A deterministic-control-based approach to fully nonlinear parabolic and elliptic equations

ROBERT V. KOHN
Courant Institute

AND

SYLVIA SERFATY
Courant Institute and
Université Pierre et Marie Curie Paris 6

Abstract

We show that a broad class of fully-nonlinear second-order parabolic or elliptic PDE's can be realized as the Hamilton-Jacobi-Bellman equations of deterministic two-person games. More precisely: given the PDE, we identify a deterministic, discrete-time, two-person game whose value function converges in the continuous-time limit to the viscosity solution of the desired equation. Our game is, roughly speaking, a deterministic analogue of the stochastic representation recently introduced by Cheridito, Soner, Touzi, and Victoir [*Comm. Pure Appl. Math.* 60, 2006, 1081-1110]. In the parabolic setting with no u-dependence, it amounts to a semidiscrete numerical scheme whose timestep is a min-max. Our result is interesting, because the usual control-based interpretations of second-order PDE's involve stochastic rather than deterministic control. © 2000 Wiley Periodicals, Inc.

1 Introduction

This paper develops a deterministic control interpretation, via "two person repeated games," for a broad class of fully nonlinear equations of elliptic or parabolic type. The equations we consider have the form

$$-u_t + f(t,x,u,Du,D^2u) = 0 \tag{1.1}$$

or

$$f(x,u,Du,D^2u) + \lambda u = 0 \tag{1.2}$$

where f is "degenerate-elliptic" in the sense that

$$f(t,x,u,p,M+N) \leq f(t,x,u,p,M) \text{ when } N \text{ is nonnegative.}$$

We need additional conditions on the continuity, growth, and u-dependence of f; they are specified at the end of this section. In the stationary setting (1.2) we focus on the Dirichlet problem, solving the equation in a domain Ω with $u = g$ at $\partial\Omega$. In the time-dependent setting (1.1) we address the Cauchy problem, solving the

Communications on Pure and Applied Mathematics, Vol. 000, 0001–0051 (2000)

equation for $t < T$ with $u = h$ at $t = T$. The PDE's and boundary conditions are always interpreted in the "viscosity sense" (for a review of what this means see Section 3). Many results about viscosity solutions require that f be nondecreasing in u. We avoid such a hypothesis, but without it we only get a subsolution and a supersolution of the PDE. To know they agree we need a comparison principle, which is generally available only when f is nondecreasing in u.

Our games have two opposing players, Helen and Mark, who always make decisions rationally and deterministically. The rules depend on the form of the equation, but there is always a small parameter ε, which governs the spatial step size and (in time-dependent problems) the time step. Helen's goal is to optimize her worst-case outcome; therefore her task is roughly speaking one of "robust control." We shall characterize her value function u^ε by a version of the principle of dynamic programming. (For the details of our games and the definition of Helen's value function, see Section 2.) Our main result, roughly speaking, is that

$\limsup_{\varepsilon\to 0} u^\varepsilon$ is a viscosity subsolution of the PDE, and
$\liminf_{\varepsilon\to 0} u^\varepsilon$ is a viscosity supersolution.

(For more precise statements see Theorem 2.2 in Section 2.3 and Theorem 2.7 in Section 2.4. For the general theory of viscosity solutions to fully nonlinear equations, we refer to [20].) This result is most interesting when the PDE has a *comparison principle*, i.e. when every subsolution must lie below any supersolution. This is the case for many (though not all) of the PDE's covered by our analysis. For such equations, we conclude that $\lim u^\varepsilon$ exists and is the unique viscosity solution of the PDE. Thus, if the PDE has a comparison principle, then its solution is the limiting value of Helen's value function as $\varepsilon \to 0$.

Our analysis provides a new type of connection between games and PDE's. More familiar connections include the following:

(a) For first-order Hamilton-Jacobi equations with concave or convex Hamiltonians, the viscosity solution is the value function of an optimal control problem (the celebrated Hopf-Lax solution formula is a special case of this result, see e.g. [3, 21]).
(b) For first-order Hamilton-Jacobi equations with more general Hamiltonians, the viscosity solution is the value function of an associated deterministic two-person game [3, 23, 24].
(c) For many second-order elliptic and parabolic equations, the solution is the value function of an associated stochastic control problem (see e.g. [25, 26]).
(d) For the infinity-Laplacian, the Dirichlet problem can be solved using a rather simple two-person stochastic game [39] (see also [1, 2, 8, 22, 34, 40, 46, 47] for related work including extensions to evolution problems and the p-Laplacian).

Until recently, the only control-based interpretations of second-order elliptic or parabolic equations involved stochasticity. That situation was changed by our recent work [35], which gave deterministic game interpretations for a variety of geometric evolution laws including motion by curvature. (For related prior work see [16, 19], and for further progress along similar lines see [14, 15, 30, 31, 33].) The interpretation in [35] can be viewed as a deterministic, robust-control analogue of the stochastic-control-based viewpoint of [11, 41, 42, 43]. The method of [35] seems restricted to geometric evolution laws, a rather special class of PDE. But in light of that development, it was natural to wonder whether other second-order equations might also have deterministic game interpretations. The answer is yes, as explained in this paper (see also [36], where our approach to the time-dependent problem (1.1) with f independent of u was sketched without proofs).

Our games are deterministic, but our approach is closely related to a recently-developed stochastic representation formula due to Cheridito, Soner, Touzi, and Victoir [17]. Their work uses a "backward stochastic differential equation" (whose structure depends on the form of the equation). Ours uses a two-person game (whose structure depends similarly on the form of the equation). The connection, loosely speaking, is this: in our game, Helen optimizes her worst-case result by making herself indifferent to Mark's choices. If, rather than being clever, Mark made his choices randomly, then Helen's task would be like controlling a diffusion and the randomized version of our game would resemble a discrete version of [17]. In brief: the present paper is to [17] as our prior paper [35] was to [11, 41, 42, 43].

It should be emphasized that while our games are related to the processes in [17], our goals and methods are entirely different. Concerning goals: ours is a deterministic-game interpretation, which amounts in many cases to a semidiscrete numerical scheme; theirs was a stochastic representation formula, which offers the possibility of solving the PDE by a Monte-Carlo method (avoiding the curse of dimensionality). Concerning methods: the players in our games use only elementary optimization to determine their moves, rather than the more sophisticated framework of backward stochastic differential equations.

Examples are enlightening, so it is natural to ask: what is the "game" interpretation of the linear heat equation? The answer, as we'll explain in Section 2.1, is related to the Black-Scholes-Merton approach to option pricing. This surprised us at first, but in retrospect it seems natural. An investor who hedges a European option using the Black-Scholes-Merton theory is indifferent to the market's movements. If the market were not random but deterministic – always moving so as to give the investor the worst possible result – it would make no difference. So an investor such as Helen – whose goal is robust control (optimizing the worst-case outcome) – will make the same choices as one who adopts the Black-Scholes-Merton viewpoint. This robust-control perspective on the Black-Scholes-Merton theory of option pricing was previously noted in [10].

A parabolic equation is typically solved by time-stepping. When f is independent of u, our game provides a semidiscrete numerical scheme (discrete in time, continuous in space) whose timestep is a min-max. When the solution is smooth the scheme reduces more or less to explicit Euler, as we'll discuss in Section 7.1.

But our scheme can be used even if the equation is highly nonlinear and/or degenerate, and our convergence results hold even when the solution is not smooth. Thus, our game provides a generalization of the explicit Euler timestepping scheme, for a broad class of (possibly degenerate) parabolic equations.

When f depends on u, we use (following [17]) a "level-set" approach. Roughly speaking, this means (in the time-dependent setting) that we obtain $u^\varepsilon(x,t)$ implicitly, by first determining a function $U^\varepsilon(x,z,t)$ defined for $x \in \mathbb{R}^n$ and $z \in \mathbb{R}$, then solving $U^\varepsilon(x,u^\varepsilon(x,t),t) = 0$. For first-order equations, such a level-set approach has been explored analytically and numerically by [32, 45] (see also [4] for such an approach to some second-order equations, [9] for related work on systems, and [28, 29, 45] for a related scheme that works for a conservation law when the solution has a shock). Our analysis focuses on u^ε, making no use of results like those in [32], but analysis based on the level-set viewpoint could be a fruitful direction for further work (see Remark 2.3). We note that while in [4, 32, 45] the level-set function U is just a tool for representing the solution, in our game-based interpretation it has a natural financial meaning.

Here is a brief summary of the paper, including our main results and the overall strategy of our analysis:

- *Section 2* presents the two-person games we associate with the PDE's (1.1) and (1.2), motivating and stating our main results. The section starts with some simple cases before addressing the general one. It is entirely elementary: understanding our games is easy, though proving a link to the PDE is more technical. When f is independent of u, our game determines a "value function" u^ε. When f depends on u it determines a pair of value functions v^ε and u^ε. Section 2 includes an informal argument linking the principle of dynamic programming to the PDE in the limit $\varepsilon \to 0$. In the time-dependent setting with f independent of u, the game can be viewed as a semidiscrete numerical scheme, and our informal argument amounts to checking consistency.
- *Section 3* addresses the link between our game and the PDE will full rigor. A landmark paper by Barles and Souganidis [7] showed that if a numerical scheme is monotone, stable, and consistent, then the associated "lower semi-relaxed limit" is a viscosity supersolution and the associated "upper semi-relaxed limit" is a viscosity subsolution. The numerical scheme associated with our game is monotone from its very definition (as is typical for schemes associated with dynamic programming). The main result in Section 3 is basically the Barles-Souganidis theorem specialized to our setting: if v^ε and u^ε remain bounded as $\varepsilon \to 0$ then the lower semirelaxed limit of v^ε is a viscosity supersolution and the upper semirelaxed limit of u^ε is a viscosity subsolution. We also have $v^\varepsilon \leq u^\varepsilon$ with no extra hypotheses in the parabolic setting, or if f is monotone in u in the elliptic setting. If the PDE has a comparison principle then it follows (as usual, c.f. [7]) that $\lim u^\varepsilon = \lim v^\varepsilon$ exists and is the unique viscosity solution of the PDE.

- The analysis in Section 3 shows that consistency and stability imply convergence. *Sections 4 and 5* provide the required consistency and stability results. As noted above, the informal derivation of the PDE in Section 2 amounts to a back-of-the-envelope check of consistency; the task of Section 4 is to make it rigorous. Concerning stability: the argument in Section 3 requires only that u^ε and v^ε be locally bounded. But the argument in Section 5 proves more, obtaining a global, uniform bound. Of course this requires appropriate hypotheses on the PDE and the data (for example, in the time dependent setting the final-time value $h(x)$ must be uniformly bounded).
- *Section 6* provides additional *a priori* estimates for u^ε, under additional hypotheses on f. The basic idea is familiar: for a translation-invariant equation such as $-u_t + f(Du, D^2u) = 0$, a global-in-space Lipschitz bound propagates in time (see e.g. [16] or [19] or Appendix B of [35]). When f depends on x, t, and u the argument gets more complicated, but if the dependence is mild enough one can still get global control on the Lipschitz continuity of u^ε and derive an independent proof of compactness of u^ε. Similar methods permit us to prove in Section 6.2 that for $T-t$ sufficiently small, Helen's two value functions v^ε and u^ε are equal and converge to a solution of the PDE.
- *Section 7* provides supplementary remarks in two directions. First, in Section 7.1, we discuss the character of our game as a timestepping scheme for time-dependent equations of the form $-\partial_t u + f(Du, D^2u) = 0$. Then, in Section 7.2, we discuss an alternative game (more like that of Section 2.1 than those of Sections 2.2-2.3) for solving $\partial_t u + \Delta u + f(t,x,u,Du) = 0$.

We close this introduction by listing our main hypotheses on the form of the PDE. The real-valued function f in (1.1) is defined on $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^n\times\mathscr{S}_n$, where $\mathscr{S}_n$ is the space of symmetric $n\times n$ matrices. It is assumed throughout to be a continuous function of all its variables, and

- f is monotone in Γ in the sense that

$$f(t,x,z,p,\Gamma+N) \le f(t,x,z,p,\Gamma) \text{ for } N \ge 0. \tag{1.3}$$

In the time-dependent setting (1.1) we permit f to grow linearly in $|z|$ (so solutions can grow exponentially, but cannot blow up). However we require uniform control in x (so solutions remain bounded as $\|x\|\to\infty$ with t fixed). In fact we assume that

- f has at most linear growth in z near $p=0$, $\Gamma=0$, in the sense that for any K we have

$$|f(t,x,z,p,\Gamma)| \le C_K(1+|z|) \tag{1.4}$$

  for all $x\in\mathbb{R}^n$ and $t,z\in\mathbb{R}$, when $\|(p,\Gamma)\|\le K$.
- f is locally Lipschitz in p and Γ in the sense that for any K, we have

$$|f(t,x,z,p,\Gamma) - f(t,x,z,p',\Gamma')| \le C_K(1+|z|)\|(p,\Gamma)-(p',\Gamma')\| \tag{1.5}$$

for all $x \in \mathbb{R}^n$ and $t,z \in \mathbb{R}$, when $\|(p,\Gamma)\| + \|(p',\Gamma')\| \le K$.

- f has controlled growth with respect to p and Γ, in the sense that for some constants $q,r \ge 1$ we have

$$|f(t,x,z,p,\Gamma)| \le C\,(1+|z|+\|p\|^q+\|\Gamma\|^r) \tag{1.6}$$

for all x, t, z, p, and Γ.

In the stationary setting (1.2) our solutions will be uniformly bounded. To prove the existence of such solutions we need the discounting to be sufficiently large. We also need analogues of (1.5)–(1.6) but they can be local in z since z will ultimately be restricted to a compact set. In fact, we assume

- There exist $\eta, C_* > 0$ such that

$$|f(x,z,0,0)| \le (\lambda-\eta)|z| + C_* \tag{1.7}$$

for all $x \in \Omega$ and $z \in \mathbb{R}$; here λ is the coefficient of u in the equation (1.2).

- f is locally Lipschitz in p and Γ in the sense that for any K,L we have

$$|f(x,z,p,\Gamma) - f(t,x,z,p',\Gamma')| \le C_{K,L}\|(p,\Gamma)-(p',\Gamma')\| \tag{1.8}$$

for all $x \in \Omega$, when $\|(p,\Gamma)\| + \|(p',\Gamma')\| \le K$ and $|z| \le L$.

- f has controlled growth with respect to p and Γ, in the sense that for some constants $q,r \ge 1$ and for any L we have

$$|f(x,z,p,\Gamma)| \le C_L\,(1+\|p\|^q+\|\Gamma\|^r) \tag{1.9}$$

for all x, p and Γ, and any $|z| \le L$.

2 The games

This section presents our games. We begin, as a warm-up, by discussing the linear heat equation. Section 2.2 addresses the time-dependent problem with f independent of u. Section 2.3 discusses the time-dependent problem with f depending nonlinearly on u; our main rigorous result for the time-dependent setting is stated there (Theorem 2.2). Finally, Section 2.4 discusses the stationary setting, and states our main rigorous results for that case (Theorem 2.7).

2.1 The linear heat equation

This section offers a deterministic two-person game approach to the linear heat equation in one space dimension

$$\begin{cases} u_t + u_{xx} = 0 & \text{for } t < T \\ u = h(x) & \text{for } t = T. \end{cases} \tag{2.1}$$

Our goal is to capture, in the simplest possible setting,

- how a deterministic game can lead to a second-order PDE, and
- how our framework provides a deterministic approach to problems usually studied using stochastic methods.

The game discussed here shares many features with the ones we'll introduce in Section 2.2-2.4, though it is not a special case. (We'll return to this game and present a generalization of it to semilinear parabolic equations in Section 7.2.)

There are two players; we call them Helen and Mark, because in the financial interpretation Helen hedges her investments and Mark represents the market. (The financial interpretation will be discussed after the game has been presented.) A small parameter $\varepsilon > 0$ is fixed throughout, as are the final time T and "Helen's payoff" (a continuous function $h : \mathbb{R} \to \mathbb{R}$). The state of the game is described by its "spatial position" $x \in \mathbb{R}$ and "Helen's score" $y \in \mathbb{R}$. We suppose the game begins at time t_0; since time steps in increments of ε^2, it convenient to assume that $T - t_0 = N\varepsilon^2$ for some N.

When the game begins, the position can have any value x_0; Helen's initial score is $y_0 = 0$. The rules are as follows: if, at time $t_j = t_0 + j\varepsilon^2$, the position is x_j and Helen's score is y_j, then

(i) Helen chooses a real number p_j.

(ii) After seeing Helen's choice, Mark chooses $b_j = \pm 1$. The position then changes to

$$x_{j+1} = x_j + \sqrt{2}\varepsilon b_j$$

and Helen's score changes to

$$y_{j+1} = y_j - \sqrt{2}\varepsilon b_j p_j.$$

(iii) The clock steps forward to $t_{j+1} = t_j + \varepsilon^2$ and the process repeats, stopping when $t_N = T$.

(iv) At the final time $t_N = T$ a bonus $h(x_N)$ is added to Helen's score, where x_N is the final-time position.

Helen's goal is to maximize her final score, while Mark's goal is to obstruct her. We are interested in Helen's "value function" $u^\varepsilon(x_0, t_0)$, defined informally as her maximum worst-case final score starting from x_0 at time t_0. It is determined by the "dynamic programming principle"

$$u^\varepsilon(x, t_j) = \max_{p \in \mathbb{R}} \min_{b = \pm 1} \left[-\sqrt{2}\varepsilon p b + u^\varepsilon(x + \sqrt{2}\varepsilon b, t_{j+1}) \right] \tag{2.2}$$

coupled with the final-time condition

$$u^\varepsilon(x, T) = h(x).$$

Evidently, if $t_0 = T - N\varepsilon^2$ then

$$u^\varepsilon(x_0, t_0) = \max_{p_{N-1} \in \mathbb{R}} \min_{b_{N-1} = \pm 1} \cdots \max_{p_0 \in \mathbb{R}} \min_{b_0 = \pm 1} \left\{ h(x_N) - \sum_{j=0}^{N-1} \sqrt{2}\varepsilon b_j p_j \right\}. \tag{2.3}$$

In calling this Helen's value function, we are using an established convention from the theory of discrete-time, two-person games (see e.g. [27]).

We now argue formally that u^ε should converge as $\varepsilon \to 0$ to the solution of the linear heat equation (2.1). Roughly speaking, our argument shows that the linear heat equation is the Hamilton-Jacobi-Bellman equation of our two-person game. The procedure for formal passage from the dynamic programming principle to an associated PDE is familiar: we assume u^ε is smooth enough to use Taylor expansion, and we suppress the dependence of u^ε on ε. Substituting

$$u(x+\sqrt{2}\varepsilon b, t+\varepsilon^2) = u(x,t) + \sqrt{2}\varepsilon b u_x(x,t) + \varepsilon^2(u_t + u_{xx})(x,t) + O(\varepsilon^3)$$

into the right hand side of (2.2), reorganizing, and dropping the term of order ε^3, we obtain

$$0 \approx \max_{p\in\mathbb{R}} \min_{b=\pm 1} \left[\sqrt{2}\varepsilon b(u_x - p) + \varepsilon^2(u_t + u_{xx})\right]. \tag{2.4}$$

If $p \neq u_x$ then Mark will make $\sqrt{2}\varepsilon b(u_x - p) = -\sqrt{2}\varepsilon|u_x - p|$. This term will then dominate the right hand side, since $\varepsilon \gg \varepsilon^2$; that's bad for Helen, since it is negative and her goal was to maximize in p. So Helen's only reasonable choice is $p = u_x(x,t)$. It makes the order-ε term drop out, and we deduce from the order-ε^2 terms in (2.4) that $u_t + u_{xx} = 0$. Notice that when Helen behaves optimally she becomes indifferent to Mark's choice; our games will always have this feature.

Now we present a financial interpretation of this game. Helen plays the role of a hedger, while Mark represents the market. The position x is a stock price which evolves as a function of time t, starting at x_0 at time t_0. The small parameter ε determines the stock price increments ($\Delta x = \pm\sqrt{2}\varepsilon$ at each time). Helen's score keeps track, as we'll see in a moment, of the profits and losses generated by her hedging activity.

If Δx were random, $x(t)$ would be a random walk, approximating the diffusion described by the stochastic differential equation $dx = \sqrt{2}\,dw$ where w is Brownian motion. However in our game $x(t)$ is not random; rather it is controlled by Mark, who is constantly working against Helen's interest.

