 0$ is the minimum delay difference between a user's equilibrium path α_i and his other choices. We then have:

Theorem 6.15. *Let $q = \sum_i \rho_i e_{i\alpha_i}$ be a strict Wardrop equilibrium of a congestion model \mathcal{C} , and assume that the users' learning rates satisfy the condition:*

Hitting times for Wardrop equilibria.

$$\lambda \sigma^2 < \Delta \omega. \quad (6.58)$$

Then, for any $\delta < 2\rho$ and any initial condition $X(0) = x \in \Delta$ with finite relative entropy $H_q(x) < \infty$, the hitting time τ_δ has finite mean:

$$\mathbf{E}_x[\tau_\delta] \leq \frac{2H_q(x)}{\Delta \omega} \frac{2\rho}{\delta(2\rho - \delta)}. \quad (6.59)$$

Proof. As in the case of Theorem 5.10, our proof hinges on the expression:

$$-\mathcal{L}H_q(x) = L_q(x) - \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 (x_{i\beta} - q_{i\gamma})(x_{i\gamma} - q_{i\gamma}), \quad (6.60)$$

where $q = \sum_i \rho_i e_{i,0}$ is the strict equilibrium in question. Therefore, write $x \in \Delta$ in the projective form $x = q + \theta z$, where $z = \sum_{i,\alpha} z_{i\alpha} e_{i\alpha} \in T_q^c \Delta$ is the ‘‘inward’’ direction:

$$z_{i,0} = -\rho_i, \quad \sum_{\mu}^i z_{i\mu} = \rho_i, z_{i\mu} \geq 0 \text{ for all } \mu = 1, 2, \dots \in \mathcal{A}_i^*. \quad (6.61)$$

Then, regarding the first term of (6.60), Lemma 6.9 readily yields:

$$\mathcal{L}_q(q + \theta z) \geq \Phi(q + \theta z) - \Phi(q) \geq \theta \sum_i \rho_i \Delta \omega_i \text{ for all } \theta \in [0, 1]. \quad (6.62)$$

Similarly, the second term of (6.60) becomes:

$$\begin{aligned} \sum_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 (x_{i\beta} - q_{i\beta})(x_{i\gamma} - q_{i\gamma}) &= \theta^2 \sum_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 z_{i\beta} z_{i\gamma} \\ &= \theta^2 \sum_r \sum_{\beta,\gamma}^i \sigma_r^2 P_{r\beta}^i P_{r\gamma}^i z_{i\beta} z_{i\gamma} = \theta^2 \sum_r \sigma_r^2 w_{ir}^2, \end{aligned} \quad (6.63)$$

where $w_i = P^i(z)$. Since $w_{ir} \leq \rho_i$ for all $r \in \mathcal{E}$, we obtain the inequality:

$$-\mathcal{L}H_q(x) \geq \theta \sum_i \rho_i \Delta \omega_i - \frac{1}{2} \sigma^2 \theta^2 \sum_i \lambda_i \rho_i = (\rho \Delta \omega) \theta - \frac{1}{2} \rho \lambda \sigma^2 \theta^2, \quad (6.64)$$

where ρ, σ^2 and $\lambda, \Delta \omega$ are the respective aggregates and averages of (6.57).

Suppose now that the rates λ_i satisfy (6.58), i.e. $\lambda \sigma^2 < \Delta \omega$. In that case, the RHS of (6.64) will be increasing for all $\theta \in [0, 1]$, and we will have:

$$-\mathcal{L}H_q(x) \geq \rho \left[\theta \Delta \omega - \frac{1}{2} \lambda \sigma^2 \theta^2 \right] > 0 \quad (6.65)$$

for all x with $\|x - q\|_1 \geq \theta \|z\|_1 = 2\theta \sum_i \rho_i = 2\rho\theta$. So, for $\|x - q\|_1 \geq \delta$, we get:

$$-\mathcal{L}H_q(x) \geq \frac{\delta}{2} \Delta \omega - \frac{1}{2} \frac{\delta^2}{4\rho^2} \sigma^2 \rho \lambda \geq \frac{\delta}{2} \Delta \omega \left(1 - \frac{\delta}{2\rho} \right) > 0, \quad (6.66)$$

and, therefore, if we denote by K_δ the compact L^1 neighborhood $K_\delta = \{x \in \Delta : \|x - q\|_1 \leq \delta\}$, we will have $\mathcal{L}H_q(x) \leq -\frac{\delta}{2} \Delta \omega \left(1 - \frac{\delta}{2\rho} \right) < 0$ for all $x \notin K_\delta$.

Consequently, the hitting time estimate of Durrett (1996, Theorem 5.3 in page 268), finally produces the bound (6.59):

$$\mathbf{E}_x[\tau_\delta] \leq \frac{2H_q(x)}{\Delta \omega(1 - \delta/2\rho)} \frac{1}{\delta}. \quad \square$$

Recall now that Theorem 6.14 ensures that a trajectory $X(t)$ which starts sufficiently close to a strict equilibrium q will converge to q with arbitrarily high probability. Thus, with Theorem 6.15 showing that $X(t)$ will come arbitrarily close to q in finite time, a tandem application of the two theorems gives:

Convergence to equilibrium.

Corollary 6.16. *If q is a strict equilibrium of \mathcal{C} and the players’ learning rates λ_i satisfy (6.58), the trajectories $X(t)$ converge to q almost surely.*

Of course, if a strict Wardrop equilibrium exists, then it is the unique equilibrium of the game (Proposition 6.7). In that case, Corollary 6.16 is the stochastic counterpart of Theorem 6.12 (and the congestion analogue of Corollary 5.12): if learning rates are soft enough compared to the noise level, then the solution orbits of the stochastic replicator dynamics (6.35) converge to a stationary traffic distribution almost surely. A few remarks are thus in order:

Remark 1 (The Critical Temperature). Just like its exponential learning sibling (5.52), the ‘‘slow learning’’ condition (6.58) shows that the replicator dynamics

reward patience: players who take their time in learning the game manage to average out the noise and converge to equilibrium. In fact, it is quite remarkable that, despite the very different stochastic dynamics, the critical learning temperature $T_c = \sigma^2 / \Delta\omega$ remains essentially the same in both cases: the aggregation over users that is inherent in (6.58) is just the result of the noise processes being correlated across paths.

Remark 2. Admittedly, the form of (6.59) is a bit opaque for practical purposes. To lighten it up, note that we are only interested in small δ , so the term $2\rho / (2\rho - \delta)$ may be ignored to leading order. Therefore, if we also assume for simplicity that all players learn at the same rate $\lambda_i = \lambda$, we get:

$$\mathbf{E}_x[\tau_\delta] \leq \frac{2h}{\lambda\Delta\omega} \frac{\rho}{\delta} \quad (6.67)$$

where $h = \rho^{-1} \sum_i \rho_i \log(\rho_i / x_i)$ is the “average” K-L distance between x and q .

This very rough estimate is pretty illuminating on its own – and its close kinship to the estimate (5.53) should not be ignored either. First and foremost, it shows that our bound for $\mathbf{E}_x[\tau_\delta]$ is inversely proportional to the learning rate λ , much the same as in the deterministic setting where λ essentially rescales time to λt . Moreover, because ρ and δ are both $\mathcal{O}(N)$, one might be tempted to think that our time estimates are *intensive*, i.e. independent of N . However, since delays increase (nonlinearly even) with the aggregate load ρ , the dependence on N is actually hidden in $\Delta\omega$.⁸

In any event, since strict equilibria do not always exist, we should return to the generic case of *interior* equilibria $q \in \text{Int}(\Delta)$. We have already seen that these equilibria are not very well-behaved in stochastic environments: they are not stationary in (6.37), and (6.50) shows that $\mathcal{L}H_q$ is actually *positive* in their vicinity. Despite all that, if the network \mathcal{Q} is irreducible and the users’ learning rates λ_i are slow enough, we will see that the replicator dynamics (6.37) admit a finite invariant measure which concentrates mass around the (necessarily) unique equilibrium of \mathcal{C} .

The trouble with interior equilibria.

To state this result precisely, recall that the projective distance $\Theta_q(x)$ of x from q is defined via the unique *projective representation* of x :

$$\Theta_q(x) = \theta \Leftrightarrow x = q + \theta z \text{ for some } z \in S_q \text{ and } 0 \leq \theta \leq 1, \quad (6.68)$$

where S_q is the “projective sphere” around q , i.e. $S_q = \{z \in T_q\Delta : q + z \in \text{bd}(\Delta)\}$. In a similar vein, we define the *essence* of $q \in \Delta$ to be:

$$\text{ess}(q) = \rho^{-1} \min \{ \|P(z)\| : z \in S_q \}, \quad (6.69)$$

where $\|\cdot\|$ denotes the ordinary Euclidean norm and the factor of ρ was included for scaling purposes. Comparably to $\text{red}(\mathcal{Q})$, $\text{ess}(q)$ measures redundancy (or rather, the lack thereof): $\text{ess}(q) = 0$ only if some direction $z \in S_q$ is null for P , i.e. only if \mathcal{Q} is reducible.

We are finally in a position to state and prove:

Theorem 6.17. *Let $q \in \text{Int}(\Delta)$ be an interior equilibrium of an irreducible congestion model \mathcal{C} , and assume that the users’ learning rates satisfy the condition:*

The invariant distribution.

$$\lambda < \frac{4}{5} \frac{m\rho\kappa^2}{\sigma^2}, \text{ where } m = \inf\{\phi'_r(y_r) : r \in \mathcal{E}, y \in P(\Delta)\} \text{ and } \kappa = \text{ess}(q). \quad (6.70)$$

⁸ This also shows that the learning rates λ_i do not have to be $\mathcal{O}(1/|\mathcal{E}|)$ -small in order to satisfy the slow-learning condition (6.58).

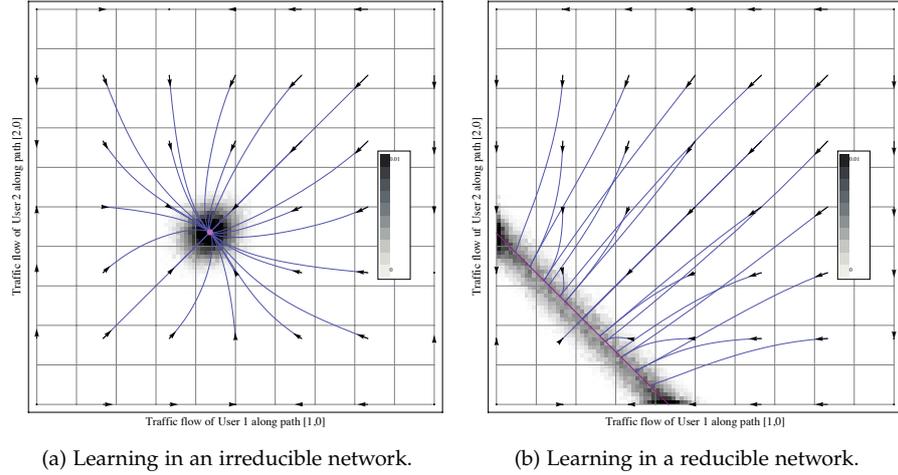


Figure 6.4: Learning in the networks of Figure 6.1 with M/M/1 latency functions $\phi_r(y_r) = (\mu_r - y_r)^{-1}$ and arbitrarily chosen capacities μ_r . The shades of gray represent the invariant distribution of (6.37) (obtained by numerical simulations) and the flow lines (blue) are the solution trajectories of (6.19) – in (b) they are actually projections because there is a third user as well.

Then, for any interior initial condition $X(0) = x \in \text{Int}(\Delta)$, the trajectories $X(t)$ are recurrent (a.s.) and their time averages are concentrated in a neighbourhood of q . Specifically, if $\Theta_q(\cdot)$ denotes the projective distance (5.34) from q , then:

$$\mathbf{E}_x \left[\frac{1}{t} \int_0^t \Theta_q^2(X(s)) ds \right] \leq \theta_\lambda^2 + \mathcal{O}(1/t), \text{ where } \theta_\lambda^2 = \frac{1}{4} \left(\frac{m\rho\kappa^2}{\lambda\sigma^2} - 1 \right)^{-1}. \quad (6.71)$$

Accordingly, the transition probabilities of $X(t)$ converge in total variation to an invariant probability measure π on Δ which concentrates mass around q . In particular, if $B_\theta = \{x \in \Delta : \Theta_q(x) \leq \theta\}$ is a “projective ball” around q , we have:

$$\pi(B_\theta) \geq 1 - \theta_\lambda^2 / \theta^2. \quad (6.72)$$

Proof. As we mentioned before, any $x \in \Delta$ may be expressed in the projective form $x = q + \theta z$, where $\theta = \Theta_q(x) \in [0, 1]$ is the projective distance of x from q , and $z \in S_q$ is a direction indicator. In this manner, (6.44) becomes:

$$\begin{aligned} -\mathcal{L}H_q(x) &= L_q(q + \theta z) - \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \theta^2 \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 z_{i\beta} z_{i\gamma} \\ &\quad - \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}). \end{aligned} \quad (6.73)$$

With regards to the first term of (6.73), Lemma 6.10 and the definition (6.69) of κ yield $L_q(q + \theta z) \geq \frac{1}{2} m \|P(z)\|^2 \theta^2 \geq \frac{1}{2} m \kappa^2 \rho^2 \theta^2$. Moreover, we have already seen in the proof of Theorem 6.15 that the second term of (6.73) is itself bounded above:

$$\frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \theta^2 \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 z_{i\beta} z_{i\gamma} \leq \frac{1}{2} \rho \lambda \sigma^2 \theta^2. \quad (6.74)$$

We are thus left to estimate the last term of (6.73). To that end, (6.50) gives:

$$\sum_{\beta, \gamma}^i q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}) \sigma_{\beta\gamma}^2 = \sum_r \sigma_r^2 y_{ir} (\rho_i - y_{ir}) \leq \frac{1}{4} \rho_i^2 \sigma^2, \quad (6.75)$$

the last inequality stemming from the trivial (and not particularly sharp) bound $y_{ir}(\rho_i - y_{ir}) \leq \frac{1}{4}\rho_i^2$ (recall that $0 \leq y_{ir} \leq \rho_i$).

Combining all of the above, we then get:

$$-\mathcal{L}H_q(x) \geq \frac{1}{2}m\kappa^2\rho^2\theta^2 - \frac{1}{2}\rho\lambda\sigma^2\theta^2 - \frac{1}{8}\rho\lambda\sigma^2 \equiv g(\theta), \quad 0 \leq \theta \leq 1. \quad (6.76)$$

As a result, if $\lambda < \frac{4}{5}\lambda_0$ where $\lambda_0 = \frac{m\rho\kappa^2}{\sigma^2}$, it is easy to see that the RHS of (6.76) will be increasing for $0 \leq \theta \leq 1$, and, moreover, it will also be positive for all θ with $\theta_\lambda < \theta \leq 1$, θ_λ being the positive root of the equation $g(\theta) = 0$: $\theta_\lambda = \frac{1}{2}(\lambda_0/\lambda - 1)^{-1/2}$.

So, pick some positive $a < g(1) = \frac{1}{2}\rho\sigma^2(\lambda_0 - \frac{5}{4}\lambda)$ and consider the set $K_a = \{q + \theta z : z \in S_q, g(\theta) \leq a\}$. By construction, K_a is a compact neighbourhood of q which does not intersect $\text{bd}(\Delta)$ and, by (6.76), we have $\mathcal{L}H_q(x) \leq -a$ outside K_a . Therefore, if $\tau_a \equiv \tau_{K_a}$ denotes the hitting time $\tau_a = \inf\{t : X(t) \in K_a\}$, the time estimate of Durrett (1996) yields:

$$\mathbf{E}_x[\tau_a] \leq \frac{H_q(x)}{a} < \infty \quad (6.77)$$

for every interior initial condition $X(0) = x \in \text{Int}(\Delta)$.

Just as in the proof of Theorem 5.13, let us now consider the transformed process $Y(t) = \Psi(X(t))$ given by $\Psi_{i\mu}(x) = \log x_{i\mu}/x_{i,0}$, $\mu \in \mathcal{A}_i^*$. Then, with $\frac{\partial \Psi_{i\mu}}{\partial x_{i\mu}} = 1/x_{i\mu}$ and $\frac{\partial \Psi_{i\mu}}{\partial x_{i,0}} = -1/x_{i,0}$, Itô's formula gives:

$$dY_{i\mu} = \mathcal{L}\Psi_{i\mu}(X) dt + dU_{i\mu} - dU_{i,0} = \mathcal{L}\Psi_{i\mu}(X) dt + \sum_r Q_{r\mu}^i \sigma_r dW_r, \quad (6.78)$$

where $Q_{r\mu}^i = P_{r\mu}^i - P_{r,0}^i$ are the components of the redundancy matrix Q of \mathcal{Q} in the basis $\tilde{e}_{i\mu} = e_{i\mu} - e_{i,0}$ of $T_q\Delta$ – recall also the expression (6.5).

We now claim that the generator of Y is elliptic. Indeed, if we drop the user index i for convenience and set $A_{\mu r} = Q_{r\mu}\sigma_r$, $\mu \in \coprod_i \mathcal{A}_i^*$, it suffices to show that the matrix AA^T is positive-definite. Sure enough, for any tangent vector $z \in Z$ expressed in the basis $\{\tilde{e}_\mu : \mu \in \coprod_i \mathcal{A}_i^*\}$ of Z as $z = \sum_\mu z_\mu \tilde{e}_\mu$, we get:

$$\langle Az, Az \rangle = \sum_{\mu, \nu} \left(AA^T \right)_{\mu\nu} z_\mu z_\nu = \sum_{\mu, \nu} \sum_r Q_{r\mu} Q_{r\nu} \sigma_r^2 z_\mu z_\nu = \sum_r \sigma_r^2 w_r^2, \quad (6.79)$$

where $w = Q(z)$. Since \mathcal{Q} is irreducible, we will have $w \neq 0$, and, in view of (6.79) above, this proves our assertion.

