Differential Kolmogorov Equations for Transiting Processes

by

Gaël Désilles

Ingénieur Diplômé de l'Ecole Polytechnique, France (1995)
Ingénieur Diplômé de l'Ecole Nationale Supérieure de l'Aéronautique et de l'Espace, France (1997)

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1998

Author .....
Department of Aeronautics and Astronautics
April 1998

Certified by .....
Eric Feron
Assistant Professor
Thesis Supervisor

Accepted by .....
Jaime Peraire
Chairman, Department Committee on Graduate Student

Differential Kolmogorov Equations for Transiting Processes

by

Gaël Désilles

Submitted to the Department of Aeronautics and Astronautics
on April 1998, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

This work develops a short theory on *transiting* stochastic processes, following standard guidelines in the theory of probabilities and stochastic processes. *Transiting* processes are processes submitted to Itô's stochastic differential equations, with drift and diffusion depending on a Poisson random process. We establish the integro-differential equations satisfied by the probability density function of such a process, and also the one satisfied by the first passage conditional probability for a stopped *transiting* process. As an application of this theoretical work, we show how existing probabilistic models of aircraft behaviour in free flight, involving mainly Poisson and Gaussian processes but also other probability distributions, lead to quite simple differential equations for the probability density function of an independent aircraft, and for the probability of collision of two aircraft. The equations are solved numerically, using finite difference or finite element methods. We also propose other methods based on Fourier analysis.

Thesis Supervisor: Eric Feron
Title: Assistant Professor

Acknowledgements

First of all, I am thankful to my advisor Eric Feron, who proposed me the exciting field of stochastic processes as applied to some security concerns in Aeronautics. As there was of course ups and downs in my work, he was keeping me on the right track, with a very supportive attitude. The lesson is, simplicity of ideas is always the trick.

I wish to thank Jaime Peraire, for he was the one who gave me this unforgettable opportunity to spend some time at MIT. My gratitude is the mirror of everything this year has brought to me.

I wish to sincerely thank Yashan Wang, my teacher in stochastic processes during Fall 1997. He made the technical and almost nowhere intuitive field of probability theory practical and intuitive, while mathematical rigor was always respected.

I wish to thank Alex Megretski, for the special spirit I tried to capture from his course during Spring 1997, not to mention the very deep discussions he is always ready to offer. His outlook on things was always to me the balance I needed to remain sighted.

I wish to thank the team of ICE Lab, and maybe especially Emilio Frazzoli, for their very nice welcome and the help they always provided when I needed it. I also thank them for not asking why I was *never* in the lab, but sometimes remotely...

I wish to thank my friends in MIT.

Contents

List of Figures

Introduction

One future perspective in sight in the Air Traffic Management is to release the hard constraints of the flight, letting the pilots make their own decisions [6, 1, 17], with the help of automated and autonomous systems aboard. The main issue here is to design alert systems to assist pilots and controllers resolve conflicts between aircraft.

One way to address the design of such systems is to imitate the behaviours of the pilots by statistical models and deduce from them predictions, such as speed, altitude, heading, etc. Among the statistics of first concern, the probability density function of an isolated airplane has shown authorizing an analytical calculation, or at least a semi-analytical one. On that statistical background, high safety standards require extreme accuracy in the computation of the probability. Two approaches have been proposed so far : one is to simplify the hypotheses of the problem, by modeling reality with Gaussian processes, on the purpose to achieve analytical solutions [1]; the other is to use Monte-Carlo simulation techniques to evaluate the probabillity by trajectory launches [17].

The analytical approach, though very efficient, does not take into account natural behaviours of the pilots, for instance taking unknown actions like quick turns or changes of altitude. These decisions are critical as a conflict arises. Continuing improvements of computation means allow to now consider the resolution of a more general model by well-known numerical techniques, similar to those used in Computational Fluid Dynamics (analogies between probability density functions and mass density of a fluid are profound). As computation power increases, the hope of solving the problem with high accuracy and efficiency becomes more and more concrete.

Poisson processes are very frequently used in the modeling of physical systems, for they provide a mathematical model of events with mean frequency of occurrence, and very few hypotheses are needed on a process to be of the Poisson type [16]. While arrival times of customers in a post office are often the academic example of such processes, Poisson processes can also model the decision times of pilots of an aircraft, in the context of aircraft behaviour modeling.

However, as soon as Poisson processes are involved in a random process, the Chapman-Kolmogorov differential equation satisfied by the probability density function of the problem can take a very awkward form, typically infinite in the number of terms, as shown in references [4]. Poisson processes are actually not the only processes to infere infinite Chapman-Kolmogorov differential equations.

Examples of such infinite differential equations show the catastrophic results of truncating Taylor's expansions at a large, but finite, rank. The solutions to the truncated equation may converge weakly to the solution of the exact equation, but the weak convergence is at the price of enormous oscillations of the approximated solution around the exact one, causing the approximated probability density function to take on negative values – among other problems.

There is therefore a need to avoiding such a global approach. Master equations, that transform the infinite expansion of the Chapman-Kolmogorov equations into an infinite-dimensional problem, are here shown to be a possible attack, as applied to a model of aircraft behaviour in free flight proposed earlier.

Chapter 1

Differential Kolmogorov equation for transiting Markov processes

The idea of the theory developed in the following lines is to find the evolution of the probability density function of a system at position $x(t)$ changing of dynamics at random times. The probability distribution of the instants of change is assumed (a decreasing) exponential, that is to say the random process of the times of change is a Poisson process. First we consider the case of a finite set of possible dynamics driving the modile between two consecutive instants of switch. Second we extend the formalism to the more general case of a continuous set of dynamics. Our point of view is fully probabilistic, in the sense that the dynamics itself is not assumed deterministic. Formally, we will identify the change of dynamics to a change of system, and with this convention, a change of system will be called a jump. The best mental representation of the behaviour of such a mobile (or, more technically, a stochastic process) is a moving body with a trajectory depending on a parameter (for instance : the heading, the speed, the curvature, any geometrical characteristics of the trajectory...), such that the parameter changes of value at "Poisson times". The mathematical background needed is exposed in the next section.

1.1 Position of the problem, Notations

Let us consider n systems $\Sigma_1, \ldots, \Sigma_n$, with n a finite integer. We assume that each of them can be completely defined by a vector-valued Markov process $x_j(t) \in \mathbb{R}^d$, thus driven by a stochastic differential equation. We assume to know for each Σ_j the transition probability $p^j(x, t \,|\, y, s)$, defined as the probability that the state $x_j(t)$ is at the position x at the instant t, given that $x_j(s) = y$ (with $s < t$) :

$$p^j(x, t \,|\, y, s) = \mathsf{P}\{\, x(t) = x \text{ in } \Sigma_j \,|\, x(s) = y \text{ in } \Sigma_j \,\}$$

The process driving Σ_j is not assumed stationary. We admit that the transition probability $p^j(x,t\,|\,y,s)$ can be developed with respect to the time delay $t-s$ as :

$$p^j(x,t\,|\,y,s) = \delta(x-y) + p_t^j(x\,|\,y;s)\,(t-s) + o(|t-s|) \tag{1.1}$$

where the term $p_t^j(x\,|y;s)$ may be formally identified to

$$p_t^j(x\,|\,y;s) = \lim_{ds\to 0,\, ds>0} \frac{1}{ds}\,(p^j(x,s+ds\,|\,y,s) - \delta(x-y)\,)$$

The expansion (1.1) can be held each time the system Σ_j has continuous sample paths. It is no evidence that $p_t^j(x|y;s)$ and $p^j(x,t\,|\,y,s)$ should be *functions* of their arguments x and y. It even simplifies a lot further analytical developments to suppose that they are distributions. Hence, we will use the formalism of bilinear brackets $\langle .\,,.\rangle$ in the following. We assume all distributions to be tempered, in the sense given by Laurent Schwarz [2]. To find the expansion (1.1), one may prefer to expand the Fourier transform of $p^j(x,t\,|\,y,s)$ (with respect to the space variable x) as a Taylor series in $t-s$, and then invert the expansion from the Fourier domain. In some major references [4], the Taylor expansion (1.1) is split into three separate statements, involving functions instead of distributions.

We now suppose that the system $x(t)$ randomly jumps from one given subsystem Σ_i to another subsystem Σ_j with probability $\mathsf{P}\{\,\Sigma_i \to \Sigma_j\,\} = \lambda_{ij}(t)\,dt$ during dt (with $\lambda_{ij}(t) \geq 0$ at any instant t. Between two consecutive jumps, $x(t)$ is thus driven by a unique subsystem *at any time.* In mathematical terms, we define :

$$\lambda_{ij}(t) = \lim_{dt\to 0} \frac{1}{dt}\,\mathsf{P}\{\,x(t+dt) \in \Sigma_j \mid x(t) \in \Sigma_i\,\}$$

It is the *Poisson hypothesis* that allows the limit to exist. There is no requirement on the functions $\lambda_{ij}(t)$ to be symmetric in the indices i and j. See Figure 1-1 for a graphic representation of the process, in terms of flows.

We want to find a partial differential equation satisfied by $W(x,t)$, the probability density function of the process $x(t)$, or the probability to find $x(t)$ at the position x and instant t :

$$W(x,t) = d\mathsf{P}\{\,x(t) = x\,\}$$

We will also precise properties of the $\lambda_{ij}(t)$, and give canonical examples of the formalism proposed.

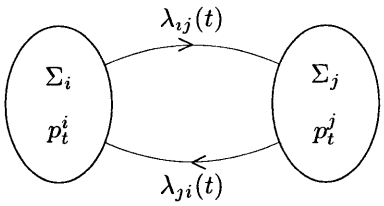

Figure 1-1: Flow chart of Poisson transfers between subsystem i and subsystem j.

1.2 Practical treatment

1.2.1 Expansion of the conditional density functions

Let $W^j(x,t)$ be the conditional probability density function defined by

$$W^j(x,t) = d\mathsf{P}\{x(t) = x \,|\, x(t) \in \Sigma_j\}$$

Using a Chapman-Kolmogorov equation, we may write

$$W^j(x,t+dt) = \left(1 - \sum_{i\neq j} \lambda_{ji}(t)\,dt\right) \langle\, p^j(x,t+dt\,|\,.\,t)\,,\, W^j(.\,,t)\,\rangle + dt \sum_{i\neq j} \lambda_{ij}(t)\,\langle\, p^i(x,t+dt\,|\,.\,,t)\,,\, W^i(.\,,t)\,\rangle \quad (1.2)$$

This relation can be easily understood if interpreted in terms of flows of particles; in that context indeed, $W^j(x,t)$ may be seen as the fraction of the total number of particles in system Σ_j that are at position x at t. Between t and $t+dt$, some particles of Σ_k, with mass proportional to $W^k(y,t)$ jump from any far y to the actual position x, and then get in Σ_j, whatever k be (j itself or $i \neq j$). The first event occurs with probability $p^k(x,t+dt\,|\,y,t)$ by definition, whereas the second one has probability $\lambda_{kj}(t)\,dt$.

Plugging the Taylor's expansion (1.1) for each $p^k(x,t+dt\,|\,.\,,t)$, and observing that by definition $\langle\, \delta(x-.)\,,\, W^k(.\,,t)\,\rangle = W^k(x,t)$, we get

$$W^j(x,t+dt) = W^j(x,t) + dt\left[\,\langle\, p_t^j(x|\,.\,;t)\,,\, W^j(.\,,t)\,\rangle - \sum_{i\neq j} \lambda_{ji}(t)\,W^j(x,t) + \sum_{i\neq j} \lambda_{ij}(t)\,W^i(x,t)\right] + o(dt)$$

Hence, by letting dt tend to zero, $W^j(x,t)$ must satisfy

$$\frac{\partial W^j}{\partial t}(x,t) = \langle\, p_t^j(x\,|\,.\,,t)\,,\, W^j(.\,,t)\rangle + \sum_{i\neq j} \left(\,\lambda_{ij}(t)\,W^i(x,t) - \lambda_{ji}(t)\,W^j(x,t)\,\right) \quad (1.3)$$

We shall comment on the formalism used in the first term later.

1.2.2 Evolution of the masses

By a straight-forward conditional expansion, the probability density function we are searching for, namely $W(x,t)$, reads :

$$W(x,t) = \sum_j W^j(x,t)\,\mathsf{P}\{x(t) \in \Sigma_j\}$$

where $\mathsf{P}\{x(t) \in \Sigma_j\} = \mu_j(t)$ may be called the *mass* of the j-th system. The requirement $\sum_j \mu_j(t) = 1$ follows from the definitions of these masses (probability 1 to be in any of the systems).

Let us consider the random variable $J(t)$ defined as the index of the system where the state $x(t)$ actually finds itself :

$$J(t) = j \;:\; x(t) \in \Sigma_j$$

The random index $J(t)$ is a well-known death-birth process [4, 14], with (infinite) transition matrix $\mathsf{P}(i \to j) = \lambda_{ij}(t)$. One can easily see that $\mu_j(t) = \mathsf{P}\{\, J(t) = j \,\}$. Therefore, the μ_j's have to satisfy

$$\frac{d\mu_j}{dt} = \sum_{i \neq j} \left(\lambda_{ij}(t)\,\mu_i(t) - \lambda_{ji}(t)\,\mu_j(t)\right) \tag{1.4}$$

that may also be seen as a law of mass conservation.

The system of ordinary differential equations satisfied by the masses μ_j is thus linear, and it is clear that, when the transfers are symmetric ($\lambda_{ij} = \lambda_{ji}\,, \forall\, i,j$), the equally weighed system of masses $\mu_j(t) = \frac{1}{n}, j = 1...n$ is a particular solution. The general case, however, is completely determined by the initial distribution of masses $\{\mu_j^0\}_{j=1...n}$, and is a uniform distribution at each instant t under the necessary and sufficient condition that the transfers be symmetric and the initial distribution uniform ($\mu_j^0 = \frac{1}{n}$ for all j).

1.3 Extension to a continuum of systems

Here, we extend the formalism shown to the case of a nondenumerable number of systems, that we will denote Σ_α further on, with α belonging to an arbitrary measurable set A. For instance, the systems Σ_α of concern could be a generic system parametrized by some real number α.

1.3.1 Writing the equations in the continuous parameter

Replacing all integer superscripts j and i by continuous superscripts α and β, we define again :

$$W^\alpha(x,t) = d\mathsf{P}\{\, x(t) = x \,|\, x(t) \in \Sigma_\alpha \,\}$$

and

$$\lambda_{\alpha\beta}(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} d\mathsf{P}\{\, x(t+\Delta t) \in \Sigma_\beta \,|\, x(t) \in \Sigma_\alpha \,\}$$

The same developments with continuous indices provide the equation :

$$\frac{\partial W^\alpha}{\partial t}(x,t) = \langle\, p_t^\alpha(x \,|\, .\,,t)\,,\, W^\alpha(\,.\,,t)\rangle \;+\; \int_A d\beta \,\Big(\, \lambda_{\beta\alpha}(t)\, W^\beta(x,t) - \lambda_{\alpha\beta}(t)\, W^\alpha(x,t)\,\Big) \tag{1.5}$$

The probability density function now reads

$$W(x,t) = \int_A d\alpha\; W^\alpha(x,t)\,\mu_\alpha(t)$$

with $\mu_\alpha(t) = \mathsf{P}\{x(t) \in \Sigma_\alpha\}$. The mass $\mu_\alpha(t)$ is again ruled by a law of mass conservation, now reading

$$\frac{d\mu_\alpha}{dt} = \int_A d\beta \,\Big(\, \lambda_{\beta\alpha}(t)\,\mu_\beta(t) - \;\lambda_{\alpha\beta}(t)\,\mu_\alpha(t)\,\Big) \tag{1.6}$$

with

$$\int_A d\beta\;\mu_\beta(t) = 1 \quad \text{and} \quad \forall\,\beta,\; \mu_\beta(t) \geq 0$$

The rule that the total mass $\int d\beta\;\mu_\beta(t)$ be 1 is consistent with any nonnegative solution to (1.6). Indeed, by simple permutation of indices α and β,

$$\int_A d\alpha\,\mu_\alpha(t) \int_A d\beta\,\lambda_{\alpha\beta}(t) = \int_A d\beta\,\mu_\beta(t) \int_A d\alpha\,\lambda_{\beta\alpha}(t)$$

Therefore the derivative of $\int_A d\alpha\,\mu_\alpha(t)$ is zero as we may derive from the integration of (1.6) with respect to α.

We can prove (1.6) in a very simple way : by the definition

$$\begin{aligned}
\mu_\alpha(t+\Delta t) &= \mathsf{P}\{x(t+\Delta t) \in \Sigma_\alpha\} \\
&= \int d\beta\,\mathsf{P}\{x(t+\Delta t) \in \Sigma_\alpha \,|\, x(t) \in \Sigma_\beta\}\,\mathsf{P}\{x(t) \in \Sigma_\beta\} \\
&= \int d\beta\,\mathsf{P}\{x(t+\Delta t) \in \Sigma_\alpha \,|\, x(t) \in \Sigma_\beta\}\,\mu_\beta(t)
\end{aligned}$$

But the conditional probability factoring μ_β is, by the Poisson hypothesis, of first order in Δt with

$$\mathsf{P}\{x(t+\Delta t) \in \Sigma_\alpha \,|\, x(t) \in \Sigma_\beta\} = \lambda_{\beta\alpha}(t)\,\Delta t \;+\; o(\Delta t)$$

Hence :

$$\mu_\alpha(t+\Delta t) = \Delta t \int d\beta\,\lambda_{\beta\alpha}(t)\,\mu_\beta(t) \;+\; o(\Delta t)$$

Regardless of the state α at time t, x is in at least one state β at time t' [1], therefore :

$$\int d\beta \, \mathsf{P}\{x(t') \in \Sigma_\beta \,|\, x(t) \in \Sigma_\alpha\} = 1$$

Then by multiplication of μ_α by 1, we have

$$\mu_\alpha(t) = \int d\beta \, \mathsf{P}\{x(t') \in \Sigma_\beta \,|\, x(t) \in \Sigma_\alpha\} \, \mu_\alpha(t)$$

With $t' = t + \Delta t$, substracting the equalities and dividing by Δt yields :

$$\frac{1}{\Delta t}\big(\mu_\alpha(t+\Delta t) - \mu_\alpha(t)\big) = \int d\beta \, \big(\lambda_{\beta\alpha}(t) \, \mu_\beta(t) - \lambda_{\alpha\beta}(t) \, \mu_\alpha(t) \big) \; + \; o(1)$$

clearly leading to the differential equation (1.6).

Now, we can check the relevance of the formalism regarding the intuitive significance of the probability density function $W(x,t)$. By the definition of the conditional probability, we have indeed :

$$\begin{aligned} \frac{d\mathsf{P}\{\, x(t) = x \;\; \textbf{and} \;\; x(t) \in \Sigma_\alpha \,\}}{d\mathsf{P}\{\, x(t) \in \Sigma_\alpha \,\}} &= d\mathsf{P}\{\, x(t) = x \,|\, x(t) \in \Sigma_\alpha \,\} \\ &= W^\alpha(x,t) \end{aligned} \tag{1.7}$$

Therefore

$$\begin{aligned} \int_{\alpha \in A} \mu_\alpha(t) \, W^\alpha(x,t) \; &= \; \int_{\alpha \in A} d\mathsf{P}\{\, x(t) = x \;\; \textbf{and} \;\; x(t) \in \Sigma_\alpha \,\} \\ &= \; d\mathsf{P}\{\, x(t) = x \,\} \\ &= \; W(x,t) \end{aligned} \tag{1.8}$$

as expected.

1.3.2 Redundancy of the definitions

It is also straight forward to see that, instead of solving two systems (1.5) and (1.6), we can replace the conditional probability density function $W^\alpha(x,t)$ by

$$\widetilde{W}^\alpha(x,t) = \mu_\alpha(t) \, W^\alpha(x,t) = d\mathsf{P}\{\, x(t) = x \;\; \textbf{and} \;\; x(t) \in \Sigma_\alpha \,\} \tag{1.9}$$

This new probability satisfies the same equation (1.5) as $W^\alpha(x,t)$. Let us consider indeed the change of unknown function $\widetilde{W}^\alpha(x,t) = \nu_\alpha(t) \, W^\alpha(x,t)$, where $\nu_\alpha(t)$ is still to be determined. By integrating (1.5) with respect to the space variable x, since $\int dx \, W^\alpha(x,t) = 1$ for each time t, the function $\nu_\alpha(t)$

[1] A better image of the fact is that the state $x(t)$ is spread over all possible states β, with given probabilities.

must verify :

$$\frac{d\nu_\alpha}{dt} = \int_A d\beta \left(\lambda_{\beta\alpha}(t)\,\nu_\beta(t) - \lambda_{\alpha\beta}(t)\,\nu_\alpha(t) \right)$$

If we impose the initial condition : $\nu_\alpha(0) = \mu_\alpha(0)$, the two functions of time ν_α and μ_α are solutions of the same system of integro-differential equations with the same initial conditions, therefore[2] they are equal. This proves the relation (1.9) between $\widetilde{W}^\alpha$ and W^α. After the change of probability (1.9), the total probability density function reads simply :

$$W(x,t) = \int_{\alpha\in A} \widetilde{W}^\alpha(x,t)$$

The integration of (1.5) with respect to the space variable is a mnemonic technique to retrieve the evolution of the masses. As it is intuitive, the integration of the joint probability defined in (1.9) with respect to x gives the probability of being in system Σ^α. Notice that the value 1 of the integral $\int dx\, W^\alpha(x,t)$ (or μ_α) is precisely an equilibrium of equation (1.6). Losing this equilibrium is the trick to solve two systems of equations (1.5) and (1.6) at the same time. From now on, we abandon the heavy tilded notation, and suppose by convention that $W^\alpha(x,t_0)$ bears the new initial condition (former, multiplied by $\mu_\alpha(t_0)$).