Helen's situation is this: she holds an option that will pay her $h(x(T))$ at time T (h could be negative). Her goal (as in the standard Black-Scholes-Merton approach to option pricing) is to hedge this position by buying or selling the stock at each time increment. She can borrow and lend money without paying or collecting any interest, and can take take any (long or short) stock position she desires. At each step, Helen chooses a real number p_j (depending on x_j and t_j), then adjusts her portfolio so it contains $-p_j$ units of stock (borrowing or lending to finance the transaction, so there is no change in her overall wealth). Mark sees Helen's choice. Taking it into account, he makes the stock go up or down (i.e. he chooses $b_j = \pm 1$), trying to degrade her outcome. The stock price changes from x_j to $x_{j+1} = x_j + \sqrt{2}\varepsilon b_j$, and Helen's wealth changes by $-\sqrt{2}\varepsilon b_j p_j$ (she has a profit if this is positive, a loss if it is negative). At the final time Helen collects her option payoff $h(x_N)$. If Helen and Mark both behave optimally at each stage, then

$$u^\varepsilon(x_0,t_0) + \sum_{j=0}^{N-1} \sqrt{2}\varepsilon b_j p_j = h(x_N).$$

by (2.3). In words,

(2.5) $$u^\varepsilon(x_0,t_0) + \text{Helen's worst-case profit/loss} = \text{option payoff}.$$

Helen's decisions are in fact identical to those of an investor hedging an option with payoff $h(x)$ in a binomial-tree market with $\Delta x = \pm\sqrt{2}\varepsilon$ at each timestep (see e.g. [18]). Since the binomial tree market is complete, such an investor takes no risk. Rather, she identifies an initial wealth (the price of the option) and a dynamic trading strategy that replicates the option payoff in every final-time state of the tree. Comparing with (2.5), we recognize that $u^\varepsilon(x_0,t_0)$ is the price of the option. This robust-control interpretation of the Black-Scholes-Merton theory was previously noted in [10].

2.2 Nonlinear parabolic equations without u-dependence

This section presents our two-person game approach to equations of the form

(2.6) $$\begin{cases} -u_t + f(Du, D^2u) = 0 & \text{for } x \in \mathbb{R}^n \text{ and } t < T \\ u = h(x) & \text{at } t = T. \end{cases}$$

The assumption that f is independent of x and t serves mainly to simplify the notation. The assumption that f is independent of u is, however, essential; the case with u-dependence requires additional ideas, presented in Section 2.3.

The game presented here does not reduce to the one of Section 2.1 when $f(Du,D^2u) = -\Delta u$. The central idea is similar, but the framework has been adapted to accept a more or less arbitrary function f. Recall that in Section 2.1 the optimal choice of Helen's hedging parameter was $p = u_x$; the games we present now have an analogous vector-valued hedging parameter, which Helen should choose by taking $p = Du$. Thus p serves as a proxy for Du. But f depends on both Du and D^2u, so we also need a proxy for D^2u. This is the role of the matrix Γ in the game we now describe.

The overall framework is the same as before: there are two players, Helen and Mark; a small parameter ε is fixed, etc. Since the PDE is to be solved in $\mathbb{R}^n$, Helen's final-time bonus h is now a function of $x \in \mathbb{R}^n$. The state of the game is described by its spatial position $x \in \mathbb{R}^n$ and Helen's score $y \in \mathbb{R}$. Helen's goal is to maximize her final score, while Mark's is to obstruct her.

The rules of the game depend on three new parameters, $\alpha, \beta, \gamma > 0$, whose presence represents no loss of generality. Their role will be clear in a moment. The requirements

(2.7) $$\alpha < 1/3$$

and

(2.8) $$\alpha + \beta < 1, \quad 2\alpha + \gamma < 2, \quad \max(\beta q, \gamma r) < 2$$

will be clear from the discussion in this section (the parameters q and r in (2.8) control the growth of f at infinity, c.f. (1.6).) But our proof of consistency in

Section 4 needs more: there we will require

$$\gamma < 1-\alpha, \quad \beta(q-1) < \alpha+1, \quad \gamma(r-1) < 2\alpha, \quad \gamma r < 1+\alpha. \tag{2.9}$$

These conditions do not restrict the class of PDE's we consider, since for any q and r there exist α, β, and γ with the desired properties.

When the game begins, at time t_0, the spatial position can have any value x_0 and Helen's score is $y_0 = 0$. The rules are as follows: if, at time $t_j = t_0 + j\varepsilon^2$, the position is x_j and Helen's score is y_j, then

(i) Helen chooses a vector $p_j \in \mathbb{R}^n$ and an $n \times n$ symmetric matrix Γ_j (depending on x_j and t_j), subject to the constraints

$$\|p\| \leq \varepsilon^{-\beta}, \quad \|\Gamma\| \leq \varepsilon^{-\gamma}. \tag{2.10}$$

(ii) Mark sees Helen's choice. Taking it into account, he picks a vector $w_j \in \mathbb{R}^n$ subject to the upper bound

$$\|w\| \leq \varepsilon^{-\alpha}. \tag{2.11}$$

The position then changes to

$$x_{j+1} = x_j + \varepsilon w_j$$

and Helen's score changes to

$$y_{j+1} = y_j - \left(\varepsilon p_j \cdot w_j + \frac{\varepsilon^2}{2} \langle \Gamma_j w_j, w_j \rangle + \varepsilon^2 f(p_j, \Gamma_j) \right).$$

(iii) The clock steps forward to $t_{j+1} = t_j + \varepsilon^2$ and the process repeats, stopping when $t_N = T$.

(iv) At the final time $t_N = T$ a bonus $h(x_N)$ is added to Helen's score, where x_N is the final-time position.

The financial interpretation is much as before. At each step Helen chooses a "hedge portfolio" associated with p_j and Γ_j; knowing them, Mark (our malevolent market) chooses the stock price increment εw_j to give Helen the worst possible result when her wealth decreases by

$$\varepsilon p_j \cdot w_j + \frac{\varepsilon^2}{2} \langle \Gamma_j w_j, w_j \rangle + \varepsilon^2 f(p_j, \Gamma_j).$$

The main differences from Section 2.1 are that

- the "stock price" is now a vector in $\mathbb{R}^n$, accounting for the prices of n different stocks;
- the price increments are still scaled by ε, but are no longer fixed in magnitude; and
- Helen's gain at time t_j depends on the form of the PDE.

We said Helen's hedge portfolio was "associated with" p_j and Γ_j. The financial interpretation of the vector p is familiar: its components determine her stock position (so an increment of εw in the stock price produces a loss of $\varepsilon p \cdot w$). The financial interpretation of the matrix Γ is less familiar, but it is something akin to a volatility derivative, since the associated change $\frac{1}{2}\varepsilon^2\langle \Gamma w, w\rangle$ in Helen's wealth depends quadratically on the price movement.

The upper bound (2.11) assures that each stock price increment $\Delta x = \varepsilon w$ is relatively small (since $\alpha < 1$). Similarly, (2.8) and (2.10) assure that each increment in Helen's wealth is relatively small. Condition (2.7) assures that $(\Delta x)^3 \ll \varepsilon^2$. We'll need this in a moment, when we use Taylor expansion to link the game with the PDE.

Helen's value function u^ε is determined by the final-time condition $u^\varepsilon(x,T) = h(x)$ and the dynamic programming principle

$$u^\varepsilon(x,t_j) = \max_{p,\Gamma} \min_{w} \left[u^\varepsilon(x+\varepsilon w, t_{j+1}) - \varepsilon p \cdot w - \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle - \varepsilon^2 f(p,\Gamma) \right]. \tag{2.12}$$

Here the minimization over w is subject to (2.11) and the maximization over p,Γ is subject to (2.10). It is easy to see that the max/min in (2.12) is achieved and $u^\varepsilon(x,t_j)$ is a continuous function of x at each discrete time. (The proof is by induction backward in time, using the fact that h and f are continuous, and noting that p,Γ and w range over bounded sets when ε is held fixed.) There is an equivalent characterization analogous to (2.3). The notation max min...max min is cumbersome, so we prefer to write this alternative characterization more informally as

$$u^\varepsilon(x_0,t_0) = \max_{\text{Helen's choices}} h(x_N) - \sum_{j=0}^{N-1} \left[\varepsilon p_j \cdot w_j + \frac{\varepsilon^2}{2}\langle \Gamma_j w_j, w_j\rangle + \varepsilon^2 f(p_j,\Gamma_j) \right]. \tag{2.13}$$

We now show that the PDE (2.6) is the formal Hamilton-Jacobi-Bellman equation associated with this game. The argument is parallel to that of the last section. Suppressing the dependence of u^ε on ε, and remembering that εw is small, we have by Taylor expansion

$$u(x+\varepsilon w, t+\varepsilon^2) \approx u(x,t) + \varepsilon Du(x,t)\cdot w + \varepsilon^2 (u_t + \tfrac{1}{2}\langle D^2 u w, w\rangle)(x,t)$$

if u is smooth. Substituting this in (2.12) and reorganizing, we obtain

$$0 \approx \max_{p,\Gamma} \min_{w} \left[\varepsilon (Du - p)\cdot w + \varepsilon^2 \left(u_t + \tfrac{1}{2}\langle (D^2 u - \Gamma) w, w\rangle - f(p,\Gamma) \right) \right] \tag{2.14}$$

where Du and D^2u are evaluated at (x,t). We have ignored the upper bounds (2.10)-(2.11) since they permit p, Γ, and w to be arbitrarily large in the limit $\varepsilon \to 0$ (we shall of course be more careful in Section 4). Evidently Helen should take $p = Du(x,t)$, since otherwise Mark can make the order-ε term negative and dominant. Similarly, she should choose $\Gamma \le D^2 u(x,t)$, since otherwise Mark can drive

$\langle (D^2u-\Gamma)w,w\rangle$ to $-\infty$ by a suitable choice of w. After recognizing this, Helen's maximization (the right hand side of (2.14)) reduces to

$$\max_{\Gamma \le D^2u(x,t)} [u_t - f(Du(x,t),\Gamma)]$$

Since the PDE is parabolic, i.e. since f satisfies (1.3), Helen's optimal choice is $\Gamma = D^2u(x,t)$, and (2.14) reduces formally to $-u_t + f(Du, D^2u) = 0$.

As in Section 2.1, when Helen chooses p and Γ optimally she is entirely indifferent to Mark's choice of w. Our games always have this feature.

2.3 General parabolic equations

This section explains what we do when f depends on u. We also permit dependence on x and t, so we are now discussing a fully-nonlinear (degenerate) parabolic equation of the form

$$\begin{cases} -u_t + f(t,x,u,Du,D^2u) = 0 & \text{for } x \in \mathbb{R}^n \text{ and } t < T \\ u = h(x) & \text{at } t = T. \end{cases} \tag{2.15}$$

When $f = f(Du, D^2u)$ the game presented here reduces to that of Section 2.2.

In the preceding sections, Helen's score y was essentially her wealth (her goal was to maximize it). In this section it is more convenient to work instead with $z = -y$, which amounts to her debt. We shall proceed in two steps. Using the language of our financial interpretation (to which the reader is, we hope, by now accustomed) we

(a) first consider $U^\varepsilon(x,z,t)$, Helen's optimal wealth at time T, if initially at time t the stock price is x and her wealth is $-z$;
(b) then we define $u^\varepsilon(x,t)$ or $v^\varepsilon(x,t)$ as, roughly speaking, the initial debt Helen should have at time t to break even at time T.

The proper definition of $U^\varepsilon(x,z,t)$ involves a game similar to that of the last section. If at time t_j Helen's debt is z_j and the stock price is x_j, then

(i) Helen chooses a vector $p_j \in \mathbb{R}^n$ and an $n \times n$ symmetric matrix Γ_j, restricted by (2.10).
(ii) Taking Helen's choice into account, Mark chooses the next stock price $x_{j+1} = x_j + \varepsilon w_j$ so as to degrade Helen's outcome. The scaled increment $w_j \in \mathbb{R}^n$ can be any vector subject to the upper bound (2.11).
(iii) Helen's debt changes to

$$z_{j+1} = z_j + \left[\varepsilon p_j \cdot w_j + \frac{\varepsilon^2}{2}\langle \Gamma w_j, w_j\rangle + \varepsilon^2 f(t_j, x_j, z_j, p_j, \Gamma_j)\right]. \tag{2.16}$$

(iv) The clock steps forward to $t_{j+1} = t_j + \varepsilon^2$ and the process repeats, stopping when $t_N = T$. At the final time Helen receives $h(x_N)$ from the option.

Helen's goal, as usual, is to maximize her worst-case score at time T, and Mark's is to work against her. Using the shorthand introduced in (2.13), her value function is

$$U^{\varepsilon}(x_0,z_0,t_0) = \max_{\text{Helen's choices}} [h(x_N) - z_N]. \tag{2.17}$$

It is characterized by the dynamic programming principle:

$$U^{\varepsilon}(x,z,t_j) = \max_{p,\Gamma} \min_{w} U^{\varepsilon}(x+\Delta x, z+\Delta z, t_{j+1}) \tag{2.18}$$

together with the final-time condition $U^{\varepsilon}(x,z,T) = h(x) - z$. Here Δx is $x_{j+1} - x_j = \varepsilon w$ and $\Delta z = z_{j+1} - z_j$ is given by (2.16), and the optimizations are constrained by (2.10) and (2.11). It is easy to see that the max/min in (2.18) is achieved and $U^{\varepsilon}(x,,z,t_j)$ is a continuous function of x and z at each discrete time. (The proof is by induction backward in time, like the argument sketched above for (2.12).)

We turn now to task (b), the definition of u^{ε} and v^{ε}. As motivation, observe that when f is independent of z, the functions $U^{\varepsilon}(x,z,t)$ defined above and $u^{\varepsilon}(x,t)$ defined in Section 2.2 are related by $U^{\varepsilon}(x,z,t) = u^{\varepsilon}(x,t) - z$. (This can be seen, for example, by comparing (2.13) to (2.17).) So $z_0 = u^{\varepsilon}(x_0,t_0)$ is the initial debt Helen should have at time t_0 to be assured of (at least) breaking even at time T. When f depends on z the function $z \mapsto U^{\varepsilon}(x,z,t)$ can be non-monotone, so we must distinguish between the minimal and maximal initial debt with which Helen breaks even at time T. Thus, we define:

$$u^{\varepsilon}(x_0,t_0) = \sup\{z_0 : U^{\varepsilon}(x_0,z_0,t_0) \geq 0\}, \tag{2.19}$$

$$v^{\varepsilon}(x_0,t_0) = \inf\{z_0 : U^{\varepsilon}(x_0,z_0,t_0) \leq 0\}. \tag{2.20}$$

with the convention that the empty set has $\sup = -\infty$ and $\inf = \infty$. (We remark that the introduction of u^{ε} and v^{ε} is analogous to the pricing of an option by super or sub replication.)

Clearly $v^{\varepsilon} \leq u^{\varepsilon}$, and $u^{\varepsilon}(x,T) = v^{\varepsilon}(x,T) = h(x)$. It is not immediately clear that $v^{\varepsilon} > -\infty$ or $u^{\varepsilon} < \infty$ for all $t < T$; however we shall establish this in Section 5 under the hypothesis that $|h|$ is uniformly bounded. It is not clear in general that u^{ε} and v^{ε} are continuous.

Since the definitions of u^{ε} and v^{ε} are now implicit, these functions can no longer be characterized by a principle of dynamic programming. However we still have two "dynamic programming inequalities:"

Proposition 2.1. *If $u^{\varepsilon}(x,t)$ is finite then*

$$u^{\varepsilon}(x,t) \leq \sup_{p,\Gamma} \inf_{w} \left[u^{\varepsilon}(x+\varepsilon w, t+\varepsilon^2) - \left(\varepsilon p \cdot w + \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle + \varepsilon^2 f(t,x,u^{\varepsilon}(x,t),p,\Gamma)\right) \right]; \tag{2.21}$$

similarly, if $v^\varepsilon(x,t)$ is finite then

(2.22)

$$v^\varepsilon(x,t) \geq \sup_{p,\Gamma}\inf_{w}\left[v^\varepsilon(x+\varepsilon w, t+\varepsilon^2) - (\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle + \varepsilon^2 f(t,x,v^\varepsilon(x,t),p,\Gamma))\right].$$

As usual, the sup and inf are constrained by (2.10) and (2.11).

Proof. We wrote sup/inf not max/min in these relations, since it is not clear that u^ε and v^ε are continuous. To prove (2.21), consider $z = u^\varepsilon(x,t)$. By the definition of u^ε (and remembering that U^ε is continuous) we have $U^\varepsilon(x,z,t) = 0$. Hence writing (2.18), we have

$$0 = \max_{p,\Gamma}\min_{w} U^\varepsilon(x+\varepsilon w, z+\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle + \varepsilon^2 f(t,x,z,p,\Gamma), t+\varepsilon^2).$$

We conclude that there exist p,Γ (constrained by (2.10)), such that for all w (constrained by (2.11)), we have

$$U^\varepsilon(x+\varepsilon w, z+\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle + \varepsilon^2 f(t,x,z,p,\Gamma), t+\varepsilon^2) \geq 0.$$

By the definition of u^ε, this implies that

$$z+\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle + \varepsilon^2 f(t,x,z,p,\Gamma) \leq u^\varepsilon(x+\varepsilon w, t+\varepsilon^2).$$

In other words, there exist p,Γ such that for every w

$$z \leq u^\varepsilon(x+\varepsilon w, t+\varepsilon^2) - \left(\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle + \varepsilon^2 f(t,x,z,p,\Gamma)\right).$$

This implies

$$z \leq \sup_{p,\Gamma}\inf_{w}\left[u^\varepsilon(x+\varepsilon w, t+\varepsilon^2) - \left(\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle + \varepsilon^2 f(t,x,z,p,\Gamma)\right)\right].$$

Recalling that $z = u^\varepsilon(x,t)$, we get (2.21). The proof of (2.22) is entirely parallel. □

Our PDE (2.15) is the formal Hamilton-Jacobi-Bellman equation associated with the dynamic programming principles (2.21)–(2.22), by essentially the same argument we used in the last section to connect (2.12) with (2.6), if one believes that $u^\varepsilon \approx v^\varepsilon$.

Rather than repeat that heuristic argument, let us state the corresponding rigorous result, which follows from the results in Sections 3 – 5. It concerns the upper and lower semi-relaxed limits, defined by

$$\overline{u}(x,t) = \limsup_{\substack{y\to x\\ t_j\to t\\ \varepsilon\to 0}} u^\varepsilon(y,t_j) \text{ and } \underline{v}(x,t) = \liminf_{\substack{y\to x\\ t_j\to t\\ \varepsilon\to 0}} v^\varepsilon(y,t_j), \tag{2.23}$$

where the discrete times are $t_j = T - j\varepsilon^2$. We shall show, under suitable hypotheses, that $\underline{v}$ and $\overline{u}$ are viscosity super and subsolutions respectively. It is natural to ask whether they are equal, and whether they give a viscosity solution of the PDE. This is a global question, which we can answer only if the PDE has a *comparison principle*. Such a principle asserts that if u is a subsolution and v is a supersolution (with some constraint on their growth at infinity if the spatial domain is unbounded) then $u \leq v$. If the PDE has such a principle (and assuming $\overline{u}$ and $\underline{v}$ satisfy the growth condition at infinity) then it follows that $\overline{u} \leq \underline{v}$. The opposite inequality is immediate from the definitions, so it follows that $\overline{u} = \underline{v}$, and we get a viscosity solution of the PDE. It is in fact the unique viscosity solution, since the comparison principle implies uniqueness. Here is a careful statement of the result just sketched:

Theorem 2.2. *Consider the final-value problem (2.15) where f satisfies (1.3)–(1.6), and h is a uniformly bounded, continuous function. Assume the parameters α, β, γ satisfy (2.7) – (2.9). Then $\overline{u}$ and $\underline{v}$ are uniformly bounded on $\mathbb{R}^n \times [t_*, T]$ for any $t_* < T$, and they are respectively a viscosity subsolution and a viscosity supersolution of (2.15). If the PDE has a comparison principle (for uniformly bounded solutions) then it follows that u^ε and v^ε converge locally uniformly to the unique viscosity solution of (2.15).*

This theorem is an immediate consequence of Propositions 3.3 and 5.1. Some sufficient conditions for the PDE to have a comparison principle can be found in Section 4.3 of [17]. Note that most comparison results require $f(t,x,z,p,\Gamma)$ to be nondecreasing in z.

Remark 2.3. If $U^\varepsilon(x,z,t)$ is a strictly decreasing function of z then $u^\varepsilon(x,t) = v^\varepsilon(x,t)$ is determined by the implicit equation $U^\varepsilon(x,u^\varepsilon(x,t),t)$. It is natural to guess that as $\varepsilon \to 0$, U^ε converges to a solution of the "level-set PDE" obtained by either (a) applying our usual formal argument to the dynamic programming principle (2.18), or (b) differentiating the relation $U(x,u(x,t),t) = 0$ (see [32]). A proof of this conjecture would seem to require new ideas, such as a comparison principle for the PDE satisfied by U (which is presently unavailable except in some rather special cases, see Remark 2.5 of [32]). The correspondence between U and u has mainly been studied for some classes of first-order equations [32, 45], but there are also some results in the second-order case [4] and for systems [9]. The analysis in the present paper does not use the level-set framework; instead we work mainly with u^ε and v^ε.

When the PDE is Burgers' equation and the solution has a shock, we do not expect either u^ε or v^ε to converge to the entropy-decreasing weak solution. Rather, we expect the locus where $U^\varepsilon = 0$ to resemble the multivalued graph of an "overturning solution." (The papers [28, 29, 45] discuss how the level-set method can be modified to avoid overturning and get the entropy weak solution; we have not explored whether a similar modification is possible in the present context.)