We have thus shown that the process $Y(t)$ hits a compact neighbourhood of $\Psi(q)$ in finite time (on average), and also that the generator of Y is elliptic. From the criteria of Bhattacharya (1978, Lemma 3.4) it follows that Y is recurrent, and since Ψ is invertible in $\text{Int}(\Delta)$, the same must hold for $X(t)$ as well. In a similar fashion, these criteria also ensure that the transition probabilities of the diffusion $X(t)$ converge in total variation to an invariant probability measure π on Δ , thus proving the first part of our theorem.

To obtain the estimate (6.71), note that Dynkin's formula applied to the quadratic growth estimate (6.76) yields:

$$\begin{aligned} \mathbf{E}_x [H_q(X(t))] &= H_q(x) + \mathbf{E}_x \left[\int_0^t \mathcal{L}H_q(X(s)) ds \right] \\ &\leq H_q(x) - \frac{1}{2}\rho\sigma^2(\lambda_0 - \lambda) \mathbf{E}_x \left[\int_0^t \Theta_q^2(X(s)) ds \right] + \frac{1}{8}\rho\lambda\sigma^2 t, \quad (6.80) \end{aligned}$$

*A hitting
time estimate.*

*Uniform
ellipticity.*

so, with $\mathbf{E}_x[H_q(X(t))] \geq 0$, we easily get:

$$\mathbf{E}_x \left[\frac{1}{t} \int_0^t \Theta_q^2(X(s)) ds \right] \leq \theta_\lambda^2 + \frac{C}{t}, \text{ where } C = \frac{2}{\rho\sigma^2} \frac{H_q(x)}{\lambda_0 - \lambda}. \quad (6.81)$$

The invariant distribution.

We are thus left to establish the bound $\pi(B_\theta) \geq 1 - \theta_\lambda^2/\theta^2$ which shows that the invariant measure π concentrates its mass around the “projective balls” B_θ . For that, we will use the ergodic property of $X(t)$, namely that:

$$\pi(B_\theta) = \lim_{t \rightarrow \infty} \mathbf{E}_x \left[\frac{1}{t} \int_0^t \chi_{B_\theta}(X(s)) ds \right], \quad (6.82)$$

where χ_{B_θ} is the indicator function of B_θ . However, with $\Theta_q^2(x)/\theta^2 \geq 1$ outside B_θ by definition, it easily follows that:

$$\mathbf{E}_x \left[\frac{1}{t} \int_0^t \chi_{B_\theta}(X(s)) ds \right] \geq \mathbf{E}_x \left[\frac{1}{t} \int_0^t \left(1 - \Theta_q^2(X(s))/\theta^2 \right) ds \right] \quad (6.83)$$

and the bound (6.72) follows by letting $t \rightarrow \infty$ in (6.81). \square

We conclude this section with a few remarks on our results so far:

Remark 1 (Comparison to Previous Results). Needless to say, the close connection between this theorem and Theorem 5.13 for potential games is quite pronounced. In fact, this intimate link is precisely what prompted our deliberately similar formulations, serving to facilitate comparisons between the two and the evolutionary results of Imhof (2005).

Some of the differences are not of great import; for example, Theorem 5.13 is stated in terms of the usual Euclidean distance, while Theorem 6.17 employs the projective “distance” $\Theta_q(x)$. By a slight tweak in our arguments, it is not too hard to rederive our results in terms of the same distance measure; that we did not do so simply reflects our desire to illustrate as wide a class of estimates as possible.

There are, however, some more significant differences, owing both to the different drift terms of the dynamics (5.47) and (6.37), but also the correlation in the martingale part of (6.37). The difference in the drift can be seen by comparing the two expressions (5.54) and (6.44) for $\mathcal{L}H_q(x)$: the quadratic term of (6.44) is antagonistic to the evolutionary index L_q , thus explaining why the critical temperature of exponential learning is lower than that of the replicator equation (6.37) – to see this, simply compare the slow-learning conditions (5.76) and (6.70) in an irreducible network with uncorrelated parallel links and, mutatis mutandis, replace $m\kappa^2$ with the minimum eigenvalue of the Hessian of the Rosenthal potential.

The effect of noise correlation across intersecting paths has equally far-reaching consequences because they are now summed over all edges that make up a path, invariably leading to the aggregate noise coefficient $\sigma^2 = \sum_r \sigma_r^2$ which is $\mathcal{O}(|\mathcal{E}|)$ times greater than the noise bounds which appear in (5.61) – see also the following remark.

The role of irreducibility.

Remark 2 (The Effect of Redundancy). On a related note, the irreducibility assumption turns out to play quite the vital part: it appears both in the “slow-learning” condition (6.70) (recall that $\text{ess}(q) = 0$ if q is an interior point of a reducible network) and also in the proof that the generator of $Y(t) = \Psi(X(t))$ is elliptic. This shows that the stochastic dynamics (6.37) are not oblivious to redundant degrees of freedom, in stark contrast with the deterministic case (that is, Theorem 6.12).

Regardless, we expect that an analogue for Theorem 6.17 still holds for reducible networks if we replace q with the entire (affine) set Δ^* . More precisely, we conjecture that under a suitably modified learning condition, the transition probabilities of $X(t)$ converge to an invariant distribution which concentrates mass around Δ^* (see Fig. 6.4b). One way to prove this claim would be to find a suitable way to “quotient out” $\ker Q$ but, since the replicator equation (6.37) is not invariant over the redundant fibres $x + \ker Q, x \in \Delta$, we have not yet been able to do so.

Remark 3 (Convergence Rate vs. Efficiency). As in Theorem 5.13, the invariant measure π converges vaguely to a point mass at q when $\lambda \rightarrow 0$. This means that if the learning temperature $T = 1/\lambda$ is high enough, then we recover Theorem 6.12, with the usual downside of using very slow learning rates: the expected time to come close to an equilibrium becomes inversely proportional to λ , so the users of the network will have to trade off speed of convergence for efficiency (measured in terms of proximity to equilibrium).

Remark 4 (Sharpness). We should also note here that the bounds we obtained are not the sharpest possible ones. For example, the learning condition (6.70) can be tightened by quite a bit (however, not as much as (5.76)), and the assumption that $\phi_r' > 0$ can actually be dropped. In that case however, the corresponding expressions would become significantly more complicated without adding much essence, so we have opted to keep our analysis focused on the simpler estimates.

6.4 SOME LOOSE ENDS

In this last section, we will attempt to touch on some loose ends that we have not been able to thoroughly address in the rest of this chapter. Truth be told, much of this section is open-ended and might well be taken as a roadmap for future directions to be explored.

LEARNING AND OPTIMALITY We have already noted that the traffic flows which minimize the aggregate latency $\omega(x) = \sum_i \rho_i \omega_i(x)$ in a network correspond precisely to the Wardrop equilibria of a congestion model which is defined over the same network and whose delay functions are given by the “marginal latencies” $\phi_r^*(y_r) = \phi_r(y_r) + y_r \phi_r'(y_r)$ (see e.g. Roughgarden and Tardos, 2002). Hence, if we set $\omega_{i\alpha}^*(x) = \sum_\alpha P_{r\alpha}^i \phi_r^*(y_r)$ and substitute $\omega_{i\alpha}^*$ instead of $\omega_{i\alpha}$ in the replicator dynamics (6.19) and (6.37), our analysis yields:

Convergence to the optimal distribution.

Theorem 6.18. *Let $\mathfrak{C} \equiv \mathfrak{C}(\mathcal{Q}, \phi)$ be a congestion model with strictly convex latency functions $\phi_r, r \in \mathcal{E}$, and assume that users follow a replicator learning scheme with cost functions $\omega_{i\alpha}^*$. Then:*

1. *In the deterministic case (6.19), players converge to a traffic flow which minimises the aggregate delay $\omega(x) = \sum_i \rho_i \omega_i(x)$.*
2. *In the stochastic case (6.37), if the network is irreducible and the players’ learning rates are slow enough, their time-averaged flows will be concentrated near the (necessarily unique) optimal distribution q which minimises ω .*

Of course, for a more precise statement one need only reformulate Theorems 6.12, 6.14, and 6.17 accordingly (the convexity of ϕ_r^* replaces the monotonicity requirement for ϕ_r). The only thing worthy of note here is that the marginal costs $\phi_r^*(y_r)$ do not really constitute “local information” that users can acquire

simply by routing their traffic and recording the delays that they experience. However, the missing components $y_r \phi'_r(y_r)$ can easily be measured by observers monitoring the edges of the network and could be subsequently publicized to all users that employ the edge $r \in \mathcal{E}$. Consequently, if the administrators of a network wish users to figure out the optimal traffic allocation on their own, they simply have to go the (small) extra distance of providing such monitors on each of the network's links.

*Exponential learning
in congestion models.*

EXPONENTIAL LEARNING In the context of Chapters 4 and 5, we have already seen that the replicator dynamics also arise as the result of an exponential learning process, itself a variant of logistic fictitious play (Mertikopoulos and Moustakas, 2010b). As such, it is not too hard to adapt this method directly to our congestion setting.

In more detail, assume that all users $i \in \mathcal{N}$ keep performance scores $V_{i\alpha}$ of the paths at their disposal as specified by the differential equation:

$$dV_{i\alpha}(t) = -\omega_{i\alpha}(X) dt + \sigma_{i\alpha} dU_{i\alpha}, \quad (6.84)$$

where, as in (6.36), $dU_{i\alpha}$ describes the total noise along the path $\alpha \in \mathcal{A}_i$, and $X(t)$ is the traffic profile at time t , defined via the Boltzmann distribution:

$$X_{i\alpha}(t) = \frac{\exp(\lambda_i V_{i\alpha}(t))}{\sum_{\beta}^i \exp(\lambda_i V_{i\beta}(t))}, \quad (6.85)$$

In this manner, Itô's lemma now gives:

$$\begin{aligned} dX_{i\alpha} = & \lambda_i X_{i\alpha} (\omega_i(X) - \omega_{i\alpha}(X)) dt + \lambda_i X_{i\alpha} \left(dU_{i\alpha} - \frac{1}{\rho_i} \sum_{\beta}^i X_{i\beta} dU_{i\beta} \right) \\ & + \frac{\lambda_i^2}{2} X_{i\alpha} \left(\sum_{\beta}^i (\rho_i \delta_{\alpha\beta} - 2X_{i\beta}) \sigma_{\alpha\beta}^2 - \frac{1}{\rho_i} \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 X_{i\gamma} (\rho_i \delta_{\beta\gamma} - 2X_{i\beta}) \right) dt. \end{aligned} \quad (6.86)$$

As far as the rationality properties of these new dynamics are concerned, a simple modification in the proof of Theorem 6.14 suffices to show that strict Wardrop equilibria are stochastically stable in (6.86). Just the same, the extra drift term in (6.86) complicates things considerably, so results containing explicit estimates of hitting times are significantly harder to obtain (as are the corresponding "slow-learning" conditions). Of course, this approach might well lead to improved convergence rates (extending those that we obtained in Chapter 5 to correlated perturbations), but since these calculations would take us too far afield, we prefer to postpone this analysis for the future.

*(Un)Interesting
equilibria.*

EQUILIBRIUM CLASSES In a certain sense, interior and strict equilibria represent the extreme ends of the Wardrop spectrum, so it was a reasonable choice to focus our analysis on them. Nevertheless, there are equilibrium classes that we did not consider: for instance, there are pure Wardrop equilibria which are not strict, or there could be "quasi-strict" equilibria q in the boundary of Δ with the property that $\omega_{i\alpha}(q) > \omega_i(q)$ for all α which are not present in q .

Strictly speaking, such equilibria are not covered by either Theorem 6.14 or Theorem 6.17. Still, by a suitable modification of our stochastic calculations, we may obtain similar convergence and stability results for these types of equilibria as well. For example, modulo a "slow-learning" condition similar to (6.70), it is easy to see that pure equilibria that are not strict are still stochastically stable. The reason we have opted not to consider all these special

cases is that it would be too much trouble for little gain: the assortment of similar-looking results that we would obtain in this way would confuse things more than it would clarify them.

THE BROWN-VON NEUMANN-NASH DYNAMICS Another powerful learning scheme is given by the Brown-von Neumann-Nash (BNN) dynamics (see e.g. [Fudenberg and Levine, 1998](#)) where users look at the “excess delays”

$$\psi_{i\alpha}(x) = [\omega_i(x) - \omega_{i\alpha}(x)]^+ = \max\{\omega_i(x) - \omega_{i\alpha}(x), 0\} \quad (6.87)$$

and update their traffic flows according to the differential equation:

$$\frac{dx_{i\alpha}}{dt} = \psi_{i\alpha}(x(t)) - \psi_i(x(t)), \quad (6.88)$$

where $\psi_i(x) = \rho_i^{-1} \sum_{\alpha} x_{i\alpha} \psi_{i\alpha}(x)$. On the negative side, these dynamics require users to monitor delays even along paths that they do not employ. On the other hand, they satisfy the pleasant property of “non-complacency” ([Sandholm, 2001](#)): the stationary states of (6.88) coincide with the game’s Wardrop equilibria and every solution trajectory converges to a connected set of such equilibria.

In terms of convergence to a Wardrop equilibrium, Theorem 6.12 shows that the replicator dynamics behave at least as well as the BNN dynamics (except perhaps on the boundary of Δ), so there is no real reason to pick the more complicated expressions (6.87), (6.88). However, this might not be true in the presence of stochastic fluctuations: in fact, virtually nothing is known about the behavior of the BNN dynamics in stochastic environments, and this question alone makes pursuing this direction a worthwhile project.

7

APPLICATIONS TO WIRELESS NETWORKS

As a result of the massive deployment of IEEE 802.11 wireless networks, and in the presence of already existing large-scale mobile third-generation systems, mobile users often have several choices of overlapping networks to connect to. In fact, mobile devices that support multiple standards already exist and, additionally, significant progress has been made towards creating flexible radio devices capable of connecting to *any* existing standard (Demestichas et al., 2004). It is thus reasonable to expect that in the near future, users will be able to connect to different networks and to switch dynamically between them.

In such a setting, the users' selfishness quickly become apparent: even though users have several choices to connect to, they still have to compete against each other for the finite resources of the combined network. Hence, this situation can be modelled by a non-cooperative game, a practice that is rapidly becoming one of the main tools in the analysis of wireless networks.¹

The scenario that we will analyze in this chapter consists of an unregulated wireless network where N heterogeneous users (e.g. mobile devices) can connect wirelessly to one of B nodes (possibly employing different standards). The users' (inherently selfish) objective is to maximize their individual downlink throughput, but, being heterogeneous, they differ vastly in their approach to obtain this goal. For example, users may have different tolerance for delay, or may wish to employ different "betting" schemes to download data at the lowest price. Thus, in general, users will have a different set of strategies to pick from, fixed at the outset of the game, and unknown to all other users.

Now, given that the users must share the nodes' resources, their objectives will lead to non-trivial competition: lacking a coordinating entity, the users seem to have no hope of reaching an organized, equilibrium state. One way to overcome this hurdle is if users base their decisions on a broadcast signal: for instance, all (or a subset) of the wireless nodes could be synchronously broadcasting a random finite length integer that it is received by all users.² Then, one might hope that, as the game is played again and again, sophisticated users develop an insight into how other users respond to the same stimulus and, eventually, learn to coordinate their actions.

From a game-theoretic standpoint, this was precisely the seminal idea behind Aumann's work on correlated games in (Aumann, 1974, 1987): players simply base their decisions on their observations of the "states of the world".

*Dynamic switches
in wireless networks.*

¹ For example, game-theoretic techniques have been used to optimize transmission probabilities (MacKenzie and Wicker, 2001), and also to calculate the optimal power allocation (Bonneau et al., 2005, 2007; Meshkati et al., 2006, 2005; Palomar et al., 2003) or even the optimal transmitting carrier (Meshkati et al., 2006). On a similar note, Chandra et al. (2004) and Shakkottai et al. (2007) considered the possibility of connecting to several access points using a single WLAN card, while the competition between service providers has been analyzed by Halldorsson et al. (2004), Zemlianov and de Veciana (2005) and Felegyhazi et al. (2007).

² This could be imposed by any public-minded authority, such as the Federal Communications Commission (FCC) in the US.

The El Farol bar and correlated equilibria.

Similar games have also been studied in econophysics, particularly after the introduction of the *El Farol* problem by Brian Arthur (1994) and the development of the *minority game* by Challet and Zhang (1997). In both these games, players “buy” or “sell” and are rewarded when they land in the minority. Again, the key idea is that in order to decide what to do, players observe and process the game’s history with the aid of some predetermined strategies. So, by employing more often the strategies that perform better, users quickly converge to a steady state that discourages unilateral deviation. And, in a rather surprising twist, it turns out that this state is actually oblivious to the source of the players’ observations (Cavagna, 1999); in fact, it was shown in Marsili et al. (2000) that what matters is simply the *amount* of feedback that players are asked to monitor and the *number* of strategies that they use to process it.

OUTLINE It should be clear that our wireless scenario stands to gain a lot in terms of coordination by such an approach. Hence, our main goal in this chapter will be to expound this scheme in a way appropriate for selfish users in an unregulated wireless network.

As a first step towards this, we generalize and adapt the minority game of Marsili et al. (2000) to our setting: this is done in Section 7.1 where we introduce the *Simplex Game*. Then, in Section 7.2, we proceed to describe the game’s equilibria and compare them to the socially optimal states. From this comparison emerges the game’s (correlated) *price of anarchy*, a fundamental notion that was first described in Koutsoupias and Papadimitriou (1999) and which measures the distance between anarchy (equilibria) and efficiency (optimal states).

Our first important result is then obtained in Section 7.3: by iterating the game based on an exponential learning scheme, we find that players quickly converge to an evolutionarily stable equilibrium that minimizes their *frustration level* (Theorem 7.13). So, having established convergence, we proceed in Section 7.4 to actually harvest the game’s price of anarchy. Quite unexpectedly, we find that *the price of anarchy is unaffected by disparities in the nodes’ characteristics* (Theorem 7.16). Moreover, we also derive an analytic expression for the price of anarchy, based on the method of *replica analysis*, borrowed from statistical physics. This expression then allows us to study the effect of the various parameters of the network on its performance, an analysis which we supplement by numerical results.

