Chapter 1

Viability Theorems for Ordinary and Stochastic Differential Equations

Introduction

This chapter is meant to be an *independent* introduction to the basic theorems of viability theory in the simple framework of ordinary differential equations $x' = f(x)$ and stochastic differential equations

$$d\xi \;=\; f(\xi(t))dt + g(\xi(t))dW(t)$$

It can be omitted by readers who are only interested in the theory for differential inclusions.

So, we begin by tackling the viability issue by isolating it in the framework of ordinary differential equations in the first section. A function $[0,T] \ni t \to x(t)$ is said to be *viable* in a given subset K on $[0,T]$ if, for any $t \in [0,T]$, the state $x(t)$ remains in K.

Actually, even in the simpler situation of differential equations, we have to be careful and make a distinction between two neighboring concepts: *viability property and invariance property*. The first one requires that, starting from any initial point of K, *at least one* *solution* to the differential equation is viable in K whereas the second one demands that *all solutions* are viable in K.

We shall characterize the first one by saying that K is a *viability domain*, i.e., that for any state x in the boundary of K, the velocity is tangent in some sense to K at x.

We may require for that purpose that K is a smooth manifold and therefore, that f is a vector field, the velocity $f(x)$ lying in the tangent space.

But first, we do not need the fact that the space of "tangent" directions (adequately defined) is a vector space. The added luxury of linearity does not compensate for its fragility, in the sense that, for instance, the intersection of two smooth manifolds is no longer smooth. Since we shall regard in most of our applications the subset K as a subset defined by constraints (and above all, inequality constraints), then it is very exceptional that such a subset is smooth. As in optimization, we are quickly led to assume that K is convex, since convex subsets are defined by linear inequality constraints. But, here again, it would be nice to dispense with this assumption if this is possible (with no added mathematical cost), for it allows us to consider also union of convex subsets, for instance.

As we know since 1942[1], one can characterize such a viability property for any closed subset K, with an adequate mathematical implementation of the concept of tangency. The one chosen is actually equivalent to the concept of "contingency" introduced ten years earlier by Bouligand. We can then define the contingent cone to K at $x \in K$ for any subset K, the price to pay being that the set of tangent directions (the contingent cone) is a closed cone instead of a vector space.

We shall only give in this introductory chapter the definition of the contingent cone and provide further properties in Chapter 5 after the presentation of the viability theorems for both ordinary differential equations, stochastic differential equations and differential inclusions.

Nagumo's Theorem states that when f is continuous, a *closed*

[1]In a seminal paper written in German by the Japanese mathematician M. Nagumo. As it could be expected, this theorem was forgotten and rediscovered (at least) fourteen times up to 1968, in different contexts, with various concepts of tangency.

subset K enjoys the viability property if and only if it is a viability domain. We prove only this theorem, and shall derive the other properties as corollaries of statements we shall prove later in the case of differential inclusions.

Many proofs of viability theorems are now available: we chose the most elementary (which is not the shortest) for several reasons: it is the prototype of the extensions of the viability theorems (to functional differential inclusions, partial differential inclusions, ...) we shall present later in this book. It is just a modification of the Euler method of approximating a solution by piecewise linear functions (polygonal lines) in order to force the solution to remain in K. Despite its "constructionist" look, this method is not a finite difference scheme (explicit or implicit.) We shall present a rudimentary numerical introduction in the third section, but are forced to postpone the proofs to chapter 10, because they use more properties of the contingent cones which are presented later in Chapter 5.

The fourth section is dedicated to the "replicator systems". This is because the most popular viability domain is the *probability simplex.* Indeed, it is often too difficult to provide a mathematical description of the state space of problems arising in biology, economics, etc. So, this difficulty is bypassed by studying instead of the evolution of the state itself, the evolution of frequencies, concentrations, probabilities, ..., of the states (without forgetting *mixed strategies* in game theory), which all range over the probability simplex $S^n \subset R^n$.

Replicator systems

$$x_i'(t) \;=\; x_i(t)(g_i(x(t)) - \tilde{u}(x(t)))$$

are the differential equations derived from evolutions

$$x_i'(t) \;=\; x_i(t)g_i(x(t)), \;\; (i = 1, \ldots, n)$$

governed by specific growth rates $g_i(\cdot)$, corrected by subtracting the closed-loop control

$$\tilde{u}(x) \;:=\; \sum_{j=1}^{n} x_j g_j(x)$$

for obeying the viability constraints. The celebrated logistic equation belongs to this class (for constant growth rates.) Dynamical models arising in population genetics, prebiotic evolution, sociobiology and population ecology devised independently are replicator systems for specific linear growth rates[2].

Finally, we conclude this introductory chapter with a brief presentation of viability and/or invariance properties of closed subsets for stochastic differential equations.

Let us consider Lipschitz maps f and g and the stochastic differential equation

$$d\xi \ = \ f(\xi(t))dt + g(\xi(t))dW(t)$$

the solution of which is given by the formula

$$\xi(t) \ = \ \xi(0) + \int_0^t f(\xi(s))ds + \int_0^t g(\xi(s))dW(s)$$

We want to characterize the *(stochastic) viability property* of a closed convex subset K of X with respect to the pair (f,g): for any random variable x in K, there exists a solution ξ to the stochastic differential equation starting at x which is *viable in* K, in the sense that

$$\forall\, t \in [0,T], \ \text{ for almost all } \omega \in \Omega, \ \ \xi_\omega(t) \ \in \ K$$

For that purpose, we adapt to the stochastic case the concept of contingent cone to a subset K at a random variable $x \in K$ as the set $\mathcal{T}_K(t,x)$ of pairs (γ, v) of random variables satisfying the following property: There exist sequences of $h_n > 0$ converging to 0 and of measurable random variables a^n and b^n satisfying for almost all $\omega \in \Omega$,

$$\forall\, n \geq 0, \ \ x_\omega + v_\omega(W_\omega(t+h_n) - W(t)) + h_n\gamma_\omega + h_n a_\omega^n + \sqrt{h_n} b_\omega^n \in K$$

and converging to 0 in some sense.

Then we shall prove in essence that the following conditions are equivalent:

1. — The subset K enjoys the viability property with respect to the pair (f,g)

[2]Replicator systems are the central theme of the monograph THE THEORY OF EVOLUTION AND DYNAMICAL SYSTEMS by J. Hofbauer and K. Sigmund.

2. — for every $\mathcal{F}_t$-random variable x viable in K,

$$(f(x), g(x)) \in \mathcal{T}_K(t, x)$$

For instance, this condition means that for every $\mathcal{F}_t$-random variable x viable in K

$$f(x) \in K \ \& \ g(x) \in K$$

when K is a vector subspace, that

$$\langle x, g(x) \rangle = 0 \ \& \ \langle x, f(x) \rangle + \frac{1}{2} \|g(x)\|^2 = 0$$

when K is the unit sphere and that

$$\langle x, g(x) \rangle = 0 \ \& \ \langle x, f(x) \rangle + \frac{1}{2} \|g(x)\|^2 \leq 0$$

when K is the unit ball.

1.1 Viability & Invariance Properties

Definition 1.1.1 (Viable functions) *Let K be a subset of a finite dimensional vector-space*[3] *X. We shall say that a function $x(\cdot)$ from $[0, T]$ to X is* viable in K on $[0, T]$ *if* $\forall\, t \in [0, T],\ \ x(t) \in K$.

Let us describe the (deterministic) dynamics of the system by a (single-valued) map f from some open subset Ω of X to X. We consider the initial value problem (or Cauchy problem) associated with the differential equation

$$\forall\, t \in [0, T], \ \ x'(t) = f(x(t)) \tag{1.1}$$

satisfying the initial condition $x(0) = x_0$.

Definition 1.1.2 (Viability & Invariance Properties) *Let K be a subset of Ω. We shall say that K* is locally viable under f *(or enjoys the local* viability property *for the map f) if for any initial state x_0 of K,* there exist *$T > 0$ and a* viable *solution on $[0, T]$ to differential*

[3] or even, a normed space.

equation (1.1) starting at x_0. *It is said to be (globally) viable under* f *(or to enjoy the* global *viability property or, simply, the viability property) if we can always take* $T = \infty$.

The subset K *is said to be* invariant under f *(or enjoy the* invariance property*) if for any initial state* x_0 *of* K, all *solutions to differential equation (1.1) (a priori defined on* Ω*) are* viable in K.

Remark — We should emphasize that the concept of invariance *depends upon the behavior of* f *on the domain* Ω *outside* K. But we observe that viability property depends only on the behavior of f on K. □

So, the viability property requires only the existence of at least one viable solution whereas the invariance property demands that all solutions are viable.

We shall begin by characterizing the subsets K which are viable under f. The idea is simple, intuitive and makes good sense: *A subset* K *is viable under* f *if at each state* x *of* K, *the velocity* $f(x)$ *is "tangent" to* K *at* x, *so to speak, for bringing back a solution to the differential equation inside* K.