1.4 Discussion

We shall now make a few comments on the equations (1.5), and the bracketed notation in direct or reverse order used in the previous paragraphs. The first remark goes to the structure of the Kolmogorov equations (1.5), seen as balances. Then the role played by the distribution p_t^α as a differential operator is pointed out. The adjoint of p_t^α is also fully specified, with the so-called *backward* version of the equations (1.5) in mind. Last but not least, the notation $\int$ used with the transition rates $\lambda_{\alpha\beta}$ is precised, since it is non-standard.

Similar comments could be formulated for the discrete case. We will discuss within the continuous context however, and this each time that a double viewpoint has no further significance than a purely formal analogy.

1.4.1 Structure of the Kolmogorov equations

The structure of the system is to be decomposed into two categories of terms : in one hand, the terms $\langle p_t^\alpha(x\,|\,.\,,t)\,,\, W^\alpha(\,.\,,t)\rangle$, describing the evolution of the probability density functions inside of each system Σ_α, and in the other hand, terms describing incoming and outcoming flows, namely $\int d\beta\, \lambda_{\beta\alpha}(t)\, W^\beta(x,t) - \lambda_{\alpha\beta}(t)\, W^\alpha(x,t)$. It is remarkable to observe that, in absence of exchange

[2] under appropriate conditions of smoothness of the transition rates $\lambda_{\alpha\beta}$

rates $\lambda_{\alpha\beta}(t)$, the only remaining term is $\langle\, p_t^\alpha(x\,|\,.\,,t)\,,\, W^\alpha(\,.\,,t)\rangle$. And this situation corresponds to a "decoupled" system Σ_α (no exchange with other systems). The equation for $W^\alpha(x,t)$ then reduces to

$$\frac{\partial W^\alpha}{\partial t}(x,t) = \langle\, p_t^\alpha(x\,|\,.\,,t)\,,\, W^\alpha(\,.\,,t)\rangle$$

But as it is extensively shown in references [4, 10, 14, 15], the probability density function satisfies a forward Fokker-Planck equation in this case :

$$\frac{\partial W^\alpha}{\partial t} = -\,\mathrm{div}\,\big(\,a^\alpha(x,t)W^\alpha(x,t)\big) \;+\; \frac{1}{2}\,\mathrm{div}\,\big(\,b^\alpha(x,t)\,\mathrm{grad}\,W^\alpha(x,t)\,\big)$$

where $a^\alpha(x,t)$ and $b^\alpha(x,t)$ have appropriate dimensions (a^α is a vector, b^α is a matrix). This clearly identifies $p_t^\alpha(x\,|\,.\,,t)$ to

$$\langle\, p_t^\alpha(x\,|\,.\,,t)\,,\, V(.)\,\rangle \equiv -\,\mathrm{div}\,\big(\,a^\alpha(x,t)V(x)\big) \;+\; \frac{1}{2}\,\mathrm{div}\,\big(\,b^\alpha(x,t)\,\mathrm{grad}\,V(x)\big) \tag{1.10}$$

Hence, the system of equations can be readily constructed, by adding and substracting proper transfer terms $\lambda_{\alpha\beta}(t)\,W^\alpha(x,t)$ to the differential equation describing a closed system Σ_α. With a mass balance requirement in mind, one can write :

$$\frac{\partial W^\alpha}{\partial t}(x,t) = \text{Fokker-Planck of closed } \Sigma_\alpha \;+\; \text{Incoming flows} \;-\; \text{Outcoming flows}$$

This practical method does not need the expansion of $p^\alpha(x,t\,|\,y,s)$. Well-known probabilistic systems with continuous sample paths often present themselves by their Fokker-Planck equation, which is exactly the first term needed on the right-hand side.

1.4.2 $p_t^\alpha(x\,|\,y;t)$ as a differential operator

The definition of the transition probability $p^\alpha(x,t\,|\,y,s)$ is such that

$$\forall t > s, \quad \int dx\, p^\alpha(x,t\,|\,y,s) \;=\; 1 \tag{1.11}$$

as a result of the conditioning. Technically, we have indeed

$$p^\alpha(x,t\,|\,y,s) \;=\; \frac{p^\alpha(x,t \;\textbf{and}\; y,s)}{\mathsf{P}\{y,s\}}$$

where the denominator is sometimes viewed as a normalization factor [16].

For we have expanded the probability $p^\alpha(x,t\,|\,y,s)$ as :

$$p^\alpha(x,t\,|\,y,s) = \delta(x-y) + p_t^\alpha(x\,|\,y;s)\,(t-s) + o(|t-s|)$$

the equality (1.11) implies :

$$\forall t, \quad \int dx\, p_t^\alpha(x \,|\, y; t) \;=\; 0$$

or more properly :

$$\forall t, \quad \langle\, p_t^\alpha(\,.\,|\, y; t)\,,\, \chi_{\mathbb{R}^d}(.)\,\rangle \;=\; 0 \tag{1.12}$$

where $\chi_{\mathbb{R}^d}$ is the characteristic function of the set $\mathbb{R}^d$.

This equality points out the nature of *differential operator* of the distribution $p_t^\alpha(\,.\,|\, y; t)$. The reader may want to keep in mind that the operation of the distribution on a smooth function W is a linear combination of first and upper order derivatives of W. The use of the bracket instead of a more detailed notation in terms of derivatives is only to allow more general properties, in the case of non-standard processes.

1.4.3 Adjoint of the operator $p_t^\alpha(x \,|\, y; t)$

Let $U(x)$ and $V(x)$ be two twice continuously differentiable functions over $\mathbb{R}^d$. According to (1.10), the usual inner product between $U(x)$ and $\langle\, p_t^\alpha(x \,|\, .; t)\,,\, V(.)\,\rangle$ is given by :

$$\langle\, U(x) \,|\, p_t^\alpha(x \,|\, .; t)\,,\, V(.)\,\rangle = \int_{\mathbb{R}^d} dx\, U(x) \left[\,-\, \mathrm{div}\left(\, a^\alpha(x,t) V(x)\right) \;+\; \frac{1}{2}\, \mathrm{div}\left(\, b^\alpha(x,t)\, \mathrm{grad}\, V(x)\right) \right]$$

By Stokes' formula, and assuming that $U(x)$ and $V(x)$ and their derivatives vanish as $|x| \to \infty$, we can write

$$\begin{aligned} \int_{\mathbb{R}^d} dx\, U(x)\, \mathrm{div}\left(\, a^\alpha(x,t) V(x)\right) &= -\int_{\mathbb{R}^d} dx\, V(x) \sum_i a_i^\alpha(x,t)\, \partial_i\, U(x) \\ \int_{\mathbb{R}^d} dx\, U(x)\, \mathrm{div}\left(\, b^\alpha(x,t)\, \mathrm{grad}\, V(x)\right) &= \int_{\mathbb{R}^d} V(x) \sum_{ij} b_{ij}^\alpha(x,t)\, \partial_{ij}\, U(x) \end{aligned} \tag{1.13}$$

By the definition of the adjoint $p_t^\alpha(x \,|\, y; t)^*$ of $p_t^\alpha(x \,|\, y; t)$, we then have

$$\langle\, p_t^\alpha(.\,|\, y; t)^*\,,\, U(.) \,|\, V(y)\,\rangle = \int_{\mathbb{R}^d} dx\, V(x) \left[\sum_i a_i^\alpha(x,t)\, \partial_i\, U(x) + \frac{1}{2} \sum_{ij} b_{ij}^\alpha(x,t)\, \partial_{ij}\, U(x) \right]$$

identifying the adjoint as the operator :

$$\langle\, p_t^\alpha(.\,|\, y; t)^*\,,\, U(.)\,\rangle = \sum_i a_i^\alpha(x,t)\, \partial_i\, U(x) + \frac{1}{2} \sum_{ij} b_{ij}^\alpha(x,t)\, \partial_{ij}\, U(x) \tag{1.14}$$

As a convention on notation, we will write $\langle\, U(.)\,,\, p_t^\alpha(.\,|\, y; t)\rangle$ instead of $\langle\, p_t^\alpha(.\,|\, y; t)^*\,,\, U(.)\,\rangle$ [3].

[3] This is similar to the "ket-bra" notation used in Quantum Mechanics

1.4.4 Integrability of the transition rates

The notation $\int d\beta\, \lambda_{\alpha\beta}$ in a common sense is proving irrelevant, since the transition rates $\lambda_{\alpha\beta}$ cannot be summable. A definition of the integral as a principal value is necessary. The case of the discrete jumps clarifies the origin of such an ill-integrability.

Indeed, regularly integrating the equality

$$\lambda_{\alpha\beta}(t) = \frac{1}{\Delta t}\, d\mathsf{P}\{\, x(t+\Delta t) \in \Sigma_\beta \,|\, x(t) \in \Sigma_\alpha \,\} \; + o(1)$$

with respect to β would lead to

$$\int d\beta\, \lambda_{\alpha\beta}(t) = \frac{1}{\Delta t} \; + o(1)$$

This equality clearly denies the convergence of the integral $\int d\beta\, \lambda_{\alpha\beta}(t)$, since the left-hand side is independent of Δt and the right-hand side diverges as $\Delta t \to 0$. The discrete case gives a more intuitive explanation of this phenomenon : writing, as in the continuous case :

$$\sum_i \lambda_{ij}(t)\, \Delta t \; = 1 + o(\Delta t)$$

where the discrete sum includes the index $i = j$, after dividing by Δt the equality becomes *nonsense* as $\Delta t \to 0$. But in the partial differential equation found (1.3), the coefficient λ_{jj} does not appear, and is therefore useless to define. By the definition of the transition rates, we must have

$$\sum_{i \neq j} \lambda_{ij}(t)\, \Delta t \leq 1$$

(the probability to jump into *another* state during Δt small is less than one). This is certainly true as $\Delta t \to 0$, if no $\lambda_{ij}(t)$ is infinite. The use of a non-defined $\lambda_{jj}(t)$ only provides a convenient start, thinking of flows when writing equation (1.2) : the factor $1 - \sum_{i \neq j} \lambda_{ij}(t)\, \Delta t$ actually represents $\lambda_{jj}(t)\, \Delta t$. However, one may want to skip this half-rigorous step and write the relation (1.3) directly, where indeed no $\lambda_{jj}(t)$ is needed.

As limits of discrete Riemann's sums, the integrals over the parameter β shall exclude the current value α in the continuous Kolmogorov differential equation – as are defined *principal value integrals* [2, 4].

1.4.5 Simultaneity of jumps and travels

One may address a legitimate question about the particle point of view. During Δt, we said that a given particle *first* travels from position y to position x, *and then* possibly jumps from α-th system to β-th system. The question indeed arises whether the order travel-jump imports. The system of partial differential equations shows, in fact, that the behaviour of the system is independent of the

order travel/jump. The reason is that we assumed continuous sample paths inside of each system. As Δt tends to zero, the distance traveled in a system before or after the jump goes to zero; it is the *way* this distance goes to zero that influences the probabilities, not the *fact* it goes to zero.

1.5 Technical proof of the Kolmogorov equation

Here we develop a simple but technical proof of the relations (1.5). This proof follows well-known references (for instance [4]), though extended to the case of a transiting process without discontinuity of the trajectories. The first step is to derive the equations (1.5), often called of the *forward* type, from a mathematical synthesis of the hypotheses of the problem. The equations (1.5) have an equivalent *backward* form, that we developp as well. The *backward* version is then used to find the equations satisfied by a probability of first passage, a very important statistics related to the problem of collision or first encounter of two objects. The notions are detailed in the corresponding paragraphs.

These technicalities complete the formalism proposed so far, as specific extensions of a standard theoretical *corpus* known as Itô's calculus [10, 15].

1.5.1 Forward expansion

We start with assuming the following first order *forward* expansion of the conditional probability with respect to the time delay Δt :

$$\begin{aligned} W(x,\alpha,t+\Delta t\,|\,z,\gamma,t) \;=\; & \delta(\alpha-\gamma)\,\delta(x-z) \;+\; \delta(\alpha-\gamma)\,p_t^{\alpha}(x\,|\,z\,,t)\,\Delta t \\ & +\; \delta(x-z)\,\lambda_{\gamma\alpha}(t)\,\Delta t \;-\; \delta(\alpha-\gamma)\,\delta(x-z)\int d\eta\,\lambda_{\alpha\eta}(t)\,\Delta t \;+\; o(\Delta t) \end{aligned} \tag{1.15}$$

where the subscripts x and α give the points the Dirac's δ-distributions are centered at. This expansion is the mathematical translation of the hypotheses of the problem, and is thereby of axiomatic nature. The integral in the last term is a principal value centered at α.

Using the fundamental Chapman-Kolmogorov relation

$$W(x,\alpha,t+\Delta t\,|\,y,\beta,s) = \int dz\,d\gamma\,W(x,\alpha,t+\Delta t\,|\,z,\gamma,t)\,W(z,\gamma,t\,|\,y,\beta,s)$$

characterizing Markov processes, we can write :

$$\begin{aligned} W(x,\alpha,t+\Delta t\,|\,y,\beta,s) = W(x,\alpha,t\,|\,y,\beta,s) \\ + \Delta t \int dz\, p_t^\alpha(x\,|\,z\,,t)\, W(z,\alpha,t\,|\,y,\beta,s) \\ + \Delta t \int d\gamma\, \lambda_{\gamma\alpha}(t)\, W(x,\gamma,t\,|\,y,\beta,s) \\ - \Delta t \int d\eta\, \lambda_{\alpha\eta}(t)\, W(x,\alpha,t\,|\,y,\beta,s) \;+\; o(\Delta t) \end{aligned} \tag{1.16}$$

The integrals are principal values, including those of the Chapman-Kolmogorov relation. The assumed summability of the probabilities ensures that principal values are equivalent to simple integrals. Again, replacing the mute index η by γ, dividing by Δt and taking the limit $\Delta t \to 0$, we get, with the bracketed notation :

$$\begin{aligned} \partial_t W(x,\alpha,t\,|\,y,\beta,s) = \langle\, p_t^\alpha(x\,|\,.\,,t)\,,\, W(.\,,\alpha,t\,|\,y,\beta,s)\,\rangle \\ + \int d\gamma\, \big[\lambda_{\gamma\alpha}(t)\, W(x,\gamma,t\,|\,y,\beta,s) \;-\; \lambda_{\alpha\gamma}(t)\, W(x,\alpha,t\,|\,y,\beta,s)\big] \end{aligned} \tag{1.17}$$

This equation is often referred to as the *forward* Kolmogorov differential equation. The initial condition follows from the *forward* expansion (1.15) :

$$W(x,\alpha,t\,|\,z,\gamma,t) \;=\; \delta(\alpha-\gamma)\,\delta(x-z)$$

The equation (1.5) is obtained by integrating (1.17) with respect to y and β, after postmultiplication by $W(y,\beta,s)$. The reader has observed that the derivation of the equation is a consequence of the expansion (1.15), which concentrates all the hypotheses we have made in the first section : in particular, continuous sample paths are responsible for Dirac's distributions δ. The expansion is a mathematical translation of the problem, and suggests how we can modify them to obtain other properties (for instance, lose the continuity requirement by replacing the δ's with other distributions with mass 1).

1.5.2 Backward expansion

Finding $\partial_s W(x,\alpha,t\,|\,y,\beta,s)$ is the purpose of the *backward* expansion of the conditional probabilities. Provided that $s < t$, the *backward* equations thus describe the evolution of the probabilities as the time goes backward (up to a simple change of sign in the convective terms).

This is again by the mean of the fundamental Kolmogorov equation that we can write for $\Delta s > 0$ and $s + \Delta s < t$:

$$W(x,\alpha,t\,|\,y,\beta,s) = \int dz\, d\gamma\, W(x,\alpha,t\,|\,z,\gamma,s+\Delta s)\, W(z,\gamma,s+\Delta s\,|\,y,\beta,s)$$

The *forward* expansion can then be applied to the second term, while the first term is assumed to possess a Taylor's series :

$$W(x,\alpha,t\,|\,z,\gamma,s+\Delta s) \;=\; W(x,\alpha,t\,|\,z,\gamma,s) \;+\; \Delta s\,\partial_s W(x,\alpha,t\,|\,z,\gamma,s) \;+\; o(\Delta s)$$

Plugging the expression for the *forward* expansion and the Taylor's series in the Kolmogorov relation provide, after rearrangements,

$$\begin{aligned} W(x,\alpha,t\,|\,y,\beta,s) \;=\; & W(x,\alpha,t\,|\,y,\beta,s) \\ & + \Delta s\,\partial_s W(x,\alpha,t\,|\,y,\beta,s) + \Delta s \int dz\, W(x,\alpha,t\,|\,z,\beta,s)\, p_t^{\beta}(z\,|\,y,s) \\ & + \int d\gamma\, \lambda_{\beta\gamma}(s)\left[W(x,\alpha,t\,|\,y,\gamma,s) - W(x,\alpha,t\,|\,y,\beta,s)\right] \;+\; o(\Delta s) \end{aligned}$$

Simplifying $W(x,\alpha,t\,|\,y,\beta,s)$ on both sides and dividing by Δs gives the *backward* differential Kolmogorov equation :

$$\begin{aligned} \partial_s W(x,\alpha,t\,|\,y,\beta,s) \;=\; & -\langle W(x,\alpha,t\,|\,.\,,\beta,s)\,,\, p_t^{\beta}(\,.\,|\,y,s)\rangle \\ & - \int d\gamma\, \lambda_{\beta\gamma}(s)\left[W(x,\alpha,t\,|\,y,\gamma,s) - W(x,\alpha,t\,|\,y,\beta,s)\right] \end{aligned} \tag{1.18}$$

One can also show that the *backward* expansion implies the *forward* expansion. Hence both expansions are equivalent. As a parallel with the first order series (1.15), it is easy to see that the *backward* expansion reads :

$$\begin{aligned} W(x,\alpha,s\,|\,z,\gamma,s+\Delta s) \;=\; & \delta(\gamma-\alpha)\,\delta(z-x) \;-\; \delta(\gamma-\alpha)\,p_t^{\gamma}(\,x\,|\,z,s)\,\Delta s \\ & - \delta(z-x)\,\big(1-\delta(\gamma-\alpha)\big)\lambda_{\beta\gamma}(s)\,\Delta s \;+\; o(\Delta s) \end{aligned} \tag{1.19}$$

1.5.3 Probability of first passage

The purpose is now to find the equations of evolution of the (conditional version) of a probability of first passage, *ie* the probability that a process with Poisson transitions crosses a fixed boundary for the first time (also known as first encounter). To this purpose, it is necessary to define a *stopped process* [15], which intuitively is the process itself up to the time τ of the first encounter, and remains stopped at the position of the encounter after τ. Like the process, the time τ is random. The equations are obtained with the use of the *backward* expansion. Some restrictive hypotheses are necessary.

Let $x(t)$ be a stochastic process, with continuous sample paths and transitions between states, as

in the previous paragraph. Let D^c be a closed subdomain of $\mathbb{R}^d$, and τ the *first hitting time* [4, 15] :

$$\tau = \inf\{t : x(t) \in D^c\}$$

with $\tau = +\infty$ if the set is empty. By continuity of the process, we can define a *stopped process* $\xi(t)$ associated with $x(t)$ [4, 15] such that

$$\xi_t = \begin{cases} x(t) & \text{if } t < \tau \\ x(\tau) & \text{if } t \geq \tau \end{cases}$$

The fundamental property of the stopped process is that if $\xi(t)$ ever encounters the boundary ∂D^c, it remains stuck on it forever. Let $\pi(y,t)$ be the conditional probability of first passage of the boundary ∂D^c, or the probability that the stopped process ξ hits the boundary ∂D^c exactly once between the instants 0 and t, starting from the position y :

$$\pi(y,t) = \mathsf{P}\{\xi(t) \in D^c \,|\, \xi(0) = y\}$$

Following well-known references [4, 5], to find what equation $\pi(y,t)$ satisfies, it is sufficient to examine the evolution of the expectation of a smooth function of the process ξ. Let indeed $f(\xi(t),t)$ be a twice continuously differentiable function of the process $\xi(t)$ and the time $t > s$. The process has the initial condition $\xi(s) = (y,\beta)$. Denoting $\xi_{y,\beta,s}(t)$ the process with fixed start, let $\Phi(y,\beta,t,s)$ be the expectation of $f(\xi_{y,\beta,s}(t),t)$ at t :

$$\Phi(y,\beta,t,s) = \mathsf{E}\big[\, f(\xi_{y,\beta,s}(t),t)\,\big] = \int dx\, d\alpha\, f(x,t,s)\, W(x,\alpha,t \,|\, y,\beta,s)\, \mu_\alpha(t)$$

where we can replace the product $W(x,\alpha,t \,|\, y,\beta,s)\, \mu_\alpha(t)$ by the simpler $\widetilde{W}(x,\alpha,t \,|\, y,\beta,s)$ after a change of convention on the function W, as hinted at the beginning of the section 1.3. By the *backward* expansion, we have :

$$\partial_s \Phi(y,\beta,t,s) = \mathsf{E}\big[\, \partial_s f(\xi_{y,\beta,s}(t),t)\big] - \langle \Phi(.\,,\beta,t,s)\,,\, p_t^\beta(.\,|\,y,s)\rangle \\ - \int d\gamma\, \lambda_{\beta\gamma}(s)\, \big[\, \Phi(y,\gamma,t,s) - \Phi(y,\beta,t,s)\,\big] \quad (1.20)$$

But if the transition probabilities of the problem do *not* depend on s (time invariance of the process, see [5], pp 173–176), the two processes $\xi_{y,\beta,s}(t)$ and $\xi_{y,\beta,0}(t-s)$ have same distribution of probability. Therefore :

$$\Phi(y,\beta,t,s) = \Phi(y,\beta,t-s,0)$$

Then $\partial_s \Phi(y,\beta,t,s) = -\partial_t \Phi(y,\beta,t-s,0)$, and denoting $\Phi(y,\beta,t-s,0) = \Phi(y,\beta,u)$ with $u = t-s$,

$$\partial_u \Phi(y,\beta,u) = -\mathsf{E}\big[\,\partial_u f(\xi_{y,\beta,0}(u),u)\big] + \langle \Phi(.\,,\beta,u)\,,\, p_t^{\beta}(.\,|\,y)\rangle + \int d\gamma\; \lambda_{\beta\gamma}\,\big[\,\Phi(y,\gamma,u) - \Phi(y,\beta,u)\,\big] \quad (1.21)$$

With $D = \mathbb{R}^d \setminus D^c$ and the characteristic function χ_D of D, we consider the sequence :

$$f_{\epsilon,n}(\xi_{y,\beta,0}(t),t) = \chi_D(\xi_{y,\beta,0}(t))\, e^{-n\int_0^t \rho_\epsilon(\xi_{y,\beta,0}(t))}$$

where $\rho_\epsilon(x)$ is a twice continuously differentiable function such that :

$$\rho_\epsilon(x) = \begin{cases} 0 & \text{if } d(x,D^c) \geq \epsilon \\ > 0 & \text{otherwise} \end{cases}$$

The expectation of the function $f_{\epsilon,n}(\xi_{y,\beta,0}(t),t)$ is an approximation of the probability

$$\mathsf{P}\{\,\xi_{y,\beta,0}(t) \in D\,\}$$

It is indeed straight forward to check the following properties :

1. If $\xi_{y,\beta,0}(t) \notin D$, $f_{\epsilon,n}(\xi_{y,\beta,0}(t),t) = 0$ because of the characteristic function χ_D factored.