Some calculational details that would detract one’s focus from the main discussion have been deferred to the end of the chapter.

The mazes of replica theory.

A WORD OF WARNING At this point, the reader should be warned that this chapter marks a departure in both style and content from the previous ones: gone are the analytic growth estimates of differential operators and the firm ground of stochastic analysis on which they stood. These are now replaced by the bewildering mazes of replica theory whose corridors physicists have been all too happy to navigate by clairvoyant senses, but which mathematicians have only recently begun to chart and survey.

This is also the source of an ironic (and, perhaps, iconic) incongruity: this most intuitive chapter is also the one with the most “engineering” content. On the face of it, this irony is not easily felt: after all, who would disagree that the engineering genius of an Archimedes, a Fourier or a Tesla was guided by anything *but* intuition? Nevertheless, the engineering field of today, mired as it is by unnavigable swamps of technical details and obscure acronyms,

seems to be plagued by “rigoritis”, that condition of the scientific mind which puts overwhelming emphasis on the merest trifles of notational consistency (extending even to the level of typography), leaving far too little room for what mathematicians call “Rigor True”.

I sincerely wish that I were able to give precise meaning to some of the replica-theoretic notions that abound in this chapter (or, equivalently, to have been able to truly understand them). However, as can be seen by the plethora of papers by Guerra, Talagrand and other modern-day giants of probability, this endeavor would require much more than a dissertation in itself. So, in ending this cautionary note, we should stress that this chapter must be taken for what it is: a collection of conjectures and the sketchy calculations that support them, waiting patiently (and unapologetically) for the firm foundations of spin-glass theory to be laid down.

7.1 THE SIMPLEX GAME

As mentioned in the introduction, our scenario consists of an unregulated wireless network where a large number N of users seek to connect wirelessly to B nodes, possibly using different wireless standards. To describe this setup, we will assign to every node $r = 1, \dots, B$ a single user spectral efficiency c_r (which may well depend on r when different nodes employ different standards). In this case, we will model the throughput of a user connected to node r by:

$$T_r = c_r / N_r \quad (7.1)$$

where N_r is the number of users connected to node r (for simplicity we assume that all users have the same transmission characteristics).

Despite the simplicity of this throughput model, it has been shown to be of the correct form for TCP and UDP protocols in IEEE 802.11b systems (Shakkottai et al., 2007), if we limit ourselves to a single class of users. Furthermore, in the case of third-generation best-effort systems, the realistic total cell-service throughput is approximately constant beyond a certain number of connected users Kogiantis et al. (2001). Thus, it is a reasonable approximation to assume that the user throughput behaves as (7.1) for single-class mobiles.

This model for the throughput is flexible enough to account for parameters that affect users’ bias towards a node; e.g. we can incorporate pricing by modifying c_r to $c_r(1 - p_r)$ where p_r reflects the price per bit. In this way, we may renormalize (7.1) to:

$$\hat{u}_r = y_r / N_r, \quad (7.2)$$

where the (non-negative) coefficients y_r are normalized to unity ($\sum_{r=1}^B y_r = 1$) and represent the “effective strength” of the node r as a function of its attributes and characteristics; in other words, if $\mathcal{B} = \{1, \dots, B\}$ is a set of B nodes, then every “strength” distribution on \mathcal{B} may be represented by a point in the simplex $\Delta_B \equiv \Delta(\mathcal{B})$. The nodes can freely manipulate the relative value of their node strength, in order to maximize their gain. However, this is assumed to take place at slower time-scales and therefore these numbers will be assumed fixed throughout our analysis.³

We may now note that the core constituents of a (finite) congestion game are all present: N players (users) are asked to choose one of B facilities (nodes), their

³ Obviously, nodes of zero strength (e.g. negligible spectral efficiency) will not appeal to any reasonable user and can be dropped from the analysis.

payoff given by the throughput of (7.2). Of course, from this game-theoretic standpoint, the “fairest” user distribution is the Nash allocation of $y_r N$ users to node r : when distributed in this way, users receive a payoff of $\hat{u}_0 = 1$ and no one could hope to earn more by a unilateral deviation to another node.⁴ As a result, the users’ discomfort can be gauged by contrasting their payoff to the Nash value, a comparison which yields:

$$\hat{u}_r - \hat{u}_0 = \frac{y_r N}{y_r N + N_r - y_r N} - 1 = \frac{y_r N - N_r}{y_r N} + \mathcal{O}(1/N). \quad (7.3)$$

We will thus focus on the leading term of (7.3) and introduce the *linearized payoffs*:

$$u_r = 1 - \frac{N_r}{y_r N}. \quad (7.4)$$

It can easily be seen that the Nash equilibria of the game remain invariant under this linearization; hence, the two payoffs (7.2) and (7.4) will be equivalent in terms of social fairness.⁵

Weighted
simplices.

More importantly for our purposes, this approximation allows us to express a user’s payoff in a particularly elegant and revealing form. To that end, we first introduce a collection of B vectors in \mathbb{R}^{B-1} (and not \mathbb{R}^B) that we will use to model the nodes:

Definition 7.1. Let $\mathcal{B} = \{1, \dots, B\}$ be a set of B nodes, and let $y = \sum_r y_r e_r \in \text{Int}(\Delta_{\mathcal{B}})$ be a strength distribution on \mathcal{B} .⁶ A y -simplex (or y -appropriate simplex) will be a collection of B vectors $q_r \in \mathbb{R}^{B-1}$, $r = 1, \dots, B$, such that:

$$q_r \cdot q_l = -1 + \frac{\delta_{rl}}{\sqrt{y_r y_l}} \quad \text{for all } r, l = 1, \dots, B. \quad (7.5)$$

Remark. Extending our notational conventions, we will identify the set of nodes \mathcal{B} with any y -appropriate simplex on \mathcal{B} once a strength distribution $y \in \Delta_{\mathcal{B}}$ has been prescribed – this is also the reason behind the somewhat sloppy choice of notation $\Delta_{\mathcal{B}} \equiv \Delta(\mathcal{B})$.

Though running a bit ahead of ourselves, it is important to note here the geometric property of y -simplices which lies beneath Definition 7.1:

Lemma 7.2. Let $\mathcal{B} = \{q_r\}_{r=1}^B$ be a y -appropriate simplex for B nodes. Then:

$$\sum_r y_r q_r = 0 \quad \text{and} \quad \sum_r y_r q_r^2 = B - 1. \quad (7.6)$$

Proof. To establish the first part, note that:

$$\left(\sum_r y_r q_r \right)^2 = \sum_{r,l} y_r y_l q_r \cdot q_l = \sum_{r,l} y_r y_l (-1 + \delta_{rl} / \sqrt{y_r y_l}) = 0, \quad (7.7)$$

and the second part follows by a similar calculation. \square

Properties of
 y -simplices.

The intuitive idea behind this lemma should be pretty clear: a y -simplex behaves just like a standard simplex whose vertices have been “weighted” by the strengths y_r . This analogy is a perfect fit for the linearized expression

⁴ This is a simple version of “water-filling” for selfish resource allocation Lai and el Gamal (2006); it is also interesting to note that this allocation is Pareto-efficient.

⁵ This is also verified by numerical experiments in figure 6.4.

⁶ To clear up any confusion: $\text{Int}(\Delta_{\mathcal{B}}) = \{y \in \mathbb{R}^B : y_r > 0 \text{ and } \sum_r y_r = 1\}$.

(7.4); indeed, if N_r players choose $q_r \in \mathcal{B}$, we may consider the aggregate $q = \sum_{l=1}^B N_l q_l$ and obtain by (7.5):

$$q_r \cdot q = \sum_l N_l q_r \cdot q_l = -N \left(1 - \frac{N_r}{y_r N} \right). \tag{7.8}$$

We then get the very useful expression for the payoffs (7.4):

$$u_r = 1 - \frac{N_r}{y_r N} = -\frac{1}{N} q_r \cdot q = -\frac{1}{N} q_r \cdot \sum_{i=1}^N q_{r_i} \tag{7.9}$$

where r_i indicates the choice of player i . Thus, by Lemma 7.2, we see that the (pure) Nash equilibria of the game will be characterized by:

$$q = \sum_i q_{r_i} = \sum_r y_r N q_r = 0 \tag{7.10}$$

i.e. the game will be at equilibrium when the players' choices balance out the weights y_r at the vertices of the simplex.

On the other hand, it is not immediately obvious whether such y -appropriate simplices exist for a given strength distribution $y \in \text{Int}(\Delta_B)$. This is guaranteed by the following lemma:

Lemma 7.3. *There exists a y -simplex $\mathcal{B} = \{q_r\}_{r=1}^B \subseteq \mathbb{R}^{B-1}$ for any $y \in \text{Int}(\Delta_B)$.*

Proof. Begin by selecting a vector $q_1 \in \mathbb{R}^B$ such that $q_1^2 = \frac{1}{y_1} - 1$ and choose $q_{r+1} \in \mathbb{R}^B$ inductively so that it satisfies (7.5) when multiplied by q_1, \dots, q_r – such a selection is always possible for $r \leq B - 1$ thanks to the dimension of \mathbb{R}^B : for a vector space of lesser dimension, this is no longer the case.⁷

In this way, we obtain B vectors $q_r \in \mathbb{R}^B$ that satisfy (7.5); our construction will be complete once we show that \mathcal{B} is contained in some $(B - 1)$ -subspace of \mathbb{R}^B . However, as in the proof of Lemma 7.2, we can see that $\sum_{r=1}^B y_r q_r = 0$; this means that \mathcal{B} is linearly dependent and completes our proof. \square

We will finish our discussion of y -simplices by presenting the following key property which plays a crucial role in the calculations of Sections A.1 and A.2:

Lemma 7.4. *Let $\mathcal{B} = \{q_r\}_{r=1}^B \subseteq \mathbb{R}^{B-1}$ be a y -simplex for some $y \in \text{Int}(\Delta_B)$. Then:*

$$\sum_r y_r (q_r \cdot x)^2 = x^2 \quad \text{for all } x \in \mathbb{R}^{B-1}. \tag{7.11}$$

Proof. Since $y \in \text{Int}(\Delta_B)$, \mathcal{B} will span \mathbb{R}^{B-1} and x may be written as a linear combination $x = \sum_{r=1}^B x_r q_r$. Thus, if we let $S = \sum_{r=1}^B x_r$ and recall that $\sum_{r=1}^B y_r = 1$, we will have: $x^2 = \sum_{l,r=1}^B x_l x_r q_l \cdot q_r = -S^2 + \sum_{r=1}^B x_r^2 / y_r$. Similarly, $(q_r \cdot x)^2 = S^2 - 2S \frac{x_r}{y_r} + \frac{x_r^2}{y_r^2}$, and an addition over r yields (7.11). \square

Unfortunately, despite these geometric characterizations, it remains unclear how the Nash allocation can be achieved in an unregulated network that lacks a central managing authority. For this reason, we will introduce a coordination mechanism akin to the one proposed by Aumann (1974, 1987). In a nutshell, Aumann's scheme is that players observe the random events γ that transpire in some sample space Γ (the "states of the world") and then place their bets based on these observations. In other words, players' decisions are ordained

⁷ Note that $\frac{1}{y_r y_l} \geq \frac{1}{y_r} + \frac{1}{y_l}$ for all r, l so that the necessary and sufficient condition $q_r^2 \cdot q_l^2 \geq (q_r \cdot q_l)^2$ holds for any B -tuple of vectors which satisfy (7.5).

by their (correlated) *strategies*, that is, functions f_i on Γ that convert events $\gamma \in \Gamma$ to betting suggestions $f_i(\gamma)$.⁸

Broadcasts and
correlated strategies.

Following [Aumann \(1987\)](#) and [Marsili et al. \(2000\)](#), we propose that a broadcast beacon transmit a *training signal* m , drawn from some (discrete) sample space \mathcal{M} . For example, the nodes could be synchronously broadcasting the same integer $m \in \{1, \dots, M\}$, drawn from a uniform random sequence that is arbitrated by a (socially minded) government agency such as the FCC in the US. The purpose of this training beacon is to help users converge faster to a socially optimal state, because it would be to the advantage of the service providers if users settle to an “organized” state with low hopping rates and lower overheads. Thus, to process this signal, user i has at his disposal S \mathcal{B} -valued random variables $c_{is} : \mathcal{M} \rightarrow \mathcal{B}$, $s = 1, \dots, S$: these are the i -th user’s *strategies*, used to convert the signal $m \in \mathcal{M}$ to a betting suggestion $c_{is}(m) \equiv c_{is}^m \in \mathcal{B}$.⁹ So, if user i selects strategy s_i , the collection of maps $c_{is_i} \in \mathcal{B}^{\mathcal{M}}$, $i \in \mathcal{N}$ will be a *correlated strategy* in the sense of [Aumann \(1987\)](#) (contrast with the f_i above and our discussion in Chapter 2).

However, unlike [Aumann \(1987\)](#), we cannot assume that users develop their strategies after careful contemplation on the “states of the world”. After all, it is quite unlikely that a user will have much time to think in the fast-paced realm of wireless networks. Consequently, we envision that when the game begins, each user randomly “preprograms” S strategies, drawn randomly from the set $\mathcal{B}^{\mathcal{M}}$ of all possible maps $\mathcal{M} \rightarrow \mathcal{B}$. Of course, since we assume users to be *heterogeneous*, they will program their strategies in wildly different ways and independently of one another. Still, rational users will exhibit a predisposition towards stronger nodes, and, to account for this, we posit that:

$$\mathbf{P}(c_{is}^m = q_r) = y_r \quad (7.12)$$

i.e. the probability that user i programs node q_r as response to the signal m is just the node’s strength y_r .¹⁰ In other words, the strategies are picked in *anticipation* of the presence of other users; if each user were expecting to play alone, he would have picked strategies that would make him bet on the strongest node.

At this point, it is probably best to collect all of the above in a formal definition:

The Simplex Game.

Definition 7.5. Let $y \in \text{Int}(\Delta_B)$ be a strength distribution for B nodes. Then, a *y-appropriate simplex game* \mathfrak{G} consists of:

1. A set of *players* $\mathcal{N} = \{1, \dots, N\}$.
2. A set of *nodes* $\mathcal{B} = \{q_r\}_{r=1}^B$, where $\mathcal{B} \subseteq \mathbb{R}^{B-1}$ is a y -simplex.
3. A set of *signals* $\mathcal{M} = \{1, \dots, M\}$, endowed with the uniform measure $q_0(m) = \frac{1}{M}$; the ratio $\lambda = M/N$ will be called the *training parameter*.
4. A set of *strategy choices* $\mathcal{S} = \{1, \dots, S\}$; also, for each player $i \in \mathcal{N}$, a probability measure $p_i = \sum_s p_{is} e_s \in \Delta(\mathcal{S})$ on \mathcal{S} , representing a *mixed strategy* of player i .

⁸ These strategies are called correlated to account for the fact that they are based on the public event $\gamma \in \Gamma$ which is common to all; see also Chapter 2.

⁹ For the sake of simplicity, we are assuming that S is the same for all users.

¹⁰ In this way, each strategy consists of a random hopping pattern *biased* towards the strongest nodes, similar to a biased hopping pattern of OFDM users.

5. A *strategy matrix* $c : \mathcal{N} \times \mathcal{S} \times \mathcal{M} \rightarrow \mathcal{B}$ where $c(i, s, m) \equiv c_{is}^m \in \mathcal{B}$ is the node that the s -th strategy of user i indicates as response to the signal $m \in \mathcal{M}$, the entries of c being drawn randomly based on (7.12) above.

Furthermore, we also endow the sample space $\Omega \equiv \mathcal{M} \times \mathcal{S}^N$ with the product probability measure $\varrho_0 \times \prod_{i=1}^N p_i$, and define the following:

6. An *instance* of \mathcal{G} will be an event $\omega = (m, s_1, \dots, s_N)$ of Ω ;
7. The *bet* of player i will be the \mathcal{B} -valued random variable $b_i(\omega) = c(i, s_i, m)$; similarly, the players' *aggregate bet* will be $b = \sum_{i=1}^N b_i$;
8. The *payoff* to player i will be given by the random variable $u_i = -\frac{1}{N} b_i \cdot b$.

Thus, similarly to the minority game of Challet and Zhang (1997) and Marsili et al. (2000), the sequence of events that we intuitively envision is as follows:¹¹

- In the “initialization” phase (steps 1-5), players program their strategies by drawing the strategy matrix c .
- In step 6, the signal m is broadcasted and, based on p_i , players pick a strategy $s \in \mathcal{S}$ to process it with: p_{is} is the probability that user i employs his s -th strategy.
- In steps 7-8, players connect to the nodes that their strategies indicate ($b_i(m, s_1, \dots, s_N) = c_{is_i}^m$) and receive the linear payoff (7.4): by (7.9), each of the N_r users that end up connecting to node q_r receives:

$$-\frac{1}{N} q_r \cdot \sum_l N_l q_l = 1 - \frac{N_r}{y_r N}. \quad (7.13)$$

- The game is then iterated by repeating steps 6-8.

As usual, the payoff that corresponds to the (mixed) profile $p = (p_1, \dots, p_N)$ will be the multilinear expectation:

$$u_i(p) = \sum_{s_1 \in \mathcal{S}_1} \cdots \sum_{s_N \in \mathcal{S}_N} p_{1,s_1} \cdots p_{N,s_N} u_i(s_1, \dots, s_N). \quad (7.14)$$

However, to avoid carrying around cumbersome sums like this, we will follow the notation of Marsili et al. (2000) and use $\langle \cdot \rangle$ to indicate expectations over a particular player's mixed strategy, that is:

$$\langle v_i \rangle = \sum_s^i p_{is} v_{is}. \quad (7.15)$$

In a similar vein, given that we will frequently need to average over the training signals as well, we will also use an overline to denote averaging over \mathcal{M} :

$$\bar{a} = \frac{1}{M} \sum_m a^m. \quad (7.16)$$

These two conventions will come in very handy, so we will use them freely in the rest of this chapter – this is also the reason that we abandoned the $\langle \cdot, \cdot \rangle$ notation that we used for the Euclidean product in the previous chapters.