We close this section with the observation that if $U^\varepsilon(x,z,t)$ is a strictly decreasing function of z then $v^\varepsilon(x,t) = u^\varepsilon(x,t)$. A sufficient condition for this to hold is that f be non-decreasing in z:

Lemma 2.4. *Suppose f is non-decreasing in z, in the sense that*

$$f(t,x,z_1,p,\Gamma) \geq f(t,x,z_0,p,\Gamma) \quad \textit{whenever } z_1 > z_0. \tag{2.24}$$

Then U^ε satisfies

$$U^\varepsilon(x,z_1,t_j) \leq U^\varepsilon(x,z_0,t_j) - (z_1 - z_0) \quad \textit{whenever } z_1 > z_0 \tag{2.25}$$

at each discrete time $t_j = T - j\varepsilon^2$. In particular, U^ε is strictly decreasing in z and $v^\varepsilon = u^\varepsilon$.

Proof. Since $U^\varepsilon(x,z,T) = h(x) - z$, (2.25) holds with equality at the final time. We shall use the dynamic programming principle (2.18) to show that if (2.25) holds at time $t+\varepsilon^2$ then it also holds at time t. For any p, Γ, and w, let $\Delta x = \varepsilon w$ and set

$$\begin{aligned} \Delta z_0 &= \varepsilon p \cdot w + \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle + \varepsilon^2 f(t,x,z_0,p,\Gamma), \\ \Delta z_1 &= \varepsilon p \cdot w + \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle + \varepsilon^2 f(t,x,z_1,p,\Gamma). \end{aligned}$$

Note that

$$\Delta z_1 \geq \Delta z_0$$

as a consequence of (2.24). Therefore (2.25) at time $t+\varepsilon^2$ (i.e., the induction hypothesis) implies

$$\begin{aligned} &U^\varepsilon(x+\Delta x, z_1+\Delta z_1, t+\varepsilon^2) \\ &\quad \leq U^\varepsilon(x+\Delta x, z_0+\Delta z_0, t+\varepsilon^2) - ((z_1+\Delta z_1) - (z_0+\Delta z_0)) \\ &\quad\quad \leq U^\varepsilon(x+\Delta x, z_0+\Delta z_0, t+\varepsilon^2) - (z_1 - z_0). \end{aligned} \tag{2.26}$$

Minimizing over w then maximizing over p,Γ, we obtain the desired inequality (2.25). The remaining assertions of the Lemma are immediate consequences of this relation. □

When f is not monotone in z the preceding argument breaks down, but we can still prove that U^ε is monotone in z for $T-t$ sufficiently small. This will be done in Section 6.2.

2.4 Nonlinear elliptic equations

This section explains how our game can be used to solve stationary boundary value problems. The framework is similar to the parabolic case, but two new issues arise:

(i) we must introduce discounting, to be sure Helen's value function is finite; and
(ii) we must be careful about the sense in which the boundary data are imposed.

Concerning (i): discounting is needed because we consider an arrival-time problem, and Helen's optimal arrival time could be infinite. The discounting affects the PDE; therefore we focus on

$$\begin{cases} f(x,u,Du,D^2u)+\lambda u=0 & \text{in } \Omega \\ u=g & \text{on } \partial\Omega. \end{cases} \tag{2.27}$$

The constant λ (which plays the role of an interest rate) must be positive, and large enough that (1.7) holds. (If f is independent of z then any $\lambda > 0$ will do.)

Concerning (ii): this issue is new only because we chose to address just Cauchy problems in Sections 2.2-2.3 (see however Remark 2.8). Our framework applies even to first-order Hamilton-Jacobi equations, where it is clear from elementary examples that data can only be imposed in the classical sense on part of the boundary. At the level of the PDE, the proper handling of this issue is well-understood: the boundary condition in (2.27) must be understood in the *viscosity sense* (see Definition 3.2 in Section 3). At the level of the game, the corresponding phenomenon is that Helen's value function may not be a continuous function of $x \in \overline{\Omega}$.

We now present the game. The main differences from Section 2.3 are the presence of discounting, and the fact that play stops when Helen gets to $\partial\Omega$ or when her score gets too large in absolute value, rather than at a fixed time. The boundary condition g is assumed to be a bounded, continuous function on $\partial\Omega$. It enters the game as an "exit bonus;" since the final position can be (slightly) outside Ω, we shall assume that g has been extended to a continuous function defined on a neighborhood of Ω. Besides the parameters α, β, γ introduced previously, in the stationary case we need two new parameters, M and m. Both are positive constants; M serves to cap the score, and m determines what happens when the cap is reached. We shall in due course choose $m = M - 1$ and require that M be sufficiently large (see Remark 2.6). Like the choices of α, β, γ, the parameters M and m are used to define the game but they do not influence the resulting PDE. As in Section 2.3, we proceed in two steps:

- (a) first we introduce $U^\varepsilon(x,z)$, the optimal worst-case present value of Helen's wealth if the initial stock price is x and her initial wealth is $-z$;
- (b) then we define $u^\varepsilon(x,t)$ and $v^\varepsilon(x,t)$ as the maximal and minimal initial debt Helen should have at time t to break even upon exit.

The definition of $U^\varepsilon(x,z)$ for $x \in \Omega$ and $z \in \mathbb{R}$ involves a game of the usual type.

- (i) Initially, at time $t_0 = 0$, the stock price is $x_0 = x$ and Helen's debt is $z_0 = z$.
- (ii) Suppose, at time $t_j = j\varepsilon^2$, the stock price is x_j and Helen's debt is z_j with $|z_j| < M$. Then Helen chooses a vector p_j and an $n \times n$ symmetric matrix Γ_j, restricted in magnitude by (2.10). Knowing these choices, Mark determines the next stock price $x_{j+1} = x_j + \varepsilon w_j$, with w_j restricted by (2.11). Helen experiences a loss at time t_j of

$$\delta_j = \left[\varepsilon p_j \cdot w_j + \frac{\varepsilon^2}{2}\langle \Gamma_j w_j, w_j\rangle + \varepsilon^2 f(x_j, z_j, p_j, \Gamma_j)\right]. \tag{2.28}$$

As a consequence, her time $t_{j+1} = t_j + \varepsilon^2$ debt becomes

$$z_{j+1} = e^{\lambda\varepsilon^2}(z_j + \delta_j)$$

(the factor $e^{\lambda\varepsilon^2}$ takes into account her interest payments).

(iii) If $z_{j+1} \geq M$ then the game terminates, and Helen pays a "termination-by-large-debt penalty" worth $e^{\lambda\varepsilon^2}(m - \delta_j)$ at time t_{j+1}. Similarly, if $z_{j+1} \leq -M$ then the game terminates, and Helen receives a "termination-by-large-wealth bonus" worth $e^{\lambda\varepsilon^2}(m + \delta_j)$ at time t_{j+1}. If the game stops this way we call t_{j+1} the "ending index" t_N.

(iv) If $|z_{j+1}| < M$ and $x_{j+1} \notin \Omega$ then the game terminates, and Helen gets an "exit payoff" worth $g(x_{j+1})$ at time t_{j+1}. If the game ends this way we call t_{j+1} the "exit index" t_E. Notice that if g is defined in a neighborhood of Ω then the exit payoff is well-defined, since the distance from x_{j+1} to Ω is at most $\|\varepsilon w_j\| \leq \varepsilon^{1-\alpha} \ll 1$.

(v) If the game hasn't terminated then Helen and Mark repeat this procedure at time $t_{j+1} = t_j + \varepsilon^2$. If the game never stops then the "exit time" t_E is $+\infty$.

Helen's goal is a bit different from before, due to the presence of discounting: she seeks to maximize the minimum *present value* of her future income, using the discount factor of $e^{-j\lambda\varepsilon^2}$ for income received at time t_j. If the game ends by exit at time t_E then the present value of her income is

$$\begin{aligned} U^\varepsilon(x_0, z_0) &= -z_0 - \delta_0 - e^{-\lambda\varepsilon^2}\delta_1 - \cdots - e^{-(E-1)\lambda\varepsilon^2}\delta_{E-1} + e^{-E\lambda\varepsilon^2} g(x_E) \\ &= e^{-E\lambda\varepsilon^2}(g(x_E) - z_E). \end{aligned}$$

If the game never ends then (since z_j and $g(x_j)$ are uniformly bounded) we can take $E = \infty$ in the preceding formula to see that the present value of her income is 0. If the game ends by capping at time t_N with $z_N \geq M$ then the present value of her income is

$$\begin{aligned} U^\varepsilon(x_0, z_0) &= -z_0 - \delta_0 - e^{-\lambda\varepsilon^2}\delta_1 - \cdots - e^{-(N-1)\lambda\varepsilon^2}\delta_{N-1} - e^{-(N-1)\lambda\varepsilon^2}(m - \delta_{N-1}) \\ &= e^{-(N-1)\lambda\varepsilon^2}(-z_{N-1} - m); \end{aligned}$$

similarly, if the game ends by capping at time t_N with $z_N \leq -M$ then the present value of her income is

$$U^\varepsilon = e^{-(N-1)\lambda\varepsilon^2}(-z_{N-1} + m).$$

Using the shorthand introduced in (2.13), Helen's value function is thus

(2.29)
$$U^\varepsilon(x_0, z_0) = \sup_{\text{Helen's choices}} \begin{cases} e^{-E\lambda\varepsilon^2}(g(x_E) - z_E) & \text{if the game ends by exit} \\ e^{-(N-1)\lambda\varepsilon^2}(-z_{N-1} - m) & \text{if it ends by capping above,} \\ e^{-(N-1)\lambda\varepsilon^2}(-z_{N-1} + m) & \text{if it ends by capping below.} \end{cases}$$

To get a dynamic programming characterization of U^ε, we observe that if $|z_0| < M$ then

$$U^\varepsilon(x_0,z_0) = \sup_{p,\Gamma} \inf_{w} \begin{cases} e^{-\lambda\varepsilon^2} U^\varepsilon(x_1,z_1) & \text{if } x_1 \in \Omega \text{ and } |z_1| < M \\ e^{-\lambda\varepsilon^2}(g(x_1)-z_1) & \text{if } x_1 \notin \Omega \text{ and } |z_1| < M \\ -z_0 - m & \text{if } z_1 \geq M \\ -z_0 + m & \text{if } z_1 \leq -M. \end{cases} \tag{2.30}$$

Since the game is stationary (nothing distinguishes time 0), the associated dynamic programming principle is that for $|z| < M$,

$$U^\varepsilon(x,z) = \sup_{p,\Gamma} \inf_{w} \begin{cases} e^{-\lambda\varepsilon^2} U^\varepsilon(x',z') & \text{if } x' \in \Omega \text{ and } |z'| < M \\ e^{-\lambda\varepsilon^2}(g(x')-z') & \text{if } x' \notin \Omega \text{ and } |z'| < M \\ -z - m & \text{if } z' \geq M \\ -z + m & \text{if } z' \leq -M \end{cases} \tag{2.31}$$

where $x' = x + \varepsilon w$ and $z' = e^{\lambda\varepsilon^2}(z+\delta)$, with δ defined as in (2.28). Here p, Γ, and w are constrained as usual by (2.10)–(2.11), and we write sup, inf rather than max, min since it is no longer clear that the optima are achieved (since the right hand side is now a discontinuous function of p, Γ, and w). The preceding discussion defines U^ε only for $|z| < M$; it is natural to extend the definition to all z by

$$\begin{aligned} U^\varepsilon(x,z) &= -z - m \text{ for } z \geq M \\ U^\varepsilon(x,z) &= -z + m \text{ for } z \leq -M \end{aligned} \tag{2.32}$$

which corresponds to play being "capped" immediately. Notice that when extended this way, U^ε is strictly negative for $z \geq M$ and strictly positive for $z \leq -M$.

Is U^ε well-defined and computable? We shall use a fixed-point argument to give an affirmative answer in Section 5.2, when M is sufficiently large and $m = M - 1$. An alternative viewpoint uses a time-dependent version of the game similar to that of Section 2.3. If, rather than letting Helen and Mark play indefinitely, we introduce a finite final time T, then our situation is similar to that of Section 2.3. There is a principle of dynamic programming (an obvious modification of (2.31)) which determines Helen's value function by stepping backward in time. Our U^ε is the function to which this time-dependent solution converges as $t \to -\infty$. (The fixed-point argument in Section 5.2 amounts to a proof that the solution has a limit as $t \to -\infty$.)

The definitions of u^ε and v^ε are slightly different from those in Section 2.3:

$$u^\varepsilon(x_0) = \sup\{z_0 : U^\varepsilon(x_0,z_0) > 0\}, \tag{2.33}$$

$$v^\varepsilon(x_0) = \inf\{z_0 : U^\varepsilon(x_0,z_0) < 0\}. \tag{2.34}$$

The change from Section 2.3 is that the inequalities in (2.33) and (2.34) are *strict*. This change seems necessary to make the proof of Proposition 2.5 work. It has, however, an important consequence: the relation $v^\varepsilon \leq u^\varepsilon$ is no longer obvious. For example, if $z \mapsto U^\varepsilon(x_0,z)$ is strictly positive for $z < a$, identically 0 for $a \leq z \leq b$, and strictly negative for $z > b$, then (2.33)-(2.34) give $u^\varepsilon(x_0) = a$ and $v^\varepsilon(x_0) = b$.

This issue does not affect our proof that $\underline{v}$ and $\overline{u}$ are viscosity super and subsolutions. But we cannot conclude (via a comparison principle) that $\underline{v} = \overline{u}$ unless we somehow know that $\underline{v} \le \overline{u}$. Our main scheme for knowing this is an elliptic version of Lemma 2.4: we will show in Section 5.2 that if f is nondecreasing in z then U^ε is strictly decreasing in z, from which it follows easily that $v^\varepsilon = u^\varepsilon$.

Note that since $U^\varepsilon > 0$ for $z \le -M$ and $U^\varepsilon < 0$ for $z \ge M$, we always have $|u^\varepsilon(x)| \le M$ and $|v^\varepsilon(x)| \le M$.

We have the following analogue of Proposition 2.1:

Proposition 2.5. *Let m_1 be a constant with $0 < m_1 < M$. Then whenever $x \in \Omega$ and $-m_1 \le u^\varepsilon(x) < M$ we have*

(2.35)
$$u^\varepsilon(x) \le \sup_{p,\Gamma} \inf_w \left[-\varepsilon p \cdot w - \frac{\varepsilon^2}{2} \langle \Gamma w, w \rangle - \varepsilon^2 f(x, u^\varepsilon(x), p, \Gamma) + e^{-\lambda \varepsilon^2} \hat{u}^\varepsilon(x + \varepsilon w) \right]$$

when ε is small enough (depending on m_1 and on the parameters of the game but not on x). Similarly, if $x \in \Omega$ and $-M < v^\varepsilon(x) \le m_1$ then when ε is small enough

(2.36)
$$v^\varepsilon(x) \ge \sup_{p,\Gamma} \inf_w \left[-\varepsilon p \cdot w - \frac{\varepsilon^2}{2} \langle \Gamma w, w \rangle - \varepsilon^2 f(x, v^\varepsilon(x), p, \Gamma) + e^{-\lambda \varepsilon^2} \hat{v}^\varepsilon(x + \varepsilon w) \right].$$

Here p, Γ, and w are constrained as usual by (2.10) and (2.11), and $\hat{u}^\varepsilon, \hat{v}^\varepsilon$ are defined to be the extensions of $u^\varepsilon, v^\varepsilon$ by g off Ω; in other words we use the convention that for any function ϕ,

$$\hat{\phi}(x) = \begin{cases} \phi(x) & \text{if } x \in \Omega \\ g(x) & \text{if } x \notin \Omega. \end{cases} \tag{2.37}$$

Proof. We shall focus on (2.35); the argument for (2.36) is entirely parallel. Since $-m_1 \le u^\varepsilon(x) < M$, there is a sequence $z^k \uparrow u^\varepsilon(x)$ such that $U^\varepsilon(x, z^k) > 0$. Since $u^\varepsilon(x)$ is bounded away from $-M$, we may suppose that z^k also stays bounded away from $-M$. Dropping the index k for simplicity of notation, consider any such $z = z^k$. The fact that $U^\varepsilon(x, z) > 0$ tells us that the right hand side of the dynamic programming principle (2.31) is positive. So there exist p, Γ satisfying (2.10) such that for any w satisfying (2.11),

$$0 < \begin{cases} e^{-\lambda \varepsilon^2} U^\varepsilon(x', z') & \text{if } x' \in \Omega \text{ and } |z'| < M \\ e^{-\lambda \varepsilon^2}(g(x') - z') & \text{if } x' \notin \Omega \text{ and } |z'| < M \\ -z - m & \text{if } z' \ge M \\ -z + m & \text{if } z' \le -M \end{cases}$$

where $x' = x + \varepsilon w$ and $z' = e^{\lambda \varepsilon^2}(z + \delta)$. Capping above (the alternative $z' \ge M$) cannot occur, since when it happens the right hand side is negative. Capping below (the alternative $z' \le -M$) cannot occur either (if ε is sufficiently small), because z is bounded away from $-M$ and δ is bounded by a positive power of ε. Therefore only two cases can occur. If $x+\varepsilon w \in \Omega$ then we have

$$0 < U^\varepsilon(x+\varepsilon w, e^{\lambda\varepsilon^2}(z+\delta))$$

whence by the definition of u^ε

$$u^\varepsilon(x+\varepsilon w) \geq e^{\lambda\varepsilon^2}(z+\delta).$$

If on the other hand $x+\varepsilon w \notin \Omega$ then we have

$$0 < g(x+\varepsilon w) - e^{\lambda\varepsilon^2}(z+\delta).$$

Combining these cases, we have shown the existence of p,Γ such that for every w,

$$z \leq -\left(\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle\Gamma w, w\rangle + \varepsilon^2 f(x,z,p,\Gamma)\right) + e^{-\lambda\varepsilon^2}\hat{u}^\varepsilon(x+\varepsilon w) \tag{2.38}$$

where $\hat{u}^\varepsilon$ is defined by (2.37). Remembering that $z = z^k \uparrow u^\varepsilon(x)$, we pass to the limit on both sides of (2.38) (with p, Γ, and w held fixed) to see that

$$u^\varepsilon(x) \leq -\left(\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle\Gamma w, w\rangle + \varepsilon^2 f(x,u^\varepsilon(x),p,\Gamma)\right) + e^{-\lambda\varepsilon^2}\hat{u}^\varepsilon(x+\varepsilon w).$$

Since this is true for some p,Γ and for every w, we have established (2.35). The proof of (2.36) is parallel. □

We shall prove in Section 5.2 that if M is sufficiently large and $m = M-1$ then $U^\varepsilon(x,z)$ exists and $V^\varepsilon(x,z) = U^\varepsilon(x,z)+z$ satisfies $|V^\varepsilon(x,z)| \leq m$. This implies, by an elementary argument (given there), that $|u^\varepsilon(x)| \leq m$ and $|v^\varepsilon(x)| \leq m$. Thus, when M is sufficiently large the hypothesis of Proposition 2.5 will hold with $m_1 = m = M-1$.

Remark 2.6. In defining the game, our choice of what happens when play ends by capping may seem rather mysterious. It was governed by our desire to have $V^\varepsilon(x,z) = U^\varepsilon(x,z)+z$ be a decreasing function of z when f is nondecreasing in z (see Lemma 5.4). Since we also want $|V^\varepsilon| \leq m$, it is natural to take $V^\varepsilon(x,-M) = m$ and $V^\varepsilon(x,M) = -m$. Our capping scheme was chosen to have this property.

The PDE (2.27) is the formal Hamilton-Jacobi-Bellman equation associated with the dynamic programming inequalities (2.35)–(2.36), by the usual Taylor-expansion based argument, if one accepts that $-M < v^\varepsilon \approx u^\varepsilon < M$. Rather than dwell on that heuristic argument, we now state our main rigorous result in the stationary setting, which follows from the results in Sections 4 and 5. It concerns the upper and lower semi-relaxed limits, defined for any $x \in \overline{\Omega}$ by

$$\overline{u}(x) = \limsup_{\substack{y\to x\\ \varepsilon\to 0}} u^\varepsilon(y) \text{ and } \underline{v}(x) = \liminf_{\substack{y\to x\\ \varepsilon\to 0}} v^\varepsilon(y). \tag{2.39}$$

with the convention that y approaches x from within Ω (since u^ε and v^ε are only defined on Ω).

Theorem 2.7. *Consider the stationary boundary-value problem (2.27) where f satisfies (1.3) and (1.7) – (1.9), and g is a uniformly bounded continuous function. Assume the parameters of the game α,β,γ satisfy (2.7) –(2.9), M is sufficiently large, and $m = M-1$. Then u^ε and v^ε are well-defined when ε is sufficiently small, and they satisfy $|u^\varepsilon| \le m$ and $|v^\varepsilon| \le m$. Their semi-relaxed limits $\overline{u}$ and $\underline{v}$ are respectively a viscosity subsolution and a viscosity supersolution of (2.15). If in addition we have $\underline{v} \le \overline{u}$ and the PDE has a comparison principle, then it follows that u^ε and v^ε converge locally uniformly in Ω to the unique viscosity solution of (2.15).*

This is an immediate consequence of Propositions 3.5 and 5.3. A sufficient condition for $\underline{v} \le \overline{u}$ is that f be nondecreasing in z (see Lemma 5.4). Sufficient conditions for the PDE to have a comparison principle can be found for example in Section 5 of [20], and (for more recent results) in [5, 6].