2. If we denote τ_ϵ the first time the process approaches the boundary within a distance less than ϵ (then $\tau_\epsilon < \tau$ by continuity of the sample paths),

$$t > \tau_\epsilon \implies \int_0^t \rho_\epsilon(\xi_{y,\beta,0}(t)) > 0 \quad \text{and} \quad \lim_{n\to+\infty} f_{\epsilon,n}(\xi_{y,\beta,0}(t),t) = 0$$

   Hence the integral keeps the memory of the first (approximated) encounter.

3. Similarly,

$$t \leq \tau_\epsilon \implies \int_0^t \rho_\epsilon(\xi_{y,\beta,0}(t)) = 0 \quad \text{and} \quad \lim_{n\to+\infty} f_{\epsilon,n}(\xi_{y,\beta,0}(t),t) = 1$$

The continuity of the sample paths of the stochastic process $\xi_{y,\beta,0}(t)$ also justifies that $\tau_\epsilon \to \tau$ as $\epsilon \to 0$.

Calling $\Phi_{\epsilon,n}(y,\beta,t)$ the corresponding expectation, and observing that [5] :

$$\mathsf{E}\big[\,\partial_t f_{\epsilon,n}(\xi_{y,0}(t),t)\,\big] = -n\,\rho_\epsilon(y)\,\Phi_{\epsilon,n}(y,\beta,t)$$

we have, for y such that $d(y, D^c) \geq \epsilon$: $\rho_\epsilon(y) = 0$ therefore

$$\mathsf{E}\big[\, \partial_t f_{\epsilon,n}(\xi_{y,0}(t), t) \,\big] = 0$$

It follows :

$$\partial_t \Phi_{\epsilon,n}(y, \beta, t) = \langle \Phi_{\epsilon,n}(.\,, \beta, t)\,,\, p_t^\beta(.\,|\,y) \rangle + \int d\gamma\; \lambda_{\beta\gamma} \left[\, \Phi_{\epsilon,n}(y, \gamma, t) - \Phi_{\epsilon,n}(y, \beta, t) \,\right] \tag{1.22}$$

because But as $n \to \infty$, using the properties enumerated it is easy to check that

$$\Phi_{\epsilon,n}(y, \beta, t) \longrightarrow \Phi_\epsilon(y, \beta, t) = \mathsf{P}\{\, \xi_{y,\beta,0}(t) \in D_\epsilon \,\}$$

with D_ϵ the ϵ-neighborhood of D : $D_\epsilon = \{\, x \in \mathbb{R}^d \;:\; d(x, D) < \epsilon \,\}$. By taking the limit $n \to \infty$ in the equation, $\Phi_\epsilon(y, \beta, t)$ must also satisfy (1.22). Finally in the limit $\epsilon \to 0$, as we hinted τ_ϵ tends to τ and

$$\Phi_\epsilon(y, \beta, t) \longrightarrow 1 - \pi(y, \beta, t) = 1 - \mathsf{P}\{\, \xi_{y,\beta,0}(t) \in D^c \,\}$$

The limit $1 - \pi(y, \beta, t)$ satisfies the limit equation, therefore $\pi(y\beta, t)$ also satisfies the limit equation :

$$\partial_t \pi(y, \beta, t) = \langle \pi(.\,, \beta, t)\,,\, p_t^\beta(.\,|\,y) \rangle + \int d\gamma\; \lambda_{\beta\gamma} \left[\, \pi(y, \gamma, t) - \pi(y, \beta, t) \,\right] \tag{1.23}$$

It has been proved in a previous section that $\langle\, 1(.)\,,\, p_t^\beta(.\,|\,y) \,\rangle = 0$ indeed.

The probability of passage is obtained by the decomposition :

$$\begin{aligned} \pi(y, t) &= \mathsf{P}\{\, \xi_{y,0}(t) \in D^c \,\} \\ &= \int d\beta\; \mathsf{P}\{\, \xi_{y,0}(t) \in D^c \,|\, \xi_{y,0}(0) \in \Sigma_\beta \,\}\, \mathsf{P}\{\, \xi_{y,0}(0) \in \Sigma_\beta \,\} \\ &= \int d\beta\; \pi(y, \beta, t)\; \mathsf{P}\{\, \xi_{y,0}(0) \in \Sigma_\beta \,\} \end{aligned} \tag{1.24}$$

1.6 Stability

This brief section is devoted to some stability aspects of the Kolmogorov equations, in a sense defined. The main issue is the convergence of numerical solutions of the corresponding numerical problems to the exact probability density functions or other statistics, such as $W(x, t)$ and $\pi(y, t)$, when it is not possible to find direct analytical expressions.

First we recall the use of the premultiplication convention (or the change of probability), that allows $W(x, t) = \int_A d\alpha\; W^\alpha(x, t)$. Now, denoting $\mathcal{W}(.,:) = \{W^\alpha(.,:)\}_{\alpha \in A}$ the time-dependent family of functions $W^\alpha(.,:)$, where the dot temporarily replaces the space variables in this section, and the

colon the time, the system can be clearly put into the form :

$$\frac{d}{dt}\mathcal{W}(.,t) = \mathbb{A}(t).\mathcal{W}(.,t) \tag{1.25}$$

where $\mathbb{A}(t)$ is the (time-dependent) integro-differential operator defined by

$$\mathbb{A}(t).\{W^\alpha(x,t)\}_{\alpha\in A} = \left\{ \langle p_t^\alpha(x\,|\,.,t)\,,\, W^\alpha(.,t)\rangle \,+\, \int_A d\beta\, \lambda_{\beta\alpha}(t)\, W^\beta(x,t) - \lambda_{\alpha\beta}(t)\, W^\alpha(x,t) \right\}_{\alpha\in A}$$

This operator is linear in $\mathcal{W}$ globally (hence a specific notation). Any solution of problem (1.25) is therefore unique, and completely determined by the initial conditions $\mathcal{W}(.,t_0)$.

We now consider the set X of all possible families $\mathcal{W}(.)$ such that

$$\int_A d\alpha \int dx\, |W^\alpha(x)| \,<\, \infty$$

and for such families we define $\|\,\mathcal{W}\,\| = \int_A d\alpha \int dx\, |W^\alpha(x)|$. Endowed with this norm, the set X is a normed $\mathbb{R}$-vector space. For families of probability density functions (with the convention of premultiplication by $\mu_\alpha(t)$), the absolute value is not necessary since all quantities are nonnegative, and the norm must be 1.

By construction, the integro-differential operator $\mathbb{A}(t)$ is the infinitesimal generator of a semigroup $\mathbb{T}(t)$ [13], that satisfies the stability condition[4]

$$\forall\, t > 0\,, \quad \|\,\mathbb{T}(t).\mathcal{W}\,\| \leq \|\,\mathcal{W}\,\| \tag{1.26}$$

This inequality even becomes an equality for true probabilities, since

$$\forall t \geq 0, \quad \int dx \int_A d\alpha\, W^\alpha(x,t) = 1$$

if $\mathcal{W}$ is a solution of (1.25). In other words, if the initial distribution $\mathcal{W}$ is a family of probability density functions, the solution to the system of integro-differential equations (1.5) is still a family of probability density functions. Such a semigroup is a particular case of *contractions.*

1.7 Important applications

Now we show two important classes of applications, namely deterministic systems and Gaussian systems Σ_α. The second case (Itô processes) is a generalization of the first one.

[4] With appropriate smoothness of the coefficients of the differential operators and the transition rates.

1.7.1 Case of deterministic systems

In the present subsection we show how to deal with determistic systems, more precisely systems Σ_α driven by ordinary differential equations $\dot{x}_\alpha = f^\alpha(x_\alpha, t)$ (not necessarily linear), where the vector fields $f^\alpha(x,t)$ are assumed to be lipschitzian[5]. In this context indeed, the transition probability $p^\alpha(x,t|y,s)$ should concentrate all its mass at the point x such that

$$x - y = \int_s^t f^\alpha(x^\alpha(\tau), \tau)\, d\tau$$

The continuity of the vector field at the point (y,t) allows to therefore expand the transition probability as

$$p^\alpha(x, t+dt \,|\, y, t) = \delta(x-y) - \delta'(x-y).f^\alpha(y,t)\, dt \; + o(dt)$$

where . denotes the scalar product. Hence, by definition, $p_t^\alpha(x \,|\, y; t) = -\delta'(x-y).f^\alpha(y,t)$. The presence of the gradient of the Dirac's δ-function $\delta'(x-y)$ allows to replace y by x in the vector field, such that

$$p_t^\alpha(x|\, y; t) = -\delta'(x-y).f^\alpha(x,t) = -\sum_{l=1}^{d} \delta'(x_l - y_l)\, f_l^\alpha(x,t)$$

The partial differential equation (1.3) satisfied by $W^j(x,t)$ takes the form:

$$\frac{\partial W^j}{\partial t}(x,t) = -\mathrm{div}\big(f^j\, W^j\big)(x,t) \; + \sum_{i \neq j} \lambda_{ij}(t)\, W^i(x,t) - \lambda_{ji}(t)\, W^j(x,t) \tag{1.27}$$

In the continuous case :

$$p^\alpha(x, t+dt \,|\, y, t) = \delta(x-y) - \delta'(x-y).f^\alpha(y,t)\, dt \; + o(dt)$$

leads to

$$\frac{\partial W^\alpha}{\partial t}(x,t) = -\mathrm{div}\big(f^\alpha\, W^\alpha\big)(x,t) \; + \int d\beta\, \lambda_{\beta\alpha}(t)\, W^\beta(x,t) - \lambda_{\alpha\beta}(t)\, W^\alpha(x,t) \tag{1.28}$$

1.7.2 Itô's stochastic partial differential equations

Here we deal with a general class of stochastic differential equations, often called Itô's stochastic partial differential equations [4, 10, 15]. This class of stochastic processes is very important in practice, for it fairly models many physical systems and allows a thorough calculation of many important statistics related to it.

[5]This allows the systems Σ_α to have continuous and unique sample paths

In the case of continuous transitions, let $x_t^\alpha \in \Sigma_\alpha$ be a stochastic process such that :

$$dx_t^\alpha \;=\; a^\alpha(x_t,t)\,dt \;+\; b^\alpha(x_t,t)\,dW_t^\alpha$$

where W_t is a n^α-dimensional standard Brownian motion [15], and $b^\alpha(x_t,t)$ a $d \times n^\alpha$ matrix. we have :

$$p^\alpha(x,t\,|\,y,s) = \delta(x-y) \;-\; \delta'(x-y).a^\alpha(x,t)\,(t-s) \;+\; \frac{1}{2}\delta''(x-y)\,b^\alpha(x,t)b^\alpha(x,t)^T\,(t-s) \;+\; o(t-s)$$

One may refer to [4], pp 96–97, for a proof. The distribution $p_t^\alpha(x\,|\,y,t)$ identifies to :

$$p_t^\alpha(x\,|\,y,t) = -\delta'(x-y).a^\alpha(x,t) \;+\; \frac{1}{2}\delta''(x-y)\,b^\alpha(x,t)b^\alpha(x,t)^T$$

and leads to the system :

$$\frac{\partial W^\alpha}{\partial t}(x,t) = -\mathrm{div}\big(a^\alpha\,W^\alpha\big)(x,t) \;+\; \frac{1}{2}\mathrm{div}\big(b^\alpha b^{\alpha T}\,\mathrm{grad}\,W^\alpha\big)(x,t) \\ +\; \int d\beta\,\lambda_{\beta\alpha}(t)\,W^\beta(x,t) - \lambda_{\alpha\beta}(t)\,W^\alpha(x,t) \quad (1.29)$$

1.8 Academic example

Here, we consider a mobile traveling along a straight line, and whose speed, though it keeps a constant magnitude, sees its sign changing with Poissonian occurences. Thus the speed remains at $\epsilon\,v$ for a random time that follows an exponential law ($\epsilon \in \{-1,1\}$), and then toggles to the value $-\epsilon\,v$, at which it remains an exponentially distributed time, and so forth. The Poisson distribution has parameter λ. We want to find the probability density function $W(x,t)$ for the position of the mobile, at each instant t, given an initial distribution $W(x,0) = W_0(x)$. We assume that initially, the probability that the mobile has speed $+v$ is equal to the probability it has $-v$.

1.8.1 Treatment

We consider the two deterministic systems

$$\begin{cases} \Sigma^+ : & \dot{x}^+ = +v \\ \Sigma^- : & \dot{x}^- = -v \end{cases} \quad (1.30)$$

and the two corresponding conditional density functions $W^+(x,t)$ and $W^-(x,t)$. According to (1.3), we have

$$\begin{aligned} \frac{\partial W^+}{\partial t} &= \quad v\,\frac{\partial W^+}{\partial x} + \lambda\,(W^- - W^+) \\ \frac{\partial W^-}{\partial t} &= -\,v\,\frac{\partial W^-}{\partial x} + \lambda\,(W^+ - W^-) \end{aligned} \tag{1.31}$$

The masses also verify

$$\begin{aligned} \frac{d\mu^+}{dt} &= \lambda\,(\mu^- - \mu^+) \\ \frac{d\mu^-}{dt} &= \lambda\,(\mu^+ - \mu^-) \end{aligned}$$

The initial conditions suggest: $\mu^+(0) = \mu^-(0) = \frac{1}{2}$. The initial equilibrium perpetuates forever, according to the later system. Then $W(x,t) = \frac{1}{2}W^+(x,t) + \frac{1}{2}W^-(x,t)$. From the former system, we have immediately

$$\begin{aligned} \frac{\partial W}{\partial t} &= v\,\frac{\partial Z}{\partial x} \\ \frac{\partial Z}{\partial t} &= v\,\frac{\partial W}{\partial x} - 2\lambda\,Z \end{aligned} \tag{1.32}$$

with the convention $Z(x,t) = \frac{1}{2}W^+(x,t) - \frac{1}{2}W^-(x,t)$. Taking the time derivative of (1) and the space derivative of (11), we can eliminate Z and get

$$\frac{\partial^2 W}{\partial t^2} = v^2\,\frac{\partial^2 W}{\partial x^2} - 2\lambda\,\frac{\partial W}{\partial t}$$

This equation may be solved[6] using the Fourier transform $w(s,t)$ of W, also called the *characteristic function* of $W(x,t)$: $w(s,t) = \mathsf{E}_W[e^{isX_t}]$, where the process X_t is supposed to have $W(x,t)$ as a probability density function. We find :

$$w(s,t) = h_1(s)\,e^{-\left(\lambda - \sqrt{\lambda^2 - v^2 s^2}\right)t} \;+\; h_2(s)\,e^{-\left(\lambda + \sqrt{\lambda^2 - v^2 s^2}\right)t}$$

with h_1 and h_2 two arbitrary functions of s. Let us identify those two functions, using the remaining information. The initial distribution $W_0(x)$ has Fourier transform $w_0(s)$, therefore $h_1(s) + h_2(s) = w_0(s)$. Moreover, the Fourier transform $z(s,t)$ of Z must satisfy

$$\begin{aligned} i\,s\,v\,z(s,t) &= \frac{\partial w}{\partial t}(s,t) \\ &= h_1(s)\left(-\lambda + \sqrt{\lambda^2 - v^2 s^2}\right) e^{-\left(\lambda - \sqrt{\lambda^2 - v^2 s^2}\right)t} \;+\; h_2(s)\left(-\lambda - \sqrt{\lambda^2 - v^2 s^2}\right) e^{-\left(\lambda + \sqrt{\lambda^2 - v^2 s^2}\right)t} \end{aligned}$$

[6] At least for the time dependent Fourier transform $w(s,t)$.

At $t = 0$, we assumed the masses equally distributed, that is $W^+(x,0) = W^-(x,0)$, or $Z(x,0) = 0$. Therefore $z(s,0) = 0$, that is

$$h_1(s)\left(-\lambda + \sqrt{\lambda^2 - v^2 s^2}\right) + h_2(s)\left(-\lambda - \sqrt{\lambda^2 - v^2 s^2}\right) = 0 \qquad \forall s$$

Hence

$$h_2(s) = h_1(s)\,\frac{-1 + \sqrt{1 - v^2 s^2/\lambda^2}}{1 + \sqrt{1 - v^2 s^2/\lambda^2}}$$

The functions h_1 and h_2 can then be completely identified to

$$h_{1,2}(s) = \frac{1}{2}\,w_0(s)\left[1 \pm \left(1 - \frac{v^2 s^2}{\lambda^2}\right)^{-1/2}\right]$$

We come up with the following expression for w :

$$w(s,t) = w_0(s)\,e^{-\lambda t}\left[\cosh\left(\lambda t\sqrt{1 - r^2 s^2}\right) + \frac{\sinh\left(\lambda t\sqrt{1 - r^2 s^2}\right)}{\sqrt{1 - r^2 s^2}}\right]$$

with $r = \frac{v}{\lambda}$. Hence $W(x,t)$ is the convolution of $W_0(x)$ by the Green function $\Gamma(x,t)$ of the problem, whose Fourier transform $\gamma(s,t)$ can be written :

$$\gamma(s,t) = \begin{cases} e^{-\lambda t}\left[\cosh\left(\lambda t\sqrt{1 - r^2 s^2}\right) + \frac{\sinh\left(\lambda t\sqrt{1 - r^2 s^2}\right)}{\sqrt{1 - r^2 s^2}}\right] & \text{when } |s| < \frac{1}{r} \\ e^{-\lambda t}\left[\cos\left(\lambda t\sqrt{r^2 s^2 - 1}\right) + \frac{\sin\left(\lambda t\sqrt{r^2 s^2 - 1}\right)}{\sqrt{r^2 s^2 - 1}}\right] & \text{when } |s| > \frac{1}{r} \end{cases}$$

The function $\gamma(s,t)$ is real, due to the symmetry of the problem (same intensity of the speed rightward and leftward, and same initial distributions).

1.8.2 Comments

It is interesting to examine the asymptotic behaviour of the characteristic function. If the initial distribution is a Dirac δ (that is to say, when $w_0(s) = 1$), it is indeed easy to check that

$$w(s,t) \sim \frac{1}{2}\,e^{-\lambda t}\left(e^{-ivst} + e^{ivst}\right)$$

Therefore, the total probability density function $W(x,t)$ is the sum of a probability with a Fourier transform vanishing when $|s| \to \infty$ and two dumped moving "peaks" (two Dirac's δ), with speed v and opposite directions. The dumping coefficient is equal to $e^{-\lambda t}$, which is the total probability that the mobile *never* changes its speed. Each peak represents an equal part of the initial distribution moving at constant speed from the origin, with no change in the direction during the move. The moving mass loses weight all along the trajectory, as it becomes harder to keep the same speed when

time goes. The loss is proportionnal to $1 - e^{-\lambda t}$.

We have run a Monte-Carlo simulation, with particles starting at the origin, and with equal probability to start leftward and rightward. At each time step, a random trial decides if a particle changes its speed, in accordance with the Poisson distribution. The time step is chosen small enough to avoid a significant bias in the distribution of the instants of changes, when time is sampled. At a given date t, the simulation is stopped, and the particles are counted in each interval $[x_k, x_{k+1})$ of the space domain. The (fast) Fourier transform of the resulting density is compared to the exact solution on figure (1-2). The distribution obtained from the simulation is also plotted on figure (1-3). The two sharp peaks are clearly visible.

We also solved the partial differential equation (1.31) numerically, by a standard first order forward Euler scheme. A comparison of the numerical solution with the simulation is plotted on figure (1-4). In this case, the initial distribution was Gaussian; this explains the significant extension of the two peaks. Higher order forward Euler schemes were then studied, with no significant improvement. On the contrary, instability was a reason for increasing the number of iterations needed. The figure (1-5) shows these results.