Some notation.

¹¹ At this point, it is important to note that for 2 identical nodes ($\mathcal{B} = \{-1, 1\}$), the simplex game reduces exactly to the original minority game of Marsili et al. (2000).

7.2 SELFISHNESS AND EFFICIENCY

Clearly now, the only way that selfish users who seek to maximize their individual throughput can come to an unmediated understanding is by reaching an equilibrial state that discourages unilateral deviation. But, since there is a palpable difference between the users' *strategic decisions* ($s \in \mathcal{S}$) and the *tactical actions* that they take based on them ($c_{is}^m \in \mathcal{B}$), one would expect the situation to be somewhat involved.

7.2.1 Notions of Equilibrium

Correlated equilibria.

Indeed, it should not come as a surprise that this dichotomy between strategies and actions is reflected on the game's equilibria. On the one hand, we have already encountered the game's tactical equilibrium: it corresponds to the Nash allocation of $y_r N$ users to node r . On the other hand, given that users only control their strategic choices, we should also examine Aumann's strategic notion of a *correlated equilibrium* as well.

To that end, recall that a correlated strategy is a collection $f \equiv \{f_i\}_{i=1}^N$ of maps $f_i : \mathcal{M} \rightarrow \mathcal{B}$ (one for each player) that convert the signal m to a betting suggestion $f_i(m) \in \mathcal{B}$. We will then say that a (pure) correlated strategy f is at *equilibrium for player i* when, for all perturbations $(f_{-i}; g_i) = (f_1, \dots, g_i, \dots, f_N)$ of f , player i gains more (on average) by sticking to f_i , i.e. $u_i(f) \geq u_i(f_{-i}; g_i)$, where u_i is the payoff of the i -th user and the expectation is taken over all $m \in \mathcal{M}$. When this is true for all players $i \in \mathcal{N}$, f will be called a *correlated equilibrium*.

As we saw before, if the i -th user picks his s_i -th strategy, the collection $\{c_{is_i} : \mathcal{M} \rightarrow \mathcal{B}\}_{i=1}^N$ is a correlated strategy, but the converse need not hold: in general, not every correlated strategy can be recovered from the limited number of preprogrammed strategic choices.¹² Thus, users will no longer be able to consider *all* perturbations of a given strategy, and we are led to:

Definition 7.6. In the setting of definition 7.5, a strategy profile $p = (p_1, \dots, p_N)$ will be called a *constrained correlated equilibrium* when, for all strategies $s \in \mathcal{S}$, and for all players $i \in \mathcal{N}$:

$$\frac{1}{M} \sum_m u_i(m, p) \geq \frac{1}{M} \sum_m u_i(m, p_{-i}; s). \quad (7.17)$$

As before, the set of all such equilibria of \mathfrak{G} will be denoted by $\Delta^* \equiv \Delta^*(\mathfrak{G})$.

In this way, a (constrained) correlated equilibrium represents *anarchy* in our setting. With no one to manage the users' selfish desires, the only thing that deters them from unilateral deviation is their expectation of (average) loss. Conceptually, this is pretty similar to the notion of a Nash equilibrium, the main difference being that, in a correlated equilibrium, one is averaging the payoff over the training signals.

This last analogy will be very useful to us, so we will make it precise by introducing the associated *correlated form* of the simplex game:

The correlated game.

Definition 7.7. The *correlated form* of a simplex game \mathfrak{G} is a game \mathfrak{G}^c with

¹² There is a total of B^{MN} correlated strategies but users can recover at most S^N of them. In fact, this is why preprogramming is so useful: it would be highly unreasonable to expect a given user to process in a timely fashion the exponentially growing number of B^M (as compared to S) strategies.

the same set of players $\mathcal{N} = \{1, \dots, N\}$, each one choosing an action from $\mathcal{S} = \{1, \dots, S\}$ for a payoff of:

$$u_i^c(s_1, \dots, s_N) = \frac{1}{M} \sum_m u_i(m, s_1, \dots, s_N) \quad (7.18)$$

In short, the payoff that players receive in the correlated game is their throughput averaged over a rotation of the training signals. Then, an important consequence of definition 7.6 is that *the constrained correlated equilibria of a simplex game \mathfrak{G} are precisely the Nash equilibria of its correlated form \mathfrak{G}^c* .

7.2.2 Harvesting the Equilibria

So, our next goal will be to understand the Nash equilibria of \mathfrak{G}^c . To begin with, a brief calculation shows that for a mixed profile $p = (p_1, \dots, p_N)$, the (expected) payoff u_i^c to player i will be:

$$u_i^c(p_1, \dots, p_N) = -\frac{1}{N} \left(\overline{\langle c_i \rangle \cdot \sum_{j \neq i} \langle c_j \rangle} + \overline{\langle c_i^2 \rangle} \right), \quad (7.19)$$

the averaging notations $\overline{(\cdot)}$ and $\langle \cdot \rangle$ being as in the end of Section 7.1. Thus, given the similarities of our game with congestion games, it might be hoped that its Nash equilibria can be harvested by means of a *potential function*, i.e. a function which measures the payoff difference between users' individual strategies. More concretely, a potential function U for \mathfrak{G}^c will satisfy:

$$u_i^c(p_{-i}; s_1) - u_i^c(p_{-i}; s_2) = U(p_{-i}; s_1) - U(p_{-i}; s_2), \quad (7.20)$$

for any mixed profile $p = (p_1, \dots, p_S)$ and any two strategic choices $s_{1,2}$ of player i . Obviously then, if a potential function exists, its local maxima will be Nash equilibria of the game – we are inverting here the usual definition (2.10) of a game's potential simply for the sake of future convenience.

Unfortunately however, since \mathfrak{G}^c does not have an exact congestion structure, it is not clear how to construct such a potential. Nevertheless, a good candidate is the game's aggregate payoff $u^c = \sum_i u_i^c$. Indeed, if player i chooses strategy s , u^c becomes:

$$u^c(p_{-i}; s) = -\frac{1}{N} \sum_{\substack{\ell, k \neq i \\ \ell \neq k}} \overline{\langle c_\ell \rangle \cdot \langle c_k \rangle} + \frac{1}{N} \sum_{k \neq i} \overline{\langle c_k^2 \rangle} + \frac{2}{N} \overline{c_{is} \cdot \sum_{k \neq i} \langle c_k \rangle} - \overline{c_{is}^2}. \quad (7.21)$$

So, after some similar algebra for $u_{is}^c(p) \equiv u_i^c(p_{-i}; s)$, we obtain the following comparison between two strategies $s_1, s_2 \in \mathcal{S}$:

$$u^c(p_{-i}; s_2) - u^c(p_{-i}; s_1) = 2 \left(u_{is_2}^c(p) - u_{is_1}^c(p) \right) + \frac{1}{N} \overline{c_{is_2}^2 - c_{is_1}^2} \quad (7.22)$$

Now, given the preprogramming (7.12) of c , we note that $(c_{is}^m)^2$ takes on the value $q_r^2 = -1 + \frac{1}{y_r}$ with probability y_r . Hence, the central limit theorem implies that $\frac{1}{M} \sum_{m=1}^M (c_{is}^m)^2$ will have mean $\sum_r y_r \left(\frac{1}{y_r} - 1 \right) = B - 1$ and variance $\frac{1}{M} \sum_r \left(\frac{1}{y_r} - B \right)$, the latter being negligible unless y is too close to the faces of Δ_B .¹³ To quantify this in more precise terms, we are led to:

¹³ Recall that $M = \lambda N = \mathcal{O}(N)$.

The game's potential.

Definition 7.8. A distribution $y \in \text{Int}(\Delta_B)$ will be called *proper* when:

$$\frac{1}{B-1} \sum_r \left(\frac{1}{y_r} - B \right) = \mathcal{O}(1); \text{ otherwise, } y \text{ will be called } \textit{degenerate}. \quad (7.23)$$

Degenerate strength profiles.

Degeneracy in a strength distribution $y \in \Delta_B$ simply indicates that certain nodes have extremely low strength scores that make any reasonable user shun them. So, in the case of such a disparity, one should proceed by removing these weaker nodes from the analysis (i.e. reduce B and modify y accordingly), until the above definition holds.¹⁴ In this way, we lose no generality by taking y to be proper and, henceforward, this will be our working assumption.

With this in mind, the last term of (7.22) will be on average 0 and with a variance of lesser order than the first term. We thus obtain:

$$u^c(p_{-i}; s_2) - u^c(p_{-i}; s_1) \sim 2 \left(u_i^c(p_{-i}; s_2) - u_i^c(p_{-i}; s_1) \right) \quad (7.24)$$

i.e. the aggregate payoff u^c is indeed a *potential function* for the game \mathfrak{G}^c in the asymptotic limit $N \rightarrow \infty$.¹⁵ In other words, we have proven:

Lemma 7.9. *Let \mathfrak{G} be a simplex game for N players. Then, as $N \rightarrow \infty$, the maxima of the averaged aggregate $u^c = \sum_{i=1}^N \bar{u}_i$ will correspond (almost surely) to correlated equilibria of \mathfrak{G} .*

7.2.3 Anarchy and Efficiency

Correlated anarchy.

Of course, one would expect quite the gulf between anarchic and efficient states: after all, players that only have their selfish interests at heart are hardly the ones to rely upon for social efficiency. In the context of networks, this contrast is frequently measured by the *price of anarchy*, a notion first introduced in [Koutsoupias and Papadimitriou \(1999\)](#) as the (coordination) ratio of the maximum attainable aggregate payoff to the one attained at the game's equilibria. Then, depending on whether we look at worst or best-case equilibria, we get the *pessimistic* or *optimistic* price of anarchy respectively.

In a simplex game \mathfrak{G} , the aggregate payoff is equal to:

$$u = \sum_{i=1}^N u_i = -\frac{1}{N} \sum_{i=1}^N b_i \cdot \sum_{j=1}^N b_j = -\frac{1}{N} b^2, \quad (7.25)$$

and attains a maximum of $u_{\max} = 0$ when $b = 0$. So, if we recall by (7.10) that a Nash equilibrium *occurs* if and only if $b = 0$, we see that *Nash anarchy does not impair efficiency*. Clearly, neither the users, nor the agencies that deploy the wireless network could hope for a better solution!

However, this also shows that the traditional definition of the price of anarchy is no longer suitable for our purposes. One reason is that $u_{\max} = 0$ and, hence, we cannot hope to get any information from ratios involving u_{\max} .¹⁶ What's more, the users' selfishness in our setting is more aptly captured by the Aumann equilibria of definition 7.6, so we should be taking the signal-averaged u^c instead of u . As a result, we are led to:

- ¹⁴ After all, the weaker nodes would not be able to serve more than $\mathcal{O}(\sqrt{N})$ users: the overall effect of this reduction is inconsequential.
- ¹⁵ It is important to note here that the large N assumption can be completely dropped for the symmetric case $y_1 = \dots = y_B$. Actually, the farther y is from the faces of Δ_B , the smaller the impact of the number of users will be.
- ¹⁶ This actually highlights a general problem with the price of anarchy: it does not behave well with respect to linear changes to the payoff functions.

Definition 7.10. Let \mathfrak{G} be a simplex game for N players and B nodes. Then, if $p = (p_1, \dots, p_N)$ is a mixed strategy profile of \mathfrak{G} , we define its *frustration level* to be:

$$R(p) = -\frac{1}{B-1}u^c(p) = \frac{1}{N(B-1)}\frac{1}{M}\sum_m b^2(p), \quad (7.26)$$

that is, the average distance from the Nash solution $b = 0$. Also, the game's *correlated price of anarchy* $R(\mathfrak{G})$ is defined as:

$$R(\mathfrak{G}) = \inf \{R(p) : p \in \Delta^*(\mathfrak{G})\} \quad (7.27)$$

i.e. the minimum value of the frustration level over the set $\Delta^*(\mathfrak{G})$ of the game's constrained correlated equilibria.

Some remarks are now in order: first and foremost, we see that the frustration level of a strategy profile measures (in)efficiency by contrasting the average aggregate payoff to the optimal case $u_{max} = 0$ (the normalization $\frac{1}{B-1}$ has been introduced for future convenience). So, with correlated equilibria representing the anarchic states of the game, we remain justified in the eyes of [Koutsoupias and Papadimitriou \(1999\)](#) by calling $R(\mathfrak{G})$ the price of anarchy: the only thing that sets us apart is that, instead of a ratio, we are taking the difference.

Moreover, one might wonder why we do not consider the *pessimistic* version by replacing the inf of the above definition with a sup. The main reason for this is that in the next section, we will present a scheme with which users will be able to converge to their *most efficient* equilibrium. Thus, there is no reason to consider worst-case equilibria as in [Koutsoupias and Papadimitriou \(1999\)](#): we only need to measure the price of *sophisticated* anarchy.

7.3 ITERATION AND EVOLUTION

As in the previous chapters, the main idea behind the learning scheme with which players will converge to a sophisticated equilibrium is that of “exponential learning”: users keep track of their strategies' performance over time and that they employ more often those that yield better payoffs. The details of this evolutionary scheme are presented in section 7.3.1 and then, in section 7.3.2, we proceed to show that the evolutionarily stable states correspond to the game's correlated equilibria. In fact, we will see that as the players' mixed strategies p_{is} evolve over time, their frustration level is minimized and they settle down to the most efficient equilibrium: on average, they will be as close as possible to the optimal solution.

7.3.1 Evolutionary Dynamics

Evidently, as the simplex game is played again and again, rational users will want to maximize the payoff they receive by employing more often the strategies that perform better. The most obvious way to accomplish this is to “reward” a strategy when it performs well and penalize it otherwise:

Learning in the simplex game.

Definition 7.11. Let \mathfrak{G} be a simplex game as in definition 7.5, and let $\omega = (m, s_1, \dots, s_N)$ be an instance of \mathfrak{G} . Then, the *reward* to the s -th strategic choice of player i is the random variable:

$$W_{is}(\omega) = \frac{1}{M}u(m, s_{-i}; s) = -\frac{1}{MN}c_{is}^m \cdot \left(b(\omega) + \left(c_{is}^m - c_{is_i}^m \right) \right) \quad (7.28)$$

In other words, the reward W_{is} that player i awards to his s -th strategy is (a fraction of) the payoff that the strategy would have garnered for the player in the given instance – the rescaling factor $\frac{1}{M}$ has been introduced so that there be no noticeable effect on a strategy's ranking until it has been checked against at least $\mathcal{O}(M)$ of the possible training signals.

A seeming problem with the above definition is that users will have to rate *all* their strategies, i.e. they must be able to calculate the payoff even of strategies they did not employ. Hence, given that the payoff is a function of the aggregate bet b , it would seem that in order to carry out this calculation, users would have to be informed at each step of every other user's bet, a prospect that downright shatters the unregulated premises of our setting. However, a more careful consideration of (7.9) reveals that it suffices for users to know the distribution of users among the nodes, and this is something that is small enough to be broadcasted by the nodes along with the signal m .¹⁷ Thus, in the language of Marsili et al. (2000), users can be considered *sophisticated* because they account properly for their own impact on the aggregate bet b .

So, let $\omega(t)$ be a sequence of instances of \mathfrak{G} ($t = 0, 1, \dots$). Then, the ranking of a player's strategic choices will be achieved via the strategy scores:

$$U_{is}(t+1) = U_{is}(t) + W_{is}(\omega(t)) \quad (7.29)$$

where we set $U_{is}(0) = 0$, reflecting the fact that there is no prior predisposition towards any given strategy. Needless to say, users will want to employ more often their highest scoring strategies, the precise model given by the familiar evolutionary scheme of exponential learning (see also Marsili et al., 2000):

$$p_{is}(t) = \frac{e^{\Gamma_i U_{is}(t)}}{\sum_{s'} e^{\Gamma_i U_{is'}(t)}} \quad (7.30)$$

where Γ_i represents the *learning rate* of player $i \in \mathcal{N}$.

Averaging and the replicator dynamics.

Now, as a first step to understand the dynamical system described by (7.30), note that the evolution of players actually takes place over the time scale $\tau = t/M$ – it takes users an average of M iterations to notice a distinct change in the scores of their strategies (7.28). In this time scale, the score of a strategy will have been modified by:

$$\delta U_{is} = \sum_{t=\tau}^{\tau+M} W_{is}(\omega(t)) = -\frac{1}{MN} \sum_{t=\tau}^{\tau+M} \left(c_{is}^{m(t)} \cdot \sum_{j \neq i} c_{js_j(t)}^{m(t)} + \left(c_{is}^{m(t)} \right)^2 \right). \quad (7.31)$$

However, by the central limit theorem, we obtain:

$$\sum_{j \neq i} c_{js_j(t)}^{m(t)} \sim \sum_{j \neq i} \langle c_j^{m(t)} \rangle, \quad (7.32)$$

and, under some mild ergodicity assumptions, we can also approximate the time average $\frac{1}{M} \sum_{t=\tau}^{\tau+M} (\cdot)$ by the ensemble average $\frac{1}{M} \sum_{m=1}^M (\cdot)$. In this way, the change in a strategy's score after M iterations will be asymptotically approximated by:

$$\delta U_{is} \sim -\frac{1}{N} \left(\overline{c_{is} \cdot \sum_{j \neq i} \langle c_j \rangle} - \overline{c_{is}^2} \right) = u_i^c(p_{-i}; s) \quad (7.33)$$

¹⁷ Actually, the signal itself could be the user distribution of the previous stage. Cavagna (1999) showed that the distinction between real and fake memory has a negligible impact.