Remark 2.8. In Section 2.3 we discussed only Cauchy problems, i.e. parabolic equations solved in the whole space $\mathbb{R}^n$. However parabolic problems on bounded domains (with a Dirichlet condition $u = g$ at the boundary) can easily be handled using ideas from the present section. Briefly: the game must stop when Helen exits, and she collects g evaluated at the exit point.

3 Convergence

This section presents our main convergence results, linking the limiting behavior of v^ε and u^ε as $\varepsilon \to 0$ to the PDE. The argument uses the framework of [7] and is basically a special case of the theorem proved there; we give the details anyway, to make this paper self-contained. Convergence is a *local* issue: in the time-dependent setting, Proposition 3.3 shows that in any region where the lower and upper semi-relaxed limits $\underline{v}$ and $\overline{u}$ are finite they are in fact viscosity super and subsolutions respectively. (The analogous statement for the stationary case is somewhat more subtle; see Remark 3.6.)

3.1 Viscosity solutions

Our PDE's can be degenerate parabolic, degenerate elliptic, or even first order. Therefore we cannot expect a classical solution, and we cannot always impose boundary data in the classical sense on the entirety of $\partial\Omega$. The theory of *viscosity solutions* provides the proper framework for handling these issues. We review the basic definitions for the reader's convenience. Consider first the final-value problem (2.15) in $\mathbb{R}^n$,

$$\begin{cases} -u_t + f(t,x,u,Du,D^2u) = 0 & \text{for } x \in \mathbb{R}^n \text{ and } t < T \\ u = h(x) & \text{at } t = T \end{cases}$$

where $f(t,x,z,p,\Gamma)$ is continuous in all its variables and satisfies the monotonicity condition (1.3) in its last variable.

Definition 3.1. A real-valued, lower-semicontinuous function $u(x,t)$ defined for $x \in \mathbb{R}^n$ and $t_* \leq t \leq T$ is a viscosity supersolution of the final-value problem (2.15) if

(a) for any (x_0,t_0) with $t_* \leq t_0 < T$, and any smooth $\phi(x,t)$ such that $u-\phi$ has a local minimum at (x_0,t_0) we have

$$\partial_t \phi(x_0,t_0) - f(t_0,x_0,u(x_0,t_0),D\phi(x_0,t_0),D^2\phi(x_0,t_0)) \leq 0, \text{ and}$$

(b) $u \geq h$ at the final time $t = T$.

Similarly, a real-valued upper-semicontinuous function $u(x,t)$ defined for $x \in \mathbb{R}^n$ and $t_* \leq t \leq T$ is a viscosity subsolution of the final-value problem (2.15) if

(a) for any (x_0,t_0) with $t_* \leq t_0 < T$, and any smooth $\phi(x,t)$ such that $u-\phi$ has a local maximum at (x_0,t_0) we have

$$\partial_t \phi(x_0,t_0) - f(t_0,x_0,u(x_0,t_0),D\phi(x_0,t_0),D^2\phi(x_0,t_0)) \geq 0, \text{ and}$$

(b) $u \leq h$ at the final time $t = T$.

A viscosity solution of (2.15) is a continuous function u that is both a subsolution and a supersolution.

Now consider the stationary problem (2.27). The definitions are similar to the time-dependent setting, however we must be careful to impose the boundary condition "in the viscosity sense:"

Definition 3.2. A real-valued, lower-semicontinuous function $u(x)$ defined on $\overline{\Omega}$ is a viscosity supersolution of the stationary problem (2.27) if

(a) for any $x_0 \in \Omega$ and any smooth $\phi(x,t)$ such that $u-\phi$ has a local minimum at x_0, we have

$$f(x_0,u(x_0),D\phi(x_0),D^2\phi(x_0)) + \lambda u(x_0) \geq 0; \text{ and}$$

(b) for any $x_0 \in \partial\Omega$ and any smooth ϕ such that $u-\phi$ has a local minimum on $\overline{\Omega}$ at x_0, we have

$$\max\left\{f(x_0,u(x_0),D\phi(x_0),D^2\phi(x_0)) + \lambda u(x_0), u(x_0) - g(x_0)\right\} \geq 0.$$

Similarly, a real-valued upper-semicontinuous function $u(x)$ defined on $\overline{\Omega}$ is a viscosity subsolution of the stationary problem (2.27) if

(a) for any $x_0 \in \Omega$ and any smooth $\phi(x,t)$ such that $u-\phi$ has a local maximum at x_0 we have

$$f(x_0,u(x_0),D\phi(x_0),D^2\phi(x_0)) + \lambda u(x_0) \leq 0, \text{ and}$$

(b) for any $x_0 \in \partial\Omega$ and any smooth ϕ such that $u-\phi$ has a local maximum on $\overline{\Omega}$ at x_0, we have

$$\min\left\{f(x_0,u(x_0),D\phi(x_0),D^2\phi(x_0)) + \lambda u(x_0), u(x_0) - g(x_0)\right\} \leq 0.$$

A viscosity solution of (2.27) is a continuous function u that is both a subsolution and a supersolution.

In stating these definitions, we have assumed that the final-time data h and the boundary data g are continuous. In Definition 3.1, the pointwise inequality in part (b) can be replaced by an apparently weaker condition analogous to part (b) of Definition 3.2. The equivalence of such a definition with the one stated above is standard; the argument uses barriers of the form $\phi(x,t) = |x-x_0|^2/\delta + (T-t)/\mu$ with δ and μ sufficiently small, and is contained in our proof of Proposition 3.3 (ii) below.

We shall be focusing on the lower and upper semi-relaxed limits of v^ε and u^ε, defined by (2.23) in the time-dependent setting and (2.39) in the stationary case.

3.2 The parabolic case

We are ready to state our main convergence result in the time-dependent setting. The proof seems at first to use only the monotonicity condition (1.3). However it also relies on the consistency of the numerical scheme, Lemma 4.1, which is proved in Section 4. So we also require that $f(t,x,z,p,\Gamma)$ satisfy (1.5)-(1.6), and that the parameters α,β,γ satisfy (2.7) – (2.9). (Actually, since consistency is a local matter, we need only local versions of (1.5)–(1.6), see Remark 3.4.)

Proposition 3.3. *Suppose f and α,β,γ satisfy the hypotheses just listed. Assume furthermore that $\overline{u}$ and $\underline{v}$ are finite for all x near x_0 and all $t \le T$ near t_0. Then*

(i) *if $t_0 < T$, then $\overline{u}$ is a viscosity subsolution at x_0 and $\underline{v}$ is a viscosity supersolution at x_0 (i.e. each satisfies part (a) of the relevant half of Definition 3.1 at x_0).*

(ii) *if $t_0 = T$, then $\overline{u}(x_0) = h(x_0)$ and $\underline{v}(x_0) = h(x_0)$. (In particular, each satisfies part (b) of the relevant half of Definition 3.1 at x_0).*

In particular, if $\overline{u}$ and $\underline{v}$ are finite for all $x \in \mathbb{R}^n$ and $t_* < t \le T$ then they are respectively a viscosity subsolution and a viscosity supersolution of (2.15) on this time interval.

Proof. We give the proof for $\overline{u}$; the argument for $\underline{v}$ is entirely parallel. Focusing first on (i), we fix $t_0 < T$ and $x_0 \in \mathbb{R}^n$, and consider a smooth function ϕ such that $\overline{u} - \phi$ has a local maximum at (x_0,t_0). Adding a constant, we can assume that $\overline{u}(x_0,t_0) = \phi(x_0,t_0)$. Replacing ϕ by $\phi + \|x-x_0\|^4 + |t-t_0|^2$ if necessary, we can assume that the local maximum is strict, i.e. that

$$\overline{u}(x,t) < \phi(x,t) \quad \text{for } 0 < \|(x,t)-(x_0,t_0)\| \le r \tag{3.1}$$

for some $r > 0$.

By the definition of $\overline{u}$ there exists a sequence ε_k, $\tilde{y}_k$, $\tilde{t}_k = T - \tilde{N}_k\varepsilon_k^2$ such that

$$\tilde{y}_k \to x_0, \quad \tilde{t}_k \to t_0, \quad u^{\varepsilon_k}(\tilde{y}_k,\tilde{t}_k) \to \overline{u}(x_0,t_0).$$

Let y_k and $t_k = T - N_k\varepsilon_k^2$ satisfy

$$(u^{\varepsilon_k} - \phi)(y_k,t_k) \ge \sup_{\|(x,t)-(x_0,t_0)\| \le r} (u^{\varepsilon_k} - \phi)(x,t) - \varepsilon_k^3.$$

(Since u^{ε_k} is defined only at discrete times, the sup is taken only over such times.) Evidently

$$(u^{\varepsilon_k}-\phi)(y_k,t_k) \geq (u^{\varepsilon_k}-\phi)(\tilde{y}_k,\tilde{t}_k)-\varepsilon_k^3$$

and the right hand side tends to 0 as $k \to \infty$. It follows using (3.1) that

$$(y_k,t_k) \to (x_0,t_0) \quad \text{and} \quad u^{\varepsilon_k}(y_k,t_k) \to \overline{u}(x_0,t_0)$$

as $k \to \infty$. Setting $\xi_k = u^{\varepsilon_k}(y_k,t_k)-\phi(y_k,t_k)$, we also have by construction that

$$\xi_k \to 0 \text{ and } u^{\varepsilon_k}(x,t) \leq \phi(x,t)+\xi_k+\varepsilon_k^3 \tag{3.2}$$
$$\text{whenever } t = T-n\varepsilon_k^2 \text{ and } \|(x,t)-(x_0,t_0)\| \leq r.$$

Now we use the dynamic programming inequality (2.21) at (y_k,t_k), which can be written concisely as

$$u^{\varepsilon}(y_k,t_k) \leq \sup_{p,\Gamma}\inf_{w}\left\{u^{\varepsilon}(y_k+\Delta x,t_k+\varepsilon_k^2)-\Delta z\right\}$$

with the conventions

$$\Delta x = \varepsilon_k w, \quad \Delta z = \varepsilon_k p\cdot w+\tfrac{1}{2}\varepsilon_k^2\langle \Gamma w,w\rangle+\varepsilon_k^2 f(t,x,u^{\varepsilon_k}(y_k,t_k),p,\Gamma).$$

Using the definition of ξ_k, (3.2), and the fact that Δx and Δz are bounded by a positive power of ε, we conclude that

$$\phi(y_k,t_k)+\xi_k \leq \sup_{p,\Gamma}\inf_{w}\left\{\phi(y_k+\Delta x,t_k+\varepsilon_k^2)+\xi_k+\varepsilon_k^3-\Delta z\right\} \tag{3.3}$$

when k is sufficiently large. Dropping ξ_k from both sides of (3.3), we apply Lemma 4.1 to evaluate the right hand side. Using the smoothness of ϕ and the Lipschitz continuity of f with respect to p and Γ, this gives

$$\phi(y_k,t_k)-\phi(y_k,t_k+\varepsilon_k^2) \leq -\varepsilon_k^2 f(t_k,y_k,u^{\varepsilon_k}(y_k,t_k),D\phi(y_k,t_k),D^2\phi(y_k,t_k))+o(\varepsilon_k^2).$$

Taylor expanding ϕ with respect to time we conclude that

$$-\varepsilon_k^2\partial_t\phi(y_k,t_k) \leq -\varepsilon_k^2 f(t_k,y_k,u^{\varepsilon_k}(y_k,t_k),D\phi(y_k,t_k),D^2\phi(y_k,t_k))+o(\varepsilon_k^2).$$

It follows in the limit $k \to \infty$ that

$$\partial_t\phi(x_0,t_0)-f(t_0,x_0,\overline{u}(x_0,t_0),D\phi(x_0,t_0),D^2\phi(x_0,t_0)) \geq 0.$$

Thus $\overline{u}$ is a viscosity subsolution at (x_0,t_0).

We turn now to (ii), i.e. the case $t_0 = T$. The preceding argument can still be used provided $t_k < T$ for all sufficiently large k. If on the other hand $t_k = T$ for a sequence of $k \to \infty$ then it follows (using the continuity of h and the fact that each u^{ε} has final value h) that $\overline{u}(x_0,T) = h(x_0)$. Thus for any smooth ϕ such that $\overline{u}-\phi$ has a local maximum at (x_0,T) we know that

(3.4) either $\overline{u}(x_0,T) = h(x_0)$ or else

$$\partial_t\phi(x_0,T)-f(T,x_0,\overline{u}(x_0,T),D\phi(x_0,T),D^2\phi(x_0,T)) \geq 0.$$

Moreover this statement applies not only at the the given point x_0, but also at any point nearby.

Now consider

$$\psi(x,t) = \overline{u}(x,t) - \frac{|x-x_0|^2}{\delta} - \frac{T-t}{\mu}$$

where the parameters δ and μ are small and positive. Suppose $\overline{u}$ is uniformly bounded on the closed half-ball $\{\|(x,t)-(x_0,T)\| \le r, t \le T\}$, and let ψ assume its maximum on this half-ball at $(x_{\delta,\mu}, t_{\delta,\mu})$. Clearly

$$(x_{\delta,\mu}, t_{\delta,\mu}) \to (x_0, T) \quad \text{as } \delta, \mu \to 0$$

since $\overline{u}$ is bounded on the half-ball. Moreover it is obvious that

$$\overline{u}(x_{\delta,\mu}, t_{\delta,\mu}) \ge \psi(x_{\delta,\mu}, t_{\delta,\mu}) \ge \overline{u}(x_0, T).$$

Taken together, these relations yield

$$\overline{u}(x_{\delta,\mu}, t_{\delta,\mu}) \to \overline{u}(x_0, T) \quad \text{as } \delta, \mu \to 0.$$

If $t_{\delta,\mu} < T$ then part (i) of the Theorem assures us that

$$(3.5) \qquad -\frac{1}{\mu} - f(t_{\delta,\mu}, x_{\delta,\mu}, \overline{u}(x_{\delta,\mu}, t_{\delta,\mu}), \frac{2(x-x_0)}{\delta}, \frac{2}{\delta}I) \ge 0.$$

If $t_{\delta,\mu} = T$ then (3.4) gives either the same conclusion or else $\overline{u}(x_{\delta,\mu}, T) = h(x_{\delta,\mu})$. But since f is continuous, for any $\delta > 0$ there exists a $\mu > 0$ such that (3.5) cannot happen. Restricting our attention to such choices of δ and μ we conclude that $t_{\delta,\mu} = T$ and $\overline{u}(x_{\delta,\mu}, T) = h(x_{\delta,\mu})$. It follows in the limit $\delta, \mu \to 0$ that $\overline{u}(x_0, T) = h(x_0)$, as asserted. □

Remark 3.4. Since the preceding result is local, only the properties of $f(t,x,z,p,\Gamma)$ near $t = t_0$, $x = x_0$ and $z = \overline{u}$ or $\underline{v}$ are relevant. Therefore while (1.5)–(1.6) assert inequalities whose constants are uniform in t,x,z, it would be enough for Theorem 3.3 that such inequalities hold locally in t, x, and z.

3.3 The elliptic case

We turn now to the stationary setting discussed in Section 2.4. As in the time-dependent setting, our convergence result depends on the fundamental consistency result Lemma 4.4. So we require that the parameters α, β, γ satisfy (2.7) – (2.9), and that $f(x,z,p,\Gamma)$ satisfy not only the monotonicity condition (1.3) but also the Lipschitz continuity and growth conditions (1.8)–(1.9). Our proof that U^ε is well-defined requires that the interest rate λ be large enough, condition (1.7), and that g be uniformly bounded. Finally, concerning the parameters m and M associated with the termination of the game, we assume that $m = M-1$ and M is sufficiently large. This hypothesis (together with the results in Section 5.2) assures the availability of the dynamic programming inequalities stated in Proposition 2.5.

Proposition 3.5. *Suppose f, g, λ and $\alpha, \beta, \gamma, m, M$ satisfy the hypotheses just listed (from which it follows that $\underline{v}$ and $\overline{u}$ are bounded away from $\pm M$ for all $x \in \overline{\Omega}$). Then $\overline{u}$ is a viscosity subsolution and $\underline{v}$ is a viscosity supersolution of (2.27) in $\overline{\Omega}$. More specifically:*

(i) *if $x_0 \in \Omega$, then each of $\overline{u}$ and $\underline{v}$ satisfies part (a) of relevant half of Definition 3.2 at x_0, and*
(ii) *if $x_0 \in \partial\Omega$, then each of $\overline{u}$ and $\underline{v}$ satisfies part (b) of the relevant half of Definition 3.2 at x_0.*

Proof. When $x_0 \in \Omega$, the proof is similar to that of Theorem 3.3 (moreover the argument is more or less included in the discussion below concerning $\underline{v}$.) Therefore we shall focus on the case when $x_0 \in \partial\Omega$.

To show that $\overline{u}$ is a viscosity subsolution at $x_0 \in \partial\Omega$, it suffices to show that $\overline{u}(x_0) \leq g(x_0)$. This is an easy consequency of the dynamic programming principle (2.35). Indeed, if $x \in \Omega$ with $\|x - x_0\| < \varepsilon^{1-\alpha}$ then for any admissible p and Γ

$$\inf_w \left[-\varepsilon p \cdot w - \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle - \varepsilon^2 f(x, u^\varepsilon(x), p, \Gamma) + e^{-\lambda\varepsilon^2} \hat{u}^\varepsilon(x + \varepsilon w) \right] \leq g(x_0) + o(1)$$

since the optimal w does at least as well as a choice such that $x + \varepsilon w \in \mathbb{R}^n \backslash \Omega$. Here the error term $o(1)$ tends to 0 as $\varepsilon \to 0$, uniformly as p and Γ range over the admissible class $\|p\| \leq \varepsilon^{-\beta}$, $\|\Gamma\| \leq \varepsilon^{-\gamma}$. Maximizing over p and Γ, we conclude from (2.35) that $u^\varepsilon(x) \leq g(x_0) + o(1)$ whenever $\|x - x_0\| < \varepsilon^{1-\alpha}$. It follows easily that $\overline{u}(x_0) \leq g(x_0)$ for every $x_0 \in \partial\Omega$, as asserted.

It remains to show that $\underline{v}$ is a viscosity supersolution at any $x_0 \in \partial\Omega$. We may assume that $\underline{v}(x_0) < g(x_0)$ since otherwise the assertion is trivial. Beginning as usual, consider a smooth function ϕ such that that $\underline{v} - \phi$ has a local minimum on $\overline{\Omega}$ at $x_0 \in \partial\Omega$. Adjusting ϕ if necessary, we can assume that $\underline{v}(x_0) = \phi(x_0)$ and that the local minimum is strict, i.e.

$$\underline{v}(x) > \phi(x) \quad \text{for } x \in \overline{\Omega} \cap \{0 < \|x - x_0\| \leq r\} \tag{3.6}$$

for some $r > 0$. By the definition of $\underline{v}$ there exists a sequence $\varepsilon_k > 0$ and $\tilde{y}_k \in \Omega$ such that

$$\tilde{y}_k \to x_0, \quad v^{\varepsilon_k}(\tilde{y}_k) \to \underline{v}(x_0)$$

We may choose $y_k \in \Omega$ such that

$$(v^{\varepsilon_k} - \phi)(y_k) \leq \inf_{\Omega \cap \{\|(x - x_0\| \leq r\}} (v^{\varepsilon_k} - \phi)(x) + \varepsilon_k^3.$$

Evidently

$$(v^{\varepsilon_k} - \phi)(y_k) \leq (v^{\varepsilon_k} - \phi)(\tilde{y}_k) + \varepsilon_k^3$$

and the right hand side tends to 0 as $k \to \infty$. It follows using (3.6) that

$$y_k \to x_0 \quad \text{and} \quad v^{\varepsilon_k}(y_k) \to \underline{v}(x_0)$$

as $k \to \infty$. Setting $\xi_k = v^{\varepsilon_k}(y_k) - \phi(y_k)$, we also have by construction that

$$\xi_k \to 0 \quad \text{and} \quad v^{\varepsilon_k}(x) \geq \phi(x) + \xi_k - \varepsilon_k^3 \text{ whenever } x \in \Omega \text{ with } \|x - x_0\| \leq r. \tag{3.7}$$

We now use the dynamic programming inequality (2.36) at y_k, which can be written concisely as

$$v^\varepsilon(y_k) \geq \sup_{p,\Gamma} \inf_w \left\{ e^{-\lambda\varepsilon^2} \hat{v}^\varepsilon(y_k + \Delta x) - \delta_k \right\}$$

with the conventions

$$\Delta x = \varepsilon_k w, \quad \delta_k = \varepsilon_k p \cdot w + \tfrac{1}{2}\varepsilon_k^2 \langle \Gamma w, w\rangle + \varepsilon_k^2 f(x, v^{\varepsilon_k}(y_k), p, \Gamma).$$

Recall that $v^{\varepsilon_k}(y_k) \to \underline{v}(x_0) < g(x_0)$ as $\varepsilon_k \to 0$, and that for any admissible w, p, Γ, $|\delta_k|$ has a uniform bound that tends to 0 with ε. Therefore a choice of w for which $y_k + \Delta x \notin \Omega$ isn't of interest (it does significantly worse than the choice $w = 0$). Combining this observation with the definition of ξ_k and (3.7), we conclude that

$$\phi(y_k) + \xi_k \geq \sup_{p,\Gamma} \inf_{x+\varepsilon w \in \Omega} \left(e^{-\lambda \varepsilon^2}[\phi(y_k + \Delta x) + \xi_k - \varepsilon_k^3] - \delta_k \right)$$

when k is sufficiently large. We may replace $e^{-\lambda\varepsilon^2}$ by $1 - \lambda\varepsilon^2$ and $e^{-\lambda\varepsilon^2}\phi(y_k + \Delta x)$ by $\phi(y_k + \Delta x) - \lambda\varepsilon^2\phi(y_k)$ while making an error which is only $o(\varepsilon^2)$. Moreover, the inequality is preserved when we drop the constraint $x + \varepsilon w \in \Omega$. Dropping ξ_k from both sides and using the fact that $\xi_k \to 0$, we conclude that

$$\phi(y_k) \geq \sup_{p,\Gamma} \inf_{w} \big(\phi(y_k + \Delta x) - \delta_k - \lambda \varepsilon_k^2 \phi(y_k)\big) + o(\varepsilon_k^2).$$

Evaluating the right hand side using Lemma 4.4, we get

$$0 \geq -\varepsilon_k^2 f(y_k, v^{\varepsilon}(y_k), D\phi(y_k), D^2\phi(y_k)) - \varepsilon_k^2 \lambda \phi(y_k) + o(\varepsilon_k^2).$$

It follows in the limit that

$$f(x_0, \underline{v}(x_0), D\phi(x_0), D^2\phi(x_0)) + \lambda \underline{v}(x_0) \geq 0.$$

Thus $\underline{v}$ is a viscosity supersolution at x_0. □

Remark 3.6. As noted in Remark 3.4, convergence is a local matter. Therefore it is tempting to assert that only the properties of $f(x,z,p,\Gamma)$ near $x = x_0$ and $z = \overline{u}$ or $\underline{v}$ are relevant to proving that $\overline{u}$ is a subsolution and $\underline{v}$ a supersolution at x_0. But to even get started we need to know that u^ε and v^ε are well-defined and satisfy the dynamic programming inequalities (2.35)–(2.36). Since our proof of that assertion is global in character (see Section 5.2), we cannot really discuss $\overline{u}$ or $\underline{v}$ without global hypotheses on f.