But we do not want to restrict ourselves to the case of *smooth* domains (i.e., differential manifolds) only for the pleasure of obtaining a vector space of tangent directions or to conform to tradition. There are many reasons for this, the first one being that simple operations on subsets — such as the intersection of manifolds — destroy their smoothness. Since we shall perform operations on viability subsets, we have to dispense with this requirement and look for other ways of implementing the idea of tangency to any subset. In economics and ecology, for instance, viability subsets are defined by a family of equality or inequality constraints. They are not differential manifolds. The best we can hope for is that they are convex, which happens, for instance, when the constraints are linear.

Naturally, by trying to define adequate concepts of tangency to nonsmooth subsets, we expect to lose some nice properties of the tangent space, and, among them, the fact that tangent spaces are vector spaces. The price to pay is then to deal with *closed cones* instead. Actually, under some regularity conditions, we shall do even better, and obtain, closed *convex*[4] cones.

We shall postpone the study of tangent cones to Chapter 5, when we will need them, after having provided a strong justification of their usefulness. An exhaustive presentation can be found in Chapter 4 of SET-VALUED ANALYSIS.

Meanwhile, we shall just provide the definition of the *contingent cone*, introduced by Bouligand in the early thirties, with which we shall characterize the viability property by following our intuitive idea.

Definition 1.1.3 *Let X be a normed space, K be a nonempty subset of X and x belong to K. The* contingent cone to K at x *is the set*

$$T_K(x) = \left\{ v \in X \;\middle|\; \liminf_{h \to 0+} \frac{d_K(x+hv)}{h} = 0 \right\}$$

where $d_K(y)$ denotes the distance of y to K, defined by

$$d_K(y) \;\; := \;\; \inf_{z \in K} \|y - z\|$$

In other words, v *belongs to* $T_K(x)$ *if and only if there exist a sequence of* $h_n > 0$ *converging to* $0+$ *and a sequence of* $v_n \in X$ *converging to* v *such that*

$$\forall\, n \geq 0, \;\; x + h_n v_n \;\in\; K$$

We see easily that

$$\forall\, x \in \mathrm{Int}(K), \;\; T_K(x) = X \tag{1.2}$$

Therefore, *when K is open, the contingent cone to K at any point $x \in K$ is always equal to the whole space.* The converse is not true.

We also observe that when K is a differential manifold, the contingent cone $T_K(x)$ coincides with the tangent space to K at x, and we shall check later that when K is convex, it coincides with the tangent cone of convex analysis. The lemma below shows right away why these cones will play a crucial role: they appear naturally whenever we wish to differentiate viable functions.

---

[4]In this case, we will be able to use duality, by associating biunivocally *polar* cones to closed convex cones, and use the *bipolar Theorem* (Theorem 2.3.3.)

Lemma 1.1.4 *Let* $x(\cdot)$ *be a differentiable* viable *function from* $[0,T]$ *to* K. *Then*

$$\forall\, t \in [0,T[,\quad x'(t) \;\in\; T_K(x(t))$$

Definition 1.1.5 (Viability Domain) *Let* K *be a subset of* Ω. *We shall say that* K *is a* viability domain *of the map* $f:\Omega \mapsto X$ *if*

$$\forall\, x \in K,\;\; f(x) \;\in\; T_K(x) \tag{1.3}$$

Example — We first give the simple example of finite dimensional vector-spaces which are viability domains of linear operators.

Definition 1.1.6 *Let* A *be a linear operator from a finite dimensional vector-space* X *to itself. We shall say that a finite dimensional vector-subspace* K *is* invariant *under* A *if*

$$A(K) \subset K$$

The following statement is naturally obvious.

Proposition 1.1.7 *Let us consider a linear operator* A *from a finite dimensional vector-space* X *to itself, elements* $b, c \in X$ *and a subspace* K *of* X.
The affine space $K + c$ *is a viability domain of the affine operator* $x \to Ax + b$ *if and only if*

$$\begin{cases} i) & K \text{ is invariant under } A \\ ii) & Ac + b \;\in\; K \quad \square \end{cases}$$

1.2 Nagumo Theorem

Nagumo was the first one to prove the viability theorem for ordinary differential equations in 1942. This theorem was apparently forgotten, for it was rediscovered many times during the next twenty years[5].

[5] This does not prove that the statement is true...

Theorem 1.2.1 (Nagumo) *Let us assume that*

$$\begin{cases} i) & K \text{ is locally compact} \\ ii) & f \text{ is continuous from } K \text{ to } X \end{cases} \tag{1.4}$$

Then K is locally viable under f if and only if K is a viability domain of f.

Since the contingent cone to an open subset is equal to the whole space (see (1.2)), an open subset is a viability domain of any map. So, it enjoys the viability property because any open subset of a finite dimensional vector-space is locally compact. The Peano existence theorem is then a consequence of Theorem 1.2.1.

Theorem 1.2.2 (Peano) *Let Ω be an open subset of a finite dimensional vector-space X and $f : \Omega \mapsto X$ be a continuous map.*

Then, for every $x_0 \in \Omega$, there exists $T > 0$ such that differential equation (1.1) has a solution on the interval $[0, T]$ starting at x_0.

The interesting case from the viability point of view is the one when the viability subset is *closed.* In this case, we derive from Theorem 1.2.1 a more precise statement.

Theorem 1.2.3 (Viability) *Let us consider a* closed *subset K of a finite dimensional vector-space X and a* continuous *map f from K to X.*

If K is a viability domain, *then for every initial state $x_0 \in K$, there exist a positive T and a* viable *solution on $[0, T]$ to differential equation (1.1) starting at x_0 such that*

$$\begin{cases} \textit{either} & T = \infty \\ \quad \textit{or} & T < \infty \quad \textit{and} \quad \limsup_{t \to T-} \|x(t)\| = \infty \end{cases} \tag{1.5}$$

Further adequate information — a priori estimates on the growth of f — allows us to exclude the case when $\limsup_{t \to T-} \|x(t)\| = \infty$.

This is the case for instance when f is bounded on K, and, in particular, when K is bounded.

More generally, we can take $T = \infty$ when f enjoys linear growth:

Theorem 1.2.4 *Let us consider a subset K of a finite dimensional vector-space X and a map f from K to X. We assume that the map f is* continuous *from K to X, that*

$$\exists\, c > 0 \quad \textit{such that } \forall\, x \in K, \;\; \|f(x)\| \;\leq\; c(\|x\| + 1)$$

and that

$$K \text{ is a closed viability domain of } f$$

Then K is viable under f: for every initial state $x_0 \in K$, there exists a viable solution on $[0, \infty]$ *to differential equation (1.1) starting at x_0.*

We shall prove only Theorem 1.2.1. The proofs of the other theorems are classical and are the same as the ones for analogous statements for differential inclusions (see Chapter 3).

Proof of Theorem 1.2.1

a) — **Necessary Condition**

Let us consider a viable solution $x(\cdot)$ to differential equation (1.1.) It is easy to check that $f(x_0) = x'(0)$ belongs to the contingent cone $T_K(x_0)$ because $x(h)$ belongs to K and consequently, the inequality

$$d_K(x_0 + hf(x_0))/h \leq \|x(0) + hx'(0) - x(h)\|/h$$

implies that

$$\lim_{h \to 0+} d_K(x_0 + hf(x_0))/h \;=\; 0$$

Hence K is a viability domain.

b) — **Sufficient Condition**

As quite often happens in analysis, the existence proof can be split into three steps. We begin by constructing approximate solutions by modifying Euler's method to take into account the viability constraints, we then deduce from available estimates that a subsequence of these solutions converges uniformly to a limit, and finally check that this limit is a viable solution to differential equation (1.1.)

1. — **Construction of Approximate Solutions**

Since K is locally compact, there exists $r > 0$ such that the ball $B_K(x_0, r) := K \cap (x_0 + rB)$ is *compact.* When C is a subset, we set

$$\|C\| := \sup_{v \in C} \|v\|$$

and

$$K_0 := B_K(x_0, r), \; C := B(f(K_0), 1), \; T := r/\|C\|$$

We observe that C *is bounded* since K_0 *is compact.* We begin by proving

Lemma 1.2.5 *For any integer m, there exists $\theta_m \in]0, 1/m[$ such that for all $x \in K_0$, there exist $h \in [\theta_m, 1/m]$ and $u \in X$ satisfying*

$$\begin{cases} i) & u \in C \\ ii) & x + hu \in K \\ iii) & (x, u) \in B(\mathrm{Graph}(f), 1/m) \end{cases}$$