Implicit in the above is the assumption that p_{is} changes very slowly with respect to t , a caveat that breaks down if the learning rates Γ_i are too high (i.e. when we freeze down to a “hard” best-response scheme).¹⁸ So, if we stay clear of this critical value, we may descend to continuous time and differentiate (7.30) to obtain the replicator equation:

$$\frac{dp_{is}}{d\tau} = \Gamma_i p_{is} \left(u_i^c(p) - u_i^c(p_{-i}; s) \right) \quad (7.34)$$

7.3.2 The Steady States

As we have already noted, the above dynamics are extremely powerful: first of all, from an evolutionary point of view, we see that they are the standard multi-population replicator dynamics for the correlated form \mathfrak{G}^c of the game, and, as such, they will adhere to the multi-population version of the folk theorem 2.16: the asymptotically stable states of (7.34) are just the (strict) Nash equilibria of the underlying game (in our case \mathfrak{G}^c). But, since a strategy profile is a Nash equilibrium for the correlated game \mathfrak{G}^c if and only if it is a correlated equilibrium for the original game \mathfrak{G} , this proves:

Lemma 7.12. *Let \mathfrak{G} be a simplex game, iterated under the exponential learning scheme of (7.30). Then, as $N \rightarrow \infty$, a mixed strategy profile $p = (p_1, \dots, p_N)$ will be asymptotically stable in the dynamics (7.30) (a.s.) if and only if it is a constrained correlated equilibrium of \mathfrak{G} .*

Stated more intuitively, the above lemma ensures that if the users reach a state where they are satisfied *on average* with their strategic choices, they will have no incentive to modify their strategies’ rankings at all. So, what remains to be seen is whether the learning scheme of (7.30) really does lead the game to such a fortuitous state.

Convergence to equilibrium.

To that end, one would expect that, as users evolve, they learn how to minimize their average frustration level and eventually settle down to a stable local minimum. On account of the existence of the (asymptotic) potential (7.24), this is indeed the case as $N \rightarrow \infty$: if we combine (7.24) and (7.34), we easily obtain:

$$\frac{du^c}{d\tau} = \frac{1}{2} \sum_i \Gamma_i \sum_s u^c(p_{-i}; s) p_{is} (u^c(p_{-i}; s) - u^c(p)) \geq 0 \quad (7.35)$$

, the last step owing to Jensen’s inequality (recall that $u^c(p) = \sum_s p_{is} u^c(p_{-i}; s)$). In other words, the frustration $R = -\frac{1}{B-1} u^c$ is a Lyapunov function for the dynamics of (7.34) and the players will converge to its global minimum:

Theorem 7.13. *If a simplex game \mathfrak{G} with a large number N of players is iterated under the exponential learning scheme (7.30), the players’ mixed strategies will converge (a.s.) to a strict constrained correlated equilibrium of \mathfrak{G} which maximizes the aggregate payoff u^c over all $p \in \Delta \equiv \prod_1^N \Delta(\mathfrak{S})$.*

7.4 MEASURING THE PRICE OF ANARCHY

So far, we have seen that the dynamics of exponential learning lead the users to an evolutionarily stable equilibrium which also maximizes (on average) their aggregate payoff (given their preprogramming). Hence, as far as measuring

¹⁸ Marsili and Challet (2001) has a lengthy discussion on this point.

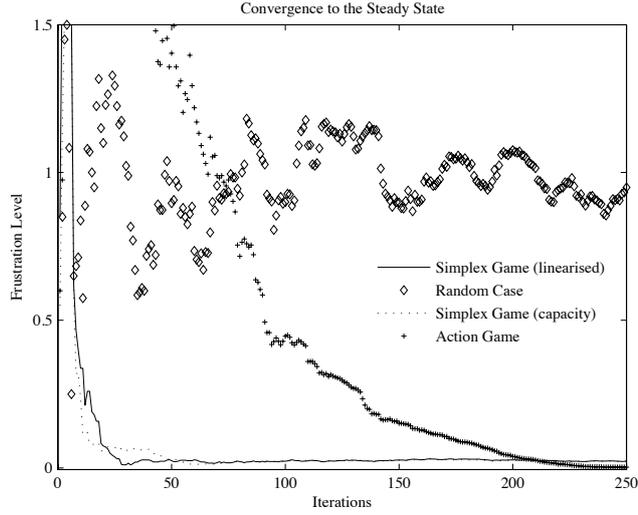


Figure 7.1: Simulation of a simplex game for $N = 50$ players that seek to connect to $B = 5$ nodes of random strengths with the help of $M = 2$ broadcasts and $S = 2$ strategies (we chose learning rates of $\Gamma_i = 20$): as predicted by theorem 7.13, players quickly converge to a steady state of minimal frustration. To justify the linearization of (7.4), we also simulated a game with the nonlinear payoff (7.1), obtaining virtually indistinguishable results. Our comparison baselines are random users who experience an average frustration $R = 1$ and action-oriented users who follow the replicator dynamics of the congestion game determined by (7.1).

Harvesting the
ground state.

anarchy is concerned, we only need to calculate the level of frustration at this steady state: rather surprisingly, it turns out that the price of anarchy is *independent* of the distribution y of the nodes' strengths. In fact, the analytic expression that we obtain shows that it is a function only of the number B of nodes in the network, the training parameter $\lambda = \frac{M}{N}$ and the number S of strategies per user.

To begin with, equation (7.19) leads to the following expression for the frustration level $R(p)$ at a mixed strategy profile p :

$$R(p) = \frac{1}{N(B-1)} \sum_i \left(\overline{\langle c_i^2 \rangle} + \sum_{j \neq i} \langle c_i \rangle \cdot \langle c_j \rangle \right) \quad (7.36)$$

So, by Definition 7.8 and (7.22), the first term of (7.36) will be asymptotically equal to $\frac{1}{N(B-1)} \sum_i \overline{\langle c_i^2 \rangle} \sim 1$. Then, as far as the second term of (7.36) is concerned, note that the aggregate bet $b(m, p) = \sum_i \langle c_i^m \rangle$ becomes:

$$b(m, p)^2 = \sum_i \sum_{s \neq s'} p_{is} p_{is'} c_{is}^m \cdot c_{is'}^m + \sum_i \sum_s p_{is}^2 (c_{is}^m)^2 + \sum_i \sum_{j \neq i} \langle c_i^m \rangle \cdot \langle c_j^m \rangle. \quad (7.37)$$

Then, to leading order in N , it is easy to see that this last expression has an (asymptotic) average of:

$$\frac{1}{M} \sum_m b(m, p)^2 \sim \sum_{\substack{i,j \\ i \neq j}} \langle c_i \rangle \cdot \langle c_j \rangle + (B-1) \sum_i \sum_s p_{is}^2, \quad (7.38)$$

and, hence, equations (7.36) and (7.38) finally yield:

$$R(p) \sim 1 + \frac{1}{MN(B-1)} \sum_m b(m, p)^2 - G(p), \quad (7.39)$$

where the term $G(p) \equiv \frac{1}{N} \sum_i \sum_s p_{is}^2$ measures the self-overlap of strategies – see e.g. Marsili et al. (2000) or Marsili and Challet (2001).

By Definition 7.10, the game's price of anarchy $R(\mathfrak{G})$ will simply be the minimum of $R(p)$ over the game's equilibria. However, since Theorem 7.13 shows that the global minimum of R is an equilibrium, we can simply take the minimum over *all* mixed profiles:

$$R(\mathfrak{G}) = \min\{R(p) : p \in \Delta\}. \quad (7.40)$$

Minimization problems of this kind are the meat and potatoes of statistical physics, where a vast array of techniques has been developed to harvest the ground states of (similar in form) energy functionals – consult for example the excellent book by Nishimori (2001). Motivated by this, we introduce the partition function:

$$\mathcal{Z}(\beta, c) = \int_{\Delta} e^{-\beta NR(p)} dp \quad (7.41)$$

where, as before, $\Delta = \prod_1^N \Delta_S$ and $dp = \prod_{i,s} dp_{is}$ denotes the Lebesgue measure of Δ .¹⁹ In this way, we may use the *steepest descent* method of asymptotic integration²⁰ to write:

$$R(\mathfrak{G}) = -\frac{1}{N} \lim_{\beta \rightarrow \infty} \frac{1}{\beta} \log \mathcal{Z}(\beta, c). \quad (7.42)$$

To proceed from here, we will make the mild (but important) assumption that, for a large number N of users, it matters little which specific strategy matrix the users actually drew during the preprogramming phase. More formally:

Assumption 7.14 (Self-averaging). For any strategy matrix c , we have:

$$\log \mathcal{Z}(\beta, c) \sim \langle \log \mathcal{Z}(\beta) \rangle_{\text{all } c} \quad \text{almost surely as } N \rightarrow \infty, \quad (7.43)$$

the averaging $\langle \cdot \rangle$ taking place over all B^{NSM} matrices c , drawn according to the probability law (7.12).

This is a fundamental assumption in statistical physics and describes the rarity of configurations which yield notable differences in macroscopically observable parameters. Under this light, we are left to calculate $\langle \log \mathcal{Z} \rangle$, a problem which we will attack with the help of the quasi-magical *replica method*.

In short, the method of replica analysis allows one to discern information about a statistical system by replicating it and studying all the copies at once.²¹ Its starting point is the fundamental identity:

$$\langle \log \mathcal{Z} \rangle = \lim_{a \rightarrow 0^+} \frac{1}{a} \log \langle \mathcal{Z}^a \rangle \quad (7.44)$$

¹⁹ \mathcal{Z} depends on the strategy matrix c through the frustration level $R(p)$.

²⁰ Essentially, this refers to the fact that $\max_D f = \lim_{x \rightarrow \infty} \frac{1}{x} \log \int_D e^{xf(t)} dt$ for any measurable function f on a compact domain D (see e.g. Dembo and Zeitouni (1998)).

²¹ See Mézard et al. (1987) or Nishimori (2001) for a general discussion of the method or Marsili et al. (2000) and de Martino and Marsili (2001) for the method's application to the minority game.

which reduces the problem to calculating powers of \mathcal{Z} :²²

$$R(\mathfrak{G}) = -\frac{1}{N} \lim_{\beta \rightarrow \infty} \lim_{a \rightarrow 0^+} \frac{1}{a\beta} \langle \mathcal{Z}^a(\beta) \rangle. \quad (7.45)$$

These last terms are much more manageable because, for $n \in \mathbb{N}$, we have:

$$\mathcal{Z}^n = \left(\int_{\Delta} e^{-\beta NR(p)} dp \right)^n = \int_{\Delta} \dots \int_{\Delta} e^{-\beta N \sum_{\mu} R(p_{\mu})} dp_1 \dots dp_N, \quad (7.46)$$

that is, $\mathcal{Z}^n = \prod_{\mu=1}^n \mathcal{Z}_{\mu}$ where

$$\mathcal{Z}_{\mu} = \int_{\Delta} \exp(-N\beta R(p_{\mu})) dp_{\mu} \quad (7.47)$$

is the partition function for the μ -th replica $p_{\mu} = \sum_{i,s} p_{is\mu} e_{is} \in \Delta$ of the system. Thus, thanks to equation (7.39), we obtain:

$$\langle \mathcal{Z}^n(\beta) \rangle = A^n \int_{\Delta^n} \left\langle e^{-\frac{\beta}{M(B-1)} \sum_{\mu} \sum_m (c_{\mu}^m)^2} \right\rangle e^{N\beta \sum_{\mu} G_{\mu\mu}(p)} dp_1 \dots dp_N, \quad (7.48)$$

where $A = e^{-N\beta}$, $c_{\mu}^m = b(m, p_{\mu}) = \sum_i \sum_s p_{is\mu} c_{is}^m$ is the aggregate bet for the mixed profile $p_{\mu} = \sum_{i,s} p_{is\mu} e_{is}$ in the μ^{th} replica (given the signal m) and $G_{\mu\nu}(p) = \frac{1}{N} \sum_i \sum_s p_{is\mu} p_{is\nu}$.

Of course, what we really need is to express $\langle \mathcal{Z}^n \rangle$ for real values of n in the vicinity of $n = 0^+$. To do that, we resort to:

Replica continuity.

Assumption 7.15 (Replica Continuity). The expression given in (7.48) for $\langle \mathcal{Z}^n \rangle$ can be continued analytically to all real values of n in the vicinity of $n = 0^+$.

At first glance, this assumption might appear as a blind leap of faith, especially since uniqueness criteria (e.g. log-convexity) are absent. However, such criteria can in some cases be established (see e.g. Mézard et al. (1987) and Coolen (2004)) and, moreover, the huge amount of literature surrounding this assumption and the agreement of our own analysis with our numerical results (see Figures 7.2, 7.3 and 7.4) makes us feel justified in employing it.

At any rate, after a rather long calculation (which we present in Appendix A.1), we are finally in a position to prove our first major result: *the price of anarchy is not affected by disparities in the nodes' strengths*. In a more precise formulation:

An isotropy result.

Theorem 7.16 (Independence on Node Strength). *Let $y, y' \in \text{Int}(\Delta_B)$ be strength distributions for B nodes, and let $\mathfrak{G}, \mathfrak{G}'$ be simplex games for y and y' respectively. Then, as $N \rightarrow \infty$:*

$$R(\mathfrak{G}) \sim R(\mathfrak{G}') \quad \text{almost surely.} \quad (7.49)$$

Remark. This theorem has important consequences for our model since it reduces it to the symmetric case of B nodes of equal strength: *the price of anarchy depends only on the number of nodes present and not on their individual strengths*.

Our next result is a quantitative criterion that expands on the above by determining the effect of (the number of) choices on the users' frustration level. To that end, if \mathfrak{G} is a simplex game for B nodes, we define its *binary*

²² To prove (7.44), simply write $\mathcal{Z}^a = e^{a \log \mathcal{Z}}$ and expand.

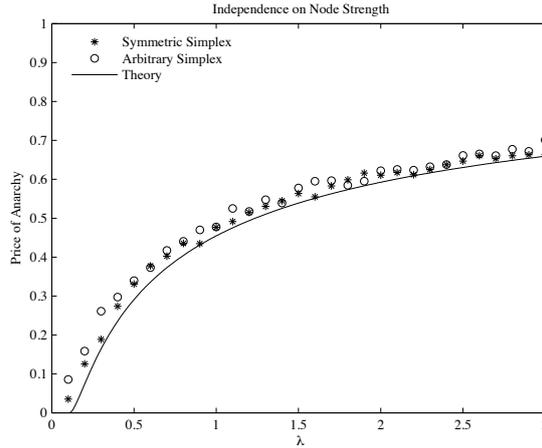


Figure 7.2: The price of anarchy (i.e. the steady-state frustration level of the game) as a function of the training parameter $\lambda = M/N$ for $B = 4$ equivalent nodes is compared to that of 4 nodes employing standards with different spectral efficiencies c_r ; in accordance with Theorem 7.16, the presence of different standards does not affect the price of anarchy.

reduction $\mathfrak{G}_{\text{eff}}$ to be a simplex game for 2 identical nodes and a training set enlarged by a factor of $B - 1$ – that is, $M_{\text{eff}} = M(B - 1)$ while everything else remains the same. Under this rescaling, the same train of calculations that is used to prove theorem 7.16 also yields:

Theorem 7.17 (Choice Reduction). *The price of anarchy for a simplex game \mathfrak{G} is asymptotically equal to that of its binary reduction $\mathfrak{G}_{\text{eff}}$; in other words, as $N \rightarrow \infty$:*

A reduction result.

$$R(\mathfrak{G}) \sim R(\mathfrak{G}_{\text{eff}}) \quad \text{almost surely.} \quad (7.50)$$

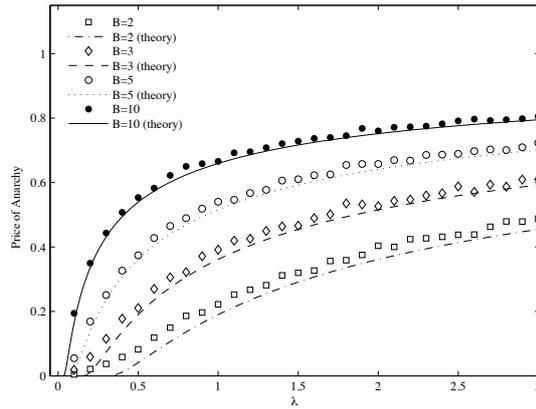
This gives us an easy way to compare performance in unregulated networks with users of the same sophistication level (i.e. same S); in particular, if $M_1(B_1 - 1) = M_2(B_2 - 1)$ for two networks $\mathfrak{G}_1, \mathfrak{G}_2$, these networks will have the same price of anarchy. So, for example, in the case of a power outage that takes out some access points, the network administrators will know exactly how much to increase the number of transmitted signals in order to maintain the same level of quality of service for the users.

Finally, it is important to note that theorem 7.17 essentially maps the simplex game to the minority game and allows us to port the vast array of results obtained in econophysics to our setting. As a simple illustration of the power of this “dictionary”, note that the price of anarchy of the simplex game as we defined it coincides with *the volatility* of the minority game (Marsili et al., 2000), rescaled by $\frac{1}{B-1}$. So, mutatis mutandis, and under the ansatz of replica symmetry, Theorem 7.17 and the calculations of Marsili et al. (2000) yield:

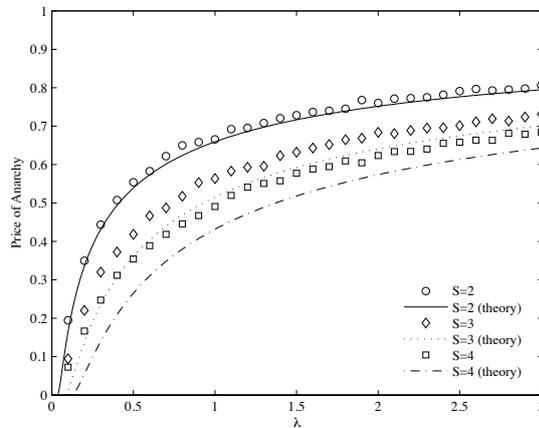
$$R(\mathfrak{G}) \sim \Theta(\lambda - \lambda_c) \left(1 - \sqrt{\lambda_c/\lambda}\right)^2 \quad (7.51)$$

where Θ is the Heaviside step function and $\lambda_c \equiv \lambda_c(S, B)$ is a critical value of the training parameter below which the price of anarchy vanishes.²³

²³ As we show in Appendix A.2, the specific form for λ_c is $\lambda_c = \zeta^2(S)/(B - 1)$, where the function ζ is itself given by: $\zeta(s) = s/2^{s-1} \sqrt{2/\pi} \int_{-\infty}^{\infty} z e^{-z^2} \operatorname{erfc}^{s-1}(z) dz$.