4 Consistency

A numerical scheme is said to be consistent if every smooth solution of the PDE satisfies it modulo an error that tends to zero with the step size. This was the essence of our formal argument in Section 2.2 linking the game to the PDE. The present section clarifies the connection between our formal argument and the consistency of the game, by discussing consistency in more conventional terms. The main point is Lemma 4.1, which simultaneously establishes the consistency of our game as a numerical scheme and justifies the handling of the sup/inf (3.3) and its stationary counterpart.

4.1 The parabolic case

Consider the game discussed in Section 2.3 for solving $-\partial_t u+f(t,x,u,Du,D^2u)=0$ in $\mathbb{R}^n$ with final-time data $u(x,T)=h(x)$. The dynamic programming principles (2.21)–(2.22) can be written as

$$u^\varepsilon(x,t)\le S_\varepsilon[x,t,u^\varepsilon(x,t),u^\varepsilon(\cdot,t+\varepsilon^2)] \qquad v^\varepsilon(x,t)\ge S_\varepsilon[x,t,v^\varepsilon(x,t),v^\varepsilon(\cdot,t+\varepsilon^2)]$$

where $S_\varepsilon[x,t,z,\phi]$ is defined for any $x\in\mathbb{R}^n$, $z\in\mathbb{R}$ and $t\le T$ and any continuous function $\phi:\mathbb{R}^n\to\mathbb{R}$ by

(4.1)
$$S_\varepsilon[x,t,z,\phi]=\max_{p,\Gamma}\min_{w}\left[\phi(x+\varepsilon w)-\varepsilon p\cdot w-\tfrac{1}{2}\varepsilon^2\langle\Gamma w,w\rangle-\varepsilon^2 f(t,x,z,p,\Gamma)\right].$$

subject to the usual constraints $\|p\|\le\varepsilon^{-\beta}$, $\|\Gamma\|\le\varepsilon^{-\gamma}$, and $\|w\|\le\varepsilon^{-\alpha}$. Recall that the time-stepping scheme $\phi(x,t)=S_\varepsilon[x,t,\phi(x,t),\phi(\cdot,t+\varepsilon^2)]$ is a consistent scheme for solving our PDE if for any C^∞ function $\phi:\mathbb{R}^n\to\mathbb{R}$ and any $x\in\mathbb{R}^n$, $z\in\mathbb{R}$, and $t\le T$, we have

$$\lim_{\varepsilon\to0}\frac{S_\varepsilon[x,t,z,\phi]-\phi}{\varepsilon^2}=-f(t,x,z,D\phi(x),D^2\phi(x)). \tag{4.2}$$

Fixing x, t, z, and ϕ, Taylor expansion shows that for any $\|w\|\le\varepsilon^{-\alpha}$,

$$\phi(x+\varepsilon w)=\phi(x)+\varepsilon D\phi(x)\cdot w+\tfrac{1}{2}\varepsilon^2\langle D^2\phi(x)w,w\rangle+C\varepsilon^{3-3\alpha}.$$

Since $\alpha<1/3$ by hypothesis, $\varepsilon^{3-3\alpha}=o(\varepsilon^2)$; therefore (4.2) is implied by the assertion that

(4.3)
$$\max_{\|p\|\le\varepsilon^{-\beta},\,\|\Gamma\|\le\varepsilon^{-\gamma}}\ \min_{\|w\|\le\varepsilon^{-\alpha}}\left[\varepsilon(D\phi(x)-p)\cdot w+\tfrac{1}{2}\varepsilon^2\langle(D^2\phi(x)-\Gamma)w,w\rangle-\varepsilon^2 f(t,x,z,p,\Gamma)\right]$$
$$=-\varepsilon^2 f(t,x,z,D\phi(x),D^2\phi(x))+o(\varepsilon^2).$$

The following lemma proves that this relation is true, provided α, β, and γ satisfy conditions (2.7)–(2.9). When we apply the lemma, the convenient choice of ϕ varies with time; therefore it is natural to let ϕ be a function of both x and t (though time enters the argument only as a parameter).

Lemma 4.1. *Let f satisfy (1.3) and (1.5)–(1.6), and assume α,β,γ satisfy (2.7)–(2.9). Then for any x,t,z and any smooth function ϕ defined near (x,t), S_ε being defined by (4.1), we have*

$$S_\varepsilon[x,t,z,\phi]-\phi=-\varepsilon^2 f(t,x,z,D\phi(x,t),D^2\phi(x,t))+o(\varepsilon^2).$$

Moreover the constant implicit in the error term is uniform as x, t, and z range over a compact subset of $\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}$.

Proof. Clearly the max-min defining $S_\varepsilon[x,t,z,\phi]$ is at least $\phi(x,t)-\varepsilon^2 f(t,x,z,D\phi(x,t),D^2\phi(x,t))$, since the choices $p=D\phi(x,t)$, $\Gamma=D^2\phi(x,t)$ make w irrelevant. Our task is therefore to demonstrate the opposite inequality, i.e. to show that for any $\|p\|\le\varepsilon^{-\beta}$

and $\|\Gamma\| \le \varepsilon^{-\gamma}$, there exists w satisfying $\|w\| \le \varepsilon^{-\alpha}$ and

$$\varepsilon(D\phi(x,t)-p)\cdot w + \tfrac{1}{2}\varepsilon^2\langle(D^2\phi-\Gamma)w,w\rangle - \varepsilon^2 f(t,x,z,p,\Gamma) \le -\varepsilon^2 f(t,x,z,D\phi(x,t),D^2\phi(x,t)) + o(\varepsilon^2) \tag{4.4}$$

with an error estimate $o(\varepsilon^2)$ that is independent of p and Γ and locally uniform in x,t,z.

In view of the conditions (2.9) we can pick $\mu > 0$ such that

$$\mu+\gamma < 1-\alpha \quad \text{and} \quad \mu+\gamma r < \alpha+1. \tag{4.5}$$

We will consider separately the three cases

(1) $\|D\phi(X)-p\| \le \varepsilon^\mu$ and $\lambda_{\min}(D^2\phi(X)-\Gamma) \ge -\varepsilon^\alpha$,
(2) $\|D\phi(X)-p\| \le \varepsilon^\mu$ and $\lambda_{\min}(D^2\phi(X)-\Gamma) \le -\varepsilon^\alpha$,
(3) $\|D\phi(X)-p\| \ge \varepsilon^\mu$,

where we write $X=(x,t)$ for notational simplicity, and $\lambda_{\min}(M)$ is the smallest eigenvalue of the symmetric matrix M.

Case (1) is easy: we choose $w=0$. Since $\lambda_{\min}(D^2\phi(X)-\Gamma) \ge -\varepsilon^\alpha$ we have $D^2\phi(X)-\Gamma+\varepsilon^\alpha I \ge 0$ and thus $\Gamma \le D^2\phi(X)+\varepsilon^\alpha I$. Using the monotonicity of f with respect to its last entry, this gives

$$f(t,x,z,p,\Gamma) \ge f(t,x,z,p,D^2\phi(X)+\varepsilon^\alpha I).$$

Since f is locally Lipschitz (1.5), we conclude that

$$f(t,x,z,p,\Gamma) \ge f(t,x,z,p,D^2\phi(X)) + O(\varepsilon^\alpha). \tag{4.6}$$

The constant in the error term is independent of p and Γ, since we are assuming in Case (1) that $\|p-D\phi(X)\| \le \varepsilon^\mu$. Moreover another application of the locally Lipschitz character of f gives

$$f(t,x,z,p,\Gamma) \ge f(t,x,z,D\phi(X),D^2\phi(X)) + O(\varepsilon^\alpha+\varepsilon^\mu). \tag{4.7}$$

Therefore the choice $w=0$ in the left hand side of (4.4) gives

$$\begin{aligned}\varepsilon(D\phi(x,t)-p)\cdot w + \tfrac{1}{2}\varepsilon^2\langle(D^2\phi-\Gamma)w,w\rangle - \varepsilon^2 f(t,x,z,p,\Gamma)\\ = -\varepsilon^2 f(t,x,z,p,\Gamma) \le -\varepsilon^2 f(t,x,z,D\phi(X),D^2\phi(X)) + o(\varepsilon^2)\end{aligned}$$

as desired.

For case (2), we choose w to be an eigenvector for the minimum eigenvalue $\lambda = \lambda_{\min}(D^2\phi(X)-\Gamma)$, of norm $\varepsilon^{-\alpha}$. Note that since f is monotone in its last entry,

$$f(t,x,z,p,\Gamma) \ge f(t,x,z,p,D^2\phi(X)-\lambda I).$$

Choosing w as announced, and changing the sign if necessary to make $(D\phi(X)-p)\cdot w \le 0$, the left hand side of (4.4) becomes

$$\begin{aligned}\varepsilon(D\phi(X)-p)\cdot w + \tfrac{1}{2}\varepsilon^2\langle(D^2\phi-\Gamma)w,w\rangle - \varepsilon^2 f(t,x,z,p,\Gamma)\\ \le \tfrac{1}{2}\varepsilon^{2-2\alpha}\lambda - \varepsilon^2 f(t,x,z,p,D^2\phi(X)-\lambda I)\end{aligned}$$

If $-1 \le \lambda \le -\varepsilon^{\alpha}$ then $\varepsilon^{2-2\alpha}\lambda \le -\varepsilon^{2-\alpha}$ and $f(t,x,z,p,D^2\phi(X)-\lambda I)$ is bounded. So for such λ we have

$$\tfrac{1}{2}\varepsilon^{2-2\alpha}\lambda - \varepsilon^2 f(t,x,z,p,D^2\phi(X)-\lambda I) \le -\tfrac{1}{2}\varepsilon^{2-\alpha} + O(\varepsilon^2). \tag{4.8}$$

In this case we are done, since the right hand side is $\le -\varepsilon^2 f(t,x,z,D\phi(X),D^2\phi(X))$ when ε is sufficiently small.

To complete case (2), suppose $\lambda \le -1$. Then using the growth hypothesis (1.6) and remembering that p is near $D\phi(X)$ we have

$$\tfrac{1}{2}\varepsilon^{2-2\alpha}\lambda - \varepsilon^2 f(t,x,z,p,D^2\phi(X)-\lambda I) \le -\tfrac{1}{2}\varepsilon^{2-2\alpha}|\lambda| + C\varepsilon^2(1+|\lambda|^r). \tag{4.9}$$

Now notice that $|\lambda| \le C(1+\|\Gamma\|) \le C\varepsilon^{-\gamma}$. Since $\gamma(r-1) < 2\alpha$ we have $\varepsilon^{2-2\alpha}|\lambda| \gg \varepsilon^2|\lambda|^r$. Therefore

$$-\frac{1}{2}\varepsilon^{2-2\alpha}|\lambda| + C\varepsilon^2|\lambda|^r \le -\frac{1}{4}\varepsilon^{2-2\alpha} \le -\varepsilon^2 f(t,x,z,D\phi(X),D^2\phi(X))$$

when ε is sufficiently small. Case (2) is now complete.

Finally consider case (3), when $\|D\phi(X)-p\| \ge \varepsilon^{\mu}$. In this case we take w to be parallel to $D\phi(X)-p$, with norm $\varepsilon^{-\alpha}$, and with the sign chosen so that

$$\varepsilon(D\phi(X)-p)\cdot w = -\varepsilon^{1-\alpha}\|D\phi(X)-p\| \le -\varepsilon^{1-\alpha+\mu}.$$

Estimating the other terms on the left hand side of (4.4), we have

$$|\varepsilon^2\langle(D^2\phi-\Gamma)w,w\rangle| \le \varepsilon^2(C+\|\Gamma\|)\|w\|^2 \le C\varepsilon^{-\gamma+2-2\alpha}$$

and

$$\varepsilon^2|f(t,x,z,p,\Gamma)| \le C\varepsilon^2\left(1+\|p\|^q+\|\Gamma\|^r\right) \le C\left(\varepsilon^2+\varepsilon^2\|p\|^q+\varepsilon^{2-\gamma r}\right). \tag{4.10}$$

Thus

$$\begin{aligned}\varepsilon(D\phi(X)-p)\cdot w &+ \tfrac{1}{2}\varepsilon^2\langle(D^2(X)\phi-\Gamma)w,w\rangle - \varepsilon^2 f(t,x,z,p,\Gamma)\\ &\le -\varepsilon^{1-\alpha}\|D\phi(X)-p\| + C\varepsilon^2\|p\|^q + O(\varepsilon^{-\gamma+2-2\alpha}+\varepsilon^{2-\gamma r}).\end{aligned}$$

Since $\varepsilon^{1-\alpha}\|D\phi(X)-p\| \ge \varepsilon^{1-\alpha+\mu}$ we have

$$\varepsilon^{-\gamma+2-2\alpha}+\varepsilon^{2-\gamma r} \ll \varepsilon^{1-\alpha}\|D\phi(X)-p\|$$

using (4.5). Thus we conclude that

$$\begin{aligned}\varepsilon(D\phi(x,t)-p)\cdot w &+ \tfrac{1}{2}\varepsilon^2\langle(D^2\phi-\Gamma)w,w\rangle - \varepsilon^2 f(t,x,z,p,\Gamma)\\ &\le -\tfrac{1}{2}\varepsilon^{1-\alpha}\|D\phi(X)-p\| + C\varepsilon^2\|p\|^q.\end{aligned}$$

If $\|p\| \le 2\|D\phi(X)\|$ then $\varepsilon^2\|p\|^q = O(\varepsilon^2) \ll \varepsilon^{1-\alpha+\mu}$, using (4.5). If $\|p\| \ge 2\|D\phi(X)\|$ then $\varepsilon^{1-\alpha}\|D\phi(X)-p\| \sim \varepsilon^{1-\alpha}\|p\| \gg \varepsilon^2\|p\|^q$, using the condition on β in (2.9). In either case the term $\varepsilon^{1-\alpha}\|D\phi(X)-p\|$ dominates and we get

$$\varepsilon(D\phi(x,t)-p)\cdot w + \tfrac{1}{2}\varepsilon^2\langle(D^2\phi-\Gamma)w,w\rangle - \varepsilon^2 f(t,x,z,p,\Gamma) \le -C\varepsilon^{1-\alpha+\mu}.$$

When ε is small this is certainly $\le -\varepsilon^2 f(t,x,z,D\phi(X),D^2\phi(X))$. Case (3) is now complete. □

Remark 4.2. Since t, x, and z are held fixed in Proposition (4.1), the proof did not need the full force of conditions (1.5) or (1.6). Rather, all we needed was that these estimates hold locally in x,t,z, with constants that are uniform if t,x,z stay in a compact set.

To prove stability in Section 5, will need the following more global variant of Lemma 4.1. This is where we need the uniformity of the constants in (1.5)–(1.6) in x and t, and the growth condition (1.6).

Lemma 4.3. *Let f satisfy conditions (1.3)-(1.6), and assume as usual that α,β,γ satisfy (2.7)–(2.9). Then, for any x,t,z and any constant m_1, we have*

$$|S_\varepsilon[x,t,z,m_1]-m_1| \leq C(1+|z|)\varepsilon^2,$$

with a constant C that is independent of x, t, z, and m_1.

Proof. The overall argument is the same as used for Lemma 4.1, but we must now pay attention to the uniformity of the constant. The lower bound

$$S_\varepsilon[x,t,z,m_1]-m_1 \geq -\varepsilon^2 f(t,x,z,0,0) \geq -C(1+|z|)\varepsilon^2$$

is obvious, by considering $w=0$ and using (1.4). For the corresponding upper bound we consider the same three cases as before. In case 1 the estimate (4.7) with $\phi \equiv m_1$ gets replaced by

$$f(t,x,z,p,\Gamma) \geq f(t,x,z,0,0)+C(1+|z|)(\varepsilon^\alpha+\varepsilon^\mu)$$

whence

$$-\varepsilon^2 f(t,x,z,p,\Gamma) \leq -C(1+|z|)\varepsilon^2.$$

In case 2 the estimate (4.8) with $\phi \equiv m_1$ gets replaced by

$$\tfrac{1}{2}\varepsilon^{2-2\alpha}\lambda - \varepsilon^2 f(t,x,z,p,-\lambda I) \leq -\tfrac{1}{2}\varepsilon^{2-\alpha}+C(1+|z|)\varepsilon^2$$

and we get an estimate of the desired form by dropping the first term. In second half of case 2 and the handling of case 3 we used the growth estimate (1.6); since z enters linearly on the right hand side of (1.6), the previous calculations still apply but we get an additional term of the form $C|z|\varepsilon^2$. Thus, in all three cases we find that when p and Γ are admissible,

$$\min_w \left[-\varepsilon p\cdot w - \tfrac{1}{2}\varepsilon^2\langle \Gamma w,w\rangle - \varepsilon^2 f(t,x,z,p,\Gamma)\right] \leq C(1+|z|)\varepsilon^2.$$

Maximizing over p and Γ gives the desired inequality $S_\varepsilon[x,t,z,m_1]-m_1 \leq C(1+|z|)\varepsilon^2$. □

Recently Caffarelli and Souganidis have shown how refined consistency results can sometimes be used to prove convergence with a rate, though the viscosity solution may not be differentiable [12, 13]. Their method requires an improved estimate for the consistency error when the test function ϕ is a quadratic polynomial in x. We remark in passing that such an estimate is relatively easy in the present context.

Indeed, if ϕ is a quadratic polynomial in x then for any x,t,z the arguments used to prove Lemma 4.1 also show that

$$(4.11)\quad \begin{aligned} 0 &\leq S_\varepsilon[x,t,z,\phi] - \phi(x) + \varepsilon^2 f(t,x,z,D\phi(x),D^2\phi) \\ &\leq \varepsilon^2 \sup_{\|\Gamma - D^2\phi\| \leq \varepsilon^\alpha,\, \|p - D\phi(x)\| \leq \varepsilon^\mu} |f(t,x,z,p,\Gamma) - f(t,x,z,D\phi(x),D^2\phi)|. \end{aligned}$$

when ε is sufficiently small. We omit the proof, since this assertion will not be used in what follows. (It remains an open question whether the method of [12, 13] can be used to prove convergence with a rate in our setting.)