In all cases, the following values were chosen :

$$\lambda = 20\ \mathrm{s}^{-1} \qquad v = .5\ \mathrm{m.s}^{-1}$$

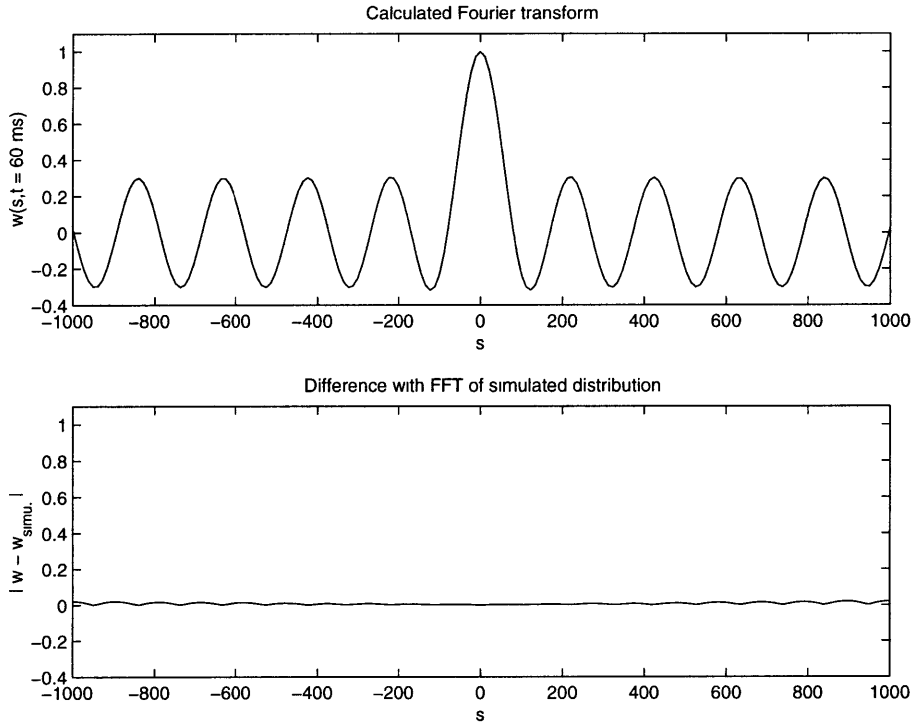

Figure 1-2: Top : exact Fourier transform of the distribution. Bottom : Comparison with the FFT of the distribution obtained by a Monte-Carlo simulation.

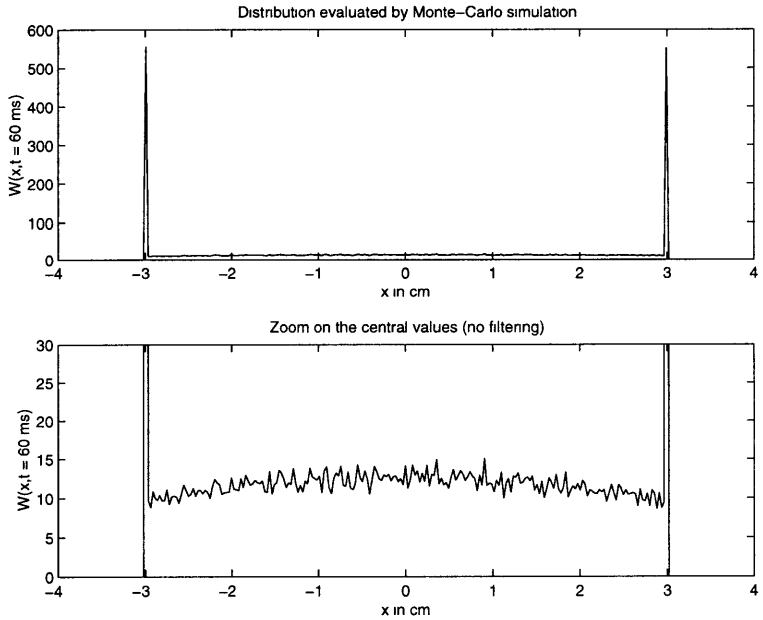

Figure 1-3: Distribution obtained from the simulation. The roughness of the curve shown on bottom is due to the use of a finite number of particles (smoothness improves as this number increases).

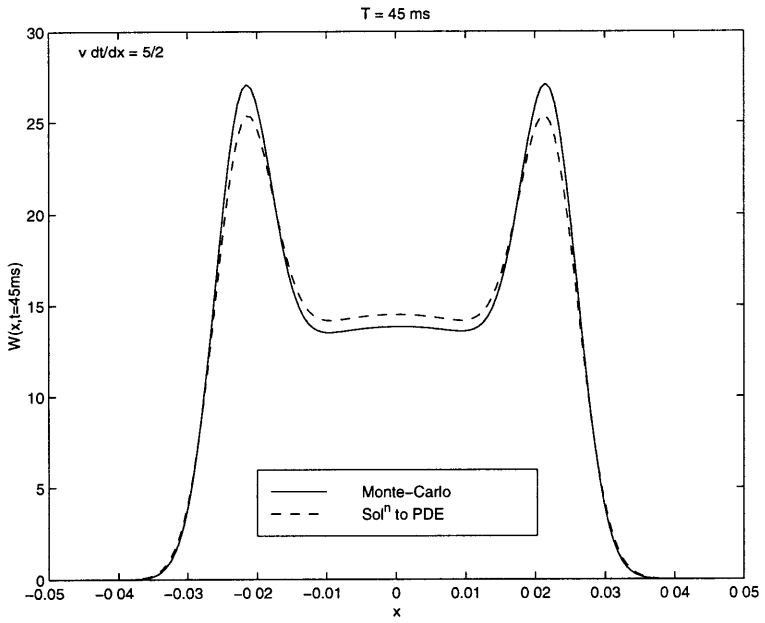

Figure 1-4: Comparison between the distribution obtained from the simulation and a numerical resolution of the differential equations (1.31). The initial distribution is a Gaussian (hence "fatter" peaks).

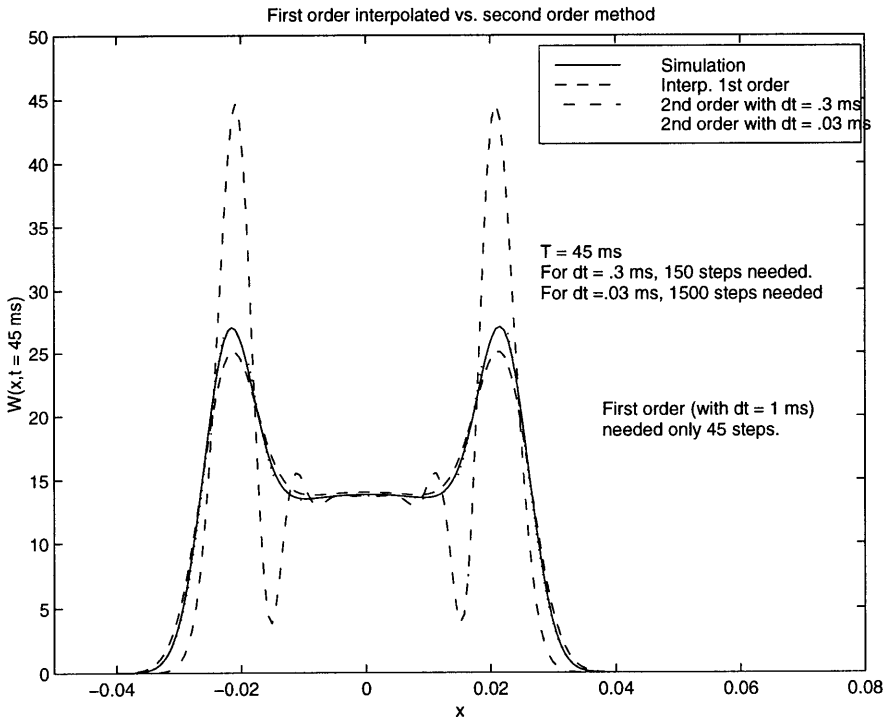

Figure 1-5: Comparisons between various cases of the Euler method. Instability is shown in the case of the second order method, with too large a time step.

Chapter 2

Probabilistic models for air traffic control

2.1 Modeling uncertainty

We may apply the framework shown in the previous chapter to the case of air traffic control, where the behaviours of pilots and aircraft can be described in terms of probability. Poisson processes play a key-role in this context, conveniently modeling events occurring with a mean frequency λ.

In the following, we use the model and the assumptions proposed earlier [6, 1, 17]. In particular, the planes are treated by pairs *host/intruder*, where the host has a well-known probabilistic behaviour while the intruder may take unexpected actions. The main issue of the modeling is to compute $W_H(x,t)$ and $W_I(x,t)$, the respective probability density functions of the host and the intruder – in fact, their evolution in time and space, as described by integro-differential equations, but also the probability of collision for the pair, as defined in [6].

Let us assume the host is flying with initial heading ψ_H, at altitude h_H and initial position (x_H, y_H) at that altitude. We denote ψ_I, h_I, and (x_I, y_I) the corresponding variables for the intruder. We now briefly recall the model used. Each plane has a nominal speed, with fluctuations assumed to be Gaussian, with standard deviation $\sigma_v = 15$ kts. Only the along-track position is affected by these fluctuations. The speed, though imperfectly known, is assumed to remain the same along the trajectory. The cross-track position is also not completely deterministic, and is assumed to be normally distributed with standard deviation $\sigma_{\text{ctp}} = 1$ nmi. The imperfect knowledge of the intruder's behaviour is modeled as follows : the plane may engage changes of altitude with a mean frequency of $\lambda = 4\ \text{h}^{-1}$, as well as changes of heading with the same mean frequency. When changing of altitude, the final altitude is random, with uniform distribution between 0 and 10,000 ft. The heading changes do not have a uniform distribution, but favor left or right turn between 5 and 20

degrees. In case of high collision risk, controllers may submit orders to the pilots of the host airplane, orders to be executed after a random time following a gamma distribution with mean 1 min and such that the orders are executed within 2 min with 95% probability (flight crew response latency) [17]. The action to take may also be suggested by an automatic system of conflict alert, aboard. As in this paragraph, we will use the subscript H for the flight parameters of the host aircraft, and the subscript I for the intruder.

We now construct the system of equations the probability density function should satisfy, following step by step the procedure described above : first derive the Fokker-Planck equation of a plane with fixed heading and altitude, second add exchange terms to obtain the system.

The variables of the problem have to split into two different groups : the variables (x, y), with Gaussian distributions, and the variables driven by Poisson processes, namely ψ and h. Observing that a plane with fixed ψ and h has a well-known (probabilistic) trajectory, we will consider from now on a continuous set of systems $\Sigma_{\psi,h}$ modeling the aircraft with heading ψ and altitude h. The possible values of the altitude are equally distributed over [0 , 10,000 ft]. In the system $\Sigma_{\psi,h}$, according to the model the aircraft follows the stochastic differential equation :

$$dX_t = V_{H,I}\,(dt + \frac{\sigma_v\sqrt{c}}{|V_{H,I}|}\,dW_{t/c}) \tag{2.1}$$

where $V_{H,I}$ denotes the vector speed, and $dW_{t/c}$ the one-dimensional Standard Brownian Motion[1]. The constant c represents the time unit, playing the role of a scaling factor; in our context, $c =$ 1 h. The nominal speed is assumed horizontal, *ie* with coordinates $V_{H,I} = v_{H,I}\,(\cos\psi, \sin\psi, 0)$. The linear, differential equation (2.1) leads to the following Fokker-Planck equation for system (ψ, h)[4, 15] :

$$\frac{\partial W^{\psi,h}}{\partial t}(X,h,t) = -V_{H,I}.\nabla W^{\psi,h}(X,h,t) + \frac{\sigma^2}{2}\nabla.B\,\nabla W^{\psi,h}(X,h,t) \tag{2.2}$$

where $X = (x, y)$, $\sigma^2 = c\,\sigma_v^2$, and B is the 2×2 symmetric, diffusion matrix

$$B = \begin{pmatrix} \cos^2\psi & \sin\psi\,\cos\psi \\ \sin\psi\,\cos\psi & \sin^2\psi \end{pmatrix}$$

Observe that $B = u\,u^T$, if u stands for the unit vector $V_{H,I}/|V_{H,I}|$ with 2D-coordinates $(\cos\psi, \sin\psi)$.

[1] As in the previous chapter, this refers to a Gaussian process with mean and standard deviation proportional to time. The coefficients of proportionality, called drift rate for the mean and variance rate for the standard deviation, are respectively 0 and 1 in arbitrary units in the case of Standard Brownian Motion. The abstract process must be scaled by physical constants to obtain a physical process, in particular a length for the standard deviation and a speed for the drift rate. This is the role of the constant c.

The last term of equation (2.2), multiplied by $\sigma^2/2$ is thus a short-hand for :

$$\cos^2\psi\,\frac{\partial^2 W}{\partial x^2} + 2\sin\psi\,\cos\psi\,\frac{\partial^2 W}{\partial x \partial y} + \sin^2\psi\,\frac{\partial^2 W}{\partial y^2} = \left(\cos\psi\,\frac{\partial}{\partial x} + \sin\psi\,\frac{\partial}{\partial y}\right)^2 W$$

In equation (2.2), the redundant superscript h is used to recall that the altitude is a random variable with Poisson behaviour, and hence deserves a particular treatment.

Also, space derivations do not hold for the altitude h. The first order term is traditionally called the drift term. The second order term is known as a diffusion term (hence the name of diffusion matrix for B). The uncertainty in cross-track position can be taken into account in the initial distribution $W^{\psi,h}(X, h, t = 0)$.

We now have to state what exchange rates the model imposes here. As for the intruder, we have

$$\textit{Intruder}: \quad \lambda(\psi, h \to \psi', h')(t) = \quad \lambda\, k(\psi' - \psi)\, u(h)$$

where $k(\phi)$ is the probability density function used by Yang and Kuchar to describe the preferred angle of turn, and $u(h)$ the uniform distribution over the possible altitudes h. Thus, $u(h)$ is a constant that will be denoted U. Note that the function $k(\phi)$ is even.

The relations of the transfer coefficients $\lambda(\psi, h \to \psi', h')(t)$ to the model are very intuitive : during dt, the probability to change the heading from ψ to ψ' and the altitude from h to h' is $k(\psi' - \psi)\, u(h)$ – since the altitude h is a uniformly distributed random variable. But this probability has to be multiplied by that of the actual decision to maneuver, namely $\lambda\, dt$. Hence the coefficients for the intruder.

Now, for the host, suppose the maneuver thought of to prevent from collision is a left turn of ϕ degrees from present heading, at a constant altitude h_H. Then only two headings are of concern, ψ_H and $\psi_H + \phi$, and we only have to describe the transition between these two angles. We then consider only W^{ψ_H, h_H} and $W^{\psi_H + \phi, h_H}$, and the corresponding transfer coefficient $\lambda(\psi_H, h_H \to \psi_H + \phi, h_H)(t)$. If t_0 denotes the time when pilots are notified to take action, we have

$$\lambda(\psi_H, h_H \to \psi_H + \phi, h_H)(t) = \begin{cases} 0 & \text{before } t_0 \\ \lambda\,\gamma(t - t_0) & \text{after } t_0 \end{cases}$$

All other transfer coefficients are zero – including $\lambda(\psi_H + \phi, h_H \to \psi_H, h_H)(t)$ (pilots do not turn right back). Remark that the instant t_0, which may be seen as an instant of switch between two models, is entirely determined by the security system (or the controller). But in turn this is an independent event (intruder plane continuing on dangerous action) that determines the decision of the system (the controller). The *host* is not a closed-loop system.

The initial distribution of the time-dependent "masses" $\int dX\, W^{\psi,h}(X, h, t)$ remains to be set up.

The initial heading is assumed to be a well-known $\psi_{H,I}$; the altitude, however, suffers from some uncertainty (due to GPS) assumed to be normally distributed around a mean $h_{H,I}$, with standard deviation 30 m. Regarding the other uncertainties (σ_v, σ_{ctp}), we chose to neglect this term : in particular $\sigma_{\text{ctp}} \gg 30$ m indeed. Hence, we take the initial distribution

$$\int dx\, W^{\psi,h}(X, h, t = 0) = \delta[(\psi, h) - (\psi_{H,I}, h_H)]$$

2.1.1 Resulting system

The model discussed in the previous paragraphs leads to the following system of integro-differential equations, for any t :

Intruder

$$\frac{\partial W^{\psi,h}}{\partial t} = -V_I.\nabla W^{\psi,h} + \frac{\sigma_v^2}{2}\nabla.B\,\nabla W^{\psi,h} + \lambda U \int d\psi' dh'\, k(\psi' - \psi)\,(W^{\psi',h'} - W^{\psi,h}) \tag{2.3}$$

Since both indices ψ and h are continuous, these equations form a global integro-differential system with continuous coupling between equations.

Host before t_0

$$\frac{\partial W^{\psi_H,h_H}}{\partial t} = -V_H.\nabla W^{\psi_H,h_H} + \frac{\sigma^2}{2}\nabla.B\,\nabla W^{\psi_H,h_H} \tag{2.4}$$

This is a simple Fokker-Planck equation, to be solved for the altitude h_H only (null initial conditions at other altitudes make the solutions trivial).

Host after t_0

Denoting $\psi_H + \phi = \psi_f$, we have *two* coupled partial differential equations

$$\frac{\partial W^{\psi_H,h_H}}{\partial t} = -V_H.\nabla W^{\psi_H,h_H} + \frac{\sigma^2}{2}\nabla.B\,\nabla W^{\psi_H,h_H} - \lambda\,\gamma(t - t_0)\,W^{\psi_H,h_H} \tag{2.5}$$

and

$$\frac{\partial W^{\psi_f,h_H}}{\partial t} = -V_H.\nabla W^{\psi_f,h_H} + \frac{\sigma^2}{2}\nabla.B\,\nabla W^{\psi_f,h_H} + \lambda\,\gamma(t - t_0)\,W^{\psi_H,h_H} \tag{2.6}$$

Recall that the nominal velocity vectors $V_{H,I}$ as well as the diffusion matrices B depend on ψ. The total probability of presence of the aircraft is, in both cases, given by

$$W(X, h, t) = \int d\psi\, W^{\psi,h}(X, t)$$

2.2 Reduction of dimensionality

The host aircraft has rather simple equations describing its probability density function. Before t_0, the heading and the altitude are fixed, and if we suppose the heading to be 0, one can express a "direct" analytical solution as :

$$W(X, h_H, t) = W(X, h_H, 0) * \frac{1}{2\pi\sigma\sqrt{t}} \exp\left[-\frac{(x - V_H\, t)^2}{2\,\sigma^2 t}\right] \tag{2.7}$$

with $X = (x, y)$ and an initial speed V_H positive (plane flying to the right). The star * is the operation of convolution with respect to the 2-D variable X. The distribution is thus convoluted with the ditribution of a one-dimensional Brownian motion, with variance rate σ^2 and drift rate V_H. This analytical solution is very useful to know with high accuracy the values of the distribution far from the plane, in regions yet close enough to constitute a danger for the intruder. The solution for $\psi_h \neq 0$ is found by rotation of the coordinates by an angle of $-\psi_H$ in the 2-D space.

After t_0, an "indirect" solution can be found. The equation (2.5) is independent of the equation (2.6), and with $\Gamma(t) = \int_{t_0}^{t} dt'\, \gamma(t')$, a change of unknown function gives a solvable equation. With again the hypothesis $\psi_H = 0$, we find

$$W^{0,h_H}(X, t) = e^{-\lambda\,\Gamma(t)}\, W^{0,h_H}(X, t_0) * \frac{1}{2\pi\sigma\sqrt{t}} \exp\left[-\frac{(x - V_H\,(t - t_0))^2}{2\,\sigma^2 (t - t_0)}\right]$$

The other component may be found through its Fourier transform $\widehat{W}^{\phi,h_H}$. With

$$A(\xi_x, \xi_y) = -iV_H\,(\xi_x \cos\phi + \xi_y \sin\phi) - \frac{\sigma^2}{2}(\xi_x \cos\phi + \xi_y \sin\phi)^2$$

coming from the Fourier transform of the differential operators of (2.6) and the change of unknown $\widehat{W}^{I\phi,h_H} = e^{-A(\xi_x,\xi_y)\,t}\, \widehat{W}^{\phi,h_H}$, we have

$$\frac{\partial \widehat{W}^{I\phi,h_H}}{\partial t}(\xi_x, \xi_y, t) = \lambda\gamma(t)\, e^{-A(\xi_x,\xi_y)\,t}\, \widehat{W}^{0,h_H}(\xi_x, \xi_y, t) \tag{2.8}$$

where $\widehat{W}^{0,h_H}(\xi_x, \xi_y, t)$ is the Fourier transform of $W^{0,h_H}(X, t)$. The numerical integration of (2.8) gives the Fourier transform of $\widehat{W}^{I\phi,h_H}$. After multiplication by $e^{A(\xi_x,\xi_y)\,t}$, one recovers $\widehat{W}^{\phi,h_H}$ that can be inverted by fast inverse Fourier transform algorithms. One may as well find a numerical solution of (2.6), knowing $\lambda\gamma(t)\, W^{0,h_H}(X, t)$.

As mentioned before, the intruder has an infinite dimensional system of integro-differential equations (2.3). This is caused by the presence of two continuous state parameters : ψ and h. One expects a very high computational complexity.