(a) The effect of choice.



(b) The effect of sophistication.

Figure 7.3: The price of anarchy (averaged over 25 realizations of the game) as a function of the training parameter $\lambda = M/N$ for $N = 50$ users and various numbers of nodes ($B = 2, 3, 5, 10$) and strategies ($S = 2, 3, 4$). Perhaps surprisingly, players seem to be confused when choices abound (their efficiency deteriorates for large B), but an increase in sophistication (large S) is always to their benefit.

This expression is one of our key results because it describes the price of anarchy in terms of the game's macroscopic parameters: the number of nodes B , the number of strategies S , and the training parameter $\lambda = \frac{M}{N}$ of the game. So, for completeness (and also to properly discuss the role of replica symmetry) we carry out the calculation for (7.51) in Appendix A.2.

7.5 A SHORT DISCUSSION

Our main goal was to analyze an unregulated network consisting of heterogeneous users that can connect to a multitude of wireless nodes with different specifications (e.g. different standards). The users' selfish desire was to maximize their individual downlink throughput (given by the simple equation (7.2)), a fact which gives rise to competition among them, as they have to share the network's finite resources. So, in the pursuit of order, and in the absence of a central overseer, we advocate the use of a training beacon (such

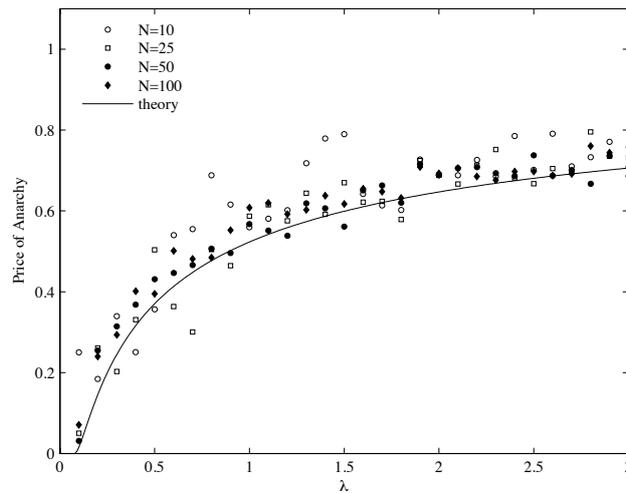


Figure 7.4: The price of anarchy as a function of the training parameter $\lambda = M/N$ for different numbers of users $N = 10, 25, 50, 100$, with $B = 5$ nodes and $S = 2$ strategies; unlike other plots, we are harvesting the price of anarchy from a *single* realization of the game. Despite the fluctuations that this entails, we see that the number of players does not seriously impact the price of anarchy (except through λ). In fact, our theoretical predictions remain valid throughout the range of typical values for N showing that the “large N ” assumption is rather mild.

as a random integer synchronously broadcasted by the nodes). This beacon will then act as a coordination stimulus: the users process it with the aid of some preprogrammed strategies and choose a node accordingly.

In this way, users that keep records of their strategies’ performance and rank them based on the evolutionary scheme of exponential learning (7.30), quickly learn to coordinate their actions and reach an evolutionarily stable state. In figure 6.4 we see that convergence occurs within tens of iterations. Thus, if each iteration is of the order of milliseconds (a reasonable transmission timescale for wideband wireless networks), this corresponds to equilibration times of tens of millisecond. This steady state is also stable in the sense that unilateral deviation is (on average) discouraged: it is a *correlated equilibrium*. Then, to measure the efficiency of users in this setting, we simply examine their distance from the optimal distribution that maximizes their aggregate throughput. Under this metric, we actually see that *exponential learning leads the users to their most efficient equilibrium* (given their preprogramming).

*The benefits
of learning.*

However, since the users’ rationality is bounded (i.e. they can only program and handle a small number of strategies), they will still remain at some distance from the social optimum. This is the *price of (correlated) anarchy* which we calculate with the method of replicas. Somewhat surprisingly, our results show that the price of anarchy does not depend on the nodes’ characteristics, but only on their number – Fig. 7.2.

In fact, our calculations provide an equivalence of our scenario with the minority game of econophysics. For starters, the analytic expression for the price of anarchy that we obtain depends only on the macroscopic parameters N, B, S, M of the system (number of users, nodes, signals and strategies per user respectively). In fact, as far as the first three parameters are concerned,

the price of anarchy is controlled by the (effective) training parameter $\lambda_{\text{eff}} = \frac{M(B-1)}{N}$.

Thanks to the above analysis, we can get quantitative results about the degree of anarchy. For example (Fig. 7.3), we see that blindly adding more nodes to a network is not a panacea: anarchy actually *increases* with the number of nodes because the users are not able to handle the extra complexity and cannot make efficient use of the added resources. On the other hand, if users become more sophisticated and employ a larger number of strategies, anarchy imposes a lesser price on the efficiency level (albeit at a slower rate of convergence to a stable state). In fact, for $S \gg 1$ $\lambda_c \approx 8(\log S)^2 / (B - 1)$; hence, within the replica-symmetric ansatz, the positive values of the price of anarchy are shifted to larger values of $\lambda = M/N$.

Finally, the number of users doesn't have to be in the regime of massively large networks: our results still hold for the much smaller numbers of users that are usually encountered in a local service area (see figure 7.4). However, when the number of users or, equivalently, the number of training signals is too small, our system does lose some of its ergodic properties. In this way, and even though training signals of smaller size generally minimize anarchy, it is generally best not to stray too far away from the critical value $\lambda_c = \frac{\xi^2(S)}{B-1}$ of the training parameter beyond which replica symmetry collapses.

CONCLUSIONS

Over the course of the last decades, the theory of non-cooperative games has rapidly grown into the *lingua franca* for investigations in the competitive side of society, where in its broadest definition, the term “society” might well account for non-human players such as animals, bacteria, or even computers. As a result, game-theoretic ideas have met with remarkable success in analyzing competitive environments in a multitude of diverse fields, ranging from economics and political science to biology and network theory.

A common underlying feature of these considerations is that the complexity of the competing players’ interactions is so high that their deductive capabilities are simply not up to the task of processing these interactions in a timely manner. As a result, players often turn to inductive reasoning instead, hoping that they will become more efficient in the game by reinforcing those strategies that seem to perform better. In this way, we are led to a question that lies at the very heart of game theory: *what kind of behavior can be expected from selfish individuals that employ adaptive learning methods in order to thrive in a competitive environment?*

This question has been at the spearhead of game-theoretic investigations for some time now, and this thesis expounds two important facets of it. One concerns the players’ environment and asks what happens if the game does not remain stationary but is itself subject to constant disturbances (caused by nature or any other unpredictable external factor). The other is geared towards the concrete applications of learning in areas such as the economics of transportation networks, where complexity rules out centralized administration in favour of distributed learning protocols that operate at the user level.

LEARNING IN THE PRESENCE OF NOISE A learning scheme which enjoys particular success in this context is that of “exponential learning”. In this procedure, players update their strategic choices based on a Boltzmann-like model which, akin to logistic fictitious play, assigns exponential weights to a player’s strategies according to their performance. In the deterministic case (where the players’ information is accurate), this procedure is equivalent to the well-understood replicator dynamics of population biology. However, if the players’ payoffs are perturbed by some exogenous random process (arising e.g. from the interference of nature with the game or from inaccuracies in the players’ observations), one obtains a new stochastic version of the replicator dynamics which is radically different from the “aggregate shocks” approach taken in evolutionary games (Fudenberg and Harris, 1992).

In stark contrast to evolutionary theory (Cabrales, 2000; Hofbauer and Imhof, 2009; Imhof, 2005), it was shown that exponential learning enables players to weed out *any* amount of noise and to identify suboptimal choices almost surely: strategies which are not rationally admissible die out at an exponential rate *irrespective of the perturbations’ magnitude*. More to the point,

similar conclusions hold for the game's (asymmetric) evolutionarily stable strategies as well: strict Nash equilibria remain stochastically stable in the face of random fluctuations of arbitrary strength (Mertikopoulos and Moustakas, 2009, 2010b).

THE ROUTING PROBLEM From an applicational perspective, one of the most important optimization criteria of the users of a network (e.g. the Internet) is that of minimising the delays that they experience. Given their concurrent minimisation efforts, this leads to a competitive scenario which has two major differences from a Nash game: delays are typically nonlinear functions of the users' traffic distributions and the leading solution concept is that of Wardrop equilibrium (related, but not equivalent to Nash's principle).

The first significant result here was to show that if users try to minimize their delays by employing a simple learning scheme inspired by the replicator dynamics, then they quickly converge to a Wardrop equilibrium. In fact, even if the users' delays fluctuate wildly due to noise in the network's links, the network's strict Wardrop equilibria remain stochastically stable regardless of the noise level; moreover, users that are patient enough converge to equilibrium almost surely. On the other hand, if there are no such equilibria and the network has no *redundancy* (an important new concept which measures the "linear dependence" between the users' paths), it was shown that the long-term averages of the users' traffic flows still become approximately stable: the users converge to a stationary distribution which is sharply concentrated in the vicinity of interior equilibria (Mertikopoulos and Moustakas, 2010a).

THE MINORITY GAME AND WIRELESS NETWORKS In an effort to further quantify the gains that can be obtained by the advent of learning methods in network design, a considerable part of this thesis was focused on the *Minority Game*, a game which can be used to model the behavior of a (thermodynamically) large number of users that share a limited number of "facilities". This called for an extension of the structure of the minority game to account for non-binary choices and nonlinear payoffs, and in this "multi-choice" generalization it was shown that players converge to a correlated equilibrium whose distance from optimum play was recognized as the *volatility* of the original minority game market model. In fact, by borrowing the method of *replica analysis* from statistical physics, we obtained a sharp analytic bound for this (correlated) *price of anarchy* and showed that it only depends on the macroscopic parameters of the game, e.g. the players-to-facilities ratio (Mertikopoulos and Moustakas, 2007, 2008).

8.1 SOME OPEN QUESTIONS

DIFFUSIONS OVER PRODUCTS OF SIMPLICES One observation that was made painfully obvious during the preparation of this dissertation was that the stochastic replicator dynamics (6.35) represent, in fact, the most general diffusion process that can be written over a product of simplices. Of course, in writing down (6.35) we have made use of an underlying congestion model, delay functions, payoffs, etc. However, *mutatis mutandis*, it is not hard to see that the deterministic drift of (6.35) is enforced by the simplicial structure of the configuration space, while the noise term (and its possible redundancy) is, again, the most general noise term that can be written for diffusion processes of this kind.

So, in essence, what we have here is a first (but important) step into understanding the behavior of these processes. Of course, in the complete generality that this question entails, we still have a long way to go: our use of a potential function was crucial to ensure convergence to the invariant distribution, and we have not been able to do away with the effects of redundancy (which will return to haunt us as the rank of the diffusion matrix, potentially ruling out ellipticity of the diffusion's generator). Nevertheless, the machinery that is already in place *does* suffice to establish local results (e.g. stochastic stability of certain vertices), so, at the very least, this approach will give us some insight into what kind of conditions are needed to obtain results of a more global nature.

LEARNING IN NONLINEAR STRATEGIC ENVIRONMENTS A limiting factor in analysis so far is that, in both the social and physical sciences, there are many competitive situations which do not adhere to the linear structure of Nash games, where strategy spaces span a finite number of "pure" strategies and the players' payoffs are multilinear over their mixed strategies. For instance, many competitive pricing models cannot be captured in this framework, and as we have already seen in the economics of transportation, most delay models are nonlinear as well.

Despite the fundamental importance of this problem, next to nothing is known about learning in such environments, mainly because it is quite difficult to establish learning schemes that respect the structure of non-affine strategy spaces. A promising first step in this direction would be to establish a theory of learning in the class of *concave games* introduced by Rosen (1965), where the strategy spaces are convex domains and the payoffs are concave. Such a theory could then act as a springboard from which to launch an investigation on the interplay between learning and rationality in nonlinear strategic environments.

TEMPORAL ISSUES From an applicational point of view, it is essential to determine the rate at which a learning scheme converges (if at all). To some extent (e.g. Theorems 5.5 and 5.6), we have already considered questions of this nature, but especially in the case of learning in environments that evolve over time (stochastically or otherwise), many issues remain open.

As far as local behavior is concerned, we are still shorn of an eigenvalue analysis for the (suitably linearized) learning dynamics near a game's equilibria. Not only would such an analysis yield a better understanding of the phase portrait near an equilibrium (and possibly highlight new equilibrium selection criteria based on the topological properties of these portraits), but it would also give an excellent estimate of convergence rates within the attraction basins of stable equilibria. Finally, given that a great number of learning schemes (e.g. regret matching) operates in discrete time (which, on many occasions, complicates their analysis), it would be very important to determine the extent to which the "continuity of time" affects rationality, a goal which can be fruitfully explored within the framework of stochastic approximation.

Part IV

APPENDIX

A

REPLICA-THEORETIC CALCULATIONS

A.1 IRRELEVANCE OF NODE STRENGTH AND REDUCTION OF CHOICES

In this section, our main goal will be to calculate the integral expression (7.48) for $\langle \mathcal{Z}^n \rangle$:

$$\langle \mathcal{Z}^n(\beta) \rangle = A^n \int_{\Delta^n} \left\langle e^{-\frac{\beta}{M(B-1)} \sum_{\mu} \sum_m (c_{\mu}^m)^2} \right\rangle e^{N\beta \sum_{\mu} G_{\mu\mu}(p)} dp_1 \cdots dp_N. \quad (7.48)$$

To do this, we begin with the fundamental property of (multi-dimensional) Gaussian integrals:

$$e^{-\frac{q^2}{2}} = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^k} e^{iq \cdot z - \frac{z^2}{2}} dz = \mathbf{E}_z e^{iq \cdot z}, \quad (A.1)$$

where \mathbf{E}_z denotes expectation over a Gaussian random vector z with values in \mathbb{R}^{B-1} and independent components $z_1, \dots, z_{B-1} \sim N(0, 1)$.¹ So, if $z_{\mu}^m = (z_{\mu,1}^m, \dots, z_{\mu,B-1}^m)$, $\mu = 1, \dots, n$, $m \in \mathcal{M}$, is such a vector, we immediately get:

$$\left\langle e^{-\frac{\beta}{M(B-1)} \sum_{\mu} \sum_m (c_{\mu}^m)^2} \right\rangle = \mathbf{E}_{z_{\mu}^m} \left\langle e^{i \sum_i \sum_s \sum_m x_{is}^m \cdot c_{is}^m} \right\rangle \quad (A.2)$$

where we have set $x_{is}^m = \sqrt{\frac{2\beta}{M(B-1)}} \sum_{\mu} p_{is\mu} z_{\mu}^m \in \mathbb{R}^{B-1}$.

Therefore, by the independence of the c_i 's (recall also (7.12)), we will be able to obtain the average $\langle \cdot \rangle$ of (A.2) over the matrices c by computing the characteristic function $\langle e^{ix \cdot q} \rangle$ for only one of them. To that end, the following lemma will prove particularly handy – in fact, it is also one of the main reasons behind the definition of y -appropriate simplices:

Lemma A.1. *Let $y \in \text{Int}(\Delta_B)$ and let $\mathcal{B} = \{q_r\}_{r=1}^B$ be a y -appropriate simplex in \mathbb{R}^{B-1} . If $x \in \mathbb{R}^{B-1}$ and q is a random vector with $\mathbf{P}(q = q_r) = y_r$, then:*

$$\langle e^{ix \cdot q} \rangle = e^{-\frac{x^2}{2}} + \mathcal{O}(|x|^3), \quad (A.3)$$

with $\langle \cdot \rangle$ denoting expectation over q .

Proof. Expanding the exponential $\langle \exp(ix \cdot q) \rangle$ readily yields:

$$\begin{aligned} \langle e^{ix \cdot q} \rangle &= \left\langle 1 + ix \cdot q - \frac{1}{2} (x \cdot q)^2 + \mathcal{O}(|x|^3) \right\rangle \\ &= 1 + ix \cdot \sum_r y_r q_r - \frac{1}{2} \sum_r y_r (q_r \cdot x)^2 + \mathcal{O}(|x|^3) \\ &= 1 - \frac{1}{2} x^2 + \mathcal{O}(|x|^3) = e^{-\frac{1}{2} x^2} + \mathcal{O}(|x|^3) \end{aligned} \quad (A.4)$$

¹ This property is also known as the *Hubbard-Stratonovich* transformation.

the last equality stemming from lemmas 7.2 and 7.4. \square

In our case, $|x_{is}^m| = \mathcal{O}(M^{-\frac{1}{2}})$, so if we apply the previous lemma to each of the random vectors c_{is}^m , the average of (A.2) becomes (to leading order in N):

$$\begin{aligned} \left\langle \exp \left(i \sum_{i,s,m} x_{is}^m \cdot c_{is}^m \right) \right\rangle &= \exp \left(-\frac{1}{2} \sum_{i,s,m} (x_{is}^m)^2 \right) + \mathcal{O}(1/N^{3/2}) \\ &\sim \exp \left(-\frac{\beta}{\lambda(B-1)} \sum_m \sum_{\mu,\nu} G_{\mu\nu}(p) z_{\mu}^m \cdot z_{\nu}^m \right) \end{aligned} \quad (\text{A.5})$$

where $\lambda = M/N$ is the game's training parameter. In this way, we may now integrate over the auxiliary variables z_{μ}^m to obtain:

$$\begin{aligned} \mathbf{E}_{z_{\mu}^m} \left\langle e^{-\frac{\beta}{M(B-1)} \sum_{\mu} \sum_m (c_{\mu}^m)^2} \right\rangle &\sim \int_{\mathbb{R}^{nM(B-1)}} \exp \left(-\frac{1}{2} \sum_{m,k} \sum_{\mu,\nu} J_{\mu\nu}(p) z_{\mu,k}^m z_{\nu,k}^m \right) \tilde{d}z \\ &= \left(\int_{\mathbb{R}^n} \exp \left(-\frac{1}{2} \sum_{\mu,\nu} J_{\mu\nu}(p) w_{\mu} w_{\nu} \right) \tilde{d}w \right)^{M(B-1)} \\ &= \exp \left(-\frac{M(B-1)}{2} \log \det(\mathbf{J}(p)) \right), \end{aligned} \quad (\text{A.6})$$

where we have introduced the $n \times n$ matrix $\mathbf{J}(p) = \mathbf{I} + \frac{2\beta}{\lambda(B-1)} \mathbf{G}(p)$ and used the Gaussian identity:

$$\int_{\mathbb{R}^n} \exp \left(-\frac{1}{2} \sum_{\mu,\nu} J_{\mu\nu} w_{\mu} w_{\nu} \right) \tilde{d}w = |\det(\mathbf{J})|^{-1/2} \quad (\text{A.7})$$

(here, tildes as in $\tilde{d}w$ denote Lebesgue measure normalized by $\sqrt{2\pi}$).