Proof of Lemma 1.2.5 — Since K is a *viability domain* of f, we know that for all $y \in K$, $f(y)$ belongs to $T_K(y)$. By definition of the contingent cone, there exists $h_y \in]0, 1/m[$ such that

$$d_K(y + h_y f(y)) < h_y/2m$$

We introduce the subsets

$$N(y) := \{x \in K \mid d_K(x + h_y f(y)) < h_y/2m\}$$

These subsets are obviously *open.* Since y belongs to $N(y)$, there exists $\eta_y \in]0, 1/m[$ such that $B(y, \eta_y) \subset N(y)$. The compactness of K_0 implies that it can be covered by q such balls $B(y_j, \eta_j)$, $j = 1, \ldots, q$. We set

$$\theta_m := \min_{j=1,\ldots,q} h_{y_j} > 0$$

Let us choose any $x \in K_0$. Since it belongs to one of the balls $B(y_j, \eta_j) \subset N(y_j)$, there exists $z_j \in K$ such that

$$\begin{aligned} & \|x + h_{y_j} f(y_j) - z_j\|/h_{y_j} \\ & \le d_K(x + h_{y_j} f(y_j))/h_{y_j} + 1/2m \le 1/m. \end{aligned}$$

Let us set $u_j := \frac{z_j - x}{h_{y_j}}$. We see that $\|x - y_j\| \leq \eta_j \leq 1/m$, that $x + h_{y_j} u_j = z_j \in K$ and that $\|u_j - f(y_j)\| \leq 1/m$. Therefore,

$$(x, u_j) \in B((y_j, f(y_j)), 1/m) \subset B(\text{Graph}(f), 1/m)$$

and $u_j \in B(f(K_0), 1/m) \subset C$. The Lemma ensues. □

We can now construct by induction a sequence of positive numbers $h_j \in]\theta_m, 1/m[$ and a sequence of elements $x_j \in K_0$ and $u_j \in C$ such that

$$\begin{cases} i) & x_{j+1} := x_j + h_j u_j \in K_0, \; u_j \in C \\ ii) & (x_j, u_j) \in B(\text{Graph}(f), 1/m) \end{cases}$$

so long as $\sum_{i=0}^{j-1} h_i \leq T$.

Indeed, the elements x_j belong to K_0, since

$$\begin{aligned} \|x_j - x\| &\leq \sum_{i=0}^{i=j-1} \|x_{i+1} - x_i\| \\ &\leq \sum_{i=0}^{i=j-1} h_i \|C\| \leq T\|C\| = r \end{aligned}$$

Since the h_j's are larger than or equal to $\theta_m > 0$, there exists J such that

$$h_1 + \cdots + h_{J-1} \leq T \leq h_1 + \cdots + h_J$$

We introduce the nodes $\tau_m^j := h_0 + \cdots + h_{j-1}, j = 1, \ldots, J+1$ and we interpolate the sequence of elements x_j at the nodes τ_m^j by the piecewise linear functions $x_m(t)$ defined on each interval $[\tau_m^j, \tau_m^{j+1}[$ by

$$\forall\, t \in [\tau_m^j, \tau_m^{j+1}[, \; x_m(t) := x_j + (t - \tau_m^j) u_j$$

We observe that this sequence satisfies the following estimates

$$\begin{cases} i) & \forall\, t \in [0, T], \; x_m(t) \in \text{co}(K_0) \\ ii) & \forall\, t \in [0, T], \; \|x'_m(t)\| \leq \|C\| \end{cases} \tag{1.6}$$

Let us fix $t \in [\tau_m^j, \tau_m^{j+1}[$. Since $\|x_m(t) - x_m(\tau_m^j)\| = h_j \|u_j\| \leq \|C\|/m$, and since (x_j, u_j) belongs to $B(\text{Graph}(f), 1/m)$ by Lemma 1.2.5, we deduce that these functions are approximate solutions in the sense that

$$\begin{cases} i) & \forall\, t \in [0,T], \;\; x_m(t) \in B(K_0, \varepsilon_m) \\ ii) & \forall\, t \in [0,T], \;\; (x_m(t), x'_m(t)) \in B(\mathrm{Graph}(f), \varepsilon_m) \end{cases} \tag{1.7}$$

where $\varepsilon_m := (\|C\| + 1)/m$ converges to 0.

2. — Convergence of the Approximate Solutions

Estimates (1.6) imply that for all $t \in [0,T]$, the sequence $x_m(t)$ remains in the compact subset $\mathrm{co}(K_0)$[6] and that the sequence $x_m(\cdot)$ is *equicontinuous*, because the derivatives $x'_m(\cdot)$ are bounded. We then deduce from Ascoli's Theorem[7] that it remains in a compact subset of the Banach space $\mathcal{C}(0,T;X)$, and thus, that a subsequence (again denoted) $x_m(\cdot)$ converges uniformly to some function $x(\cdot)$. Furthermore, the sequence $x'_m(\cdot)$ also converges to $x'(\cdot)$ because $x'_m(t) = f(x_m(t))$ and f is uniformly continuous on the compact $\mathrm{co}(K_0)$.

3. — The Limit is a Solution

Condition (1.7)i) implies that

$$\forall\, t \in [0,T], \;\; x(t) \in K_0$$

[6] The *(closed) convex hull* of a subset is the intersection of the (closed) convex subsets which contain it. The convex hull of a compact subset is also compact.

[7] Let us recall that a subset $\mathcal{H}$ of continuous functions of $\mathcal{C}(0,T;X)$ is *equicontinuous* if and only if

$$\forall\, t \in [0,T], \; \forall\, \varepsilon > 0, \; \exists\, \eta \;:=\; \eta(\mathcal{H}, t, \eta) \mid \forall\, s \in [t-\eta, t+\eta], \;\; \sup_{x(\cdot)\in\mathcal{H}} \|x(t) - x(s)\| \leq \varepsilon$$

Locally Lipschitz functions with the same Lipschitz constant form an equicontinuous set of functions. In particular, a subset of differentiable functions satisfying

$$\sup_{t\in[0,T]} \|x'(t)\| \leq c < +\infty$$

is equicontinuous.

Ascoli's Theorem states that a subset $\mathcal{H}$ of functions is *relatively compact* in $\mathcal{C}(0,T;X)$ if and only if it is equicontinuous and satisfies

$$\forall\, t \in [0,T], \; \mathcal{H}(t) \;:=\; \{x(t)\}_{x(\cdot)\in\mathcal{H}} \text{ is compact.}$$

i.e., that $x(\cdot)$ is viable.

Since f is uniformly continuous on K_0, then for all $\varepsilon > 0$, there exists $\eta \in]0, \varepsilon[$ such that

$$\|f(x) - f(y)\| \leq \varepsilon \text{ whenever } \|x - y\| \leq \eta$$

Since the sequence $x_m(\cdot)$ converges uniformly to $x(\cdot)$ and since property (1.7)ii) holds true, we deduce that for large m and for all $t \in [0, T]$, there exists $u_m^t \in X$ such that

$$\begin{cases} \|x'_m(t) - f(x(t))\| \\ \leq \|x'_m(t) - f(u_m^t)\| + \|f(u_m^t) - f(x_m(t))\| + \|f(x_m(t)) - f(x(t))\| \\ \leq 3\varepsilon \end{cases}$$

so that

$$\left\| x_m(t) - x_0 - \int_0^t f(x(s))ds \right\| \leq \int_0^t \|x'_m(s) - f(x(s))\| ds \leq 3\varepsilon t$$

By letting m go to ∞, these inequalities imply that

$$\forall t \in [0, T], \quad x(t) = x_0 + \int_0^t f(x(s))ds$$

Hence the limit $x(\cdot)$ is a solution to differential equation (1.1), and thus, K enjoys the viability property. □

1.3 Numerical Schemes

A natural approximation scheme for approximating viable solutions to differential equations is the projected explicit difference scheme

$$x_{j+1} = \pi_K(x_j + hf(x_j))$$

where $h > 0$ is fixed and where π_K denotes a selection of the projector of best approximation Π_K defined by

$$y \in \Pi_K(x) \Longleftrightarrow y \in K \;\&\; \|y - x\| = d_K(x)$$

Let us observe that π_K satisfies the property

$$\forall z \in K, \forall x \in X, \quad \|\pi_K(x) - z\| \leq 2\|x - z\|$$

because, whenever $y \in \Pi_K(x)$,

$$\|y - z\| \le \|y - x\| + \|x - z\| = d_K(x) + \|x - z\| \le 2\|x - z\|$$

When K is convex, Π_K is a Lipschitz single-valued map (with Lipschitz constant equal to 1.)

Projectors of best approximation are instances of *quasi-projectors*:

Definition 1.3.1 *We shall say that a map Γ_K from X onto K satisfying*

$$\begin{cases} i) & \forall\, z \in K, \ \ \Gamma_K(z) = z \\ & \\ ii) & \exists\, \lambda > 0 \quad \textit{such that} \\ & \forall\, x \in X, \forall\, z \in K, \ \ \|\Gamma_K(x) - z\| \ \le \ \lambda\|x - z\| \end{cases}$$

is a quasi-projector *onto K.*

There are many other examples of quasi-projectors. They enjoy the following property:

Lemma 1.3.2 *Let Γ_K be a quasi-projector from X onto K. Then $\|\Gamma_K(x + hv) - x - hv\| \le (\lambda + 1)d_K(x + hv)$ so that, for instance,*

$$\forall\, v \in T_K(x), \ \ \liminf_{h \to 0+} \frac{\|\Gamma_K(x + hv) - x - hv\|}{h} = 0$$

We can associate with any quasi-projector a projected explicit difference scheme providing a sequence x_j starting from x_0 and defined by

$$x_{j+1} \ := \ \Gamma_K(x_j + hf(x_j)) \tag{1.8}$$

and an approximate viable solution $x_h(\cdot)$ which is the piecewise linear function interpolating this sequence on the nodes $\tau_h^j := jh$ defined by $x_h(t) := x_j + (t - jh)(x_{j+1} - x_j)/h$ on the intervals $[jh, (j + 1)h[$.