4.2 The elliptic case

For the game corresponding to the stationary equation, we consider the operator S_ε defined, for any $x \in \Omega$, $z \in \mathbb{R}$, and any continuous function $\phi : \Omega \to \mathbb{R}$, by

(4.12)

$$S_\varepsilon[x,z,\phi] = \sup_{p,\Gamma} \inf_w \left[e^{-\lambda\varepsilon^2} \hat{\phi}(x+\varepsilon w) - \varepsilon p \cdot w - \tfrac{1}{2}\varepsilon^2 \langle \Gamma w, w \rangle - \varepsilon^2 f(x,z,p,\Gamma) \right].$$

with the usual conventions that p,Γ are constrained by (2.10), w is constrained by (2.11), and $\hat{\phi}$ is defined by (2.37). The dynamic programming principles (2.35)–(2.36) can be written as

$$u^\varepsilon(x) \leq S_\varepsilon[x, u^\varepsilon(x), u^\varepsilon] \quad \text{and} \quad v^\varepsilon(x) \geq S_\varepsilon[x, v^\varepsilon(x), v^\varepsilon].$$

The analogue of Lemma 4.1 is:

Lemma 4.4. *Let f satisfy conditions (1.3) and (1.8) – (1.9). Assume α, β, γ satisfy (2.7) – (2.9). Then for any $x \in \Omega$, $z \in \mathbb{R}$ and any smooth function ϕ defined near x, S_ε being defined by (4.12), we have*

$$(4.13)\qquad S_\varepsilon[x,z,\phi] - \phi = -\varepsilon^2 (f(x,z,D\phi(x),D^2\phi(x)) + \lambda\phi(x)) + o(\varepsilon^2).$$

Moreover the constants implicit in the error term are uniform as x and z range over a compact subset of $\Omega \times \mathbb{R}$.

Proof. The argument is entirely parallel to the proof of Lemma 4.1. (Note that since x is an interior point, the extension of ϕ outside Ω is irrelevant: $\hat{\phi}(x+\varepsilon w) = \phi(x+\varepsilon w)$ when ε is sufficiently small.) □

For stability, we will need a variant of the preceding lemma. This is where we use the hypothesis (1.7) on the z-dependence of f, i.e. the condition that $|f(x,z,0,0)| \leq (\lambda - \eta)|z| + C_*$.

Lemma 4.5. *Let f satisfy (1.3) and (1.7)–(1.9), and assume as always that α, β, γ satisfy (2.7)–(2.9). Fix $M > 0$, and let $m \leq M$ be a positive constant such that $m < \|g\|_{L^\infty}$. Then for any $|z| \leq M$ and any $x \in \Omega$ we have*

$$S_\varepsilon[x,z,m] \leq m + \varepsilon^2(1 + C_* + (\lambda - \eta)|z|) - \varepsilon^2 \lambda m.$$

for all sufficiently small ε. (The smallness condition on ε depends on M, but not on x.)

Proof. The hypothesis $|z| \leq M$ assures that the constants in (1.8) and (1.9) are uniform. There the constant implicit in the error term of (4.13) is uniform for ε sufficiently small, and the smallness condition depends only on M. The hypothesis $m > \|g\|_{L^\infty}$ (resp. $m < \|g\|_{L^\infty}$) assures that $\hat{m} \leq m$ (resp. $\hat{m} \geq m$), so the appearance of $\hat{\phi}$ rather than ϕ on the right hand side of (4.12) does not interfere with the desired estimate. Therefore the usual argument applied to the constant function $\phi = m$ gives

$$S_\varepsilon[x,z,m] \leq e^{-\lambda\varepsilon^2} m - \varepsilon^2 f(x,z,0,0) + o(\varepsilon^2).$$

Estimating $f(x,z,0,0)$ by (1.7) and noting that $e^{-\lambda\varepsilon^2} m = (1-\lambda\varepsilon^2)m + O(\varepsilon^4 m)$, we easily deduce the desired estimate. □

5 Stability

In the time-dependent setting, we showed in Section 3.2 that if v^ε and u^ε remain bounded as $\varepsilon \to 0$ then $\underline{v}$ is a supersolution and $\overline{u}$ is a subsolution. The argument was local, using mainly the consistency of the game as a numerical solution scheme. It remains to prove that v^ε and u^ε are indeed bounded; this is achieved in Section 5.1.

For the stationary setting, we must do more. Even the existence of $U^\varepsilon(x,z)$ remains to be proved. We also need to show that the associated functions u^ε and v^ε are bounded away from M, so that (by Lemma 2.5) they satisfy the dynamic programming inequalities at each $x \in \Omega$. These goals will be achieved in Section 5.2, provided the parameters M and m associated with the termination of the game satisfy (i) $m = M-1$ and (ii) M is sufficiently large. We also show in Section 5.2 that if f is a non-decreasing function of z then U^ε is strictly decreasing in z; when it applies, this result assures that $\underline{v} \leq \overline{u}$, permitting us to conclude that $\underline{v} = \overline{u}$ is the unique viscosity solution if the boundary value problem has a comparison principle.

5.1 The parabolic case

Since f grows at most linearly in z, we expect u to grow at most exponentially in $T-t$. The following result confirms this expectation.

Proposition 5.1. *Assume the hypotheses of Lemma 4.3, and suppose furthermore that the final-time data are uniformly bounded:*

$$|h(x)| \leq N \quad \text{for all } x \in \mathbb{R}^n. \tag{5.1}$$

Then there exists a constant $c > 0$ *(independent of* ε*) such that*

$$u^\varepsilon(x,t) \leq Nc^{T-t} \quad \text{and} \quad v^\varepsilon(x,t) \geq -Nc^{T-t} \text{ for all } x \in \mathbb{R}^n \tag{5.2}$$

for every $t < T$.

Proof. We shall demonstrate the upper bound on u^ε; the proof of the lower bound on v^ε is entirely parallel. The argument proceeds backward in time $t_k = T - k\varepsilon^2$. At $k = 0$ we have a uniform bound $u^\varepsilon(x,T) = h(x) \leq N_0$ by hypothesis, and we may assume without loss of generality that $N_0 \geq 1$.

Now suppose that for fixed $k \geq 0$ we already know a bound $u^\varepsilon(x,t_k) \leq N_k$, where $N_k \geq 1$. By (2.21) we have

$$u^\varepsilon(x,t_k-\varepsilon^2) \leq S_\varepsilon[x,t_k-\varepsilon^2,u^\varepsilon(x,t_k-\varepsilon^2),u^\varepsilon(\cdot,t_k)].$$

Since S_ε is monotone in its last argument, we have

$$S_\varepsilon[x,t_k-\varepsilon^2,u^\varepsilon(x,t_k-\varepsilon^2),u^\varepsilon(\cdot,t_k)] \leq S_\varepsilon[x,t_k-\varepsilon^2,u^\varepsilon(x,t_k-\varepsilon^2),N_k].$$

But by Lemma 4.3 we have

$$S_\varepsilon[x,t_k-\varepsilon^2,u^\varepsilon(x,t_k-\varepsilon^2),N_k] \leq N_k + C(1+|u^\varepsilon(x,t_k-\varepsilon^2)|)\varepsilon^2.$$

If $u^\varepsilon(x,t_k-\varepsilon^2) \leq 0$ then we are done (recall we are looking for an upper bound $N_{k+1} \geq 1$); if not then we have

$$u^\varepsilon(x,t_k-\varepsilon^2) \leq N_k + C(1+u^\varepsilon(x,t_k-\varepsilon^2))\varepsilon^2.$$

It follows that $u^\varepsilon(x,t_k-\varepsilon^2) = u^\varepsilon(x,T-(k+1)\varepsilon^2)$ is bounded by N_{k+1} where

$$N_{k+1}(1-C\varepsilon^2) = N_k + C\varepsilon^2 \leq N_k(1+C\varepsilon^2).$$

Combining these estimates, we easily deduce that $u^\varepsilon(x,T-k\varepsilon^2) \leq \tilde{N}_k$ for all k with

$$\tilde{N}_k = N_0\left(\frac{1+C\varepsilon^2}{1-C\varepsilon^2}\right)^k.$$

Since $k = (T-t)/\varepsilon^2$, we have shown that

$$u^\varepsilon(x,t) \leq N_0 c_\varepsilon^{T-t}$$

with

$$c_\varepsilon = \left(\frac{1+C\varepsilon^2}{1-C\varepsilon^2}\right)^{1/\varepsilon^2}.$$

Since c_ε has a finite limit as $\varepsilon \to 0$ we obtain a bound on u^ε of the desired form (5.2). □

5.2 The elliptic case

We shall assume throughout this section that the parameters M and m controlling the termination of the game are related by $m = M - 1$; in addition, we need to assume that M is sufficiently large. Our plan is to show, using a fixed point argument, the existence of a function $U^\varepsilon(x,z)$ (defined for all $x \in \Omega$ and $|z| < M$) satisfying (2.31) and also

$$-z - m \leq U^\varepsilon(x,z) \leq -z + m. \tag{5.3}$$

This implies that $U^\varepsilon(x,z)<0$ when $z>m$, and $U^\varepsilon(x,z)>0$ when $z<-m$. Recalling the definitions (2.33)-(2.34) of u^ε and v^ε, it follows from (5.3) that

$$|v^\varepsilon(x)| \le m, \quad |u^\varepsilon(x)| \le m \tag{5.4}$$

for all $x\in\Omega$.

It is convenient to work with $V^\varepsilon(x,z)=U^\varepsilon(x,z)+z$ rather than U^ε, since this turns (5.3) into the statement that

$$|V^\varepsilon(x,z)| \le m,$$

whose right-hand side is constant. The dynamic programming principle (2.31) for U^ε is equivalent (after a bit of manipulation) to the statement that for all $x\in\Omega$ and all $|z|<M$,

$$V^\varepsilon(x,z)=\sup_{p,\Gamma}\inf_w \begin{cases} e^{-\lambda\varepsilon^2}V^\varepsilon(x',z')-\delta & \text{if } x'\in\Omega \text{ and } |z'|<M \\ e^{-\lambda\varepsilon^2}g(x')-\delta & \text{if } x'\notin\Omega \text{ and } |z'|<M \\ -m & \text{if } z'\ge M \\ +m & \text{if } z'\le -M \end{cases} \tag{5.5}$$

where $x'=x+\varepsilon w$ and $z'=e^{\lambda\varepsilon^2}(z+\delta)$, with δ defined as in (2.28):

$$\delta=\varepsilon p\cdot w+\frac{\varepsilon^2}{2}\langle\Gamma w,w\rangle+\varepsilon^2 f(x,z,p,\Gamma).$$

Here p,Γ, and w are constrained as usual by (2.10)–(2.11).

Preparing for our fixed point argument, consider for any L^∞ function ϕ defined on $\Omega\times(-M,M)$, the function $(x,z)\mapsto R_\varepsilon[x,z,\phi]$, defined by

$$R_\varepsilon[x,z,\phi]=\sup_{p,\Gamma}\inf_w \begin{cases} e^{-\lambda\varepsilon^2}\phi(x',z')-\delta & \text{if } x'\in\Omega \text{ and } |z'|<M \\ e^{-\lambda\varepsilon^2}g(x')-\delta & \text{if } x'\notin\Omega \text{ and } |z'|<M \\ -m & \text{if } z'\ge M \\ +m & \text{if } z'\le -M. \end{cases} \tag{5.6}$$

We shall identify V^ε as the unique fixed point of the mapping $\phi(\cdot,\cdot)\mapsto R_\varepsilon[\cdot,\cdot,\phi]$.

Proposition 5.2. *Assume the hypotheses of Lemma 4.5. Suppose further that $m=M-1$, with M chosen large enough to satisfy condition (5.9) below. Then for all sufficiently small ε, the map $\phi(\cdot,\cdot)\mapsto R_\varepsilon[\cdot,\cdot,\phi]$ is a contraction in the L^∞ norm, which preserves the ball $\|\phi\|_{L^\infty(\Omega\times(-M,M))}\le m$. In particular it has a unique fixed point, which solves (5.5) and has L^∞ norm at most m.*

Proof. First we show that the map is a contraction. (This part of the proof works for any m and M). Let ϕ_i, $i=1,2$ be two L^∞ functions defined on $\Omega\times(-M,M)$ to $\mathbb{R}$. If we hold p, Γ, and w fixed, it is clear that

$$e^{-\lambda\varepsilon^2}\phi_1(x',z')-\delta \le e^{-\lambda\varepsilon^2}\phi_2(x',z')-\delta+e^{-\lambda\varepsilon^2}\|\phi_1-\phi_2\|_{L^\infty}.$$

Since the second, third, and fourth alternatives on the right hand side of (5.6) are independent of ϕ, it follows (after minimizing over w and maximizing over p,Γ) that

$$R_\varepsilon[x,z,\phi_1] \leq R_\varepsilon[x,z,\phi_2] + e^{-\lambda\varepsilon^2}\|\phi_1-\phi_2\|_{L^\infty}$$

for each x,z. Reversing the roles of ϕ_1 and ϕ_2, we conclude that

$$\|R_\varepsilon[\cdot,\cdot,\phi_1]-R_\varepsilon[\cdot,\cdot,\phi_2]\|_{L^\infty} \leq e^{-\lambda\varepsilon^2}\|\phi_1-\phi_2\|_{L^\infty}.$$

Thus the map is a contraction for any ε.

Now we prove that if M is large enough and $m = M-1$, the map preserves the ball $\|\phi\|_{L^\infty(\Omega\times(-M,M))} \leq m$. Since $R_\varepsilon[x,z,\phi]$ is monotone in its last argument it suffices to show that

$$R_\varepsilon[x,z,m] \leq m \quad \text{and} \quad R_\varepsilon[x,z,-m] \geq -m. \tag{5.7}$$

For the first half of (5.7), let p and Γ be fixed, and consider

$$\inf_w \begin{cases} e^{-\lambda\varepsilon^2}m-\delta & \text{if } x'\in\Omega \text{ and } |z'|<M \\ e^{-\lambda\varepsilon^2}g(x')-\delta & \text{if } x'\notin\Omega \text{ and } |z'|<M \\ -m & \text{if } z'\geq M \\ +m & \text{if } z'\leq -M. \end{cases} \tag{5.8}$$

If a minimizing sequence uses the second alternative then the inf is less than m, since we have assumed that $\|g\|_{L^\infty} < m$, and δ is bounded by a positive power of ε. If a minimizing sequence uses the third or fourth alternatives then the inf is $\pm m$. In the remaining case, when all minimizing sequences use the first alternative, we apply Lemma 4.5 to see that (5.8) is bounded above by

$$m+\varepsilon^2(1+C_*+(\lambda-\eta)|z|)-\varepsilon^2\lambda m$$

Since $m = M-1$ and $|z|\leq M$, this is at most

$$m+\varepsilon^2(1+C_*+(\lambda-\eta)M)-\varepsilon^2\lambda(M-1) = m-\varepsilon^2\eta M+\varepsilon^2(1+C_*+\lambda).$$

Thus, assuming

$$M > (1+C_*+\lambda)/\eta \tag{5.9}$$

we deduce that (5.8) is bounded above by m. Taking the supremum over p and Γ, it follows that $R_\varepsilon[x,z,m]\leq m$, as asserted.

For the second half of (5.7), the argument is similar but a bit easier. Consider the choice $p=0$, $\Gamma=0$. We claim that the analogue of (5.8) with m replaced by $-m$ is bounded below by $-m$. If a minimizing sequence uses the second, third, or fourth alternative this is clear. If it uses the first alternative, then the value is

$$-e^{-\lambda\varepsilon^2}m-\delta = -e^{-\lambda\varepsilon^2}m-\varepsilon^2 f(x,z,0,0)$$

(the value of w is irrelevant since $p=\Gamma=0$). By (1.7) this is at most

$$-e^{-\lambda\varepsilon^2}m-\varepsilon^2[(\lambda-\eta)|z|+C_*] \geq -(1-\lambda\varepsilon^2)m-\varepsilon^2(1+C_*+(\lambda-\eta)M).$$

Arguing as before, this is bounded below by $-m$ when M satisfies (5.9). The preceding calculation was for $p = \Gamma = 0$, but the sup over p and Γ can only be larger. Thus it follows that $R_\varepsilon[x,z,-m] \geq -m$, completing the proof of (5.7).

We have shown that the map $\phi(\cdot,\cdot) \mapsto R_\varepsilon[\cdot,\cdot,\phi]$ preserves the ball

$$\|\phi\|_{L^\infty(\Omega\times(-M,M))} \leq m.$$

Since it is also a contraction, the map has a unique fixed point, which lies in this ball. □

These results justify the discussion of the stationary case given in Section 2, by showing that (i) the value functions u^ε and v^ε are well-defined, and bounded independent of ε, and (ii) they satisfy the dynamic programming inequalities:

Proposition 5.3. *Suppose f satisfies (1.3) and (1.7) – (1.9), and suppose the boundary condition g is uniformly bounded. Assume the parameters α,β,γ,m,M determining the stationary version of the game satisfy (2.7) – (2.9) and $m = M-1$, and assume M is large enough that (5.9) holds and $m > \|g\|_{L^\infty}$. Let V^ε be the solution of (5.5) obtained by Proposition 5.2, and let $U^\varepsilon(x,z) = V^\varepsilon(x,z) - z$. Then the associated functions $u^\varepsilon, v^\varepsilon$ defined by (2.33) – (2.34) satisfy $|u^\varepsilon| \leq m$ and $|v^\varepsilon| \leq m$ for all sufficiently small ε, and they satisfy the dynamic programming inequalities (2.35) and (2.36) at all points $x \in \Omega$.*

Proof. The bounds on u^ε and v^ε were demonstrated in (5.4). These bounds assure that the dynamic programming inequalities hold for all $x \in \Omega$, as a consequence of Proposition. 2.5). □

We close this section with the stationary analogue of Lemma 2.4.

Lemma 5.4. *Under the hypotheses of Proposition 5.2, suppose in addition that*

$$f(x,z_1,p,\Gamma) \geq f(x,z_0,p,\Gamma) \quad \text{whenever } z_1 > z_0. \tag{5.10}$$

Then U^ε satisfies

$$U^\varepsilon(x,z_1) \leq U^\varepsilon(x,z_0) - (z_1 - z_0) \quad \text{whenever } z_1 > z_0. \tag{5.11}$$

In particular, U^ε is strictly decreasing in z and $v^\varepsilon = u^\varepsilon$.

Proof. In terms of V^ε, (5.11) asserts that

$$V^\varepsilon(x,z_1) \leq V^\varepsilon(x,z_0) \quad \text{whenever } z_1 > z_0.$$

Since V^ε is a fixed point of the map taking $\phi(x,z)$ to $\tilde{\phi}(x,z) = R_\varepsilon[x,z,\phi]$, it suffices to show that if

$$\phi(x,z_1) \leq \phi(x,z_0) \quad \text{whenever } z_1 > z_0$$

then

$$R_\varepsilon[x,z_1,\phi] \leq R_\varepsilon[x,z_0,\phi] \quad \text{whenever } z_1 > z_0.$$

Let $z_1 > z_0$ be fixed, and consider any p, Γ, w. For $i = 0, 1$ let

$$\delta_i = \varepsilon p \cdot w + \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle + \varepsilon^2 f(x, z_i, p, \Gamma) \quad \text{and} \quad z_i' = e^{\lambda\varepsilon^2}(z_i + \delta_i)$$

so that $R_\varepsilon[x, z_i, \phi]$ is defined by (5.6) with z' replaced by z_i'. Note that our hypothesis on f gives

(5.12) $$\delta_1 \geq \delta_0 \text{ and } z_1' > z_0' \text{ whenever } z_1 > z_0.$$

Our goal is to show that

(5.13)
$$\left.\begin{array}{ll} e^{-\lambda\varepsilon^2}\phi(x', z_1') - \delta_1 & \text{if } x' \in \Omega \text{ and } |z_1'| < M \\ e^{-\lambda\varepsilon^2} g(x') - \delta_1 & \text{if } x' \notin \Omega \text{ and } |z_1'| < M \\ -m & \text{if } z_1' \geq M \\ +m & \text{if } z_1' \leq -M \end{array}\right\} \leq \left\{\begin{array}{ll} e^{-\lambda\varepsilon^2}\phi(x', z_0') - \delta_0 & \text{if } x' \in \Omega \text{ and } |z_0'| < M \\ e^{-\lambda\varepsilon^2} g(x') - \delta_0 & \text{if } x' \notin \Omega \text{ and } |z_0'| < M \\ -m & \text{if } z_0' \geq M \\ +m & \text{if } z_0' \leq -M. \end{array}\right.$$

If $z_0' \geq M$ then we must also have $z_1' \geq M$ by (5.12), so both sides of (5.13) are equal to $-m$. If $z_0' \leq -M$ then the right hand side is m, whereas the left hand side is less than or equal to m by the proof of Proposition 5.2. If $|z_0'| < M$ then by (5.12) either the left hand side is in the corresponding regime or else $z_0' < M < z_1'$. In the former situation the desired inequality is immediate from (5.12); in the latter situation the left hand side is $-m$ whereas the right hand side is greater than or equal to $-m$ by the proof of Proposition 5.2. Since we have considered all the possible cases, (5.13) has been verified and the Lemma is proved. □

6 Uniform continuity of U^ε and u^ε

We have now achieved the main goals of this paper: the stability of the scheme, and the convergence of the associated value functions to viscosity super and subsolutions of the associated PDE. If the PDE has a comparison principle (and if $\underline{v} \leq \overline{u}$, which is always true in the time-dependent setting) it follows that $\lim u^\varepsilon = \lim v^\varepsilon$ exists and is the unique viscosity solution of the PDE.