So, after these calculations, equation (7.48) finally becomes:

$$\begin{aligned} \frac{1}{A^n} \langle z^n(\beta) \rangle &\sim \int_{\Delta^n} e^{N\beta \left\{ \text{tr}(\mathbf{G}(p)) - \frac{\lambda(B-1)}{2\beta} \log \det \left(\mathbf{I} + \frac{2\beta}{\lambda(B-1)} \mathbf{G}(p) \right) \right\}} dp_1 \cdots dp_n \end{aligned} \quad (\text{A.8})$$

Clearly, this last expression is independent of y , a fact which proves Theorem 7.16. In addition, we observe that (A.8) remains invariant when we pass from the game \mathfrak{G} to its binary reduction $\mathfrak{G}_{\text{eff}}$ with the rescaled training parameter $\lambda_{\text{eff}} = \lambda(B-1)$, thus proving Theorem 7.17 as well.

A.2 AN ANALYTIC EXPRESSION FOR THE PRICE OF ANARCHY

To actually calculate the integral (A.8), we will introduce a number of δ -functions to isolate the integration over the strategy profiles p_{is} . First, we will use the integral representation of δ -functions:

$$\delta(\mathbf{Q} - \mathbf{G}(p)) = \left(\frac{N\beta}{2\pi} \right)^{n^2} \int e^{iN\beta \sum_{\mu,\nu} k_{\mu\nu} (Q_{\mu\nu} - G_{\mu\nu}(p))} \prod_{\mu,\nu} dk_{\mu\nu} \quad (\text{A.9})$$

In this way, the integral in (A.8) becomes:

$$\int e^{-N\beta \left[\frac{\lambda(B-1)}{2\beta} \log \det \left(\mathbf{I} + \frac{2\beta}{\lambda(B-1)} \mathbf{Q} \right) - \text{tr}(\mathbf{Q}) - i \sum_{\mu,\nu} k_{\mu\nu} (Q_{\mu\nu} - G_{\mu\nu}(p)) \right]} d\sigma, \quad (\text{A.10})$$

where $d\sigma = \prod_{\mu} dp_{\mu} \times \prod_{\mu,\nu} dk_{\mu\nu} \times \prod_{\mu,\nu} dQ_{\mu\nu}$ is the standard product measure on $\Delta^n \times \mathbb{R}^{n^2} \times \mathbb{R}^{n^2}$.

However, p only appears in the last term of the above expression, which can hence be integrated separately to yield:

$$\begin{aligned} & \int_{\Delta^n} \exp\left(-iN\beta \sum_{\mu,\nu} k_{\mu\nu} G_{\mu\nu}(p)\right) dp_1 \cdots dp_n \\ &= \prod_i \int_{\Delta(\mathcal{S})^n} \exp\left(-i\beta \sum_{\mu,\nu} k_{\mu\nu} \sum_s p_{is\mu} p_{is\nu}\right) \prod_{s,\mu} dp_{is\mu} \\ &= \exp\left(N \log \int_{\Delta(\mathcal{S})^n} e^{-i\beta \sum_{\mu,\nu} k_{\mu\nu} \sum_s p_{s\mu} p_{s\nu}} \prod_{s,\mu} dp_{s\mu}\right) \quad (\text{A.11}) \end{aligned}$$

(recall that $G_{\mu\nu}(p) = \frac{1}{N} \sum_i \sum_s p_{is\mu} p_{is\nu}$ and $\Delta^n = (\Delta_{\mathcal{S}})^{N \times n}$). So, by descending steeply to the limit $N \rightarrow \infty$ in order to perform the asymptotic integration over \mathbf{Q} and \mathbf{k} , we find:

$$\begin{aligned} \frac{1}{N} \log \langle \mathcal{Z}^n(\beta) \rangle &\sim -\beta \left(n + \frac{\lambda(B-1)}{2\beta} \log \det \left(\mathbf{I} + \frac{2\beta}{\lambda(B-1)} \mathbf{Q} \right) - \text{tr}(\mathbf{Q}) - \right. \\ &\quad \left. -i \sum_{\mu,\nu} k_{\mu\nu} Q_{\mu\nu} - \frac{1}{\beta} \log \int_{\Delta^n} e^{-i\beta \sum_{\mu,\nu} k_{\mu\nu} \sum_s p_{s\mu} p_{s\nu}} \prod_{s,\mu} dp_{s\mu} \right) \equiv -\beta \Lambda \quad (\text{A.12}) \end{aligned}$$

where the matrices \mathbf{Q} and \mathbf{k} have been chosen so as to extremize the function Λ within the brackets.

To proceed from here, we will assume a very powerful conjecture (de Martino and Marsili, 2001; Mézard et al., 1987):

Conjecture A.2 (Replica Symmetry). *The saddle-points of Λ are of the form:*

$$Q_{\mu\nu} = q + (Q - q) \delta_{\mu\nu}; \quad k_{\mu\nu} = i\lambda\beta \frac{B-1}{2} (r + (R - r) \delta_{\mu\nu}). \quad (\text{A.13})$$

In other words, we seek saddle-point matrices that are symmetric in replica space (the scaling factors are there for future convenience). Of course, since there is no a priori reason for the replicas to converge to the same state, this is quite the leap of faith. In fact, it has been shown by de Martino and Marsili (2001) that this assumption is not strictly valid and the authors perform the first step of *replica symmetry breaking* (1-RSB) in order to locate saddle points within the setting of the Parisi solution. Still, it is also shown there that replica-symmetric saddle points yield a value for Λ that does not differ significantly from the 1-RSB analysis, so the assumption inherent in A.2 is not expected to introduce a significant error to our computations.

Under this ansatz, we obtain:

$$\begin{aligned} \Lambda &= n + \frac{\lambda(B-1)}{2\beta} \log \det \left(\frac{2\beta}{\lambda} \frac{q}{B-1} + \left(1 + \frac{2\beta}{\lambda} \frac{Q-q}{B-1}\right) \delta_{\mu\nu} \right) - nQ \\ &\quad + n\lambda\beta \frac{B-1}{2} (QR - qr) + n^2 \lambda\beta \frac{B-1}{2} qr \\ &\quad - \frac{1}{\beta} \log \int_{\Delta^n} e^{\lambda\beta^2 \frac{B-1}{2} \left((R-r) \sum_{\mu} \mathbf{p}_{\mu}^2 + r (\sum_{\mu} \mathbf{p}_{\mu})^2 \right)} \prod_{s,\mu} dp_{s\mu}, \quad (\text{A.14}) \end{aligned}$$

where, as before, p_{μ} is the generic profile $(p_{1,\mu}, \dots, p_{S,\mu})$ in the μ -th replica. Therefore, by noting that $\det(q + p\delta_{\mu\nu}) = p^n \left(1 + n \frac{q}{p}\right)$, we see that the second term of (A.14) becomes:

$$\frac{q}{1+\chi} + \frac{\lambda}{2\beta} (B-1) \log(1 + \chi) + o(n), \quad (\text{A.15})$$

where $\chi = \frac{2\beta}{\lambda} \frac{Q-q}{B-1}$.

Now, if we use the Hubbard-Stratonovich transformation (A.2) with a canonical Gaussian variable z in \mathbb{R}^S , the last term of (A.14) may be written as:

$$e^{r\lambda\beta^2 \frac{B-1}{2} (\sum_{\mu} \mathbf{p}_{\mu})^2} = \mathbf{E}_z e^{-\beta \sqrt{r\lambda(B-1)} z \cdot \sum_{\mu} \mathbf{p}_{\mu}}. \quad (\text{A.16})$$

For notational convenience, we also let:

$$V(z, \mathbf{p}) = \sqrt{r\lambda(B-1)} z \cdot \mathbf{p} - \lambda\beta \frac{B-1}{2} (R-r) \mathbf{p}^2. \quad (\text{A.17})$$

Then, in this way, the integral of (A.14) becomes:

$$\begin{aligned} \log \mathbf{E}_z \int_{\Delta(S)^n} e^{-\beta \sum_{\mu} V(z, \mathbf{p}_{\mu})} \prod_{s, \mu} dp_{s\mu} \\ = \log \mathbf{E}_z \exp \left(n \log \int_{\Delta_S} e^{-\beta V(z, \mathbf{p})} dp_1 \cdots dp_S \right) \\ = n \mathbf{E}_z \log \int_{\Delta_S} e^{-\beta V(z, \mathbf{p})} dp + o(n), \end{aligned} \quad (\text{A.18})$$

with $dp = \prod_{s=1}^S dp_s$ denoting here Lebesgue measure on $\Delta(S)$.

From (7.45) and the premises of replica continuity (Conjecture 7.15), what we really need to calculate is $\Lambda_0 = \lim_{n \rightarrow 0} \frac{1}{n} \Lambda$. Hence, by ignoring all terms of (A.14) that are $o(n)$, we obtain:

$$\begin{aligned} \Lambda_0 = 1 + \frac{q}{1+\chi} + \frac{\lambda}{2\beta} (B-1) \log(1+\chi) - Q \\ + \lambda\beta \frac{B-1}{2} (QR - qr) - \frac{1}{\beta} \mathbf{E}_z \left[\log \int_{\Delta_S} e^{-\beta V(z, \mathbf{p})} dp \right] \end{aligned} \quad (\text{A.19})$$

where Q, q, R, r have been chosen so as to satisfy the *replica-symmetric* saddle-point equations: $\frac{\partial \Lambda_0}{\partial Q} = 0$, etc.

The first thing that can be shown is that both $Q - q$ and $R - r$ are of order $\mathcal{O}(1/\beta)$, i.e. χ must remain finite as $\beta \rightarrow \infty$. So, to obtain a solution in this limiting case, we will again take advantage of the steepest descent method to rid ourselves of the integrals that appear in $\frac{\partial \Lambda_0}{\partial R} = 0$ and $\frac{\partial \Lambda_0}{\partial r} = 0$. In its turn, this leads us to consider the vertex $p^*(z)$ of $\Delta(S)$ which minimizes the harmonic function $V(z, \cdot)$ and, hence, we finally obtain the replica-symmetric solutions:

$$\begin{aligned} Q &\sim \phi & R &= r + \frac{1}{\beta} \frac{2}{\lambda(B-1)} \frac{\chi}{1+\chi} \\ q &\sim \phi + \frac{1}{\beta} \frac{\zeta}{\sqrt{\lambda(B-1)}r} & r &= \frac{4}{\lambda^2(B-1)^2} \frac{1}{(1+\chi)^2} \end{aligned} \quad (\text{A.20})$$

where $\phi = \mathbf{E}_z[\mathbf{p}_*^2(z)]$ and $\zeta = \mathbf{E}_z[\mathbf{p}_*(z) \cdot z]$.

Now, if we let $\beta \rightarrow \infty$ and substitute (A.20) in (A.19), we get:

$$\Lambda_0 = 1 - \phi + \phi \left(1 + \zeta(S) / \sqrt{\phi\lambda(B-1)} \right)^2, \quad (\text{A.21})$$

and, after some elementary geometry, we also obtain $\phi = 1$ and:

$$\begin{aligned}\zeta &\equiv \zeta(S) = \mathbf{E}_z[\min\{z_1, \dots, z_S\}] \\ &= \frac{S}{2^{S-1}} \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} z e^{-z^2} \operatorname{erfc}^{S-1}(z) dz. \quad (\text{A.22})\end{aligned}$$

Thus, for finite χ (i.e. for $\lambda \geq \lambda_c = \zeta^2(S)/(B-1)$), we finally acquire expression (7.51) for the game's *price of anarchy*:

$$R(\mathfrak{G}) \sim \Lambda_0 \sim \Theta(\lambda - \lambda_c) \left(1 - \sqrt{\lambda_c/\lambda}\right)^2. \quad (7.51)$$

BIBLIOGRAPHY

- Ali Khan, M., 1986: Equilibrium points of nonatomic games over a Banach space. *Transactions of the American Mathematical Society*, **293** (2), 737–749. (Cited on page 25.)
- Altman, E., T. Boulogne, R. el Azouzi, T. Jiménez, and L. Wynter, 2006: A survey on networking games in telecommunications. *Computers and Operations Research*, **33** (2), 286–311. (Cited on page 4.)
- Arnold, L., 1974: *Stochastic Differential Equations: Theory and Applications*. Wiley. (Cited on page 51.)
- Aumann, R. J., 1974: Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, **1** (1), 67–96. (Cited on pages 3, 26, 27, 29, 127, and 131.)
- Aumann, R. J., 1987: Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, **55** (1), 1–18. (Cited on pages 27, 29, 30, 127, 131, and 132.)
- Beckmann, M., C. B. McGuire, and C. Winsten, 1956: *Studies in the Economics of Transportation*. Yale University Press. (Cited on pages 97, 98, 104, and 105.)
- Benaïm, M., 1999: Dynamics of stochastic approximation algorithms. *Séminaire de probabilités de Strasbourg*, **33**. (Cited on page 57.)
- Benaïm, M., J. Hofbauer, and W. H. Sandholm, 2008: Robust permanence and impermanence for stochastic replicator dynamics. *Journal of Biological Dynamics*, **2** (2), 180–195. (Cited on page 98.)
- Bhattacharya, R. N., 1978: Criteria for recurrence and existence of invariant measures for multidimensional diffusions. *The Annals of Probability*, **6**, 541–553. (Cited on pages 51, 52, 92, 93, and 121.)
- Bomze, I. M., 1986: Non-cooperative two-person games in biology: A classification. *International Journal of Game Theory*, **15**, 31–57. (Cited on page 38.)
- Bonneau, N., E. Altman, M. Debbah, and G. Caire, 2005: An evolutionary game perspective to ALOHA with power control. *Proceedings of the 19th International Teletraffic Congress*. (Cited on page 127.)
- Bonneau, N., M. Debbah, E. Altman, and A. Hjørungnes, 2007: Wardrop equilibrium for CDMA systems. *RAWNET '07: Proceedings of the 3rd workshop on Resource Allocation in Wireless Networks*. (Cited on page 127.)
- Börgers, T. and R. Sarin, 1997: Learning through reinforcement and replicator dynamics. *Journal of Economic Theory*, **77**, 1–14. (Cited on page 108.)
- Braess, D., 1968: Über ein paradoxon der verkehrsplanung. *Unternehmensforschung*, **12**, 258–268. (Cited on page 106.)
- Braess, D., A. Nagurney, and T. Wakolbinger, 2005: On a paradox of traffic planning. *Transportation Science*, **39** (4), 446–450. (Cited on page 101.)

- Brian Arthur, W., 1994: Inductive reasoning and bounded rationality (the El Farol problem). *American Economic Review: Papers and Proceedings*, **84**, 406–411. (Cited on pages 27, 69, and 128.)
- Bringhurst, R., 2002: *The Elements of Typographic Style*. Hartley & Marks. (Cited on page 167.)
- Brown, G. W. and J. von Neumann, 1950: Solutions of games by differential equations. *Contributions to the Theory of Games I*, H. W. Kuhn and A. W. Tucker, Eds., Princeton University Press, Princeton, NJ, *Annals of Mathematics Studies*, Vol. 24. (Cited on page 36.)
- Cabrales, A., 2000: Stochastic replicator dynamics. *International Economic Review*, **41** (2), 451–81. (Cited on pages 7, 8, 55, 62, 64, 65, 75, 98, 114, and 147.)
- Cavagna, A., 1999: Irrelevance of memory in the minority game. *Physical Review E*, **59**, 3783–3786. (Cited on pages 128 and 138.)
- Challet, D. and Y.-C. Zhang, 1997: On the minority game: analytical and numerical studies. *Physica A*, **246**, 407–418. (Cited on pages 128 and 133.)
- Chandra, R., P. Bahl, and P. Bahl, 2004: MultiNet: Connecting to multiple IEEE 802.11 networks using a single wireless card. *Infocom '04: Proceedings of the 23rd conference of the IEEE Communications Society*. (Cited on page 127.)
- Coolen, A. C. C., 2004: *The mathematical theory of minority games*. Oxford University Press. (Cited on page 142.)
- Cressman, R., 1992: *The stability concept of evolutionary game theory*. Springer-Verlag, Berlin. (Cited on page 31.)
- Cressman, R., 2003: *Evolutionary Dynamics and Extensive Form Games*. The MIT Press. (Cited on page 6.)
- Cressman, R., 2005: Stability of the replicator equation with continuous strategy space. *Mathematical Social Sciences*, **50** (2), 127–147. (Cited on pages 19 and 26.)
- Dafermos, S. C. and F. T. Sparrow, 1969: The traffic assignment problem for a general network. *Journal of Research of the National Bureau of Standards*, **73B** (2), 91–118. (Cited on pages 97 and 104.)
- Daskalakis, C., 2008: The complexity of Nash equilibria. Ph.D. thesis, University of California, Berkeley. (Cited on pages 1 and 26.)
- Daskalakis, C., P. W. Goldberg, and C. Papadimitriou, 2006: The complexity of computing a Nash equilibrium. *STOC '06: Proceedings of the 38th annual ACM symposium on the Theory of Computing*. (Cited on page 26.)
- de Martino, A. and M. Marsili, 2001: Replica symmetry breaking in the minority game. *Journal of Physics A: Mathematical and General*, **34** (12), 2525–2537, URL <http://stacks.iop.org/0305-4470/34/2525>. (Cited on pages 141 and 155.)
- Dembo, A. and O. Zeitouni, 1998: *Large Deviations Techniques and Applications*. Springer-Verlag, Berlin. (Cited on page 141.)