Theorem 1.3.3 *Let us consider a continuous map f from a compact subset $K \subset X$ to X such that, for every $x \in K, f(x) \in T_K(x)$, and a quasi-projector Γ_K. Then, starting from $x_0 \in K$, the solutions to the projected explicit difference scheme (1.8) converge to a viable solution to differential equation $x' = f(x)$ when $h \to 0+$, in the sense that a subsequence of the piecewise linear functions x_h which interpolates the x_j's on the nodes jh converges uniformly to a viable solution $x(\cdot)$.*

This is a corollary of the set-valued version Theorem 10.3.2 of our statement.

Remark — Actually, when f is not continuous, the proof shows that the solutions to the projected explicit scheme converge to viable solutions so long as property

$$\forall\, x \in K, \quad \lim_{h\to 0+, K\ni y\to x} \frac{d_K(y+hf(y))}{h} = 0 \quad \square \tag{1.9}$$

(which is a consequence of the continuity of f) holds true.

Remark — When the viability domain K of f is convex and compact, we can derive from the Equilibrium Theorem 3.7.6 below that there exists a viable solution to the implicit finite difference scheme

$$x_{j+1} \;=\; x_j + hf(x_{j+1}) \;\;\&\;\; x_{j+1} \in K$$

starting from x_0. $\square$

1.4 Replicator Systems

We begin by studying the viability property of the probability simplex

$$S^n := \left\{ x \in \mathbf{R}^n_+ \mid \sum_{i=1}^{n} x_i = 1 \right\}$$

This is the most important instance of a viability set, because, in many problems, it is too difficult to describe the state of the system mathematically. We shall provide examples later in this section.

But for recognizing whether the simplex is the viability domain of some differential equation, we need to compute its contingent cones.

Lemma 1.4.1 *The contingent cone $T_{S^n}(x)$ to S^n at $x \in S^n$ is the cone of elements $v \in \mathbf{R}^n$ satisfying*

$$\sum_{i=1}^{n} v_i = 0 \;\;\&\;\; v_i \geq 0 \;\text{ whenever }\; x_i = 0 \tag{1.10}$$

We provide a direct proof of this lemma, which is a consequence of the calculus of contingent cones.

Proof — Let us take $v \in T_{S^n}(x)$. There exist sequences $h_p > 0$ converging to 0 and v_p converging to v such that $y_p := x + h_p v_p$ belongs to S^n for any $p \geq 0$. Then

$$\sum_{i=1}^{n} v_{p_i} = \frac{1}{h_p}\left(\sum_{i=1}^{n} y_{p_i} - \sum_{i=1}^{n} x_{p_i}\right) = 0$$

so that $\sum_{i=1}^{n} v_i = 0$. On the other hand, if $x_i = 0$, then $v_{p_i} = y_{p_i}/h_p \geq 0$, so that $v_i \geq 0$.

Conversely, let us take v satisfying (1.10) and deduce that $y := x + hv$ belongs to the simplex for h small enough. First, the sum of the y_i is obviously equal to 1. Second, $y_i \geq 0$, either when $x_i = 0$ because in this case v_i is nonnegative, or when $x_i > 0$, because it is sufficient to take $h < x_i/|v_i|$ for having $y_i \geq 0$. Hence y does belong to the simplex. □

We shall investigate now how to make viable the evolution of a system for which we know the growth rates $g_i(\cdot)$ of the evolution without constraints (also called "specific growth rates"):

$$\forall\ i = 1, \ldots, n,\ \ x_i'(t) = x_i(t) g_i(x(t))$$

There are no reasons[8] for the solutions to this system of differential equations to be viable in the probability simplex.

But we can correct it by subtracting to each initial growth rate the common "feedback control $\tilde{u}(\cdot)$" (also called "global flux" in many applications) defined as the weighted mean of the specific growth rates

$$\forall\ x \in S^n,\ \ \tilde{u}(x) := \sum_{j=1}^{n} x_j g_j(x)$$

Indeed, the probability simplex S^n is obviously a viability domain of the new dynamical system, called *replicator system* (or system *under*

[8] By Nagumo's Theorem and Lemma 1.4.1, the functions g_i should be continuous and satisfy:

$$\forall\, x \in S^n, \quad \sum_{i=1}^{n} x_i g_i(x) = 0$$

constant organization):

$$\begin{cases} \forall\ i=1,\ldots,n,\ x_i'(t) = x_i(t)(g_i(x(t)) - \tilde{u}(x(t))) \\ = x_i(t)(g_i(x(t)) - \sum_{j=1}^n x_j(t)g_j(x(t))) \end{cases} \quad (1.11)$$

As we shall see at the end of the section, these equations come up in many biological models related to the concept of "replicator" in the sense of Dawkins, who coined the term. They lead to many mathematical problems.

Remark — There are other methods for correcting a dynamical system to make a given closed subset a viability domain. A general method consists in projecting the dynamics onto the contingent cone (see chapter 10.) Here, we have taken advantage of the particular nature of the simplex. □

An equilibrium α of the replicator system (1.11) is a solution to the system

$$\forall\, i=1,\ldots,n,\quad \alpha_i(g_i(\alpha) - \tilde{u}(\alpha)) = 0$$

(Such an equilibrium does exist, thanks to Equilibrium Theorem 3.7.6 below.) These equations imply that either $\alpha_i = 0$ or $g_i(\alpha) = \tilde{u}(\alpha)$ or both, and that $g_{i_0}(\alpha) = \tilde{u}(\alpha)$ holds true for at least one i_0. We shall say that an *equilibrium α is nondegenerate* if

$$\forall\, i=1,\ldots,n,\quad g_i(\alpha) = \tilde{u}(\alpha) \quad (1.12)$$

Equilibria α which are strongly positive (this means that $\alpha_i > 0$ for all $i = 1,\ldots,n$) are naturally non degenerate.

We associate with any $\alpha \in S^n$ the function V_α defined[9] on the simplex S^n by

$$V_\alpha(x) := \prod_{i=1}^n x_i^{\alpha_i} := \prod_{i\in I_\alpha} x_i^{\alpha_i}$$

[9]The reason why we introduce this function is that α is the unique maximizer of V_α on the simplex S^n. This follows from the concavity of the function $\varphi :=$ log: Setting $0\log 0 = 0 \log\infty = 0$, we get

$$\sum_{i=1}^n \alpha_i \log\frac{x_i}{\alpha_i} = \sum_{\alpha_i>0} \alpha_i\log\frac{x_i}{\alpha_i} \le \log\left(\sum_{\alpha_i>0} x_i\right) \le \log 1 = 0$$

where we set $0^0 := 1$ and $I_\alpha := \{i = 1, \ldots, n \mid \alpha_i > 0\}$.

Let us denote by S^I the subsimplex of elements $x \in S^n$ such that $x_i > 0$ if and only if $i \in I$.

Theorem 1.4.2 *Let us consider n continuous growth rates g_i. For every initial state $x_0 \in S^n$, there exists a solution to replicator system (1.11) starting from x_0 and which is viable in the subsimplex $S^{I_{x_0}}$.*

The viable solutions satisfy

$$\forall\, t \geq 0, \ \sum_{i=1}^{n} g_i(x(t))x_i'(t) \ \geq \ 0 \tag{1.13}$$

and, whenever $\alpha \in S^n$ is a nondegenerate equilibrium,

$$\frac{d}{dt}V_\alpha(x(t)) \ = \ -V_\alpha(x(t))\sum_{i=1}^{n}(x_i(t) - \alpha_i)(g_i(x(t)) - g_i(\alpha)) \tag{1.14}$$

Proof — We first observe that

$$\forall\, x \in S^{I_{x_0}}, \ \sum_{i \in I_{x_0}} x_i(g_i(x) - \tilde{u}(x)) \ = \ 0$$

because, $x_i = 0$ whenever $i \notin I_{x_0}$, i.e., whenever $x_{0_i} = 0$. Therefore, the subsimplex $S^{I_{x_0}}$ is a viability domain of the replicator system (1.11.)