Nonetheless, the altitude h plays a particular role : first, the velocity V_I of the airplane depends

only on ψ, not h. Second, as a parameter of the state describing the trajectory of the airplane, it has a uniform probability distribution U over a (bounded) interval. These particularities allow to easily reduce the dimensionality of our problem, without modification of the equations and/or solutions. It will make the resolution easier, either by numerical or possibly analytical methods.

Here is how dimensionality can be reduced. If we multiply the equation (2.3) by the constant U and integrate with respect to the altitude h, we get for $Z^{\psi}(X,t) \triangleq U \int dh\, W^{\psi,h}(X,t)$:

$$\frac{\partial Z^{\psi}}{\partial t} = -V_I.\nabla Z^{\psi} + \frac{\sigma^2}{2}\nabla.B\,\nabla Z^{\psi} + \lambda \int d\psi'\, k(\psi'-\psi)\,(Z^{\psi'} - Z^{\psi}) \tag{2.9}$$

We can indeed commute the two integrals with respect to the altitudes h, h' in the last term, then we have to evaluate $\int dh'\, U\, Z^{\psi'}(X,t)$ where $Z^{\psi'}$ does not depend on h' at all : it can be factored out, and the remaining integral is 1 by definition of U.

Equation (2.9), as can be easily seen, is a 2D problem, for which any numerical computation is expected to cost much less than a 3D. On one hand, for an altitude h fixed, we may write equation (2.3) as

$$\frac{\partial W^{\psi,h}}{\partial t} + \lambda\, W^{\psi,h} + V_I.\nabla W^{\psi,h} = \lambda \int d\psi'\, k(\psi'-\psi)\, Z^{\psi'} \tag{2.10}$$

For this we use the property $U \int dh' d\psi'\, k(\psi'-\psi) = 1$, for every ψ. In (2.10), the right-hand side may be interpreted as a shared inhomogeneous term. Then knowing $Z^{\psi}(X,t)$ for every X, t, ψ, *ie* solving the 2D problem (2.9), allows to completely ***decouple*** the integro-differential equations (2.10) satisfied by the family $W^{\psi,h}$ – and these even become classical partial differential equations.

On the other hand, the reader may now remind we chose deterministic initial conditions in the parameters ψ and h, so to speak the plane has well-known initial altitude h_I and initial heading ψ_I. Then, as far as the parameter h is concerned, there are only two cases to consider : 1. the altitude h where we solve (2.10) is actually h_I, 2. the altitude h is not h_I. Indeed, neither the equation to solve nor the initial conditions for $W^{\psi,h}$ distinguish two altitudes h_1, h_2 that have in common to be in case 2. Therefore, for all h_1, h_2,

$$h_1 \neq h_I \textbf{ and } h_2 \neq h_I \implies W^{\psi,h_1} \equiv W^{\psi,h_2},$$

by uniqueness of the solutions to the well-posed inhomogeneous problem (2.10).

The remarkable consequence is that we only have to solve three 2D problems in order to solve the entire 3D problem :

i. Find $Z^{\psi}(X,t)$, for all ψ, X, t, by solving the independent system (2.9),

ii. Solve a classical partial differential equation (2.10) for $h = h_I$, with Dirac initial distribution

in ψ,

iii. Solve a classical partial differential equation (2.10) for $h \neq h_I$, with zero initial distribution in ψ.

Note that solving (2.10) with zero initial condition does not provide a zero solution, because of the nonzero inhomogeneous term in Z. Here, by the change of unknown function $\widetilde{W}^{\psi,h}(X,t) = e^{\lambda t} W^{\psi,h}(X + V_I^{\psi}.t, t)$, we have

$$\begin{aligned}\frac{\partial \widetilde{W}^{\psi,h}}{\partial t} &= e^{\lambda t}\left(\frac{\partial W^{\psi,h}}{\partial t} + \lambda W^{\psi,h} + V_I^{\psi}.\nabla W^{\psi,h}\right)(X + V_I^{\psi}.t, t) \\ &= \lambda\, e^{\lambda t} \int d\psi'\, k(\psi' - \psi)\, Z^{\psi'}(X - V_I^{\psi}.t, t)\end{aligned}$$

that makes the resolution reduce to an integration.

Our efforts have thus concentrated on the integro-differential equation (2.9). We present in the next sections two possible numerical procedures for the solution if this equation.

2.3 Numerical method based on finite differences

2.3.1 Description of the scheme used

Finite difference methods [3, 11] can be mixed with interpolation methods to obtain from system (2.9) a numerical approximation of $Z^{\psi}(X,t)$. To approximate the equation, one must keep in mind that we deal with probabilities, therefore all approximations made should be conservative, in the sense that the numerical solution must have a total mass equal to 1. The integro-differential equation is thus approximated in three ways. First the continuous variables X and ψ are discretized, with respective steps ΔX and $\Delta\psi$. Calling ψ_i the points of the grid used for ψ, the integral term is replaced by a Riemann sum, in which the probability $k(\psi' - \psi)$ is approximated with

$$\tilde{k}(\psi_j - \psi_i) = s_{N_\psi}\, k(\psi_j - \psi_i)$$

where s_{N_ψ} is a constant depending on the number N_ψ of points ψ_i in the grid, and insuring $\sum_i \Delta\psi\, \tilde{k}(\psi_i - \psi_j) = 1$ for every ψ_j. As N_ψ becomes larger, s_{N_ψ} tends to 1. The second order derivatives are approximated by finite differences computed on a grid X_{lm}, such that $X_{l+1,m} - X_{l,m} = X_{l,m+1} - X_{l,m} = \Delta X$. The finite differences used here have first order of accuracy. As for the gradient of the first term, we use instead the property :

$$\lim_{\Delta t \to 0} \frac{Z^{\psi}(X - V_{H,I}\,\Delta t, t) - Z^{\psi}(X,t)}{\Delta t} = -V_{H,I}.\nabla Z^{\psi}(X,t) \tag{2.11}$$

To evaluate the first term of the denominator, in the case when $X - V_{H,I}\,\Delta t$ is not a point of the space grid, we linearly interpolate the values of Z^{ψ} taken on at the three nearest vertices of the grid.

Also discretizing time with a step Δt, we come up with the following approximation of the problem (2.9) :

$$Z^{\psi_i}(X_{lm}, t_{r+1}) = \bar{Z}^{\psi_i}(X_{lm} - V_{H,I}\,\Delta t, t_r) + \Delta t\,\big(\text{Diffusion terms} + \text{Coupling terms}\big) \tag{2.12}$$

where $\bar{Z}^{\psi_i}(X_{lm} - V_{H,I}\,\Delta t, t_r)$ is the triangle-based linear interpolation of Z^{ψ_i} at the point $X_{lm} - V_{H,I}\,\Delta t$, and with

$$\begin{aligned} \text{Diffusion terms} &= \tfrac{\sigma_v^2}{2}\big(\cos^2\psi_i\, Z_{xx}^{\psi_i}(X_{lm}, t_r) + 2\sin\psi_i\cos\psi_i\, Z_{xy}^{\psi_i}(X_{lm}, t_r) + \sin^2\psi_i\, Z_{yy}^{\psi_i}(X_{lm}, t_r)\big) \\ \text{Coupling terms} &= \lambda\Delta\psi\textstyle\sum_j \tilde{k}(\psi_j - \psi_i)\big(Z^{\psi_j}(X_{lm}, t_r) - Z^{\psi_i}(X_{lm}, t_r)\big) \end{aligned} \tag{2.13}$$

if we denote $Z_{xx}^{\psi_i}$, $Z_{xy}^{\psi_i}$ and $Z_{yy}^{\psi_i}$ the second-order differences approximating the second-order derivatives of Z^{ψ} with respect to X.

Since on one hand the linear interpolation satisfies with excellent approximation

$$\sum_{lm} \bar{Z}^{\psi_i}(X_{lm} - V_{H,I}\,\Delta t, t_r) \approx \sum_{lm} Z^{\psi_i}(X_{lm}, t_r)$$

for any ψ_i and any t_r, and, on the other hand,

$$\sum_{lm} \text{Diffusion terms}(X_{lm}) \approx 0$$

with also excellent approximation if we choose a domain of resolution large enough (then the probability Z^{ψ} has very small values on the boundaries), we can verify that, with excellent approximation, the total mass of the probability is conserved by the scheme :

$$\forall r, \quad \sum_{i,lm} \Delta\psi\,\Delta X^2\, Z^{\psi_i}(X_{lm}, t_r) \approx 1 \tag{2.14}$$

The sum is in fact slightly less than 1, due to losses at the boundaries, as is required for convergence.

2.3.2 Stability and convergence

It is well-known that a correct choice of the parameters of the mesh ΔX, Δt and $\Delta\psi$ allows the stability of the scheme. In particular, Δt must be chosen such that the coupling terms and the diffusion terms, that may have negative values, remain smaller in magnitude than the interpolated value $\bar{Z}^{\psi_i}(X_{lm})$. With such Δt, the positivity of the probability distribution is therefore conserved

by the scheme. We have translated the equation by $-V_{H,I}(\psi = 0)\, t$ to prevent from very high drift terms and thus obtain more stability. Then the resolution is made in a moving referential, whose velocity is exactly the nominal speed of the plane, and it is easy to perform a change of coordinates to recover the solution in the fixed referential (or in the referential of one plane). An approximate analysis (following classical methods [11]) gives a rough range of stability for the scheme : with $V = |V_{H,I}|$, the interpolated term can be, in first approximation, assimilated to

$$\bar{Z}^{\psi_i}(X_{lm}) \approx Z^{\psi_i}(X_{lm}) - V\,\Delta t\,\big((1-\cos\psi_i)(Z^{\psi_i}(X_{l,m}) - Z^{\psi_i}(X_{l-1,m})) \\ + \sin\psi_i(Z^{\psi_i}(X_{l,m}) - Z^{\psi_i}(X_{l,m-1}))\big)$$

Suppose that the ratio of positive numbers $Z^{\psi_i}(X_{l,m})/Z^{\psi_i}(X_{l,m-1})$ remains between $\frac{1}{\epsilon}$ and $\epsilon > 0$, then the stability condition reads

$$\Delta t\,\big(\frac{V}{\Delta X}\sup_{\psi}((1-\cos\psi) + \frac{1}{\epsilon}\sin|\psi|) + \lambda + \frac{2\sigma^2}{\Delta X^2}\big) \leq 1$$

using the fact $\sum_i \Delta\psi\,\tilde{k}(\psi_i - \psi_j) = 1$ for every ψ_j by the choice of $\tilde{k}$. The factor $\sup_\psi(1-\cos\psi + \frac{1}{\epsilon}\sin|\psi|)$ is large, and can be bounded by $2 + \frac{1}{\epsilon}$. For $\lambda = 4\,\mathrm{h}^{-1}$, $\sigma^2/\Delta X^2 = 225\,\mathrm{h}^{-2}$, $\Delta X = 1\,\mathrm{nmi}$ as in previous works[1, 17], and an ϵ equal to .1, the stability condition gives $\Delta t \leq .58\,\mathrm{s}$. The inefficiency of the method is clear. With ϵ close to 1, the condition releases to a better, but still bad $\Delta t \leq 2.6\,\mathrm{s}$. The Figure (2-1) shows occurrence of instability at $t = 285\,\mathrm{s}$ with $\Delta t = 5\,\mathrm{s}$.

If stability is fulfilled, one can bound the maximum pointwise error between the exact solution and the numerical one. An estimate of the error is found through the computation of the truncation error $\tau(\Delta X, \Delta t, \Delta\psi)$ of the scheme. We find :

$$|\tau(\Delta X, \Delta t, \Delta\psi)| \leq C_1\Delta t + C_2\,\Delta X^2 + C_3\Delta\psi \tag{2.15}$$

where C_1,C_2 and C_3 are three positive constants depending on the maximum values of the initial condition $Z^\psi(X, t = 0)$ as well as its derivatives with respect to time, space and heading. Here the Dirac initial distribution in ψ is a problem, because C_3 should be infinite. In practice, we can replace the Dirac $\delta(\psi)$ by a narrow peak, with a finite C_3. From the estimate of the truncation error and the conservative property of the scheme (2.14), the following estimate of the maximum error is derived

$$\max_{l,m,i,r} |Z^{\psi_i}_{\text{num.}}(X_{lm}, t_r) - Z^{\psi_i}_{\text{exact}}(X_{lm}, t_r)| \leq (C_1\Delta t + C_2\,\Delta X^2 + C_3\,\Delta\psi)\,T$$

over the bounded time interval $[0, T]$. This distance therefore tends to zero as the parameters of the mesh go to zero, and this means convergence of the numerical solution. In practice, the constant C_3 has proven very high, not to mention $C_{1,2}$: typically $C_3 \approx 150\,\mathrm{rad}^{-1}$ in our context. Even with the

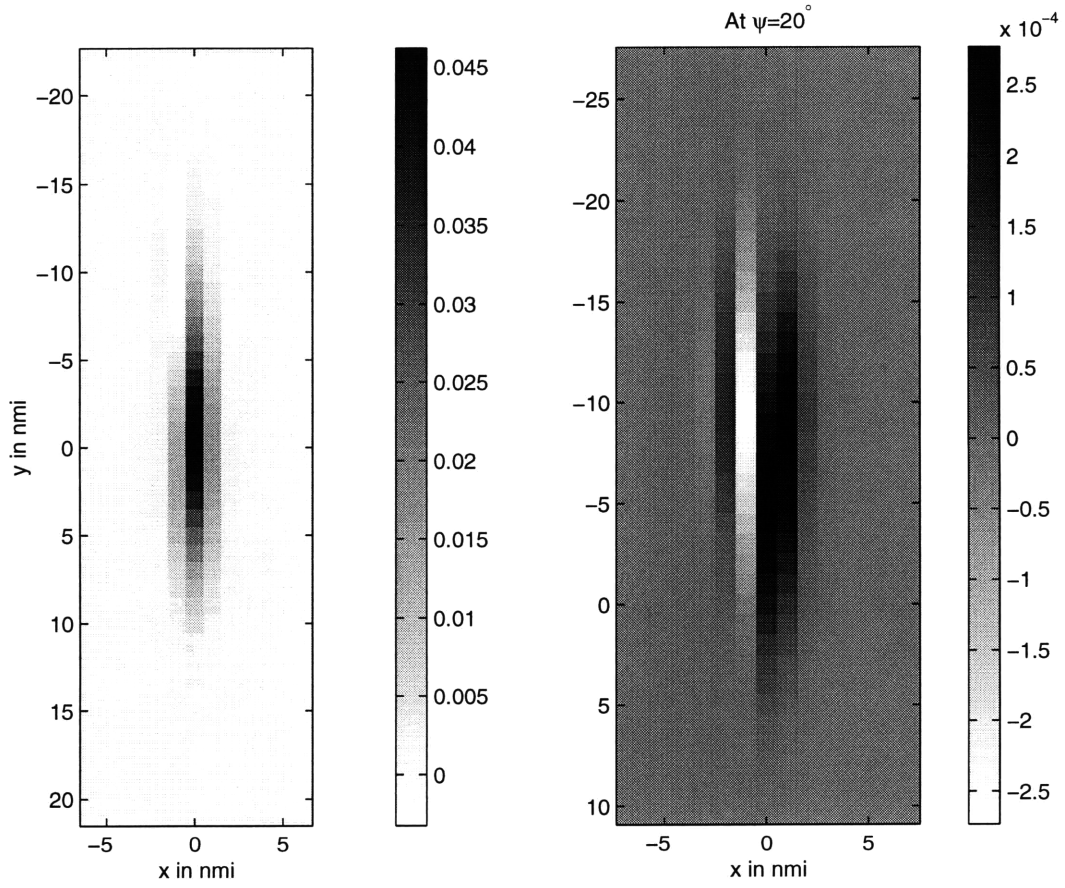

Figure 2-1: Probability density function of the intruder at $t = 285\,\mathrm{s}$, with instability of the numerical scheme. Here $\Delta t = 5\,\mathrm{s}$. The instability is particularly important for the component at $\psi = 20°$, as the strip of negative values shown on the right testifies.

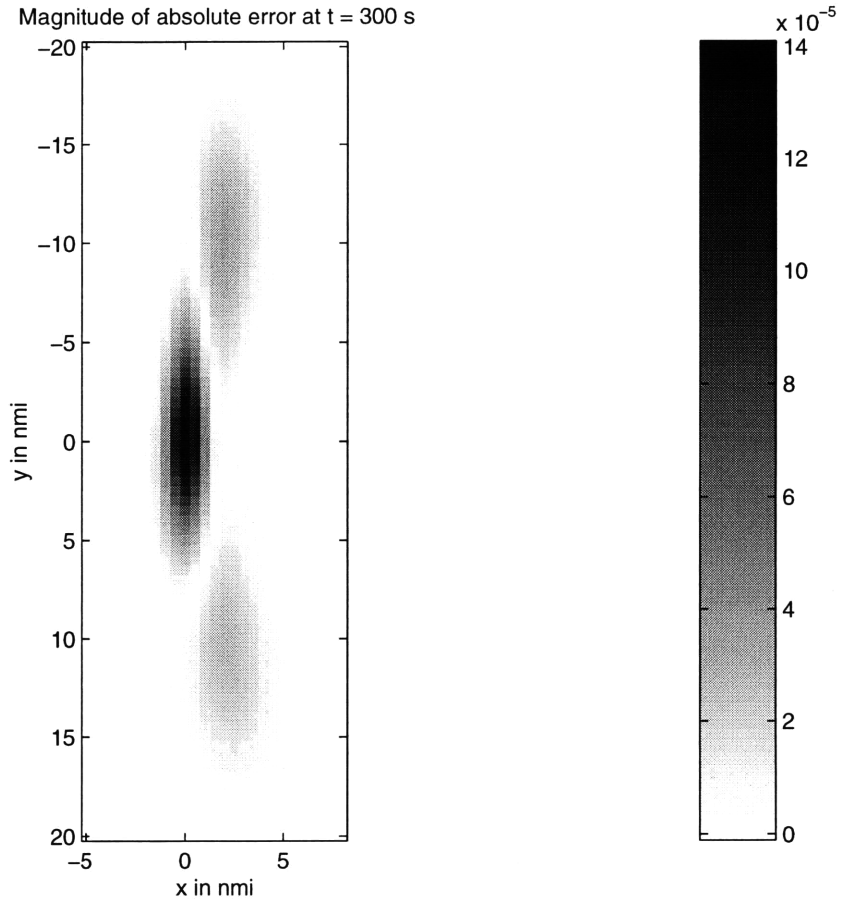

Figure 2-2: Comparison of the computation of the probability density function of the intruder at $t = 300\,\mathrm{s}$, with a coarse and a narrow grid of headings ψ (respectively 61 and 241 points). The maximum absolute error is $1.4\,10^{-4}$. It is therefore useless to compute the probability with the narrow grid.

result of convergence, the method fails to prove accurate. To illustrate this fact, we have computed the same solution for a narrow grid and a coarse grid of headings, with the same choice of Δt and ΔX in both cases (see Figure (2-2)). Though the absolute error between the two numerical solutions seems very small (less than $2\,10^{-5}$ at peak values), the relative error is still large : up to 40 % near the boundaries.

2.3.3 Numerical solutions

In practice, to obtain more stability, we have decreased the standard deviation σ to $5\,\mathrm{nmi.h}^{-1}$. One gains indeed on the stability bound. An argument is that the precision of position sensors shall improve in the future – as suggested by authors before [1].

We have achieved good stability by choosing values such that $\Delta t = 5\,\mathrm{s}$, $\Delta X = .5\,\mathrm{nmi}$, and various cases of $\Delta\psi$. The reason for diminishing ΔX instead of increasing it as the stability condition would suggest, is that the smaller ΔX, the closer ϵ to 1 – and this is the leading term of the bound.

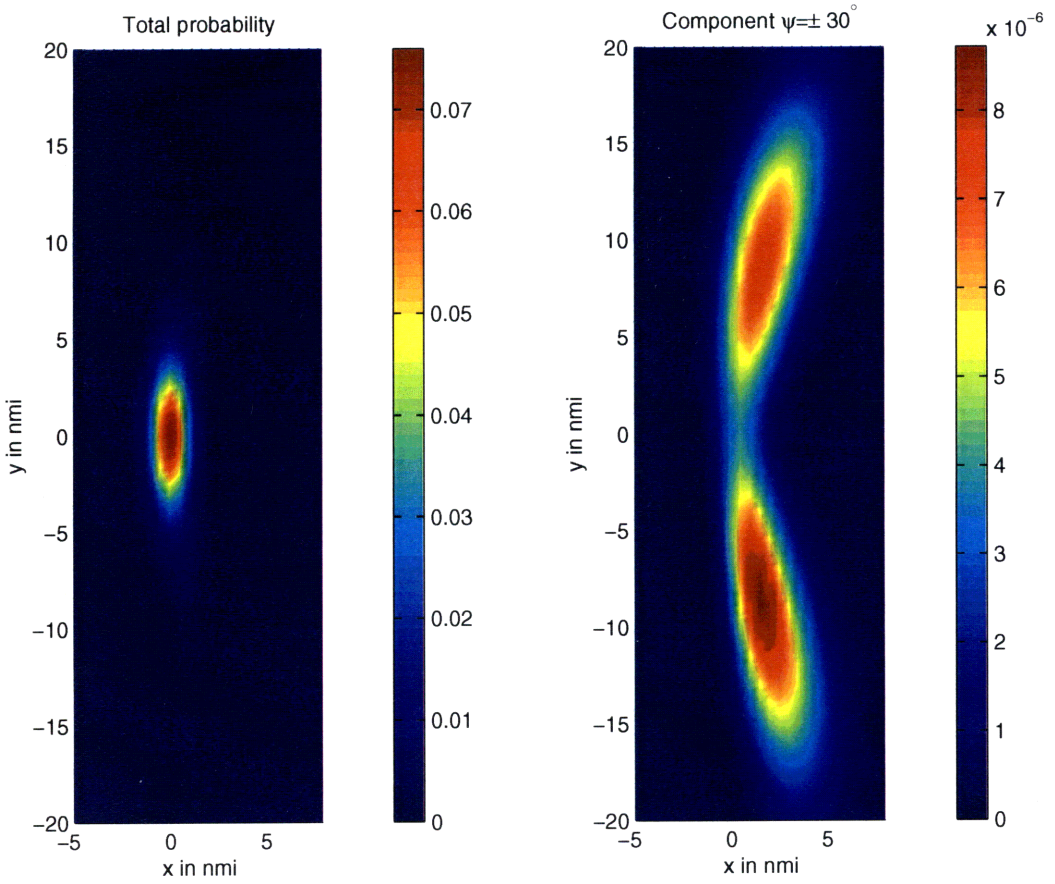

Figure 2-3: Probability density function of the intruder at $t = 300\,\mathrm{s}$. On the right : component $Z^{-30^\circ}(X, t = 300\,\mathrm{s})$. Nominal speed parallel to the x-axis, oriented to the right.