- Demestichas, P., G. Vivier, K. El-Khazen, and M. Theologou, 2004: Evolution in wireless systems management concepts: from composite radio environments to reconfigurability. *IEEE Communications Magazine*, **42** (5), 90–98. (Cited on page 127.)
- Durrett, R., 1996: *Stochastic Calculus*. CRC Press, Boca Raton, FL. (Cited on pages 91, 93, 118, and 121.)
- Einstein, A., 1905: Über die von molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annullen der Physik*, **322** (8), 549–560. (Cited on page 40.)
- Einstein, A., 1956: *Investigations on the Theory of Brownian Movement*. Dover, New York. (Cited on page 40.)
- Felegyhazi, M., M. Cagalj, D. Dufour, and J.-P. Hubaux, 2007: Border games in cellular networks. *Infocom '07: Proceedings of the 26th conference of the IEEE Communications Society*. (Cited on page 127.)
- Fischer, S. and B. Vöcking, 2004: On the evolution of selfish routing. *Proceedings of the 12th European Symposium on Algorithms*. (Cited on pages 98, 104, and 112.)
- Flood, M., 1951: A preference experiment. Tech. rep., The RAND corporation. (Cited on page 2.)
- Friedman, J. W., 1971: A non-cooperative equilibrium for supergames. *Review of Economic Studies*, **38**, 1–12. (Cited on page 6.)
- Fudenberg, D. and C. Harris, 1992: Evolutionary dynamics with aggregate shocks. *Journal of Economic Theory*, **57** (2), 420–441. (Cited on pages 7, 55, 60, 75, 98, 99, 113, and 147.)
- Fudenberg, D. and D. K. Levine, 1998: *The Theory of Learning in Games*, MIT Press Series on Economic Learning and Social Evolution, Vol. 2. The MIT Press. (Cited on pages 4, 6, 15, 57, 58, and 125.)
- Fudenberg, D. and J. Tirole, 1991: *Game Theory*. The MIT Press. (Cited on pages 15 and 30.)
- Gikhman, I. I. and A. V. Skorokhod, 1971: *Stochastische Differentialgleichungen*. Akademie-Verlag. (Cited on pages 51, 52, 69, and 116.)
- Halldorsson, M. M., J. Y. Halpern, L. E. Li, and V. S. Mirrokni, 2004: On spectrum sharing games. *PODC' 04: Proceedings of the 23d annual ACM SIGACT-SIGOPS symposium on Principles of Distributed Computing*. (Cited on page 127.)
- Harsányi, J., 1967-1968: Games with incomplete information played by Bayesian players I, II, and III. *Management Science*, **14**, 159–182, 320–334, 486–502. (Cited on pages 3 and 15.)
- Hart, S., 2005: Adaptive heuristics. *Econometrica*, **73** (5), 1401–1430. (Cited on page 4.)
- Hida, T., 1980: *Brownian Motion*. Springer-Verlag, Berlin. (Cited on pages 42 and 43.)

- Hofbauer, J. and L. A. Imhof, 2009: Time averages, recurrence and transience in the stochastic replicator dynamics. *The Annals of Applied Probability*, **19** (4), 1347–1368. (Cited on pages 7, 56, 61, 62, 65, 68, 72, 98, 114, and 147.)
- Hofbauer, J., J. Oechssler, and F. Riedel, 2009: Brown-von Neumann-Nash dynamics: the continuous strategy case. *Games and Economic Behavior*, **65** (2), 406–429. (Cited on pages 19 and 26.)
- Hofbauer, J. and K. Sigmund, 1998: *Evolutionary Games and Population Dynamics*. Cambridge University Press. (Cited on page 4.)
- Hofbauer, J. and K. Sigmund, 2003: Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, **40** (4), 479–519. (Cited on pages 4, 6, and 33.)
- Imhof, L. A., 2005: The long-run behavior of the stochastic replicator dynamics. *The Annals of Applied Probability*, **15** (1B), 1019–1045. (Cited on pages 7, 8, 55, 56, 61, 62, 64, 72, 73, 75, 92, 93, 98, 114, 122, and 147.)
- Itô, K., 1944: Stochastic integral. *Proceedings of the Imperial Academy of Tokyo*, **20**, 519–524. (Cited on pages 39, 40, and 47.)
- Itô, K., 1946: On a stochastic integral equation. *Proceedings of the Imperial Academy of Tokyo*, **22**, 32–35. (Cited on pages 39 and 40.)
- Itô, K., 1951: On stochastic differential equations. *Memoirs of the American Mathematical Society*, **4**, 1–51. (Cited on pages 39 and 40.)
- Johanssen, W., 1903: Om arvelighed i samfund og i rene linier. *Proceedings of the Royal Danish Academy of Sciences and Letters*, **3**, 247–270. (Cited on page 5.)
- Karatzas, I. and S. E. Shreve, 1998: *Brownian Motion and Stochastic Calculus*. Springer-Verlag, Berlin. (Cited on page 63.)
- Khas'minskii, R. Z., 1960: Ergodic properties of recurrent diffusion processes and stabilization of the solution of the cauchy problem for parabolic equations. *Theory of Probability and Applications*, **5**, 179–196. (Cited on page 52.)
- Khas'minskii, R. Z. and N. Potsepun, 2006: On the replicator dynamics behavior under Stratonovich type random perturbations. *Stochastic Dynamics*, **6**, 197–211. (Cited on page 62.)
- Klebaner, F. C., 2005: *Introduction to Stochastic Calculus with Applications*. Imperial College Press, London. (Cited on pages 46 and 48.)
- Knight, F. B., 1981: *Essentials of Brownian Motion and Diffusion*, Mathematical Surveys and Monographs, Vol. 18. American Mathematical Society. (Cited on pages 41 and 42.)
- Kogiantis, A. G., N. Joshi, and O. Sunay, 2001: On transmit diversity and scheduling in wireless packet data. *ICC '01: Proceedings of the 2001 International Conference on Communications*. (Cited on page 129.)
- Koutsoupias, E. and C. Papadimitriou, 1999: Worst-case equilibria. *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, 404–413. (Cited on pages 97, 106, 128, 136, and 137.)
- Kuo, H.-H., 2006: *Introduction to Stochastic Integration*. Springer, Berlin. (Cited on pages 39, 41, 42, 43, 44, 45, 46, 47, and 48.)

- Lai, L. and H. el Gamal, 2006: Fading multiple access channels: a game-theoretic perspective. *ISIT '06: Proceedings of the 2006 International Symposium on Information Theory*. (Cited on page 130.)
- Lamperti, J., 1977: *Stochastic Processes*. Applied Mathematical Sciences, Springer-Verlag, Berlin. (Cited on page 64.)
- Leslie, D. S. and E. J. Collins, 2003: Convergent multiple-timescales reinforcement learning algorithms in normal form games. *The Annals of Applied Probability*, **13** (4), 1231–1251. (Cited on page 35.)
- Lewontin, R. C., 1974: *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York, NY. (Cited on page 5.)
- MacKenzie, A. B. and S. B. Wicker, 2001: Selfish users in ALOHA: A game theoretic approach. *IEEE Communications Magazine*, **126**, 126. (Cited on page 127.)
- Marsili, M. and D. Challet, 2001: Continuum time limit and stationary states of the minority game. *Physical Review E*, **64**. (Cited on pages 139 and 141.)
- Marsili, M., D. Challet, and R. Zecchina, 2000: Exact solution of a modified El Farol's bar problem: Efficiency and the role of market impact. *Physica A*, **280**, 522–553. (Cited on pages 35, 69, 128, 132, 133, 138, 141, and 143.)
- Maruyama, G. and H. Tanaka, 1959: Ergodic property of N -dimensional recurrent Markov processes. *Memoirs of the Faculty of Science of Kyushu University, Series A*, **13**, 157–72. (Cited on page 52.)
- Maynard Smith, J., 1974: The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology*, **47** (1), 209–221. (Cited on page 27.)
- Maynard Smith, J. and G. R. Price, 1973: The logic of animal conflict. *Nature*, **246**, 15–18. (Cited on pages 3, 4, and 32.)
- Mertikopoulos, P. and A. L. Moustakas, 2007: The simplex game: Can selfish users learn to operate efficiently in wireless networks? *ValueTools '07: Proceedings of the 2nd International Conference on Performance Evaluation Methodologies and Tools*. (Cited on page 148.)
- Mertikopoulos, P. and A. L. Moustakas, 2008: Correlated anarchy in overlapping wireless networks. *IEEE J. Sel. Areas Commun.*, **26** (7), 1160–1169. (Cited on page 148.)
- Mertikopoulos, P. and A. L. Moustakas, 2009: Learning in the presence of noise. *GameNets '09: Proceedings of the 1st International Conference on Game Theory for Networks*. (Cited on pages 55, 98, 113, and 148.)
- Mertikopoulos, P. and A. L. Moustakas, 2010a: Balancing traffic in networks: redundancy, learning and the effect of stochastic fluctuations. *submitted*. (Cited on page 148.)
- Mertikopoulos, P. and A. L. Moustakas, 2010b: The emergence of rational behavior in the presence of stochastic perturbations. *The Annals of Applied Probability*, **20** (4), 1359–1388. (Cited on pages 55, 98, 113, 124, and 148.)
- Meshkati, F., M. Chiang, V. Poor, and S. C. Schwartz, 2006: A game-theoretic approach to energy-efficient power control in multicarrier CDMA systems. *IEEE Journal on selected areas in communications*, **24**, 1115–1129. (Cited on page 127.)

- Meshkati, F., V. Poor, and S. C. Schwartz, 2005: An energy-efficient approach to power control and receiver design in wireless data networks. *IEEE Transactions on Communications*, **53** (11), 1885. (Cited on page 127.)
- Mézard, M., G. Parisi, and M. A. Virasoro, 1987: *Spin Glass Theory and Beyond*. World Scientific. (Cited on pages 141, 142, and 155.)
- Milchtaich, I., 1996: Congestion games with player-specific payoff functions. *Games and Economic Behavior*, **13**, 111–124. (Cited on page 69.)
- Milchtaich, I., 2000: Generic uniqueness of equilibrium in large crowding games. *Mathematics of Operations Research*, **25** (3), 349–364. (Cited on page 25.)
- Monderer, D. and L. S. Shapley, 1996: Potential games. *Games and Economic Behavior*, **14** (1), 124 – 143. (Cited on pages 21, 23, and 104.)
- Nachbar, J. H., 1990: Evolutionary selection dynamics in games. *International Journal of Game Theory*, **19**, 59–89. (Cited on pages 6 and 38.)
- Nash, J. F., 1950a: Equilibrium points in n -person games. *Proceedings of the National Academy of Sciences (USA)*, Vol. 36, 48–49. (Cited on pages 15 and 25.)
- Nash, J. F., 1950b: Non-cooperative games. Ph.D. thesis, Princeton University. (Cited on pages 15, 25, and 27.)
- Nash, J. F., 1951: Non-cooperative games. *The Annals of Mathematics*, **54** (2), 286–295. (Cited on pages 10, 15, 17, and 25.)
- Nishimori, H., 2001: *Statistical Physics of Spin Glasses and Information Processing*. Oxford University Press. (Cited on page 141.)
- Nowak, M. and K. Page, 2002: Unifying evolutionary dynamics. *Journal of Theoretical Biology*, **219**, 93–98. (Cited on page 6.)
- Øksendal, B., 2007: *Stochastic Differential Equations*. 6th ed., Springer-Verlag, Berlin. (Cited on pages 39, 41, 45, 46, 47, 48, 49, 50, 51, and 93.)
- Palomar, D. P., J. M. Cioffi, and M. Lagunas, 2003: Uniform power allocation in MIMO channels: a game-theoretic approach. *IEEE Transactions on Information Theory*, **49** (7), 1707. (Cited on page 127.)
- Papadimitriou, C., 2001: Algorithms, games, and the internet. *Proceedings of the 33rd Annual ACM Symposium on the Theory of Computing*. (Cited on page 97.)
- Revuz, D. and M. Yor, 1999: *Continuous Martingales and Brownian Motion*. 3d ed., Springer, Berlin. (Cited on page 42.)
- Risken, H., 1989: *The Fokker-Planck Equation*. 2d ed., Springer, Berlin. (Cited on page 51.)
- Ritzberger, K. and K. Vogelsberger, 1990: The Nash field. Tech. rep., Institute for Advanced Studies, Vienna. (Cited on page 79.)
- Rosen, J. B., 1965: Existence and uniqueness of equilibrium points for concave n -person games. *Econometrica*, **33** (3), 520–534. (Cited on pages 17, 25, 103, and 149.)

- Rosenthal, R. W., 1973: A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, **2**, 65–67. (Cited on pages [21](#), [23](#), [98](#), and [104](#).)
- Roughgarden, T. and E. Tardos, 2002: How bad is selfish routing? *Journal of the ACM*, **49** (2), 236–259. (Cited on pages [97](#), [100](#), [104](#), [106](#), and [123](#).)
- Roughgarden, T. and E. Tardos, 2004: Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior*, **47** (2), 389–403. (Cited on pages [97](#), [100](#), and [105](#).)
- Samuelson, L. and J. Zhang, 1992: Evolutionary stability in asymmetric games. *Journal of Economic Theory*, **57**, 363–391. (Cited on pages [6](#), [37](#), and [65](#).)
- Sandholm, W. H., 2001: Potential games with continuous player sets. *Journal of Economic Theory*, **97**, 81–108. (Cited on pages [20](#), [26](#), [35](#), [36](#), [75](#), [77](#), [97](#), [98](#), [104](#), [108](#), [112](#), and [125](#).)
- Sandholm, W. H., 2005: Excess payoff dynamics and other well-behaved evolutionary dynamics. *Journal of Economic Theory*, **124**, 149–170. (Cited on pages [36](#) and [75](#).)
- Sandholm, W. H., 2010: *Population Games and Evolutionary Dynamics*. MIT Press, Cambridge, MA. (Cited on page [33](#).)
- Schmeidler, D., 1973: Equilibrium points of nonatomic games. *Journal of Statistical Physics*, **7** (4), 295–300. (Cited on pages [18](#) and [25](#).)
- Selten, R., 1980: A note on evolutionarily stable strategies in asymmetric animal conflicts. *Journal of Theoretical Biology*, **84** (93-101). (Cited on page [32](#).)
- Shakkottai, S., E. Altman, and A. Kumar, 2007: Multihoming of users to access points in WLANs: A population game perspective. *IEEE Journal on selected areas in communications*, **25** (6), 1207. (Cited on pages [127](#) and [129](#).)
- Shiryayev, A. N., 1995: *Probability*. 2d ed., Springer, Berlin. (Cited on page [42](#).)
- Stratonovich, R. L., 1966: A new representation for stochastic integrals and equations. *SIAM Journal on Control and Optimization*, **4**, 362–371. (Cited on page [40](#).)
- Swinkels, J. M., 1992: Evolutionary stability with equilibrium entrants. *Journal of Economic Theory*, **57** (2), 306–332. (Cited on page [31](#).)
- Taylor, P. D., 1979: Evolutionarily stable strategies with two types of player. *Journal of Applied Probability*, **16**, 76–83. (Cited on page [31](#).)
- Taylor, P. D. and L. B. Jonker, 1978: Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, **40** (1-2), 145–156. (Cited on pages [vii](#), [4](#), and [33](#).)
- Tversky, A., 2004: *Preference, Belief, and Similarity: Selected Writings*. The MIT Press. (Cited on page [2](#).)
- von Neumann, J. and O. Morgenstern, 1944: *Theory of Games and Economic Behavior*. Princeton University Press. (Cited on page [15](#).)
- Wardrop, J. G., 1952: Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers, Part II*, Vol. 1, 325–78. (Cited on page [97](#).)

Weibull, J. W., 1995: *Evolutionary Game Theory*. MIT Press, Cambridge, MA. (Cited on pages [3](#), [4](#), [6](#), [15](#), [32](#), [33](#), [37](#), [38](#), [70](#), [79](#), [80](#), and [112](#).)

Wiener, N., 1923: Differential space. *Journal of Mathematics and Physics*, **58**, 131–174. (Cited on page [42](#).)

Wong, E. and M. Zakai, 1965: On the convergence of ordinary integrals to stochastic integrals. *The Annals of Mathematical Statistics*, **36 (5)**, 1560–1564. (Cited on page [49](#).)

Zeeman, E. C., 1981: Dynamics of the evolution of animal conflicts. *Journal of Theoretical Biology*. (Cited on page [80](#).)

Zemlianov, A. and G. de Veciana, 2005: Cooperation and decision making in wireless multi-provider settings. *Infocom '05: Proceedings of the 24th conference of the IEEE Communications Society*. (Cited on page [127](#).)

COLOPHON

This thesis was typeset with $\text{\LaTeX} 2_{\epsilon}$ using Hermann Zapf's beautiful *Palatino* type face (the actual Type 1 PostScript fonts used were *URW Palladio L* and *FPL*). The monospaced text (hyperlinks, etc.) was typeset in *Bera Mono*, originally developed by Bitstream, Inc. as "Bitstream Vera". (Type 1 PostScript fonts were made available by Malte Rosenau and Ulrich Dirr.)

The typographic style of this dissertation was inspired by [Bringhurst's](#) genius as presented in *The Elements of Typographic Style* ([Bringhurst, 2002](#)), and made available for \LaTeX (via CTAN) thanks to the vigor and dedication of André Miede, designer of the "[classicthesis](#)" template. Any unsightly deviations from these exquisite works should be attributed solely to the tweakings of the author and to his (not always successful) efforts to conform to the awkward A4 paper size.

Stochastic Perturbations in Game Theory and Applications to Networks

© Panayotis Mertikopoulos 2010

Athens, June 2010

Panayotis Mertikopoulos