It is natural to ask whether we can get some additional control on the value functions directly from the game. The present section provides two results of this type for the time-dependent version of our game. The first, presented in Section 6.1, is a uniform Lipschitz estimate. The proof assumes that $u^\varepsilon = v^\varepsilon$, and that f is globally Lipschitz in t, x, z. When it applies, this result gives an alternative proof that the functions $\{u^\varepsilon\}$ have a limit, at least along a subsequence $\varepsilon_j \to 0$. (If the PDE has a comparison principle then the limit is the viscosity solution and there is no need for a subsequence; however our compactness result applies even when no comparison principle is known.)

Our second result, presented in Section 6.2, concerns a sufficient condition for knowing that $u^\varepsilon = v^\varepsilon$. We showed in Lemma 2.4 that this relation is always true if f is a non-decreasing function of z. Lemma 6.4 complements that result by showing that if f is merely Lipschitz continuous in z, then $u^\varepsilon = v^\varepsilon$ when $T - t < C$ (with C independent of ε).

6.1 Compactness of u^ε

The basic idea is familiar: for a translation-invariant equation such as $-u_t + f(Du, D^2u) = 0$, a global-in-space Lipschitz bound propagates in time (see e.g. [16] or [19] or Appendix B of [35]). When f depends on x, t, and u the argument gets more complicated, but it is not fundamentally different.

We shall assume throughout this subsection that $u^\varepsilon = v^\varepsilon$. Indeed, our main tool will be the dynamic programming equality

$$u^\varepsilon(x,t) = S_\varepsilon[x,t,u^\varepsilon(x,t),u^\varepsilon(\cdot,t+\varepsilon^2)] \tag{6.1}$$

where S_ε is defined by (4.1). The validity of (6.1) follows immediately from the inequalities (2.21)–(2.22) and the assumption that $u^\varepsilon = v^\varepsilon$.

We shall also assume throughout this section certain conditions on the PDE and its final-time data in addition to those introduced in the prior sections. First: f is uniformly Lipschitz in t, x, and z; in other words, there is a constant C such that for all t,t',x,x',z,z' and all p,Γ,

$$|f(t,x,z,p,\Gamma) - f(t',x',z',p,\Gamma)| \le C(|z-z'| + |x-x'| + |t-t'|). \tag{6.2}$$

Second: we assume that h satisfies a uniform Lipschitz condition

$$|h(x) - h(x')| \le C|x-x'| \tag{6.3}$$

for all $x,x' \in \mathbb{R}^n$ (notice that this implies $\|Dh\|_{L^\infty} \le C$). Third: we assume that h is a C^2 function, with a uniform bound on its second derivatives:

$$\|D^2h\|_{L^\infty(\mathbb{R}^n)} \le C. \tag{6.4}$$

We begin with a lemma quantifying the stability of the process of stepping backward in time.

Lemma 6.1. *Assume f satisfies the assumptions (1.3)—(1.6) and in addition (6.2). Let ϕ_i ($i = 1,2$) be two continuous functions on $\mathbb{R}^n$ and let $\tilde\phi_i$ be functions which satisfy for some t_i, x_i and for all x,*

$$\tilde\phi_i(x) = \max_{p,\Gamma}\min_{w}\left[\phi_i(x+\varepsilon w) - \left(\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle\Gamma w,w\rangle + \varepsilon^2 f(t_i, x+x_i, \tilde\phi_i(x), p, \Gamma)\right)\right]$$

where p,Γ and w are constrained respectively by (2.10), (2.11). Then there exists a constant C (depending only on the constant in (6.2)) such that

$$\|\tilde\phi_1 - \tilde\phi_2\|_{L^\infty(\mathbb{R}^n)} \le C\varepsilon^2(|t_1-t_2| + |x_1-x_2|) + (1+C\varepsilon^2)\|\phi_1-\phi_2\|_{L^\infty(\mathbb{R}^n)}.$$

Proof. If we hold p,Γ,w fixed, it is clear that

$$\begin{aligned}
&\phi_1(x+\varepsilon w) - \left[\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle\Gamma w,w\rangle + \varepsilon^2 f(t_1,x+x_1,\tilde\phi_1(x),p,\Gamma)\right]\\
&\quad\le \phi_2(x+\varepsilon w) - \left[\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle\Gamma w,w\rangle + \varepsilon^2 f(t_2,x+x_2,\tilde\phi_2(x),p,\Gamma)\right]\\
&\quad+ \|\phi_1-\phi_2\|_{L^\infty} + \varepsilon^2\left[f(t_2,x+x_2,\tilde\phi_2(x),p,\Gamma) - f(t_1,x+x_1,\tilde\phi_1(x),p,\Gamma)\right].
\end{aligned}$$

But, by assumption (6.2), we have

$f(t_2,x+x_2,\tilde{\phi}_2(x),p,\Gamma)-f(t_1,x+x_1,\tilde{\phi}_1(x),p,\Gamma)\le C(|\tilde{\phi}_1(x)-\tilde{\phi}_2(x)|+|x_1-x_2|+|t_1-t_2|)$.

It follows that

$$\begin{aligned}(6.5)\quad \phi_1(x+\varepsilon w)-&\left[\varepsilon p\cdot w+\frac{\varepsilon^2}{2}\langle\Gamma w,w\rangle+\varepsilon^2 f(t_1,x+x_1,\tilde{\phi}_1(x),p,\Gamma)\right]\\ &\le \phi_2(x+\varepsilon w)-\left[\varepsilon p\cdot w+\frac{\varepsilon^2}{2}\langle\Gamma w,w\rangle+\varepsilon^2 f(t_2,x+x_2,\tilde{\phi}_2(x),p,\Gamma)\right]\\ &\qquad\le \|\phi_1-\phi_2\|_{L^\infty}+C\varepsilon^2(|\tilde{\phi}_1(x)-\tilde{\phi}_2(x)|+|x_1-x_2|+|t_1-t_2|).\end{aligned}$$

Minimizing over w then maximizing over p,Γ we find

$$\tilde{\phi}_1(x)-\tilde{\phi}_2(x)\le \|\phi_1-\phi_2\|_{L^\infty}+C\varepsilon^2(|\tilde{\phi}_1(x)-\tilde{\phi}_2(x)|+|x_1-x_2|+|t_1-t_2|).$$

Reversing the roles of ϕ_1 and ϕ_2 and modifying the constant C we arrive at the result. ☐

The next lemma demonstrates stability with respect to the final time objective.

Lemma 6.2. *Assume the hypotheses of Lemma 4.1. Suppose in addition the final-time data satisfy (5.1) and (6.4). Defining*

$$\tilde{h}(x,z)=\max_{p,\Gamma}\min_{w}\left[h(x+\varepsilon w)-\left(z+\varepsilon p\cdot w+\frac{\varepsilon^2}{2}\langle\Gamma w,w\rangle+\varepsilon^2 f(T-\varepsilon^2,x,z,p,\Gamma)\right)\right],$$

we have

$$\|\max\{z:\tilde{h}(x,z)\ge 0\}-h(x)\|_{L^\infty(\mathbb{R}^n)}\le C\varepsilon^2$$

where C depends on h only through the constants in (5.1), (6.3), and (6.4).

Proof. Clearly $\tilde{h}(x,z)=S_\varepsilon[x,T-\varepsilon^2,z,h]-z$. Using the Taylor expansion of h, we have that

$$\begin{aligned}(6.6)\quad &|S_\varepsilon[x,T-\varepsilon^2,z,h]-h(x)|\\ &\le \max_{p,\Gamma}\min_{w}\left[\varepsilon(Dh-p)\cdot w+\frac{\varepsilon^2}{2}\langle(D^2h-\Gamma)w,w\rangle+\varepsilon^2 f(T-\varepsilon^2,x,z,p,\Gamma)\right]+C\varepsilon^2\end{aligned}$$

where C depends only on the constant in (6.4). Moreover, arguing as in the proof of Lemma 4.3 we have

$$\max_{p,\Gamma}\min_{w}\left[\varepsilon(Dh-p)\cdot w+\frac{\varepsilon^2}{2}\langle(D^2h-\Gamma)w,w\rangle+\varepsilon^2 f(T-\varepsilon^2,x,z,p,\Gamma)\right]\le C\varepsilon^2(1+|z|)$$

(the constant now depends on (6.3) as well as (6.4), since the optimal p and Γ are formally $Dh(x)$ and $D^2h(x)$ respectively). We conclude that

$$(6.7)\qquad |\tilde{h}(x,z)+z-h(x)|\le C\varepsilon^2(1+|z|).$$

Since $\tilde{h}$ is continuous, there exists z_0 such that

$$\max\{z:\tilde{h}(x,z)\ge 0\}=z_0\quad\text{and}\quad \tilde{h}(x,z_0)=0.$$

Taking $z = z_0$ in (6.7) we get

$$|z_0 - h(x)| \le C\varepsilon^2(1 + |z_0|),$$

from which it follows easily that $|z_0 - h(x)| \le C\varepsilon^2$ (with a different value for C).

□

These Lemmas imply a uniform Lipschitz bound for u^ε:

Proposition 6.3. *Assume the hypotheses of Lemma 4.1, and in addition that f satisfies (6.2) and h satisfies (5.1), (6.3), and (6.4). Suppose furthermore that $u^\varepsilon = v^\varepsilon$. Then for any $t_* < T$, there exists a constant C_{t_*} such that*

$$|u^\varepsilon(x,t) - u^\varepsilon(x',t')| \le C_{t_*}(|x - x'| + |t - t'|)$$

for all $t, t' \in [t_*, T]$ and all $x, x' \in \mathbb{R}^n$.

Proof. First we prove equicontinuity in time. The main idea is that there are two equivalent ways to think of $U^\varepsilon(x,z,t)$. The definition (2.18) can be encapsulated as saying that $U^\varepsilon(x,z,t)$ is Helen's optimal value if she starts at time t and position x, her initial debt is z, the horizon time is T, and her objective is $h(x) - z$. It is also Helen's optimal value if she starts at the same time t, position x and debt z but the horizon time is $T - \varepsilon^2$ and the objective is $\tilde{h}(x,z)$.

But we want to work with u^ε not U^ε. By definition

$$u^\varepsilon(x,t) = \sup\{z : U^\varepsilon(x,z,t) \ge 0\}.$$

We will compare it to two related functions, $\overline{u}^\varepsilon(x,t)$ and $w^\varepsilon(x,t)$, defined as follows:

- The function $\overline{u}^\varepsilon(x,t)$ is equal to $u^\varepsilon(x, t + \varepsilon^2)$, but we prefer to think of it differently. This is the analogue of u^ε for a modified game in which the objective is $h(x) - z$, the horizon time is $T - \varepsilon^2$, and the z increments are given by

$$\varepsilon p \cdot w + \frac{\varepsilon^2}{2}\langle \Gamma w, w\rangle + \varepsilon^2 f(t + \varepsilon^2, x, z, p, \Gamma),$$

  (thus: the function f is replaced by $f(\cdot + \varepsilon^2, \cdot, \cdot, \cdot)$).
- The function $w^\varepsilon(x,t)$ is the analogue of u^ε for the usual game with horizon time $T - \varepsilon^2$. (The objective is still $h(x) - z$ and the z increments are defined using $f(t,x,z,p,\Gamma)$.) It satisfies the same dynamic programming principle as u^ε,

$$w^\varepsilon(x,t) = S_\varepsilon[x, t, w^\varepsilon(x,t), w^\varepsilon(\cdot, t + \varepsilon^2)],$$

  but the two functions have different values at the horizon time:

$$w^\varepsilon(x, T - \varepsilon^2) = h(x) \quad \text{whereas} \quad u^\varepsilon(x, T - \varepsilon^2) = \sup\{z : \tilde{h}(x,z) \ge 0\}.$$

By a simple induction on time, we deduce from Lemmas 6.2 and 6.1 that

$$\|w^\varepsilon(\cdot,t)-u^\varepsilon(\cdot,t)\|_{L^\infty(\mathbb{R}^n)} \le (1+C\varepsilon^2)^{\frac{T-t}{\varepsilon^2}} \|\sup\{z : \tilde{h}(x,z)\ge 0\}-h(x)\|_{L^\infty(\mathbb{R}^n)} \le C\varepsilon^2(1+C\varepsilon^2)^{\frac{T-t}{\varepsilon^2}}.$$

Since $(1+C\varepsilon^2)^{\frac{T-t}{\varepsilon^2}} \to e^{C(T-t)}$ as $\varepsilon \to 0$, it follows that

$$\|w^\varepsilon(\cdot,t)-u^\varepsilon(\cdot,t)\|_{L^\infty(\mathbb{R}^n)} \le C^{1+T-t}\varepsilon^2. \tag{6.8}$$

On the other hand, using Lemma 6.1, an easy induction gives

$$|\overline{u}^\varepsilon(x,t)-w^\varepsilon(x,t)| \le M_k$$

where $k = \frac{T-\varepsilon^2-t}{\varepsilon^2}$ and M_k satisfies

$$M_{k+1} \le (1+C\varepsilon^2)M_k + C\varepsilon^4.$$

One can easily calculate that

$$M_k \le (1+C\varepsilon^2)^k M_0 + C\varepsilon^2((1+C\varepsilon^2)^k-1). \tag{6.9}$$

Since $M_0 = 0$ it follows as above that

$$|\overline{u}^\varepsilon(x,t)-w^\varepsilon(x,t)| \le C(e^{C(T-t)}-1)\varepsilon^2. \tag{6.10}$$

Combining (6.10) with (6.8) we deduce that

$$|u^\varepsilon(x,t+\varepsilon^2)-u^\varepsilon(x,t)| \le e^{C(T-t+1)}\varepsilon^2$$

where C is independent of t (in a finite interval). Since this is valid for all t, we may add up such relations for $t+k\varepsilon^2$ and find that for $t,t' \in [t_*,T]$,

$$|u^\varepsilon(x,t')-u^\varepsilon(x,t)| \le C_{t_*}|t'-t|.$$

This establishes the desired Lipschitz bound in time.

For equicontinuity in space, we apply Lemma 6.1 to $\phi_1(x) = u^\varepsilon(x,t)$ and $\phi_2(x) = u^\varepsilon(x+a,t)$, with $x_1 = a$ and $x_2 = 0$. It follows that

$$\|u^\varepsilon(\cdot,t+\varepsilon^2)-u^\varepsilon(\cdot+a,t+\varepsilon^2)\|_{L^\infty(\mathbb{R}^n)} \le (1+C\varepsilon^2)\|u^\varepsilon(\cdot,t)-u^\varepsilon(\cdot+a,t)\|_{L^\infty(\mathbb{R}^n)} + C\varepsilon^2|a|.$$

Then by induction, for any $a \in \mathbb{R}^n$,

(6.11)

$$\begin{aligned}\|u^\varepsilon(\cdot,t)-u^\varepsilon(\cdot+a,t)\|_{L^\infty(\mathbb{R}^n)} &\le (1+C\varepsilon^2)^{\frac{T-t}{\varepsilon^2}}\|u^\varepsilon(\cdot,T)-u^\varepsilon(\cdot+a,T)\|_{L^\infty}+C(T-t)|a| \\ &\le e^{C(T-t)}\|h(\cdot)-h(\cdot+a)\|_{L^\infty(\mathbb{R}^n)}+C(T-t)|a| \le \left(e^{C(T-t+1)}+C(T-t)\right)|a|\end{aligned}$$

using assumption (6.3). This establishes the desired Lipschitz bound in space. □

If the PDE has a comparison principle, then we easily conclude from the preceding result (when it applies) that $\lim u^\varepsilon(x,t)$ exists and is the unique viscosity solution of (2.15). Moreover, we also conclude that the viscosity solution is Lipschitz continuous.

6.2 The uniform Lipschitz character of U^ε

We assumed in Section 6.1 that $u^\varepsilon = v^\varepsilon$. This is true when f is non-decreasing in z, by Lemma 2.4. For more general f we can prove it only locally in time, when $T-t$ is sufficiently small. The proof is elementary, using only the assumption that f is globally Lipschitz in z.

Lemma 6.4. *Assume f satisfies*

$$|f(t,x,z',p,\Gamma) - f(t,x,z,p,\Gamma)| \le C|z-z'| \tag{6.12}$$

where C is independent of t,x,p, and Γ. Then for all x,z,z',t, we have

$$\left|U^\varepsilon(x,z,t) - U^\varepsilon(x,z',t) - (z'-z)\right| \le (e^{C(T-t)}-1)|z-z'|.$$

In particular, there is a constant c (independent of ε) such that $U^\varepsilon(x,z,t)$ is a monotone function of z when $T-t<c$. For such t we have $u^\varepsilon = v^\varepsilon$.

Proof. Arguing by induction on $k = \frac{T-t}{\varepsilon^2}$, assume that

$$\left|U^\varepsilon(x,z,T-k\varepsilon^2) - U^\varepsilon(x,z',T-k\varepsilon^2) - (z'-z)\right| \le M_k|z-z'|. \tag{6.13}$$

When $k=0$ we may take $M_0 = 0$, since $U^\varepsilon(x,z,T) - U^\varepsilon(x,z',T) = z'-z$. We shall find the form of M_{k+1} as we do the inductive step. Let $t = T-(k+1)\varepsilon^2$, and let p,Γ,w be fixed. Assuming the result is true for k, we may write

$$\begin{aligned}
&U^\varepsilon\left(x+\varepsilon w, z+(\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle\Gamma w,w\rangle + \varepsilon^2 f(t,x,z,p,\Gamma)), t+\varepsilon^2\right)\\
&\quad - U^\varepsilon\left(x+\varepsilon w, z'+(\varepsilon p\cdot w + \frac{\varepsilon^2}{2}\langle\Gamma w,w\rangle + \varepsilon^2 f(t,x,z',p,\Gamma)), t+\varepsilon^2\right)\\
&\le z'-z+\varepsilon^2\left(f(t,x,z',p,\Gamma)-f(t,x,z,p,\Gamma)\right)+M_k|z-z'+\varepsilon^2\left(f(t,x,z',p,\Gamma)-f(t,x,z,p,\Gamma)\right)|\\
&\qquad \le z'-z+C\varepsilon^2|z-z'|+M_k(|z-z'|+C\varepsilon^2|z-z'|),
\end{aligned}$$

where we have used (6.12). Minimizing over w then maximizing over p,Γ and using the dynamic programming principle (2.18), we deduce that

$$U^\varepsilon(x,z,t) \le U^\varepsilon(x,z',t) + z'-z + (C\varepsilon^2(1+M_k)+M_k)|z-z'|.$$

Reversing the roles of z and z' it follows that

$$M_{k+1} \le M_k(1+C\varepsilon^2) + C\varepsilon^2.$$

We deduce (as in the proof of (6.9)) that $M_k \le e^{C(T-t)}-1$, whence

$$|U^\varepsilon(x,z',t) - U^\varepsilon(x,z,t) - (z'-z)| \le (e^{C(T-t)}-1)|z-z'|.$$

The remaining assertions of the Lemma follow trivially. □

7 Additional remarks

This paper began with a long, informal discussion motivating and describing our games. Then it got more technical, providing bounds on the resulting value functions and demonstrating their convergence to viscosity sub and supersolutions of the underlying PDE. Now we return to the more informal style of Section 2. Section 7.1 offers some comments on the interpretation of the game as a numerical scheme, and Section 7.2 discusses how the approach of Section 2.1 can be extended to a broad class of semilinear parabolic equations.

7.1 Comments on the timestep problem

The central result about numerical methods for viscosity solutions of second-order equations is the theorem of Barles-Souganidis [7], which says that every monotone, stable, and consistent scheme is convergent.

That result is important, but it does not provide us with any particular schemes. There is of course a large body of literature in that direction, most of it aimed at specific classes of equations; examples include [19] (for equations of the form $u_t - Tr(\theta(x,Du)\theta(x,Du)^T D^2u) = 0$), [37] (for motion by curvature), and [38] (for the infinity-Laplacian).

For equations of the form $-u_t + f(t,x,Du,D^2u) = 0$, the dynamic programming principle (2.12) amounts to a semidiscrete scheme for stepping the PDE backward in time. It seems natural to ask: what kind of numerical method is this?

The answer is simple when u is smooth. Then our dynamic programming principle amounts more or less to the explicit Euler scheme

$$u(x,t_j) = u(x,t_{j+1}) - \varepsilon^2 f\left(Du(x,t_{j+1}), D^2u(x,t_{j+1})\right). \tag{7.1}$$

Indeed, consider the minimization over w_j on the right hand side of (2.12). Writing $\Delta x = \varepsilon w$ and expanding $u(x+\Delta x, t_{j+1})$ as a Taylor series, it becomes

$$\min_{\|\Delta x\| \le \varepsilon^{1-\alpha}} \Big[(Du(x,t_{j+1}) - p)\cdot \Delta x + \tfrac{1}{2}\langle (D^2u(x,t_{j+1}) - \Gamma)\Delta x, \Delta x\rangle + u(x,t_{j+1}) - \varepsilon^2 f(p,\Gamma) + O(\|\Delta x\|^3) \Big].$$

The assumption (2.7) ensures that $\|\Delta x\|^3 \ll \varepsilon^2$ so the remainder is negligible. We explained after (2.14) (and established rigorously in Lemma 4.1) that Mark's minimization forces Helen to choose $p = Du(x,t_{j+1})$ and $\Gamma \le D^2u(x,t_{j+1})$ (or at least nearly so), since otherwise Mark can take advantage of her mistake. With these choices, using the monotonicity of f and ignoring the error term, the dynamic programming principle reduces to (7.1).