Figures (2-3)–(2-4) show what density $Z(X,t)$ is obtained at $t = 300\,\mathrm{s}$ (60 iterations, requiring about 60 minutes of computation on a 300 MHz PC, with vectorized code under Matlab V). The distribution of a component $Z^{\psi}(X,t)$ with $\psi = -30^\circ$ is also shown. This computation was made with 61 components, from -30° to 30°. On the 3-D plot, two tails can be seen, having developed from the initial Gaussian distribution under the effect of the Poisson transfers. The mean direction of those tails is roughly perpendicular to the direction of the flight, since at $\psi = \pm 90^\circ$, the y-component $V \sin\psi$ of the drift is maximum, occurring the most intense evasion of probability in that direction.

2.3.4 Fourier Transform

As we stated before, the system (2.9) is linear. We can therefore considerably reduce its complexity using Fourier transforms, and references show how this method is related to the search for the eigenvectors and eigenvalues of the linear operator of the problem. We shall abandon the superscript notation for ψ and will write it as a variable like X instead.

Here we group the variables of the problem in the following way : on one hand the equation is

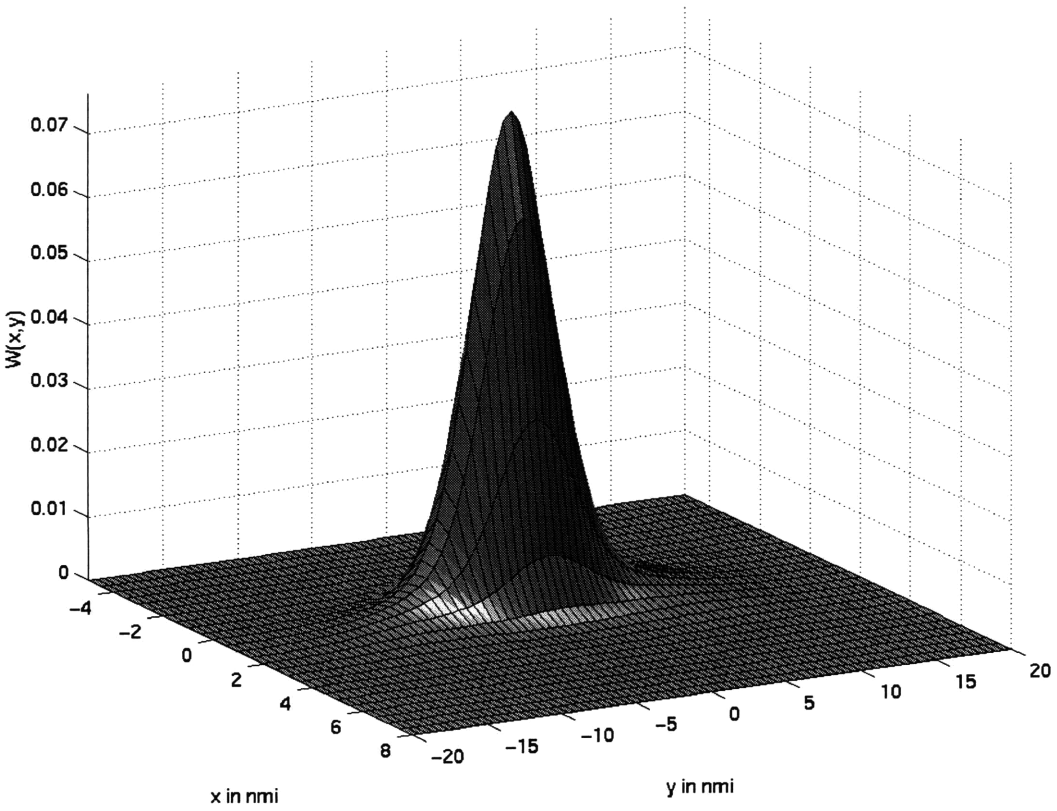

Figure 2-4: Probability density function of the intruder at $t = 300\,\mathrm{s}$ – 3-D plot.

2π-periodic in the parameter ψ and we shall apply an adapted Fourier transform to this variable, on the other hand we are in the presence of X with respect to which the solution is not expected to be periodic at all. We will apply to X a classical Fourier transform. As usual, time is kept apart any transform.

Let us denote $\widehat{Z}(\xi,t)$ the mixed Fourier transform of $Z(X,t,\psi)$ defined by

$$\widehat{Z}(\xi,t) = \{\hat{z}_n(\xi,t)\}_{n\in\mathbb{Z}}$$

where

$$\hat{z}_n(\xi,t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\psi \int dX\ Z(X,t,\psi)\, e^{-i(\xi.X+n\psi)}$$

The space integration is performed over the whole space $\mathbb{R}^2$. The dot . stands for the inner product in $\mathbb{R}^2$. We have the converse formula

$$Z(X,t,\psi) = \sum_{n\in\mathbb{Z}} \int d\xi\ \hat{z}_n(\xi,t)\, e^{i(\xi.X+n\psi)}$$

The probability density function $k(\phi)$ describing possible turns $\phi = \psi' - \psi$ from actual heading ψ must be 2π-periodic as well. It therefore has the Fourier series :

$$k(\phi) = \sum_{n\in\mathbb{Z}} \hat{k}_n\, e^{i\,n\,\phi}$$

with $\hat{k}_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi\, k(\phi)\, e^{-i\,n\,\phi}$. Observing that in (2.9) the last term is the convolution of $Z(X,t,.)$ by $k(.)$ at heading ψ, we have, by mixed Fourier transform and identification term by term

$$\frac{\partial \hat{z}_n}{\partial t}(\xi,t) + \lambda\,\hat{z}_n(\xi,t) = -i\xi.\left(\widehat{V} * \widehat{Z}(\xi,t)\right)_n + \lambda\,\hat{z}_n(\xi,t)\,\hat{k}_n \qquad \forall n\in\mathbb{Z}$$

where $\widehat{V}$ denotes the (two-dimensional) sequence of Fourier coefficients of the ψ-dependent speed ($*$ is the operation of sequence convolution). The speed for parameter ψ reads simply $V(\psi) = |V|\,(\cos(\psi),\sin(\psi))$, therefore we can explicit the sequence $\widehat{V}$:

$$\widehat{V}_m = \begin{cases} \frac{|V|}{2}(1,-i) & \text{if } m = 1 \\ \frac{|V|}{2}(1,i) & \text{if } m = -1 \\ 0 & \text{otherwise} \end{cases}$$

Hence

$$\left(\widehat{V} * \widehat{Z}(\xi,t)\right)_n = \frac{|V|}{2}\begin{pmatrix} \hat{z}_{n+1}(\xi,t) + \hat{z}_{n-1}(\xi,t) \\ -i\hat{z}_{n+1}(\xi,t) + i\hat{z}_{n-1}(\xi,t) \end{pmatrix}$$

If we denote $\xi = (\xi_x, \xi_y)$ when relevant, we get the following infinite coupled system :

$$\frac{\partial \hat{z}_n}{\partial t} = -\frac{|V|}{2}\left[(i\xi_x + \xi_y)\hat{z}_{n+1} + (i\xi_x - \xi_y)\hat{z}_{n-1}\right] + \lambda\,(\hat{k}_n - 1)\,\hat{z}_n \qquad \forall n \in \mathbb{Z} \tag{2.16}$$

that has to be solved for the family of functions $\{\hat{z}_n(\xi,t)\}_{n\in\mathbb{Z}}$. We may complete the system by computing the coefficients $\hat{k}_n$; between $-\pi$ and $-\pi$, the function k takes on the following values :

$$k(\phi) = c.\begin{cases} 1 & \text{for } \phi \in [-\phi_2, -\phi_1] \cup [\phi_1, \phi_2] \\ |\phi|/\phi_1 & \text{for } \phi \in [-\phi_1, \phi_1] \\ 0 & \text{otherwise} \end{cases}$$

where $\phi_1 = 5°$, $\phi_2 = 20°$, and c is a normalization constant (with $1/c = 2\phi_2 - \phi_1$). By symmetry of the real function k, it is not difficult to check that $\hat{k}_{-n} = \bar{\hat{k}}_n = \hat{k}_n$, therefore the $\hat{k}_n$'s are real and symmetric in n. Then

$$\hat{k}_n = \begin{cases} 1 & \text{if } n = 0 \\ \frac{c}{n\pi}\left(\sin(n\phi_2) + \frac{\cos(n\phi_1)-1}{n\phi_1}\right) & \text{if } n \neq 0 \end{cases}$$

(observe that $\hat{k}_0$ can be obtained from the general term when $n \neq 0$ by making n go to 0 in the expression). The leading term of $\hat{k}_n$ is thus $O(\frac{1}{n})$, and $\hat{k}_n$ is bounded. If we assume the function $Z(X,t,\psi)$ infinitely many times continuously differentiable, then the functions $\hat{z}_n(\xi,t)$ are rapidly decreasing as $|\xi|$ and n tend to ∞ (decrease faster than any positive polynomial in $|\xi|$ and n). The bounded coefficient $\hat{k}_n$ does not influence the fast decreasing of all quantities in equation (2.16).

Solving completely the system (2.16) in an analytical way seems out of range. However, the fast decreasing of all functions as n and $|\xi|$ tend to ∞ allows to truncate the system at ranks $\pm N$, for N large enough, and solve numerically on a bounded box for the variable ξ.

Back to the 2D problem we try to solve, recall that we assumed a Dirac initial distribution for the parameter ψ. Suppose, with no restriction on generality, that the Dirac is centered at $\psi_0 = 0$. Then the initial Fourier coefficients $\hat{z}_n(\xi, t = t_0)$ are independent of n : to be more precise, if $Z(X, t_0, \psi) = \delta(\psi).Z'(X)$, we have

$$\hat{z}_n(\xi, t = t_0) = \int dX\, Z'(X)\, e^{-i\,\xi.X} = \widehat{Z'}(\xi)$$

because the integration of the Dirac with respect to ψ only "selects" the value $\psi = 0$. But the system (2.16) is linear in the functions $\hat{z}_n$. Therefore the solution is directly proportional to $\widehat{Z'}(\xi)$, the coefficient of proportionality being the solution with uniform initial conditions 1.

The advantage of this property is that the family of function $\hat{z}_n(\xi,t)$ is then very easy to compute : it is sufficient to compute and keep *before hand* the family $\hat{u}_n(\xi,t)$ that is solution to the system

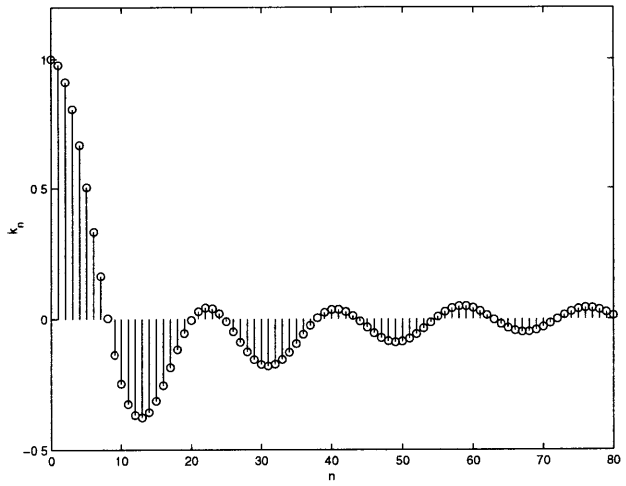

Figure 2-5: Fourier coefficients of the probability density function k, for $n \geq 0$ and up to 60. The decay rate is rather slow, due to a leading term $O(\frac{1}{n})$.

with initial conditions 1, and then to perform the simple multiplication $\hat{z}_n(\xi, t) = \widehat{Z'}(\xi)\, \hat{u}_n(\xi, t)$. We insist that the family $\hat{u}_n(\xi, t)$ is totally independent of the initial distribution of concern, and can be computed and memorized once for all. Of course, the family depends on $|V|, \lambda, \phi_1, \phi_2$.

Now, the advantage of this method becomes clear. To find the distribution $Z(X, t, \psi)$ on a grid, given the initial distribution $Z'(X)$ at $\psi = 0$, we need :

1. Perform the two-dimensional Fourier transform of $Z'(X)$, and store it in $\widehat{Z'}(\xi)$.
2. Perform the products $\widehat{Z'}(\xi)\, \hat{u}_n(\xi, t)$ for each ξ, and each t available.
3. Inverse Fourier transform the result in 2D, to recover $Z(X, t, \psi)$.

Direct and inverse Fourier transforms can be computed extremely fast, especially when the number of vertices in the grid is a power of 2. The multiplication stage is a very elementary operation that is not to take long time as well. The method must then be very quick. As for the accuracy reached, it is clear that the size of the mesh is the primary factor, since we can compute the functions $\hat{u}_n(\xi, t)$ with an accuracy as high as wanted (and for instance, sample a very fine solution, computed on a very narrow grid, to get the values taken by these functions on a coarser grid). The computation cost of the method is directly proportional to the number of vertices in the grid, times the number of instants at which we compute. Also remark that these instants are not submitted to be equally distributed over the time interval $[0, T]$, and therefore could be adapted at constant computational cost to get more accuracy.

The method, however, has a main drawback : this is the cost of the precomputation of the coefficients $\hat{u}_n(\xi, t)$. Indeed, to compute them, we need

1. Find the eigenvectors and eigenvalues of the truncated operator of the chain equation (2.16), and this for every ξ of the grid of interest.

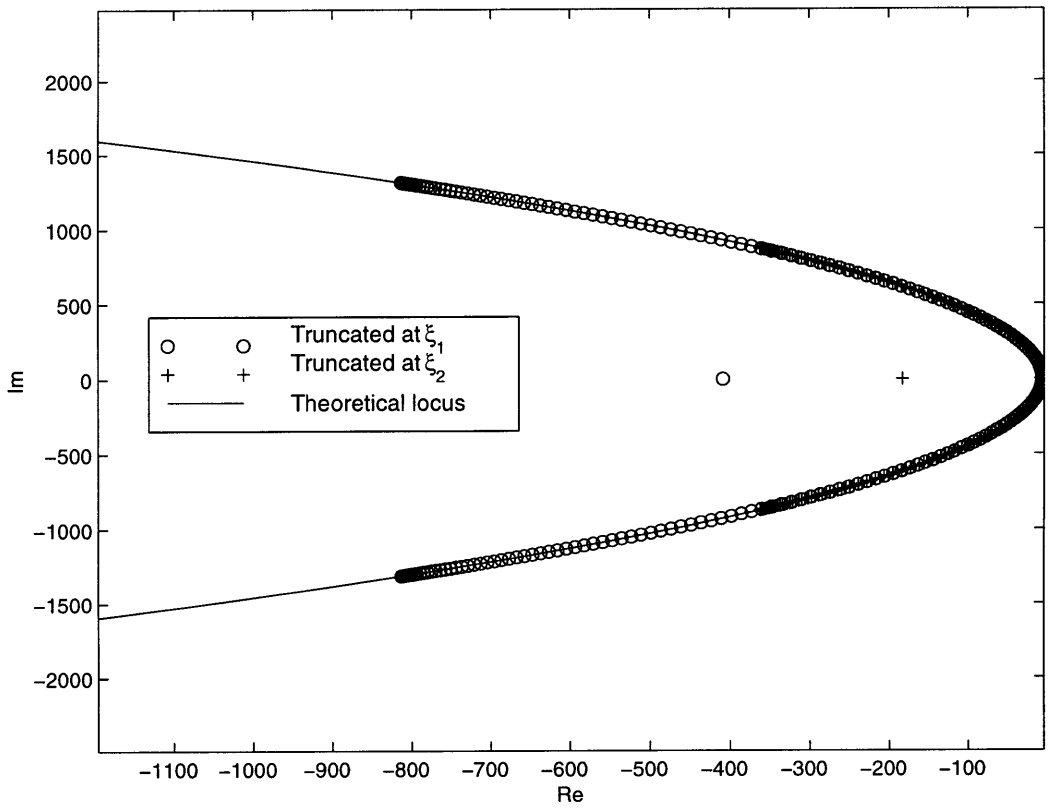

Figure 2-6: Distribution of the eigenvalues on a parabolic curve, independently of the point ξ of computation

2. Writing the $\hat{u}_n(\xi, t)$ as linear combinations of the eigenvectors, weighed by a decreasing exponential, find the coefficients of the combination, using the initial conditions. This requires the inversion of the matrix of the eigenvectors, again at each point ξ of the grid.

The first stage can be performed pretty quickly, because the matrix of system (2.16) is tridiagonal. The inversion of the matrix of the eigenvectors, however, is a heavy computation, because this matrix is neither symmetric nor sparse. One shows on figure (2-6) the distribution of the eigenvalues at an arbitrary point (ξ_x, ξ_y).

Moreover, the solutions behave quite differently whether we approximate the system in a way or another. We have indeed two choices : truncate the system at rank $\pm N$, with N large enough, or truncate the system at that N *and* make it circular : this would be done by adding to equation N a term in $\hat{z}_{-N}$ to replace the absent term $\hat{z}_{N+1}$, and by similarly adding to equation $-N$ a term in $\hat{z}_N$. This completion of the system, however, deeply perturbes the eigenvalues, even providing unstable eigenvalues if N is not large enough. This clearly endangers accuracy of solutions.

2.4 Collision probability

The modeling as seen in the previous sections had a major goal : find the probability density function of each plane of the scenery, as considered separately. The density function can be filtered, for the sake of position estimation, or other applications. However, this approach does not allow to compute the probability of collision for an arbitrary pair of planes, since in a current pair, the two aircraft are assumed *independent* of one another. The independence is lost whenever the two planes collide. Authors [6] have suggested the correct approach, that we may now explain.

2.4.1 Definition

Continuing on the same notation as before, let us define the difference position vector for two airplanes H and I :

$$\Delta(X_t, h_t) = (X_H(t) - X_I(t), h_H(t) - h_I(t))$$

Now, for a fixed, bounded, closed region D^c in the space, that will be called the *security region*, a collision occurs after $t_0 = 0$ whenever $\Delta(X_t, h_t)$ "hits" the boundary ∂D^c at some instant t. After the collision, the process $\Delta(X_t, h_t)$ remains inside the region forever. There therefore exists a random *first hitting time process* $\tau = \inf\{t : \Delta(X_t, h_t) \in D^c\}$ (with $\tau = +\infty$ if the set is empty). From the two random processes $\Delta(X_t, h_t)$ and τ, we define the *stopped process* ξ_t as [15] :

$$\xi_t = \begin{cases} \Delta(X_t, h_t) & \text{if } t < \tau \\ \Delta(X_\tau, h_\tau) & \text{if } t \geq \tau \end{cases}$$

with the property $t \geq \tau \Rightarrow \xi_t \in D^c$.

In this context, the collision probability of the two planes between the instants $t_1 < t_2$ is

$$\mathsf{P}\{\, \xi_{t_2} \in D^c \text{ and } \xi_{t_1} \notin D^c \,\}$$

provided that the planes at instant t_1 are separated by the security distance at least (diameter of the region D^c).

2.4.2 Evolution in time

To find what equations the collision probability satisfies, one may write [6] :

$$\begin{aligned} \mathsf{P}\{\, \xi_{t_2} \in D^c \text{ and } \xi_{t_1} \notin D^c \,\} &= \mathsf{P}\{\, \xi_{t_2} \in D^c\} - \mathsf{P}\{\, \xi_{t_2} \in D^c,\ \xi_{t_1} \in D^c \,\} \\ &= \mathsf{P}\{\, \xi_{t_2} \in D^c\} - \mathsf{P}\{\, \xi_{t_1} \in D^c\} \end{aligned}$$

decomposing the event $\{\xi_{t_2} \in D^c\}$ as the disjoint union of the sets $\{\xi_{t_2} \in D^c\} \cap \{\xi_{t_1} \in D^c\}$ and $\{\xi_{t_2} \in D^c\} \cap \{\xi_{t_1} \notin D^c\}$, and using the property : $\xi_{t_1} \in D^c \Rightarrow \xi_{t_2} \in D^c$.

Using the conditional expansion, we have for $t \geq 0$

$$\mathsf{P}\{\xi_t \in D^c\} = \int d\xi \, \mathsf{P}\{\xi_t \in D^c \mid \xi_0 = \xi\} \, \mathsf{P}\{\xi_0 = \xi\}$$

where $\mathsf{P}\{\xi_0 = \xi\}$ is the initial distribution of the separation vector, directly dependent on the knowledge of the respective initial positions of the planes. This distribution is completely determined by the position uncertainty as induced by the position sensors (GPS, etc...), taken into account in the model of aircraft behaviour. Therefore the only quantity to evaluate is $\mathsf{P}\{\xi_t \in D^c \mid \xi_0 = \xi\}$.

