This summary briefly describes the background to my paper ‘Cycles of cooperation and defection in imperfect learning’ [1] and its main results. I also provide a short outlook on applications of the method developed.

Game theory traditionally used to focus on the investigation of equilibrium points, the outcome of interaction between fully rational players. Dynamical pictures have replaced these static views, coining the term ‘evolutionary game theory’. The focus is here on populations of individuals, each one endowed with a strategy inherited from their parents. Populations then evolve through birth–death processes, subject to natural selection and mutation. A large part of current work in evolutionary game theory is concerned with the effects of demographic noise in finite populations. Such stochasticity can render deterministic approaches, based for example on replicator equations, invalid. It leads to effects purely driven by noise, such as coherent quasi-cycles or fixation. Methods from statistical physics have proven to be of use in characterizing these effects, and tools from the theory of stochastic processes, interacting many-particle systems and from nonlinear and off-equilibrium dynamics are now widely used.

In [1] we take a third view on game theory, and study the learning of two fixed players from repeated interaction. In particular we focus on the effects of noise resulting from imperfect sampling of the opponent’s moves. Specifically, we investigate a model of learning the iterated prisoner’s dilemma game. Players have the choice between three strategies, always defect (ALLD), always cooperate (ALLC) and tit-for-tat (TFT). The only strict Nash equilibrium in this situation is ALLD. When players learn to play this game, convergence to the equilibrium is not guaranteed, for example we find cooperative behaviour if players discount observations in the distant past. When agents use small
samples of observed moves to estimate their opponent’s strategy, the learning process is stochastic. Players cannot obtain a clear picture of their opponent’s strategic propensities, and instead have to rely on a finite and hence limited number of moves drawn at random from the opponent’s unknown mixed strategy. We show that sustained oscillations between cooperation and defection can emerge from this stochasticity, see also [2]. An example of such coherent quasi-cycles is shown in figure 1.

These cycles are similar to those found in stochastic evolutionary processes of the same game [3, 4], but the origin of the noise sustaining the oscillations is different and lies in the imperfect sampling of the opponent’s strategy. Based on a systematic expansion technique, we are able to predict the properties of these learning cycles, providing an analytical tool with which the outcome of more general stochastic adaptation processes can be characterized. We anticipate that these methods can be extended to analyse learning algorithms in computer science, and other adaptation dynamics in the social sciences, for example more intricate models of reinforcement and belief learning. Ultimately we hope that the understanding of stochastic effects in game learning will help to link theoretical models of strategic decision making more closely with real-world experiments in behavioural game theory [5].

References

[1] Galla T, 2011 J. Stat. Mech. P08007
[2] Galla T, Intrinsic noise in game dynamical learning, 2009 Phys. Rev. Lett. 103 198702
[3] Imhof L A, Fudenberg D and Nowak M A, 2005 Proc. Nat. Acad. Sci. USA 102 10797
[4] Bladon A J, Galla T and McKane A J, 2010 Phys. Rev. E 81 066122
[5] Camerer C F, 2003 Behavioural Game Theory—Experiments in Strategic Interaction (Princeton, NJ: Princeton University Press)

Figure 1. Sustained oscillations in stochastic learning of the iterated prisoner’s dilemma. The figure shows the frequency with which TFT is used by one of the interacting players. A modified version of this figure is also contained in [1].