Inequality (1.13) follows from the Cauchy-Schwarz inequality because

$$\left(\sum_{i=1}^{n} x_i g_i(x)\right)^2 \ \leq \ \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{j=1}^{n} x_i g_i(x)^2\right) \ = \ \sum_{i=1}^{n} x_i g_i(x)^2$$

We deduce formula (1.14) from

$$\begin{cases} \frac{d}{dt}V_\alpha(x(t)) \ = \ \sum_{i \in I_\alpha} \frac{\partial}{\partial x_i} V_\alpha(x(t))x_i'(t) \\ \\ = \ V_\alpha(x(t)) \sum_{i \in I_\alpha} \alpha_i \frac{x_i'(t)}{x_i(t)} \end{cases}$$

---

so that

$$\sum_{i=1}^{n} \alpha_i \log x_i \ \leq \ \sum_{i=1}^{n} \alpha_i \log \alpha_i$$

and thus, $V_\alpha(x) \leq V_\alpha(\alpha)$ with equality if and only if $x = \alpha$.

and from

$$\sum_{i=1}^{n} \alpha_i \frac{x_i'(t)}{x_i(t)} = \sum_{i=1}^{n} (\alpha_i - x_i(t)) g_i(x(t))$$

Then we take into account that α being a non degenerate equilibrium, equation (1.12) implies that

$$\sum_{i=1}^{n} (\alpha_i - x_i(t)) g_i(\alpha) = 0 \quad \square$$

Remark — When the specific growth rates are derived from a differentiable potential function U by

$$\forall\, i = 1, \ldots, n, \quad g_i(x) := \frac{\partial U}{\partial x_i}(x)$$

condition (1.13) implies that

$$\forall\, t \geq 0, \quad \frac{dU}{dt}(x(t)) \geq 0$$

because

$$\frac{dU}{dt}(x(t)) = \sum_{i=1}^{n} \frac{\partial U}{\partial x_i}(x(t)) x_i'(t) = \sum_{i=1}^{n} g_i(x(t)) x_i'(t) \geq 0$$

Therefore *the potential function U does not decrease along the viable solutions to the replicator system* (1.11.)

Furthermore, when this potential function U is homogeneous with degree p, Euler's formula implies that

$$\tilde{u}(x) = pU(x)$$

(because $\sum_{i=1}^{n} x_i \frac{\partial}{\partial x_i} U(x) = pU(x)$) so that in this case, *the global flux $\tilde{u}(x(t))$ also does not decrease along the viable solutions to the replicator system* (1.11.)

On the other hand, if we assume that the growth rates $-g_i$ are "monotone " in the sense that

$$\forall\, x, y \in S^n, \quad \sum_{i=1}^{n} (x_i - y_i)(g_i(x) - g_i(y)) \leq 0$$

then inequality (1.14) implies that for any non degenerate equilibrium $\alpha \in S^n$,

$$\forall\, t \geq 0, \quad \frac{dV_\alpha}{dt}(x(t)) \;\geq\; 0$$

When $g(x) := U'(x)$ is derived from a concave differentiable potential U, it is decreasing so that, *for a concave potential, both $U(x(\cdot))$ and $V_\alpha(x(\cdot))$ are not decreasing.* □

Example: Replicator systems for constant growth rates. The simplest example is the one where the specific growth rates $g_i(\cdot) \equiv a_i$ are constant. Hence we correct constant growth systems $x_i' = a_i x_i$ whose solutions are exponential $x_{0_i} e^{a_i t}$, by the 0-order replicator system

$$\forall\, i = 1, \ldots, n, \;\; x_i'(t) \;=\; x_i(t)(a_i - \sum_{j=1}^{n} a_j x_j(t))$$

whose solutions are given explicitly by:

$$x_i(t) \;=\; \frac{x_{0_i} e^{a_i t}}{\sum_{j=1}^{n} x_{0_j} e^{a_j t}} \;\; \text{whenever } x_{0_i} > 0$$

(and $x_i(t) \equiv 0$ whenever $x_{0_i} = 0$.)

Furthermore, *the functions $\sum_{i=1}^{n} a_i x_i(\cdot)$ are increasing* and converge to α defined by

$$\alpha_i \;=\; \begin{cases} 0 & \text{if } i \notin J^a \\ x_{0_i} / \sum_{j \in J^a} x_{0_j} & \text{if } i \in J^a \end{cases}$$

where $J^a := \{i = 1, \ldots, n, \mid a_i := \max_{j=1,\ldots,n} a_j\}$. Indeed, set $a_0 := \max_{j=1,\ldots,n} a_j$; the above claim follows obviously from formula

$$x_i(t) \;=\; \frac{x_{0_i} e^{-(a_0 - a_i)t}}{\sum_{j=1}^{n} x_{0_j}} e^{-(a_0 - a_j)t}$$

Observe that the limit points of the viable solutions achieve the maximum of the function $x \to \sum_{i=1}^{n} a_i x_i$ on S^n, since any $\alpha \in S^{J^a}$

achieves the maximum of this linear functional[10]. Observe also that the elements $\alpha \in S^{J^a}$ are equilibria of the 0-order replicator system. Actually, the equilibria of the 0-order replicator system are the elements of the each subsimplex S^{L_k} where $L_k := \{j \mid a_j = a_k\}$[11].

When $n = 2$, after setting $x(\cdot) := x_1(\cdot)$ and $r := a_1 - a_2$, we obtain the celebrated Verhust-Pearl's *logistic equation*

$$\forall\, t \geq 0, \;\; x'(t) \;=\; rx(t)(1 - x(t))$$

the solutions of which are the logistic curves (the S-curves)

$$x(t) \;:=\; \frac{1}{1 + ce^{-bt}}$$

The logistic equation played an important role in *population dynamics*. In the simplest case, assume that the growth rate of an organism is constant whenever there are no constraints on the resources needed for growth. This is no longer valid when the resources are limited for whatever reason. This fact is translated by saying that the growth rate becomes negative for large populations: the larger the population, the more severe the inhibition on further growth. The simplest growth rate fitting these requirements is the function $r(1 - x)$, so that the evolution of the population obeys the logistic equation. □

[10]Since $a_0 := \max_{j=1,\ldots,n} a_j$, we deduce that

$$\begin{cases} \sum_{i=1}^n a_i(\alpha_i - x_i) \;=\; \sum_{i\in J^a} a_i(\alpha_i - x_i) - \sum_{i\notin J^a} a_i(\alpha_i - x_i) \\ \;=\; a_0(1 - \sum_{i\in J^a} x_i) - \sum_{i\notin J^a} a_i x_i = \sum_{i\in J^a}(a_0 - a_i)x_i \geq 0 \end{cases}$$

[11]Indeed, if $\alpha \in S^{L_k}$, then $\sum_{i=1}^n a_j\alpha_j = \sum_{j\in L_k} a_j\alpha_j = a_k$ and thus,

$$\alpha_i(a_i - \sum_{j=1}^n a_j\alpha_j) \;=\; 0$$

for any $i = 1, \ldots, n$.

Example: Replicator systems for linear growth rates. The next class of examples is provided by linear growth rates

$$\forall\, i = 1, \ldots, n, \;\; g_i(x) \;:=\; \sum_{j=1}^{n} a_{ij} x_j$$

Let A denote the matrix the entries of which are the above a_{ij}'s. Hence the global flux can be written

$$\forall\, x \in S^n, \;\; \tilde{u}(x) \;=\; \sum_{k,l=1}^{n} a_{kl} x_k x_l \;=\; < Ax, x >$$

Therefore, first order replicator systems can be written[12].

$$\forall\, i = 1, \ldots, n, \;\; x_i'(t) \;=\; x_i(t)(\sum_{j=1}^{n} a_{ij} x_j(t) - \sum_{k,l=1}^{n} a_{kl} x_k(t) x_l(t))$$

Such systems have been investigated independently in

— *population genetics* (allele frequencies in a gene pool)

— *theory of prebiotic evolution* of self replicating polymers (concentrations of polynucleotides in a dialysis reactor)

— *sociobiological studies* of evolutionary stable traits of animal behavior (distributions of behavioral phenotypes in a given species)

— *population ecology* (densities of interacting species)

In population genetics, *Fisher-Wright-Haldane's model* regards the state $x \in S^n$ as the frequencies of alleles in a gene pool and the matrix $A := (a_{ij})_{i,j=1,\ldots,n}$ as the *fitness matrix*, where a_{ij} represents the fitness of the genotype (i, j). In this case, the matrix A is obviously symmetric and we denote by

$$\tilde{u}(x) := < Ax, x > \quad \textit{the average fitness}$$

Since the growth rate can be derived from the potential $U(x) := \tilde{u}(x)/2$, we conclude that whenever A is positive-definite, the *average adaptability does not decrease*[13] *along viable solutions.*

[12]Observe that if for each i, all the a_{ij} are equal to b_i, we find 0-order replicator systems

[13]This property is known as the fundamental theorem of natural selection in population genetics.

In the theory of *prebiotic evolution*, the state represents the concentrations of polynucleotides. It is assumed in *Eigen-Schuster's "hypercycle"* that the growth rate of the i^{th}-polynucleotide is proportional to the concentration of the preceding one:

$$\forall\, i = 1, \ldots, n, \quad g_i(x) \;=\; c_i x_{i-1} \;\text{ where }\; x_{-1} := x_n$$

In other words, the growth of polynucleotide i is catalyzed by its predecessor by Michaelis-Menten type chemical reactions.