The stopped process ξ_t has continuous sample paths. When the headings of the two planes are fixed, and up to time τ, the stochastic differential equation (Itô's equation) that ξ_t follows is simply, with the same notation as above :

$$\begin{aligned} d\xi_t &= dX_I(t) - dX_H(t) \\ &= (V_I - V_H)\, dt + \sigma_v \sqrt{c}(u_I \, dW^1_{t/c} - u_H \, dW^2_{t/c}) \\ &= (V_I - V_H)\, dt + \sigma_v \sqrt{c} \begin{pmatrix} \cos\psi_I & -\cos\psi_H \\ \sin\psi_I & -\sin\psi_H \end{pmatrix} \begin{pmatrix} dW^1_{t/c} \\ dW^2_{t/c} \end{pmatrix} \end{aligned} \tag{2.17}$$

where u_H and u_I are the two unit vectors supporting the velocities of the aircraft, and $dW^{1,2}$ are two independent one-dimensional Brownian motions. After τ, the process, as we defined, is stopped in the region D^c. According to authors [6, 5], the continuity of the sample paths and the existence of an Itô's equation for the process are sufficient to prove that the probability $p(\xi, t) = \mathsf{P}\{\xi_t \in D^c \mid \xi_0 = \xi\}$, also known as a passage probability, is subject to the partial differential equation (with contracted notations allowed by the independence of the coefficients of the space variables) :

$$\frac{\partial p}{\partial t}(\xi, t) = (V_I - V_H).\nabla p(\xi, t) + \frac{\sigma^2}{2} \nabla . B(\psi_H, \psi_I) B(\psi_H, \psi_I)^T \nabla p(\xi, t) \tag{2.18}$$

where $B(\psi_H, \psi_I)$ is the matrix of equation (2.17). The product reads :

$$\mathcal{B} = B(\psi_H, \psi_I) B(\psi_H, \psi_I)^T = \begin{pmatrix} \cos^2\psi_H + \cos^2\psi_I & \sin\psi_H \cos\psi_H + \sin\psi_I \cos\psi_I \\ \sin\psi_H \cos\psi_H + \sin\psi_I \cos\psi_I & \sin^2\psi_H + \sin^2\psi_I \end{pmatrix}$$

Back to the problem, the process ξ_t with fixed *heading* – and by this, we mean the difference $\psi = \psi_I - \psi_H$ of the headings of the two planes – follows a regular stochastic differential equation associated with a (*forward*) Fokker-Planck equation. We have shown above how one is then led to

the following *backward* Kolmogorov equations :

$$\frac{\partial p^{\psi,h}}{\partial t}(\xi,t) = (V_H - V_I).\nabla p^{\psi,h}(\xi,t) + \frac{\sigma^2}{2}\nabla.\mathcal{B}\nabla p^{\psi,h}(\xi,t) + \int d\psi' dh' \lambda(\psi,h \to \psi',h') \left[p^{\psi',h'}(\xi,t) - p^{\psi,h}(\xi,t) \right] \quad (2.19)$$

as soon as the transition rates $\lambda(\psi,h \to \psi',h')$ can be explicited.

A particular case of the sceneries described above is when the host follows a straight route : ψ_H = constant, while the intruder has the behaviour depicted before. Then the transition rates $\lambda(\psi,h \to \psi',h')$ are those of the intruder, and do not depend on time in this case. If we consider the 2-D problem – on simplification purposes only – it is now easy to check that the passage probability satisfies :

$$\frac{\partial p^{\psi}}{\partial t}(\xi,t) = (V_I - V_H).\nabla p^{\psi}(\xi,t) + \frac{\sigma^2}{2}\nabla.\mathcal{B}\nabla p^{\psi}(\xi,t) + \lambda \int d\psi' k(\psi - \psi') \left[p^{\psi'} - p^{\psi} \right](\xi,t) \quad (2.20)$$

With no restriction on generality, we can suppose $\psi_H = 0$, whereby $\psi = \psi_I$ and the matrix $\mathcal{B}$ shall be written

$$\mathcal{B} = \begin{pmatrix} 1 + \cos^2\psi & \sin\psi\cos\psi \\ \sin\psi\cos\psi & \sin^2\psi \end{pmatrix}$$

For the separation process starts from outside the security domain D^c, and stops in any case as soon as it reaches its boundary, we have the following Dirichlet type boundary conditions [6] :

$$\begin{aligned} \lim_{\xi \to \partial D,\, \xi \in D} p^{\psi}(\xi,t) &= 1, \qquad \text{for } t > 0 \\ \lim_{t \downarrow 0} p^{\psi}(\xi,t) &= 0, \qquad \text{for } \xi \in D \text{ fixed} \end{aligned} \quad (2.21)$$

As we have shown in section 1.5.3, the unconditional passage probability at time t is

$$p(\xi,t) = \int d\psi\, p^{\psi}(\xi,t)\, \mathsf{P}\{\, \xi_{\xi_0,0}(0) \in \Sigma_{\psi} \,\}$$

(with obviously $\xi_{\xi_0,0}(0) = \xi_0$). Here, the initial state ψ_0 is well-known, and $\mathsf{P}\{\, \xi_{\xi_0,0}(0) \in \Sigma_{\psi} \,\} = \delta(\psi - \psi_0)$. Hence

$$p(\xi,t) = p^{\psi=\psi_0}(\xi,t)$$

2.4.3 Numerical computation by the finite element method

We have implemented a finite element method [11] to solve the system of equations (2.20) with the boundary conditions (2.21) described in the previous section for the probability of collision. We now briefly describe the method.

The domain $D = \mathbb{R}^d \setminus D^c$ is first restricted to a bounded subdomain D' such that $D^c \subset D' \subset D$, and sufficiently large to contain the region of interest. We then decompose D' into a finite *triangulation* $\mathcal{T}$, defined by a set of *node points* P_j and triangles T_k having the P_j's as vertices. Multiplying by some test function $\phi_i(\xi)$ and integrating over $\Omega = D'$, we can write :

$$\frac{\partial}{\partial t}\int_\Omega p^\psi\,\phi_i = \int_\Omega (V_I - V_H).\nabla p^\psi\,\phi_i + \frac{\sigma^2}{2}\int_\Omega \nabla.\mathcal{B}\nabla p^\psi\,\phi_i \\ + \lambda\int_\Omega\int d\psi' k(\psi-\psi')\left[p^{\psi'} - p^\psi\right]\phi_i \quad (2.22)$$

A test function $\phi_j(\xi)$ is by convention a continuous, real function of ξ such that $\phi_j(P_i) = \delta_i^j$ with the Kronecker notation, and is linear over each triangle T_k of $\mathcal{T}$. The fundamental idea of the method is to project the problem (2.20) onto a finite dimensional vector space of functions $\eta_j(\xi)$. Solving the problem is then defined as finding the linear combination

$$\bar{p}^\psi(\xi,t) = \sum_j \bar{p}_j(t)\,\eta_j(\xi) \quad (2.23)$$

that minimizes a distance from the exact solution $p^\psi(\xi,t)^*$ to the equation (2.20) at t. The distance used is traditionally inherited from the L^2-norm.

We denote v_x^ψ and v_y^ψ the coordinates of $V_I - V_H$, and $\mathcal{B}_{ab}^\psi$ the 4 components of the matrix $\mathcal{B}$, with $a, b \in \{x, y\}$. Using various forms of Stokes' equality [2], we can transform the differential terms of (2.22) as follows :

$$\int_\Omega \frac{\partial p^\psi}{\partial a}\,\phi_i = \int_{\partial\Omega} p^\psi\,\phi_i\,n_a\,ds \;-\; \int_\Omega p^\psi\,\frac{\partial\phi_i}{\partial a} \\ \int_\Omega \frac{\partial^2 p^\psi}{\partial a\partial b}\,\phi_i = \int_{\partial\Omega} \frac{\partial p^\psi}{\partial b}\,\phi_i\,n_a\,ds \;-\; \int_\Omega \frac{\partial p^\psi}{\partial b}\,\frac{\partial\phi_i}{\partial a} \quad (2.24)$$

with again $a, b \in \{x, y\}$, and n_a the corresponding coordinate of the outward normal vector along $\partial\Omega$. Since p^ψ must be equal to 1 along ∂D, and is left "free" along $\partial D' \setminus \partial D$, the first term of the first equality simplifies to

$$\int_{\partial D}\phi_i\,n_a\,ds \;+\; \int_{\partial D'\setminus\partial D} p^\psi\,\phi_i\,n_a\,ds$$

Here we take $\eta_j(\xi) = \phi_j(\xi)$. It is easy but tedious to check that the differential problem can be

written as the finite dimensional ordinary differential equation :

$$M\,\frac{d}{dt}\bar{P}^{\psi}(t) \;=\; \left(v_x^{\psi}\,(H^x - E^x) + v_y^{\psi}\,(H^y - E^y)\right)\bar{P}^{\psi}(t) + \frac{\sigma^2}{2}\sum_{a,b\in\{x,y\}} \mathcal{B}_{ab}^{\psi}\left(L_{ab} - R_{ab}\right)\bar{P}^{\psi}(t) + v_x^{\psi}\,G^x + v_y^{\psi}\,G^y + F^{\psi}(t) \quad (2.25)$$

in which $\bar{P}^{\psi}(t)$ is the vector of the coefficients $\bar{p}_j(t)$ of (2.23), and $M, E^x, E^y, H^x, H^y, L^{ab}, R^{ab}$ are matrices and G^x, G^y vectors such that

$$\begin{aligned} M_{ij} &= \int_{\Omega} \phi_i\,\phi_j & H_{ij}^a &= \int_{\partial D'\setminus\partial D} \phi_i\,n_a\,\phi_j\,ds \\ E_{ij}^a &= \int_{\Omega} \frac{\partial\phi_i}{\partial a}\,\phi_j & L_{ij}^{ab} &= \int_{\partial D} \phi_i\,n_a\,\frac{\partial\phi_j}{\partial b}\,ds \\ R_{ij}^{ab} &= \int_{\Omega} \frac{\partial\phi_i}{\partial a}\,\frac{\partial\phi_j}{\partial b} & G_i^a &= \int_{\partial D} \phi_i\,n_a\,ds \end{aligned} \quad (2.26)$$

for $a, b \in \{x, y\}$ and i, j indexing the vertices of $\mathcal{T}$. The vector $F^{\psi}(t)$ (the same length as $\bar{P}(t)$) is an approximation of the coupling term, obtained by discretization of the integral over the parameter ψ and the values of $p^{\psi'}$ at t.

To solve (2.25), we used the *Backward Euler Method* [11] : with a discretization of time $t_n = n\Delta t$ and ${\bar{P}^{\psi}}_n = \bar{P}^{\psi}(t_n)$, the equation is approximated by

$$\left[M - \frac{\sigma^2}{2}\,(L - R)\,\Delta t\right]{\bar{P}^{\psi}}_{n+1} \;=\; \left[M + (H - E)\,\Delta t\right]{\bar{P}^{\psi}}_{n} \;+\; \left(F^{\psi}(t_n) + G\right) \quad (2.27)$$

for each ψ of a discretization of the segment $[-\pi, \pi]$, and the ψ-dependent arrays :

$$\begin{aligned} E &= v_x^{\psi}\,E^x + v_y^{\psi}\,E^y & R &= \sum_{a,b\in\{x,y\}} \mathcal{B}_{ab}^{\psi}\,R_{ab} \\ H &= v_x^{\psi}\,H^x + v_y^{\psi}\,H^y & & \\ G &= v_x^{\psi}\,G^x + v_y^{\psi}\,G^y & L &= \sum_{a,b\in\{x,y\}} \mathcal{B}_{ab}^{\psi}\,L_{ab} \end{aligned} \quad (2.28)$$

The reason for separating x- and y-components of these arrays is to quicken the precomputation of the matrices before solving (2.27) – also known as *assembling* stage [11, 12, 8]. Indeed, the components shown do not depend on the value of ψ, while the coefficients $v_x^{\psi}, v_y^{\psi}, \mathcal{B}_{xx}^{\psi}, \mathcal{B}_{xy}^{\psi}, \mathcal{B}_{yx}^{\psi}, \mathcal{B}_{yy}^{\psi}$ are scalar.

As matrices of an assembled problem, E, H, L and R are sparse. Thus so is the premultiplicating matrix of the left-hand side of (2.27), whose inversion can be therefore performed by a fast method (Gaussian elimination for instance). To force the fulfillment of the boundary condition $p^{\psi}(\xi, t) = 1$ along ∂D, we have used a system reduction method, which consists of solving the system (2.27) only for the vertices in $\Omega \setminus \partial D$. The method is fully described in [12], chapter 4. The set of modules

implemented to perform the computation are listed in the Appendix.

On figure (2-7), the collision probability between 0 and 90 s is shown. The host is flying from West to East, while the intruder initially comes from the North ($\psi_H = 0, \psi_I^0 = -90°$). The velocities of the aircraft were taken equal to 300 kts. The computation was stable for a time step of 15 s (6 iterations), and needed about 550 s to complete on a 300 MHz PC under Matlab V. This stability was enforced by an artificial diffusion term, aimed at making the matrix $\mathcal{B}$ invertible for every ψ's, while it is normally singular at $\psi = 0$. The amplitude of the diffusion was limited to 10% of the total $\sigma_v = 30$ kts. The mesh generated had 2244 node points and 4320 triangles. The number of systems was 51.

The diffusion term added to the right-hand side of the equation (2.20) corresponds to an isotropic Brownian motion, with zero drift rate and a small (but nonzero) variance rate. This Brownian motion w_t is perturbing the stopped process ξ_t (the former adds to the latter), but one expects negligible changes of the probability of collision : indeed since w_t has no drift,

$$\mathsf{E}[\xi_t + w_t] = \mathsf{E}[\xi_t]$$

such that *in average* the perturbed trajectories behave as the nonperturbed ones. Studying the equation (2.20) with an artificial, isotropic diffusion term, one may easily convince oneself that the solution with artificial Brownian motion is not the same as for the pure process. The perturbation thus needs be small, and the variance rate is a good parameter for tuning it up.

The figure (2-7) testifies for the convection of a high probability of collision along the bisecting line of the initial angle between the two planes (45° in our case). This fact is in full agreement with the physical reality : when starting from two points located on the same bisecting line with the same speed, the two planes have a maximum probability of colliding, due to a rather small mean frequency of turns of the intruder. Of course, causality explains why there is a dyssymmetry of the distribution of probabilities about an axis orthogonal to the bisecting line : the two planes may collide only if their respective vector speeds are oriented such that the instant of collision would be in the future. There is a nonzero probability of collision however, for those planes starting with *uncausal* initial speeds, due to the Gaussian along-track effect of the wind (anisotropic diffusion term in (2.20)). Even with uncausal initial ground speeds, the planes driven by the wind may fly backward to one another, causing a collision.

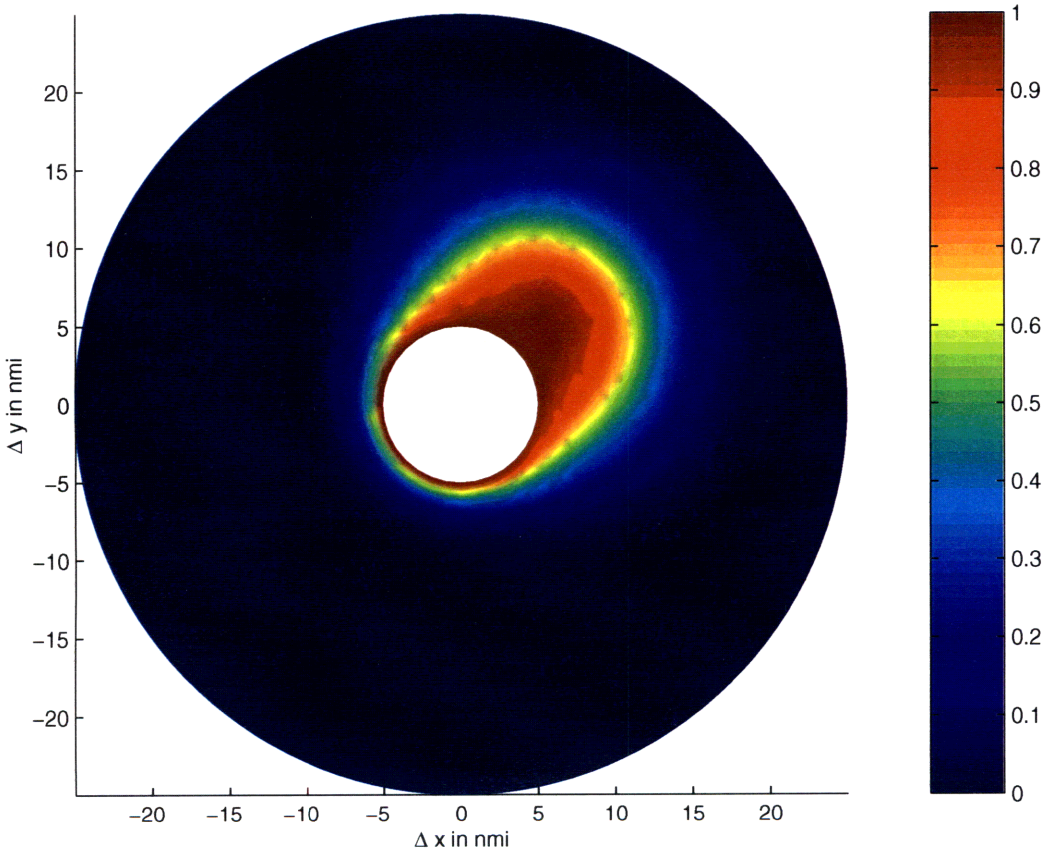

Figure 2-7: Collision probability at 90 s, computed with a time step of 15 s in about 550 s. The host is flying from West to East, the intruder from North to South.

Chapter 3

Perspectives

The theory exposed of posing differential equations to find conditionnal or *transition* probabilities of a transiting system is part of the more general theory of filtering. As usually admitted, filtering consists of finding estimates of the state of a system, taking measurements into account. In a first section, we clarify the stages of filtering, and show what is the position of our work in this frame. We also mention other approaches. Despite the capital role played by filtering in every practical fields of experimental sciences and in industry, it was out of our reach to explore every aspects of the question.

In a second part of this chapter, we rise questions more specific to probabilities in security systems – though the fundamental material remains of general interest. With again practical issues in mind, there is of course a need in the high-standard requirements of ATM to compute important statistics with high accuracy. In the case of a collision probability and the alert systems that are designed upon them [17], an upper bound instead of an approximate solution is for instance sufficient to determine secure thresholds of alert. Similarly, in many other fields of stochastic processes, only upper or lower bounds are wanted for the probabilities. One possible, but seemingly difficult perspective of this work thus consists of developing a practical and efficient technique of computing such bounds instead of the exact solution. Very trivial bounds (for instance the maximum of the probability) are of course useless; the main requirement is a convergence of the bound to the probability, as some parameter of the bound goes to zero. In two sections, we give a linear and a nonlinear approach. The style is deliberatly rather abstract, as a few examples are convincing that each step on the way to the technique sought needs much wariness.

3.1 Theory of filtering with incorporation of measurements

We briefly present the theory of filtering, in the particular case of continuous time systems with discrete time observations. An argument for keeping a continuous time in real-life systems, that are

yet always sampled, is the existence of differential equations satisfied by the conditionnal probabilities. These equations are solved approximately between two instants t_k and t_{k+1} or even at a higher sampling rate, but we ought not exclude from the study possible *exact* integrations between t_k and t_{k+1} in particular cases. Note that the equations we write in this section have a form adapted to recursive computations, to allow real-time evaluation of estimates.

Let $x(t)$ be a Markov process in $\mathbb{R}^d$, and t_k a sequence of consecutive instants of measurements y_k, sometimes called *observations* of the state $x(t)$. This terminology hides that the observation may be partial, *ie* some components of the state $x(t)$ cannot be measured. Criteria and issues of observability will not be examined here. We suppose no feed-back loop in the dynamical system, that is to say the observations y_k do *not* influence the dynamics of the system. Systems that would be controlled through an estimate are therefore excluded.

Another important hypothesis is that the measurements involve only noises that are independent in time (for instance Gaussian). The vector of all observations up to time t_k is denoted $y_{[0,k]}$. The same notation is used for the process $x(t)$, sampled at times t_k.