The answer is almost as simple when u is not smooth: then our scheme reduces (in the same somewhat formal sense as above) to

$$u(x,t_j) = \max_{(p,\Gamma)\in J_2^- u(x,t_{j+1})} \left[u(x,t_{j+1}) - \varepsilon^2 f(p,\Gamma)) \right] \tag{7.2}$$

where $J_2^- u(x,t_{j+1})$ is the *second-order lower semijet* of u, viewed as a function of the spatial variable only, at time t_{j+1} and position x. To be clear about the definition:

(7.3)
$$(p,\Gamma) \in J_2^- u(x,t_{j+1}) \Longleftrightarrow u(x+\Delta x,t_{j+1}) \geq u(x,t_{j+1}) + p\cdot\Delta x + \tfrac{1}{2}\langle \Gamma\Delta x, \Delta x\rangle + o(\|\Delta x\|^2)$$

as $\Delta x \to 0$. Loosely speaking, the semijet $J_2^- u$ identifies the quadratic functions that touch the graph of u at x but stay locally below it.

To see why our scheme amounts roughly to (7.2), consider again Mark's minimization on the right hand side of (2.12): writing $\Delta x = \varepsilon w_j$ as before but not doing any Taylor expansion, we get

$$\min_{\|\Delta x\| \leq \varepsilon^{1-\alpha}} \left[u^\varepsilon(x+\Delta x, t_{j+1}) - p\cdot\Delta x - \tfrac{1}{2}\langle \Gamma\Delta x, \Delta x\rangle - \varepsilon^2 f(p,\Gamma)\right]. \tag{7.4}$$

Again Helen should choose $(p,\Gamma) \in J_2^- u(x,t_{j+1})$; then Mark observes that

$$u^\varepsilon(x+\Delta x,t_{j+1}) - p\cdot\Delta x - \tfrac{1}{2}\langle \Gamma\Delta x,\Delta x\rangle \geq u(x,t_{j+1}) + o(\|\Delta x\|^2)$$

and (ignoring the error term) he concludes that his best choice is $\Delta x = 0$. The minimum in (7.4) is then $u^\varepsilon(x,t_{j+1}) - \varepsilon^2 f(p,\Gamma)$ and the outcome of the max/min is (7.2).

The timestepping scheme (7.2) seems quite natural, but to the best of our knowledge nothing like it has previously been proposed for second-order equations. A similar idea (using first-order semijets) was however considered by Taras'ev for a class of Hamilton-Jacobi equations [44]. His work includes a discussion of spatial discretization and a proof of convergence as $\varepsilon \to 0$ (the error is of order ε). He also identifies a relationship between this approach and the more conventional Lax-Wendroff and Godunov schemes.

Equation (7.2) amounts to a *semidiscrete numerical scheme*

$$\frac{u(x,t_{j+1}) - u(x,t_j)}{\varepsilon^2} = \min_{(p,\Gamma)\in J_2^- u(x,t_{j+1})} f(p,\Gamma). \tag{7.5}$$

When u is defined on a grid, the obvious discretization of the right hand side is to minimize f over all p and Γ such that

$$p\cdot\Delta x + \tfrac{1}{2}\langle \Gamma\Delta x,\Delta x\rangle \leq u(x+\Delta x,t_{j+1}) - u(x,t_{j+1})$$

with Δx restricted to a stencil. We wonder how well this *fully discrete numerical scheme* would work in practice.

7.2 An alternative game for some nonlinear heat equations

The games discussed in this paper are naturally related to the form of the PDE. However there can be multiple "two-person game" approaches to the same PDE.

In this section we discuss an alternative "two-person game" approach to solving semilinear parabolic equations of the form

$$\begin{cases} \partial_t u + \Delta u + f(x,t,u,Du) = 0 & \text{for } x \in \mathbb{R}^n \text{ and } t < T \\ u(x,T) = h(x) & \text{at } t = T. \end{cases} \tag{7.6}$$

This discussion generalizes the game of Section 2.1 to higher spatial dimensions and to nonlinear equations. There are still two opposing players, Helen and Mark. The game begins at an arbitrary position $x_0 \in \mathbb{R}^n$, and Helen's debt is set to $z_0 \in \mathbb{R}$. The rules are as follows: if at time $t_j = t_0 + j\varepsilon^2$, the position is x_j and Helen's debt is z_j, then

(i) Helen chooses an orthonormal frame $(v_1^{(j)}, \cdots, v_n^{(j)})$, and a vector $p_j \in \mathbb{R}^n$.
(ii) After seeing Helen's choice, Mark chooses numbers $b_1^{(j)}, \cdots, b_n^{(j)} \in \{1, -1\}$.
(iii) Denoting $w_j = \sqrt{2}\sum_{i=1}^n b_i^{(j)} v_i^{(j)}$, the position changes to
$$x_{j+1} = x_j + \varepsilon w_j$$
and Helen's debt gets changed to
$$z_{j+1} = z_j + \varepsilon p \cdot w_j - \varepsilon^2 f(x_j, t_j, z_j, p_j).$$
(iv) The clock steps forward to $t_{j+1} = t_j + \varepsilon^2$ and the process repeats, stopping when $t_N = T$.
(v) At the final time t_N, Helen collects $h(x_N)$, where x_N is the final-time position.

Helen's final wealth is $h(x_N) - z_N$. Her goal is to maximize her final time wealth, while Mark's is to obstruct her. Helen's value function, starting from x_0 with score z_0 at time t_0, is defined by
$$U^\varepsilon(x_0, z_0, t_0) = \max_{\text{Helen's choices}} [h(x_N) - z_N]$$
using our usual convention (2.13) about the meaning of this expression.

We then define (as in Section 2.3)
$$u^\varepsilon(x_0, t_0) = \sup\{z_0 : U^\varepsilon(x_0, z_0, t_0) \geq 0\}, \quad v^\varepsilon(x_0, t_0) = \inf\{z_0 : U^\varepsilon(x_0, z_0, t_0) \leq 0\}.$$
Arguing as for Proposition 2.1, one finds that u^ε and v^ε satisfy the dynamic programming inequalities
$$u^\varepsilon(x,t) \leq \sup_{\{v_i\},p} \inf_{b_i=\pm 1} \left[u^\varepsilon(x+\varepsilon w, t+\varepsilon^2) - (\varepsilon p \cdot w - \varepsilon^2 f(x,t,u^\varepsilon(x,t),p))\right], \tag{7.7}$$
$$v^\varepsilon(x,t) \geq \sup_{\{v_i\},p} \inf_{b_i=\pm 1} \left[v^\varepsilon(x+\varepsilon w, t+\varepsilon^2) - (\varepsilon p \cdot w - \varepsilon^2 f(x,t,v^\varepsilon(x,t),p))\right], \tag{7.8}$$
where $w = \sqrt{2}\sum_{i=1}^n b_i v_i$.

We believe that, after placing appropriate constraints on p and f, an analysis based on these dynamic programming relations (analogous to the one in this paper) would prove the convergence of u^ε and v^ε as $\varepsilon \to 0$ to the unique viscosity solution of (7.6). (Note that an equation of this form has a comparison principle, under very mild conditions on f [20].)

To support this conjecture, let us present the heuristic argument explaining why the PDE (7.6) is the formal Hamilton-Jacobi-Bellman equation associated with this game. We restrict our attention to the case where U^ε is monotone in z (guaranteed if for example f is) which ensures that $u^\varepsilon = v^\varepsilon$, and the dynamic programming inequalities (7.7)–(7.8) reduce to the equality

$$u^\varepsilon = \max_{\{v_i\},p} \min_{b_i=\pm 1} \left[u^\varepsilon(x+\varepsilon w, t+\varepsilon^2) - (\varepsilon p\cdot w - \varepsilon^2 f(x,t,u^\varepsilon(x,t),p)) \right]. \tag{7.9}$$

Suppressing the dependence of u^ε on ε, assuming it is smooth enough, and noting that εw is small, we have by Taylor expansion

$$0 \approx \max_{\{v_i\},p} \min_{b_i=\pm 1} \left[\varepsilon(Du-p)\cdot w + \varepsilon^2(\partial_t u + \tfrac{1}{2}\langle D^2 u w, w\rangle + f(x,t,u,p)) \right].$$

Keeping first only the leading order ε term in this relation, we have to compute

$$\max_{\{v_i\},p} \min_{b_i=\pm 1} \left[\sqrt{2}\varepsilon \sum_{i=1}^n b_i (Du-p)\cdot v_i \right] = \max_{\{v_i\},p} \left[-\sqrt{2}\varepsilon \sum_{i=1}^n |(Du-p)\cdot v_i| \right].$$

The term to be maximized is always ≤ 0, and since the v_i's form an orthonormal frame, it is 0 if and only if $p = Du$. This imposes Helen's choice of p as $p = Du$. The first order term then vanishes and (substituting Du for p) there remains

$$0 \approx \max_{\{v_i\}} \min_{b_i=\pm 1} \left[\partial_t u + \tfrac{1}{2}\langle D^2 u w, w\rangle + f(x,t,u,Du)) \right].$$

Using the fact that $\{v_i\}$ form an orthonormal frame, we have

$$\tfrac{1}{2}\langle D^2 u w, w\rangle = \sum_{i=1}^n \langle D^2 u v_i, v_i\rangle + \sum_{i\neq j} b_i b_j \langle D^2 u v_i, v_j\rangle = \Delta u + \sum_{i\neq j} b_i b_j \langle D^2 u v_i, v_j\rangle.$$

We are thus led to the relation

$$0 \approx \partial_t u + f(x,t,u,Du) + \Delta u + \max_{\{v_i\}} \min_{b_i=\pm 1} \left(\sum_{i\neq j} b_i b_j \langle D^2 u v_i, v_j\rangle \right). \tag{7.10}$$

We claim that

$$\max_{\{v_i\}} \min_{b_i=\pm 1} \sum_{i\neq j} b_i b_j \langle D^2 u v_i, v_j\rangle = 0, \tag{7.11}$$

with equality achieved when $\{v_i\}$ is chosen to be an orthonormal frame which diagonalizes $D^2 u$. Once this is observed (7.10) and thus (7.9) reduce formally to the equation (7.6).

We conclude by proving the claim: first it is easy to see that the max in (7.11) is always ≥ 0, by choosing $\{v_i\}$ to be an orthonormal frame which diagonalizes $D^2 u$. Then to show that it is always ≤ 0, we show that given any numbers a_{ij} we may always choose $b_i = \pm 1$ such that $\sum_{i\neq j} a_{ij} b_i b_j \le 0$. This can be done by induction on the dimension n: it is clearly true for $n = 1, 2$. Assuming it is true up to $n-1$, we write

$$\sum_{i\neq j} a_{ij} b_i b_j = \sum_{i\neq j\le n-1} a_{ij} b_i b_j + 2b_n (\sum_{i=1}^{n-1} a_{in} b_i).$$

The first term on the right-hand side can be made nonpositive by the induction hypothesis, and the second one can be made nonpositive by choosing $b_n = -sgn(\sum_{i=1}^{n-1} a_{in} b_i)$.

In fact, the whole sum can be made negative unless $a_{ij} = 0$ for all $i \neq j$. Setting $a_{ij} = \langle D^2 u v_i, v_j \rangle$, this proves the claim (and equality is achieved if and only if (v_i) diagonalizes $D^2 u$).

Acknowledgment. We thank H. Mete Soner for suggesting the treatment of the linear heat equation presented in Section 2.1, and for suggesting that ideas from [17] could be useful for handling more general equations. We also thank Guy Barles, Yoshikazu Giga, and Takis Souganidis for helpful comments. Finally, we gratefully acknowledge support from the National Science Foundation through grants DMS-0313744 (RVK), DMS-0807347 (RVK), and DMS-0239121 (SS), and from the European Science Foundation through a EURYI award (SS).

Bibliography

[1] Armstrong, S. N.; Smart, C. K. A finite-difference approach to the infinity Laplace equation and tug-of-war games. *Trans. Amer. Math. Soc.*, in press.

[2] Armstrong, S. N.; Smart, C. K.; Somersille, S. J. An infinity Laplace equation with gradient term and mixed boundary conditions. Preprint (arxiv:0910.3744).

[3] Bardi, M.; Capuzzo-Dolcetta, I. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Birkhaüser Boston, Inc., Boston, MA, 1997.

[4] Barles, G.; Biton, S.; Ley, O. A geometrical approach to the study of unbounded solutions of quasilinear parabolic equations. *Arch. Rational Mech. Anal.* **162** (2002), 287–325.

[5] Barles G.; Busca, J. Existence and comparison results for fully nonlinear degenerate elliptic equations without zeroth-order term. *Comm. PDE* **26** (2001), no. 11–12, 2323–2337.

[6] Barles G.; Rouy, E. A strong comparison result for the Bellman equation arising in stochastic exit time control problems and its application. *Comm. PDE* **23** (1998), no. 11–12, 1995–2033.

[7] Barles, G.; Souganidis, P. E. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.* **4** (1991), no. 3, 271–283.

[8] Barron, N. E.; Evans, L. C.; Jensen, R. The infinity Laplacian, Aronsson's equation and their generalizations. *Trans. Amer. Math. Soc.* **360** (2008), no. 1, 77–101.

[9] Bellettini, G.; Chermisi, H.; Novaga, M. The level set method for systems of PDEs, *Comm. PDE* **32** (2007), no. 7–9, 1043–1064.

[10] Bernhard, P.; El Farouq, N.; Thiery, S. Robust control approach to option pricing: a representation theorem and fast algorithm. *SIAM J. Control Optim.* **46** (2007), no. 6, 2280–2302.

[11] Buckdahn, R.; Cardaliaguet, P.; Quincampoix, M. A representation formula for the mean curvature motion. *SIAM J. Math. Anal.* **33** (2001), no. 4, 827–846.

[12] Caffarelli, L. A.; Souganidis, P. E. A rate of convergence for monotone finite difference approximations of fully nonlinear, uniformly elliptic PDEs. *Comm. Pure Appl. Math* **61** (2007), no. 1, 1–17.

[13] Caffarelli, L. A.; Souganidis, P. E. Rates of convergence for the homogenization of fully nonlinear uniformly elliptic PDE in random media. Preprint.

[14] Carlini, E. A single-pass scheme for the mean curvature motion of convex curves. *Numerical Mathematics and Advanced Applications*, 671–678. Edited by K. Kunisch, G. Of, and O. Steinbach, Springer-Verlag, 2008.

[15] Carlini, E.; Falcone, M.; Ferretti, R. A semi-Lagrangian scheme for the curve shortening flow in codimension-2. *J. Comput. Phys.* **225** (2007), no. 2, 1388–1408.

[16] Catté, F.; Dibos, F.; Koepfler, G. A morphological scheme for mean curvature motion and applications to anisotropic diffusion and motion of level sets. *SIAM J. Numer. Anal.* **32** (1995), no. 6, 1895–1909.

[17] Cheridito, P.; Soner, H. M.; Touzi, N.; Victoir, N. Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs. *Comm. Pure Appl. Math.* **60** (2007), no. 7, 1081–1110.

[18] Cox, J.; Ross, S.; Rubinstein, M. Option pricing: a simplified approach. *J. Fin. Econ.* **7** (1979), 229–264.

[19] Crandall, M. G.; Lions, P.-L. Convergent difference schemes for nonlinear parabolic equations and mean curvature motion. *Numer. Math.* **75** (1996), 17–41.

[20] Crandall, M. G.; Ishii, H.; Lions, P.-L. User's guide to viscosity solutions. *Bull. Amer. Math. Soc.* **27** (1992), no. 1, 1–67.

[21] Evans, L. C. *Partial Differential Equations*, American Mathematical Society, Providence, RI, 1997.

[22] Evans, L. C. The 1-Laplacian, the ∞-Laplacian and differential games. *Perspectives in nonlinear partial differential equations, in honor of Haim Brezis*, 245–254. Contemp. Math. **446**. Edited by H. Berestycki, M. Bertsch, L. Nirenberg, L. Peletier, and L. Véron. American Mathematical Society, 2007.

[23] Evans, L. C. Some min-max methods for the Hamilton-Jacobi equation. *Indiana Univ. Math. J.* **33** (1984), no. 1, 31–50.

[24] Evans, L. C.; Souganidis, P. E. Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations. *Indiana Univ. Math. J.* **33** (1984), no. 5, 773–797.

[25] Fleming, W. H.; Soner, H. M. *Controlled Markov processes and viscosity solutions*, Springer-Verlag, 2005.

[26] Fleming, W. H.; Souganidis, P. E. On the existence of value functions of two-player, zero-sum stochastic differential games. *Indiana Univ. Math. J.* **38** (1989), no. 2, 293–314.

[27] Friedman, A. Differential games. *Handbook of game theory with economic applications*, Vol. II, 781–799. Handbooks in Economics **11**, North-Holland, Amsterdam, 1994.

[28] Giga, M.-H.; Giga, Y. Minimal vertical singular diffusion preventing overturning for the Burgers equation. *Recent advances in scientific computing and partial differential equations (Hong Kong, 2002)*, 73–88. Contemp. Math. **330**, Edited by S. Y. Cheng, C.-W. Shu, and T. Tang. American Mathematical Society, Providence, RI, 2003.

[29] Giga, Y. Viscosity solutions with shocks, *Comm. Pure Appl. Math.* **55** (2002), no. 4, 431–480.

[30] Giga, Y.; Liu, Q. A remark on the discrete deterministic game approach for curvature flow equations. *Nonlinear phenomena with energy dissipation*, 103–115. GAKUTO Internat. Ser. Math. Sci. Appl. **29**, Gakkotosho, Tokyo, 2008.

[31] Giga, Y.; Liu, Q. A billiard-based game interpretation of the Neumann problem for the curve shortening equation. *Adv. Diff. Eqns.* **14** (2009), no. 3–4, 201–240.

[32] Giga, Y.; Sato, M.-H. A level set approach to semicontinuous viscosity solutions for Cauchy problems, *Comm. PDE* **26** (2001), no. 5, 813–839.

[33] Imbert, C.; Serfaty, S. Repeated games for eikonal equations, integral curvature flows, and nonlinear parabolic integro-differential equations, preprint.

[34] Juutinen, P.; Kawohl, B. On the evolution governed by the infinity Laplacian. *Math. Ann.* **335** (2006), no. 4, 819–851.

[35] Kohn, R. V.; Serfaty, S. (with an appendix by G. Barles and F. Da Lio). A deterministic-control-based approach to motion by curvature. *Comm. Pure Appl. Math.* **59** (2006), no. 3, 344–407.

[36] Kohn, R. V.; Serfaty, S. Second-order PDE's and deterministic games. *6th International Congress on Industrial and Applied Mathematics – Zurich, Switzerland, 16-20 July 2007 – Invited Lectures*, 239–249. Edited by R. Jeltsch and G. Wanner, European Mathematical Society, 2009.

[37] Oberman, A. M. A convergent monotone difference scheme for motion of level sets by mean curvature. *Numer. Math.* **99** (2004), no. 2, 365–379.

[38] Oberman, A. M. A convergent difference scheme for the infinity Laplacian: construction of absolutely minimizing Lipschitz extensions. *Math. Comp.* **74** (2005), 1217–1230.

[39] Peres, Y.; Schramm, O.; Sheffield, S.; Wilson, D. B. Tug-of-war and the infinity Laplacian. *J. Amer. Math. Soc.* **22** (2009), 167–210.

[40] Peres, Y.; Sheffield, S. Tug of war with noise: a game theoretic view of the p-Laplacian. *Duke Math. J.* **145** (2008), no. 1, 91–120.

[41] Soner, H. M.; Touzi, N. Dynamic programming for stochastic target problems and geometric flows. *J. Euro. Math. Soc.* **4** (2002), 201–236.

[42] Soner, H. M.; Touzi, N. A stochastic representation for level set equations. *Comm. PDE* **27** (2002), nos. 9 & 10, 2031–2053.

[43] Soner, H. M.; Touzi, N. A stochastic representation for mean curvature type flows. *Ann. Prob.* **31** (2003), no. 3, 1145–1165.

[44] Taras'ev, A.M. Approximation schemes for constructing minimax solutions of Hamilton-Jacobi equation. *J. Appl. Math. Mech.* **58** (1994), no. 2, 207–221.

[45] Tsai, Y.-H. R.; Giga, Y.; Osher, S. A level set approach for computing discontinuous solutions of Hamilton-Jacobi equations, *Math. Comp.* **72** (2002), no. 241, 159–181.

[46] Yu, Y. Uniqueness of values of Aronsson operators and running costs in "tug-of-war" games. *Ann. Inst. H. Poincaré Anal. Nonlinéaire* **26** (2009), no. 4, 1299–1308.

[47] Yu, Y. Maximal and minimal solutions of an Aronsson equation: L^∞ variational problems versus the game theory. *Calc. Var. PDE* **37** (2010), no. 1–2, 73–74.

Received Month 200X.