The feedback $\tilde{u}(x) = \sum_{i=1}^{n} c_i x_i x_{i-1}$ can be regarded as a selective pressure to maintain the concentration.

The equilibrium α of such a system is equal to

$$\forall\, i = 1, \ldots, n, \quad \alpha_i \;=\; \frac{1}{c_{i+1}} \left(\sum_{j=1}^{n} \frac{1}{c_j} \right)^{-1} \quad \text{where } \; c_{n+1} := c_1$$

First order replicator systems also offer a quite interesting model of *dynamic game* theory proposed in 1974 by J. Maynard-Smith to explain the evolution of genetically programmed behaviors of individuals of an animal species.

We denote by $i = 1, \ldots, n$ the n possible "strategies" used in interindividual competition in the species and denote by a_{ij} the "gain" when strategy i is played against strategy j. The state of the system is described by the "mixed strategies" $x \in S^n$, which are the probabilities with which the strategies are implemented. Hence the growth rate $g_i(x) := \sum_{j=1}^{n} a_{ij} x_j$ is the gain obtained by playing strategy i against the mixed strategy x and $\tilde{u}(x) := \sum_{i,j=1}^{n} a_{ij} x_i x_j$ can be interpreted as the *average gain* .

So the growth rate of the strategy i in the replicator system is equal to the difference between the gain of i and the average gain (a behavior which had been proposed in 1978 by Taylor and Jonker.)

We shall say that an equilibrium α is *evolutionary stable* if and only if the property

$$\exists\, \eta > 0 \quad \text{such that } \; x \neq \alpha, \; \sum_{i=1}^{n} g_i(x)(\alpha_i - x_i) > 0$$

holds true in a neighborhood of α.

This implies that

$$\frac{d}{dt}V_\alpha(x(t)) = -V_\alpha(x(t))\sum_{i=1}^{n}(x_i(t)-\alpha_i)(g_i(x(t))-g_i(\alpha)) > 0$$

in a neighborhood of α.

It is interesting to observe that first order replicator systems can be used at the two extremes of biological evolution, prebiotic evolution at the molecular level and behavioral evolution in ethology (animal behavior.)

In ecology, the main models are elaborations of the *Lotka-Volterra equations*

$$\forall\, i = 1,\ldots,n, \;\; x_i'(t) = x_i(t)\left(a_{i0} + \sum_{j=1}^{n} a_{ij}x_j(t)\right)$$

where the growth rate of each species depend in an affine way upon the number of organisms of the other species. A very simple transformation replaces this system by a first order replicator system. We compactify $\mathbf{R}_+^n$ by introducing homogeneous coordinates. We set $x_0 := 1$ and we introduce the map

$$\forall\, i = 0,\ldots,n, \;\; y_i := \frac{x_i}{\sum_{j=1}^{n} x_j}$$

from $\mathbf{R}_+^n$ onto S^{n+1}, the inverse of which is defined by $x_i := y_i/y_0$.

We set $a_{0j} = 0$ for all j, so that Lotka-Volterra's equation becomes

$$\forall\, i = 1,\ldots,n, \;\; y_i' = \frac{y_i}{y_0}\left(\sum_{j=0}^{n} a_{ij}y_j - \sum_{k,l=1}^{n} a_{kl}y_l y_k\right)$$

because

$$y_i' = \frac{x_i'}{\sum x_j} - \frac{x_i\sum x_j'}{(\sum x_j)^2} = x_i\left(\sum_{j=0}^{n} a_{ij}x_j\right)y_0 - x_i\left(\sum_{k,l=0}^{n} a_{kl}x_l x_k\right)y_0^2$$

This is, up to the multiplication by $\frac{1}{y_0}$, i.e., up to a modification of the time scale, a $(n+1)$-dimensional first order replicator system.

So, first-order replicator systems appear as a common denominator underlying these four biological processes.

1.5 Stochastic Viability and Invariance

The aim of this section is to extend to the stochastic case Nagumo's Theorem on viability and/or invariance properties of closed subsets with respect to a differential equation.

1.5.1 Stochastic Tangent Sets

Let us consider a σ-complete probability space $(\Omega, \mathcal{F}, P)$, an increasing family of σ-sub-algebras $\mathcal{F}_t \subset \mathcal{F}$ and a finite dimensional vector-space $X := \mathbf{R}^n$.

The constraints are defined by closed subsets $K_\omega \subset X$, where the set-valued map

$$K : \omega \in \Omega \rightsquigarrow K_\omega \subset X$$

is assumed to be $\mathcal{F}_0$-measurable (which can be regarded as a *random set-valued variable.*)

We denote by $\mathcal{K}$ the subset

$$\mathcal{K} \ := \ \{u \in L^2(\Omega, \mathcal{F}, P) \mid \text{for almost all } \omega \in \Omega, \ \ u_\omega \in K_\omega\}$$

For simplicity, we restrict ourselves to scalar $\mathcal{F}_t$-Wiener processes $W(t)$.

Definition 1.5.1 (Stochastic Contingent Set) *Let us consider an $\mathcal{F}_t$-random variable $x \in K$ (i.e., an $\mathcal{F}_t$-measurable selection of K.)*

We define the stochastic contingent set $\mathcal{T}_K(t, x)$ *to K at x (with respect to $\mathcal{F}_t$) as the set of pairs (γ, v) of $\mathcal{F}_t$-random variables satisfying the following property: There exist sequences of $h_n > 0$ converging to 0 and of $\mathcal{F}_{t+h_n}$-random variables a^n and b^n such that*

$$\left\{ \begin{array}{ll} i) & \mathbf{E}(\|a^n\|^2) \to 0 \\ ii) & \mathbf{E}(\|b^n\|^2) \to 0 \\ iii) & \mathbf{E}(b^n) \ = \ 0 \\ iv) & b^n \ \text{ is independent of } \ \mathcal{F}_t \end{array} \right. \tag{1.15}$$

and satisfying

$$\forall\, n \geq 0, \ \ x + v(W(t+h_n) - W(t)) + h_n\gamma + h_n a^n + \sqrt{h_n} b^n \ \in \ \mathcal{K} \tag{1.16}$$

1.5.2 Stochastic Viability

We consider the stochastic differential equation

$$d\xi \ = \ f(\xi(t))dt + g(\xi(t))dW(t) \tag{1.17}$$

where f and g are Lipschitz.

We say that a stochastic process $\xi(t)$ defined by

$$\xi(t) = \xi(0) + \int_0^t f(\xi(s))ds + \int_0^t g(\xi(s))dW(s) \qquad (1.18)$$

is a solution to the stochastic differential equation (1.17) if the functions f and g satisfy:

$$\text{for almost all } \omega \in \Omega, \ \ f(\xi(\cdot)) \in L^1(0,T;X) \ \ \& \ \ g(\xi(\cdot)) \in L^2(0,T;X)$$

Definition 1.5.2 *We shall say that a stochastic process $x(\cdot)$ is* viable in *K if and only if*

$$\forall\, t \in [0,T], \ \ x(t) \in \mathcal{K} \qquad (1.19)$$

i.e., if and only if

$$\forall\, t \in [0,T], \ \ \text{for almost all } \omega \in \Omega, \ \ \xi_\omega(t) \in K_\omega$$

The random set-valued variable K is said to be (stochastically) invariant *by the pair (f,g) if every solution ξ to the stochastic differential equation starting at a random variable $x \in K$ is* viable in *K.*

When K is a subset of X (i.e., a constant set-valued random variable) and when the maps (f,g) are defined on K, we shall say that K enjoys the (stochastic) viability property *with respect to the pair (f,g) if for any random variable x in K, there exists a solution ξ to the stochastic differential equation starting at a x which is* viable in *K.*

Since K_ω and $\xi_\omega(0)$ are $\mathcal{F}_0$ measurable, the projection $\Pi_{K_\omega}(\xi_\omega(0))$ is also a $\mathcal{F}_0$-measurable map (see Theorem 8.2.13, p. 317 of SET-VALUED ANALYSIS.) Then there exists a $\mathcal{F}_0$-measurable selection $y_\omega \in \Pi_{K_\omega}(\xi_\omega(0))$, which we call a *projection of the random variable $\xi(0)$ onto the random set-valued variable K.*

Theorem 1.5.3 (Stochastic Viability) *Let K be a closed convex subset of X. Then the following conditions are equivalent:*

1. — The subset K enjoys the stochastic viability property with respect to the pair (f,g)

2. — for every $\mathcal{F}_t$-random variable x viable in K,

$$(f(x), g(x)) \in \mathcal{T}_K(t,x) \qquad (1.20)$$

We shall deduce this theorem from more general Theorems 1.5.4 and Theorems 1.5.5 below dealing with set-valued random variables instead of closed convex subsets.

1.5.3 Necessary Conditions

Let K be a set-valued random variable.

Theorem 1.5.4 *If the random set-valued variable K is invariant by the pair (f, g), then for every $\mathcal{F}_t$-random variable x viable in K,*

$$(f(x), g(x)) \in \mathcal{T}_K(t, x) \tag{1.21}$$

Proof — We consider the viable stochastic process $\xi(t)$

$$\xi(h) = x + \int_0^h f(\xi(s))ds + \int_0^h g(\xi(s))dW(s) \tag{1.22}$$

which is a solution to the stochastic differential equation (1.17) starting at x.

We can write it in the form

$$\xi(t) = \xi(0) + hf(\xi(0)) + g(\xi(0))W(h) + \int_0^h a(s)ds + \int_0^h b(s)dW(s)$$

where

$$\begin{cases} a(s) = f(\xi(s)) - f(\xi(0)) \\ b(s) = g(\xi(s)) - g(\xi(0)) \end{cases}$$

converge to 0 with s.

We set

$$a^h := \frac{1}{h}\int_t^{t+h} a(s)ds$$

and

$$b^h := \frac{1}{\sqrt{h}}\int_t^{t+h} b(s)dW(s)$$

and we observe that

$$\begin{cases} \mathbf{E}\left(\|a^h\|^2\right) = \frac{1}{h^2}\mathbf{E}\left(\left\|\int_t^{t+h} a(s)ds\right\|^2\right) \\ \leq \frac{1}{h}\int_t^{t+h} \mathbf{E}\left(\|(a(s)\|^2)\right) ds \end{cases}$$

converges to 0 because $\mathbf{E}\left(\left\|\int_0^t \varphi(s)ds\right\|^2\right) \leq t\int_0^t \mathbf{E}(\|\varphi(s)\|^2)ds$.

In the same way,

$$\begin{cases} \mathbf{E}\left(\|b^h\|^2\right) = \frac{1}{h}\mathbf{E}\left(\left\|\int_t^{t+h} b(s)dW(s)\right\|^2\right) \\ = \frac{1}{h}\int_t^{t+h} \mathbf{E}\left(\|b(s)\|^2\right) ds \end{cases}$$

also converges to 0 because $\mathbf{E}\left(\left\|\int_0^t \varphi(s)dW(s)\right\|^2\right) = \int_0^t \mathbf{E}(\|\varphi(s)\|^2)ds$.

The expectation of b^h is obviously equal to 0 and b^h is independent of $\mathcal{F}_t$. Since $\xi(h)_\omega$ belongs to K_ω for almost all ω, we deduce that the pair $(f(x), g(x))$ belongs to $\mathcal{T}_K(t, x)$. □

1.5.4 Sufficient Conditions for Stochastic Invariance

Theorem 1.5.5 (Stochastic Invariance) *Assume that the set-valued random variable K satisfies the following property: for every $\mathcal{F}_t$-random variable x, there exists an $\mathcal{F}_t$- measurable projection $y \in \Pi_K(x)$ such that*

$$(f(x), g(x)) \in \mathcal{T}_K(t, y) \tag{1.23}$$

Then the set-valued random variable K is invariant by (f, g)

Remark — Observe that the sufficient condition of invariance requires the verification of the "stochastic tangential condition" (1.23) for every stochastic process y, including stochastic processes which are not viable in K. □

In order to prove Theorem 1.5.5, we need the following:

Lemma 1.5.6 *Let K be a random set-valued variable, $\xi(0)$ a $\mathcal{F}_0$-adapted stochastic process.*

We define

$$\xi(t) = \xi(0) + \int_0^t f(\xi(s))ds + \int_0^t g(\xi(s))dW(s)$$

and we choose a $\mathcal{F}_0$- measurable projection $y \in \Pi_K(\xi(0))$.

Then, for any pair of $\mathcal{F}_0$-random variable (γ, v) in the stochastic contingent set $\mathcal{T}_K(0, y)$, the following estimate

$$\begin{cases} \liminf_{t_n \to 0} \left(\mathbf{E}(d_K^2(\xi(t_n)) - \mathbf{E}(d_K^2(\xi(0)))\right)/t_n \\ \leq 2\mathbf{E}\left(\langle \xi(0) - y, f(\xi(0)) - \gamma\rangle\right) + \mathbf{E}(\|g(\xi(0)) - v\|^2) \end{cases}$$

holds true.

Proof — Let us set $x = \xi(0)$, choose a projection $y \in \Pi_K(x)$ and take (γ, v) in the stochastic contingent set $\mathcal{T}_K(0, y)$. This means that there exist sequences $t_n > 0$ converging to 0 and $\mathcal{F}_{t_n}$-measurable a^n and b^n satisfying the assumptions (1.15) and

$$\forall\, n \geq 0, \text{ for almost all } \omega \in \Omega,\ y_\omega + v_\omega W_\omega(t_n) + \gamma_\omega t_n + t_n a^n_\omega + \sqrt{t_n} b^n_\omega \in K_\omega$$

Therefore

$$\begin{cases} d_K^2(\xi(t_n)) - d_K^2(\xi(0)) \ \leq \\ \left\| x + \int_0^{t_n} f(\xi(s))ds + \int_0^{t_n} g(\xi(s))dW(s) - y - vW(t_n) - \gamma t_n - t_n a^n - \sqrt{t_n} b^n \right\|^2 \\ - \|x-y\|^2 \\ = \left\| (x-y) + \int_0^{t_n} (f(\xi(s)) - \gamma)ds + \int_0^{t_n} (g(\xi(s)) - v)dW(s) - t_n a^n - \sqrt{t_n} b^n \right\|^2 \\ - \|x-y\|^2 \\ =: \ I \end{cases}$$

The latter term can be split in the following way:

$$\begin{cases} I \ = \ 2 < x-y, \int_0^{t_n} (g(\xi(s)) - v)dW(s) > & I_1 \\ + 2 < x-y, \int_0^{t_n} (f(\xi(s)) - \gamma)ds > & I_2 \\ + \| \int_0^{t_n} (g(\xi(s)) - v)dW(s)) \|^2 & I_3 \\ + \| \int_0^{t_n} (f(\xi(s)) - \gamma)ds \|^2 & I_4 \\ +2 \left\langle \int_0^{t_n} (g(\xi(s)) - v)dW(s), \int_0^{t_n} (f(\xi(s)) - \gamma)ds \right\rangle & I_5 \\ -2 \left\langle x - y + \int_0^{t_n} (f(\xi(s)) - \gamma)ds + \int_0^{t_n} (g(\xi(s)) - v)dW(s), t_n a^n \right\rangle & I_6 \\ -2 \left\langle x - y + \int_0^{t_n} (f(\xi(s)) - \gamma)ds + \int_0^{t_n} (g(\xi(s)) - v)dW(s), \sqrt{t_n} b^n \right\rangle & I_7 \\ + \|t_n a^n + \sqrt{t_n} b^n\|^2 & I_8 \end{cases}$$

We take the expectation in both sides of this inequality and estimate each term of the right hand-side. First, we observe that

$$\mathbf{E}\left(\left\langle x-y, \int_0^{t_n} (g(\xi(s)) - v)dW(s)\right\rangle\right) \ = \ 0$$

so that the expectation of the first term I_1 of the right-hand side of the above inequality vanishes.

The second term I_2 is estimated by $2t_n\alpha_n$ where

$$\alpha_n \ := \ \mathbf{E}\left(\left\langle x-y, \frac{1}{t_n}\int_0^{t_n} (f(\xi(s)) - \gamma)ds\right\rangle\right)$$

converges to

$$\alpha := \mathbf{E}\left(\langle x-y, f(\xi(0))-\gamma\rangle\right)$$

The third term I_3 is estimated by $t_n\beta_n$ where

$$\begin{cases} \beta_n := \frac{1}{t_n}\mathbf{E}\left(\| \int_0^{t_n}(g(\xi(s))-v)dW(s))\|^2\right) \\ = \frac{1}{t_n}\int_0^{t_n}\mathbf{E}\left(\|g(\xi(s))-v\|^2\right)ds \end{cases}$$

converges to

$$\beta := \mathbf{E}(\|g(\xi(0))-v\|^2)$$

because $\mathbf{E}\left(\left\|\int_0^t \varphi(s)dW(s)\right\|^2\right) = \int_0^t \mathbf{E}(\|\varphi(s)\|^2)ds$. The fourth term I_4 is easily estimated by $t_n\delta_n$ where

$$\begin{cases} \delta_n := \frac{1}{t_n}\mathbf{E}\left(\| \int_0^{t_n}(f(\xi(s))-\gamma)ds\|^2\right) \\ \leq \int_0^{t_n}\mathbf{E}\left(\|f(\xi(s))-\gamma\|^2\right)ds \leq ct_n \end{cases}$$

because $\mathbf{E}\left(\left\|\int_0^t \varphi(s)ds\right\|^2\right) \leq t\int_0^t \mathbf{E}(\|\varphi(s)\|^2)ds$.

By the Cauchy-Schwarz inequality, the term I_5 is estimated by $2t_n\eta_n$ where

$$\begin{cases} \eta_n := \frac{1}{t_n}\mathbf{E}\left(\left\langle \int_0^{t_n}(g(\xi(s))-v)dW(s), \int_0^{t_n}(f(\xi(s))-\gamma)ds\right\rangle\right) \\ \leq \frac{1}{t_n}\mathbf{E}(\| \int_0^{t_n}(g(\xi(s))-v)dW(s)\|^2)^{1/2}\mathbf{E}(\| \int_0^{t_n}(f(\xi(s))-\gamma)ds\|^2)^{1/2} \\ = \frac{1}{t_n}\left(\int_0^{t_n}\mathbf{E}(\|g(\xi(s))-v\|^2 ds)\right)^{1/2}(\mathbf{E}(\| \int_0^{t_n}(f(\xi(s))-\gamma)ds\|^2)^{1/2} \\ \leq \sqrt{t_n}\left(\frac{1}{t_n}\int_0^{t_n}\mathbf{E}(\|g(\xi(s))-v\|^2)\right)^{1/2}\left(t_n\int_0^{t_n}(\mathbf{E}(\|f(\xi(s))-\gamma\|^2)\right)^{1/2} \\ \leq ct_n^{\frac{1}{2}} \end{cases}$$

We now estimate the three latter terms involving the errors a^n and b^n. We begin with I_6. First,

$$\mathbf{E}\left(\langle x-y, a^n\rangle\right) \leq \mathbf{E}\left(\|x-y\|^2\right)^{\frac{1}{2}}\left(\mathbf{E}(\|a^n\|^2)\right)^{\frac{1}{2}}$$

which converges to 0 by assumption (1.15)i).

Now, the Cauchy-Schwarz inequality implies that

$$\begin{cases} \mathbf{E}\left(\left\langle \int_0^{t_n} (f(\xi(s)) - \gamma)ds, a^n \right\rangle\right) \\ \\ \leq \mathbf{E}\left(\left\| \int_0^{t_n} (f(\xi(s)) - \gamma)ds \right\|^2\right)^{\frac{1}{2}} \mathbf{E}\left(\|a^n\|^2\right)^{\frac{1}{2}} \end{cases}$$

Finally, the stochastic term is estimated in the following way:

$$\begin{cases} \mathbf{E}\left(\left\langle \int_0^{t_n} (g(\xi(s)) - v)dW(s), a^n \right\rangle\right) \\ \\ \leq \mathbf{E}\left(\left\| \int_0^{t_n} (g(\xi(s)) - v)dW(s) \right\|^2\right)^{\frac{1}{2}} \mathbf{E}\left(\|a^n\|^2\right)^{\frac{1}{2}} \\ \\ = \left(\int_0^{t_n} \mathbf{E}(\|g(\xi(s)) - v)\|^2 ds\right)^{\frac{1}{2}} \mathbf{E}\left(\|a^n\|^2\right)^{\frac{1}{2}} \end{cases}$$

which obviously converges to 0.

We continue with I_7. We have

$$\mathbf{E}\left\langle x - y, \frac{1}{\sqrt{t_n}} b^n \right\rangle = 0$$

since b^n is independent of $x - y$ and $\mathbf{E}(b^n) = 0$.

The Cauchy-Schwarz inequality implies that

$$\begin{cases} \mathbf{E}\left(\left\langle \int_0^{t_n} (f(\xi(s)) - \gamma)ds, \frac{1}{\sqrt{t_n}} b^n \right\rangle\right) \\ \\ \leq \mathbf{E}\left(\left\| \int_0^{t_n} (f(\xi(s)) - \gamma)ds \right\|^2\right)^{\frac{1}{2}} \mathbf{E}\left(\|\frac{1}{\sqrt{t_n}} b^n\|^2\right)^{\frac{1}{2}} \\ \\ = \frac{1}{\sqrt{t_n}}\left(t_n \int_0^{t_n} \mathbf{E}(\|(f(\xi(s)) - \gamma)\|^2) ds\right)^{\frac{1}{2}} \mathbf{E}\left(\|b^n\|^2\right)^{\frac{1}{2}} \end{cases}$$

Finally, the worst term of I_7 is estimated in the following way:

$$\begin{cases} \mathbf{E}\left(\left\langle \int_0^{t_n} (g(\xi(s)) - v)dW(s), \frac{1}{\sqrt{t_n}} b^n \right\rangle\right) \\ \\ \leq \frac{1}{\sqrt{t_n}} \sqrt{\mathbf{E}\left(\left\| \int_0^{t_n} (g(\xi(s)) - v)dW(s) \right\|^2\right)} \sqrt{\mathbf{E}\left(\|b^n\|^2\right)} \\ \\ = \left(\frac{1}{t_n} \int_0^{t_n} \mathbf{E}(\|g(\xi(s)) - v)\|^2 ds\right)^{\frac{1}{2}} \mathbf{E}\left(\|b^n\|^2\right)^{\frac{1}{2}} \end{cases}$$

which converges to 0 by assumption (1.15)ii).

It remains to estimate the last term of I_8. There is no difficulty because

$$\frac{1}{t_n}\mathbf{E}\left(\left\|t_n a^n + \sqrt{t_n} b^n\right\|^2\right) = \mathbf{E}\left(\left\|\sqrt{t_n} a^n + b^n\right\|^2\right)$$

converges to 0.

Putting everything together, we deduce the inequality of the lemma. □

Proof of Theorem 1.5.5 Since the solution to the stochastic differential equation can be written for any $h \geq 0$

$$\xi(t+h) = \xi(t) + \int_t^{t+h} f(\xi(s))ds + \int_t^{t+h} g(\xi(s))dW(s)$$

we deduce from Lemma 1.5.6 that

$$\begin{cases} \liminf_{h\to 0+} \left(\mathbf{E}(d_K^2(\xi(t+h)) - \mathbf{E}(d_K^2(\xi(t)))\right)/h \\ \leq 2\mathbf{E}\left(\langle \xi(t) - y(t), g(\xi(t)) - \gamma\rangle\right) + \mathbf{E}(\|g(\xi(t)) - v\|^2) \end{cases}$$

for any $\mathcal{F}_t$-measurable selection $y(t)$ of $\Pi_K(\xi(t))$ and any $(v(t), \gamma(t)) \in \mathcal{T}_K(t, y(t))$.

Since there exists a selection $y(t)$ of $\Pi_K(\xi(t))$ such that we can take $v(t) := g(\xi(t))$ and $\gamma(t) := f(\xi(t))$ by assumption, we infer that setting

$$\varphi(t) := \mathbf{E}\left(d_K^2(\xi(t))\right)$$

the contingent epiderivative

$$D_\uparrow\varphi(t)(1) := \liminf_{h\to 0+} \frac{\varphi(t+h) - \varphi(t)}{h}$$

is non positive.

This implies that $\varphi(t) \leq 0$ for all $t \in [0,T]$. If not, there would exist $T > 0$ such that $\varphi(T) > 0$. Since φ is continuous, there exists $\eta \in]0,T[$ such that

$$\forall\, t \in]T-\eta, T], \;\; \varphi(t) > 0$$

Let us introduce the subset

$$A := \{s \in [0,T] \mid \forall\, t \in]s,t], \;\; \varphi(t) > 0\}$$

and $t_0 := \inf A$.

We observe that for any $t \in]t_0, T], \varphi(t) > 0$ and that $\varphi(t_0) = 0$. Indeed, if $\varphi(t_0) > 0$, there would exist $t_1 \in]t_1, t_0[$ such that $\varphi(t) > 0$ for all $t \in]t_1, t_0]$, i.e., $t_1 \in A$, so that t_0 would no be an infimum.

Therefore, $D_\uparrow \varphi(t)(1) \leq 0$ for any $t \in]t_0, T]$ and thus, we obtain the contradiction

$$0 < \varphi(T) = \varphi(T) - \varphi(t_0) \leq 0$$

Consequently, for every $t \in [0, T]$, we have

$$\mathbf{E}(d_K^2(\xi(t))) = \int_\Omega d_{K_\omega}^2(\xi_\omega(t)) dP(\omega) = 0$$

Since the integrand is nonnegative, we infer that $d_{K_\omega}(\xi_\omega(t)) = 0$ almost surely, i.e., that the stochastic process ξ is viable in K. □

Proof of Theorem 1.5.3 The necessary condition following obviously from Theorem 1.5.4, it remains to prove that it is sufficient. For that purpose, we extend the maps f and g defined on K by the maps $\widetilde{f}$ and $\widetilde{g}$ defined on the whole space by

$$\widetilde{f}(x) := f(\pi_K(x)) \ \& \ \widetilde{g}(x) := g(\pi_K(x))$$

Then the pair $(\widetilde{f}, \widetilde{g})$ satisfies obviously condition

$$(\widetilde{f}(x), \widetilde{g}(x)) \in \mathcal{T}_K(t, \pi_K(x))$$

so that K is invariant by $(\widetilde{f}, \widetilde{g})$ thanks to Theorem 1.5.5. Since these maps do coincide on K, we infer that K is a viability domain of (f, g). □