The purpose of the following calculus is to find a recursive expression of the conditionnal probability $d\mathsf{P}(x_{[0,k]}|y_{[0,k]})$, used to compute the estimate $\widehat{x}_{[0,k]|[0,k]}$ given the observation $y_{[0,k]}$. The result follows from two elementary steps :

1. Prediction

$$\begin{aligned} d\mathsf{P}(x_{[0,k]}|y_{[0,k-1]}) &= d\mathsf{P}(x_{[0,k-1]}, x_k|y_{[0,k-1]}) \\ &= d\mathsf{P}(x_k|x_{[0,k-1]}, y_{[0,k-1]})\, d\mathsf{P}(x_{[0,k-1]}|y_{[0,k-1]}) \end{aligned} \tag{3.1}$$

The second equality uses the definition of a conditionnal probability. Since the observation process $y(t)$ does not control the process $x(t)$, x_k does not depend on $y_{[0,k-1]}$, therefore

$$d\mathsf{P}(x_k|x_{[0,k-1]}, y_{[0,k-1]} = d\mathsf{P}(x_k|x_{[0,k-1]})$$

But the new conditionnal probability reduces to $d\mathsf{P}(x_k|x_{k-1})$, because $x(t)$ is a Markov process. We have then

$$d\mathsf{P}(x_{[0,k]}|y_{[0,k-1]}) = d\mathsf{P}(x_k|x_{k-1})\, d\mathsf{P}(x_{[0,k-1]}|y_{[0,k-1]}) \tag{3.2}$$

2. Correction through the measurement y_k

Using Bayès' rule for conditionnal probability, we have

$$\begin{aligned} d\mathsf{P}(x_{[0,k]}|y_{[0,k]}) &= d\mathsf{P}(x_{[0,k]}|y_{[0,k-1]}, y_k) \\ &= \frac{d\mathsf{P}(y_k|x_{[0,k]}, y_{[0,k-1]})\, d\mathsf{P}(x_{[0,k]}|y_{[0,k-1]})}{d\mathsf{P}(y_k|y_{[0,k-1]})} \end{aligned} \tag{3.3}$$

Since x_k contains all the information actually "seen" by y_k because of the independence of the measurements, we can simplify the first term. The equality reads :

$$d\mathsf{P}(x_{[0,k]}|y_{[0,k]}) = \frac{d\mathsf{P}(y_k|x_k)}{d\mathsf{P}(y_k|y_{[0,k-1]})}\, d\mathsf{P}(x_{[0,k]}|y_{[0,k-1]})$$

We get the recursive relation :

$$d\mathsf{P}(x_{[0,k]}|y_{[0,k]}) \;=\; z_k\; d\mathsf{P}(x_{[0,k-1]}|y_{[0,k-1]}) \tag{3.4}$$

with

$$z_k = \frac{d\mathsf{P}(y_k|x_k)\, d\mathsf{P}(x_k|x_{k-1})}{d\mathsf{P}(y_k|y_{[0,k-1]})} \tag{3.5}$$

In the definition (3.5) of the linear coefficient z_k, the term $d\mathsf{P}(y_k|x_k)$ is known through the modeling of the sensor. The denominator, independent of x_k and x_{k-1}, is a normalization factor with respect to y_k. The remaining term : $d\mathsf{P}(x_k|x_{k-1})$ is the transition probability of the process $x(t)$ between the state x_{k-1} at time t_{k-1} and the state x_k at time t_k. The main issue of filtering is often to evaluate this transition probability, since the other terms defining the coefficient are known. For instance, if the sensor is modeled with some uncorrelated, additive noise ν_k : $y_k = H(x_k) + \nu_k$ we have simply

$$d\mathsf{P}(y_k|x_k) = d\mathsf{P}^{\nu}\big(y_k - H(x_k)\big)$$

The value y_k is known by definition, and x_k is a variable, such that the probability is a function of x_k.

The position of the work developed in the first chapter is clear : we have precisely found the evolution equation of the transition probability of a transiting process. Hence, one could conceive a full filter bank based on the numerical methods we have proposed. The filter would first compute the transition probability $\mathsf{P}(x_k|x_{k-1})$ as a solution of the differential equations (1.5) with the initial condition $\delta(x_k - x_{k-1})$ (Green's function). Then the coefficient z_k would be evaluated, and the recursive relation (3.4) would be applied, to find the conditionnal probability. Applying well-known methods (least variance of error, maximum likelihood, etc...) it is then easy to compute various estimates of the state $x(t)$ at t_k.

The complexity of the differential equations seems to prevent from analytical forms of these

estimates, but this is no new difficulty : completely Gaussian systems are the only case known for allowing closed forms of optimal estimates with no approximation, by the method called Kalman filtering. An approximation method, called Particle filtering [7][1], is based on "smart" Monte-Carlo simulations of the system to find the transition probability. In both cases, the convergence of the estimates computed (in a suitable sense) can be proven in rather restrictive hypotheses for Kalman filtering, and more general hypotheses for Particle filtering. In our case, the convergence of the estimate would be a direct consequence of the convergence of the numerical solutions of the differential equations, no matter what method is used to find an approximate solution.

3.2 Pessimistic Numerical Schemes for Solving PDE's

In this section, we propose a linear approach to the finding of an upper bound $\overline{u_n}$ for the solution u of a partial differential equation. The bound is sought a solution of a recursive, linear equation : $\overline{u_{n+1}} = G_\eta\, \overline{u_n}$. The inequality $u \leq \overline{u_n}$ is valid on a discrete mesh only, for it would be meaningless elsewhere : thus we compare the mathematical solution of a differential equation with the solution of a recursive algorithm computing values on a space mesh. The idea of the approach is to find what sufficient conditions G_η shall meet, in relation with the Cauchy problem (identified to a semigroup G). On this purpose, we define *pessimistic* numerical schemes.

3.2.1 Definition of the problem

We consider, in the Banach space $(X, |.|)$, the linear Cauchy problem

$$\begin{cases} \frac{du}{dt} = A\,u, & t > 0 \\ u(0) = u_0 \end{cases} \tag{3.6}$$

to which u is called a *classical* solution if $u : \mathbb{R}^+ \longrightarrow X$ satisfies the ordinary differential equation for $t > 0$ and has the initial condition $u(0) = u_0 \in X$. A is assumed to be a linear, continuous operator of X, and we define

$$\| A \| = \sup_{v \in X, v \neq 0} \frac{|Av|}{|v|}$$

The theory of semigroups provides the following result : Cauchy initial value problem has a unique classical solution $u : \mathbb{R}^+ \longrightarrow X$, and it reads

$$u(t) = e^{A\,t}.u_0$$

[1] Unfortunately the reference cited in the text is in French. There is still no references in English on this efficient and fructuous method.

where $e^{A\,t}$ is the exponential of the linear, continuous (and thus bounded) operator $A.t$ at t : $e^{A\,t} = \sum_0^\infty (At)^n/n!$.

3.2.2 Definitions, Estimation of operators

Let Y be any Banach space. A family of continuous, linear operators $\{G_\eta\}_{\eta>0}$ is called an approximation of a fixed $G \in \mathcal{L}(Y)$ if

$$\exists \lim_{\eta\to 0} G_\eta = G \quad \Longleftrightarrow \quad \exists \lim_{\eta\to 0} \| G_\eta - G \| = 0$$

Now, suppose Y has *nonnegative* and *nonpositive* elements, and let Y^+ be the subset of all *nonnegative* elements (Y^+ is a cone). Let Z be another Banach space, with nonempty subsets Z^+ and Z^- of *nonnegative* and *nonpositive* elements. $G \in \mathcal{L}(Y, Z)$ is called nonnegative, with the notation $G \geq 0$ if

$$\forall x \in Y^+, \qquad G.x \in Z^+$$

In $\mathcal{L}(Y, Z)$, G is said to be greater than or equal to H if $G - H \geq 0$; in that case, we write $G \geq H$. In that context, a family $\{G_\eta\}_{\eta>0}$ in $\mathcal{L}(Y)$ will be called an *estimation from above* of a fixed G, if it is an approximation of G and

$$\forall \eta > 0, \qquad G_\eta \geq G$$

3.2.3 Numerical schemes

As proposed in [13]. For an easier reading, we recall the definitions and theorems used.

Let $\{(X_\eta, |.|_\eta)\}_{\eta>0}$ be a family of Banach spaces, and $\{P_\eta : X \longrightarrow X_\eta\}_\eta$ and $\{E_\eta : X_\eta \longrightarrow X\}_\eta$ be continuous, linear operators defined as follows :

1. The families $\{P_\eta\}_\eta$ and $\{E_\eta\}_\eta$ are uniformly bounded : exist N and N' such that $\forall \eta > 0$, $\| P_\eta \| \leq N$, and $\| E_\eta \| \leq N'$.
2. For any $x \in X$, $\exists \lim_{\eta\to 0} |P_\eta x|_\eta = |x|$.
3. For any $x \in X$, $\exists \lim_{\eta\to 0} |E_\eta\, P_\eta x - x| = 0$.
4. For all $\eta > 0$, $P_\eta\, E_\eta = I_\eta$.

DEFINITION : A family $\{A_\eta : D(A_\eta) \subseteq X_\eta \longrightarrow X_\eta\}_\eta$ of linear, continuous operators in respective subspaces $D(A_\eta)$ of the Banach spaces X_η is said to be convergent to A if there exists a domain

$$D(A) = \{x \in X : P_\eta x \in D(A_\eta), \text{ and } A_\eta\, P_\eta\, x \text{ converges as } \eta \to 0\}$$

in X. In that case, we define $A\,x$ as to be $\lim_{\eta\to 0} A_\eta\, P_\eta\, x$, for $x \in D(A)$. It is clear that $D(A)$ is a sub-vectorspace of X, and the operator A defined is linear.

The notion of limit in that definition needs also be defined : we say that any family $\{x_\eta\}_{\eta>0}$ of elements of X_η converges to $x \in X$ if $|x_\eta - P_\eta\, x|_\eta \to 0$ as $\eta \to 0$.

The convergence of A_η to A in that sense is denoted by $A_\eta \to\!\!\to A$. Each A_η in X_η will be considered a numerical scheme; in this frame, the parameter η may be seen as a mesh size.

THEOREM - DEFINITION :
Let $\{F_\eta\}_\eta$ be a family of bounded, linear operators F_η (numerical scheme), such that :

1. There exists a positive constant M and a real ω such that $\forall k \in \mathbb{N}, \| F_\eta^k \|_\eta \leq M\, e^{\omega \eta k}$.
2. $A_\eta = \frac{1}{\eta}(F_\eta - I_\eta) \to\!\!\to A$

If $D(A)$ is dense in X, and there exists $\lambda_0 \in \mathbb{C}$, with $\mathrm{Re}(\lambda_0) > \omega$ such that $\mathrm{Rg}(\lambda_0 I - A)$ be dense in X, then $\bar{A}$, the closure of the linear operator A is the infinitesimal generator of a C_0-semigroup $\{S(t)\}_{t\geq 0}$ on X (then $S(t) = \exp(t.A)$). If $k.\eta \to t$ as $\eta \to 0$ for a fixed and arbitrary $t \geq 0$, we have moreover :

$$F_\eta^k \to\!\!\to S(t) \qquad \text{as } \eta \to 0$$

and $D(S(t))$ is entire X.

In this context, the scheme F_η is called *pessimistic* if it satisfies the inequality :

$$\forall \eta > 0, \qquad F_\eta\, P_\eta \geq P_\eta\, e^{\eta A} \tag{3.7}$$

This implies, in particular, that $(F_\eta)^{t/\eta}$ converges to $S(t)$ *from above* as $\eta \to 0$.

3.2.4 Accuracy

DEFINITION : A numerical scheme $\{F_\eta : X_\eta \longrightarrow X_\eta\}_\eta$ is said to estimate a bounded, linear operator $S : X \longrightarrow X$ with accuracy $p > 0$, if there exists a positive constant C such that

$$\| F_\eta\, P_\eta - P_\eta\, S \|_{\mathcal{L}(X,X_\eta)} \leq C\, \eta^{p+1}$$

for all $\eta > 0$. The subscript $\mathcal{L}(X, X_\eta)$ is to recall that the operators $F_\eta\, P_\eta$ and $P_\eta\, S$ work from X to X_η. The norm used is in this case

$$\| B \|_{\mathcal{L}(X,X_\eta)} = \sup_{x\in X, x\neq 0} \frac{|Bx|_\eta}{|x|}$$

3.2.5 Approach proposed

At this point, it is important to note that the problem to solve is in the discrete space X_η, not in the continuous space X, for the simple reason that a comparison between an exact solution of the partial differential equation and a numerical solution has a sense only at node points (the points of X_η by definition). In the notations defined, *stricto sensu* a scheme F_η is pessimistic when

$$F_\eta \geq P_\eta \, e^{\eta A} \, E_\eta$$

For $E_\eta \, P_\eta$ is not I in X but only in the limit $\eta \to 0$, this definition is clearly distinct from (3.7). Here, we have slightly shifted the definition, to prepare for further work, where we believe the solution might be.

Now, we propose a strategy to find a pessimistic numerical scheme that would converge, in the sense defined earlier, to a given semigroup $S(t)$. This strategy is, basically, to solve the problem in the continuous space X, which is expected to be much easier, and then find a bridge between numerical schemes in X_η and approximations (from above) of $S(t)$ in X. Using the bridge, we would transport the approximation from X into X_η and get a numerical scheme with the expected properties. The hard part of the work in X (believed easier than in X_η) is to find how to reduce the condition on the approximation from above (that possibly also carries the notion of accuracy) to a condition easier to check (for instance on eigenvalues). Indeed, the previous definitions of "pessimism" – such as (3.7) – are inequalities between operators in infinite dimension.

The most obvious drawback of the approach is that it is well-adapted to finite difference methods, but possibly nothing else.

3.2.6 Bridge between estimations and numerical schemes

Suppose that $\{(\tau_{\eta,m}, \tilde{\tau}_{\eta,m})\}_{\eta>0, m\in\mathbb{Z}}$ is a family of linear operators in X and X_η respectively, such that for every $\eta > 0$ and every m,

$$P_\eta \, \tau_{\eta,m} \; = \; \tilde{\tau}_{\eta,m} \, P_\eta$$

(think of $\tau_{\eta,m}$ and $\tilde{\tau}_{\eta,m}$ as translations by m times a given vector in the continuous space X and the discrete space X_η).

Let G_η be a (formal) series of the $\tau_{\eta,m}$:

$$G_\eta = \sum_{m\in\mathbb{Z}} g_m \, \tau_{\eta,m}$$

This is a linear operator (if well defined) in X. Then let F_η be the transported operator into X_η :

$$F_\eta = \sum_{m\in\mathbb{Z}} g_m \, \tilde{\tau}_{\eta,m}$$

It is clear that

$$P_\eta \, G_\eta \;=\; F_\eta \, P_\eta$$

and it is not difficult to check that the formal

$$\Psi : G_\eta = \sum_{m\in\mathbb{Z}} g_m \, \tau_{\eta,m} \;\longmapsto\; F_\eta = \sum_{m\in\mathbb{Z}} g_m \, \tilde{\tau}_{\eta,m} = \Psi(G_\eta)$$

is a linear, one-to-one operator. Further properties satisfied by the g_m's and the $\tau_{\eta,m}$'s would allow to infere properties for Ψ itself. This is not the aim to do so here. Ψ is what we would call a *bridge* between approximations in X and numerical schemes in X_η. Note that the elements in those two sets, for which we define a correspondence, are very peculiar.

Now, suppose we know a G_η in $\mathcal{L}(X)$ satisfying

$$e^{\eta A} \;\leq\; G_\eta \;\leq\; e^{\eta A} + C\,\eta^{p+1}\, I \tag{3.8}$$

where I is the identity in X. Let us assume that for any $H \in \mathcal{L}(X), H \geq 0 \Rightarrow P_\eta \, H \geq 0$. We can plunge then the previous inequalities into X_η using P_η :

$$P_\eta \, e^{\eta A} \;\leq\; P_\eta \, G_\eta \;\leq\; P_\eta \, e^{\eta A} + C\,\eta^{p+1}\, P_\eta$$

If G_η was a series of the $\tau_{\eta,m}$'s, we can use the bridge and replace $P_\eta \, G_\eta$ by $\Psi(G_\eta) \, P_\eta$ in these inequalities, and obtain

$$P_\eta \, e^{\eta A} \;\leq\; F_\eta \, P_\eta \;\leq\; P_\eta \, e^{\eta A} + C\,\eta^{p+1}\, P_\eta \tag{3.9}$$

denoting again $F_\eta = \Psi(G_\eta)$. We recognize, written in terms of estimations of linear operators, the requirements of accuracy and "pessimism" the numerical scheme F_η should meet.

3.3 Variational approach

Following recent work [9], we describe what possible extension could be sought for, on the way to finding upper (lower) bounds for the solution to a linear system. The method would be based upon the Minimax Theorem, as applied to an augmented Lagrangian. We will not recall the definitions of the notations used, since they belong to a well-known corpus of Analysis.

Let V be a real Hilbert space, L a continuous, linear form on V, and a a continuous, *coercive*, bilinear form on V. We define the linear problem to solve as :

$$\text{Find } u \in V \text{ such that } \forall v \in V, \quad a(u,v) = L(v)$$

$L(v)$ is often the inner product of a fixed function f by v in V, and indeed we will assume in the following :

$$L(v) = \langle f, v \rangle$$

Similarly, we suppose that a is symmetric and there exists a linear operator A such that

$$\forall v \in V, \quad a(v,v) = \langle Av, v \rangle$$

By the Lax-Milgram theorem, the problem (3.3) has a unique solution u in V. Note that even for the existence of a solution, the coercivity of a is crucial.

Let ϕ^* be a continuous, linear form on V. We define the linear output of the problem as to be $\phi^*(u)$, for u the solution to the problem (3.3). For instance, $\phi^*(v) = \langle \phi, v \rangle$, where ϕ is a C^∞ function with a very narrow support around x, and an integral equal to 1, that is to say, an approximation of δ_x.

The trick of the method is to find an *augmented* Lagrangian for the problem, that reduces to the output $\phi^*(u)$ at the saddle-point. We denote V^* the dual of V (with the requirement that the elements of V^* are continuous). Let κ be a positive real parameter, and consider in V the functional :

$$\mathcal{S}^{\pm}(v) = \kappa \left(a(v,v) - L(v) \right) \pm \phi^*(v)$$

By the uniqueness of the solution to (3.3), we have [9]

$$\pm \phi^*(u) = \inf_{\{v : L(v) = L(u)\}} \mathcal{S}^{\pm}(v)$$

and this constitutes a constrained minimization problem. One can show that, by the Minimax Theorem,

$$\begin{aligned} \inf_{\{v:L(v)=L(u)\}} \mathcal{S}^{\pm}(v) &= \inf_{\{v:L(v)=L(u)\}} \sup_{\{\mu^* \in V^*\}} \mathcal{L}^{\pm}(v,\mu^*) \\ &= \sup_{\{\mu^* \in V^*\}} \inf_{\{v:L(v)=L(u)\}} \mathcal{L}^{\pm}(v,\mu^*) \end{aligned} \tag{3.10}$$

where $\mathcal{L}^{\pm}$ is a functional on $V \times V^*$ defined by

$$\mathcal{L}^{\pm}(v,\mu^*) = \mathcal{S}^{\pm}(v) \; + \; \mu^*(Av - f)$$

We come up with the inequalities

$$\forall \mu^*,\ \forall \nu^*,\quad \inf_{\{v:L(v)=L(u)\}} \mathcal{L}^+(v,\mu^*) \leq \phi^*(u) \leq -\inf_{\{v:L(v)=L(u)\}} \mathcal{L}^-(v,\nu^*) \tag{3.11}$$

for u the solution of the problem (3.3). Denoting

$$\mathcal{R}^{\pm}(\mu^*) = \pm \inf_{\{v:L(v)=L(u)\}} \mathcal{L}^{\pm}(v,\mu^*)$$

the upper and lower bounds for the output are respectively $\mathcal{R}^-(\nu^*)$ and $\mathcal{R}^+(\mu^*)$. An optimization on the paramater κ allows to obtain bounds as tight as possible.

The estimation (3.11) holds in the continuous space V. As in the section (3.2), we can project the space V onto a finite dimensional subspace V_η, with $\eta > 0$ the characteristic size of the discretization. V_η may be a space of finite elements for instance. Under reasonnable hypotheses and for fixed ν^*, μ^*, we can define in V_η : $\mathcal{R}_\eta^-$ and $\mathcal{R}_\eta^+$ with $\nu_\eta^* \in V_\eta^*$ and $\mu_\eta^* \in V_\eta^*$, such that

$$\begin{aligned} \lim_{\eta\to 0} \nu_\eta^* &= \nu^* & \lim_{\eta\to 0} \mu_\eta^* &= \mu^* \\ \lim_{\eta\to 0} \mathcal{R}_\eta^-(\nu_\eta^*) &= \mathcal{R}^-(\nu^*) & \lim_{\eta\to 0} \mathcal{R}_\eta^+(\mu_\eta^*) &= \mathcal{R}^+(\mu^*) \end{aligned} \tag{3.12}$$

The notion of convergence of a sequence $\nu_\eta^* \in V_\eta$ to an element $\nu^* \in V$ is identical to that defined in the section (3.2.3). For $\epsilon > 0$ such that

$$\epsilon < \min\left\{ \mathcal{R}^-(\nu^*) - \phi^*(u)\,,\ \phi^*(u) - \mathcal{R}^+(\mu^*) \right\} \tag{3.13}$$

there exists η_0 depending on ϵ such that $\eta < \eta_0$ implies

$$\mathcal{R}_\eta^+(\mu_\eta^*) \leq \phi^*(u) \leq \mathcal{R}_\eta^-(\nu_\eta^*) \tag{3.14}$$

which is exactly the estimation wanted, since the bounds are directly derived from numerical solutions [9].

Though very tempful, the method has weaknesses. First, since $\phi^*(u)$ is unknown, one has to estimate the distances needed in (3.13) to obtain the bounds. Second, the parameter η_0 depends on ϵ, and this means one needs to know what is the dependence to actually compute the bounds in (3.14). Third, as an approximation of the Dirac's δ, ϕ^* also *estimates* $u(x)$ for fixed x, and this is another source of uncertainty.

On top of those technical problems, the computational cost of the method is obvious : for each point x of the mesh, a complete recomputation of the bounds must be performed, and this computation itself has a high price (computation and inversion of a matrix). In 2-D, the complexity

of the computation is such that it has encouraged authors [9] to estimate the bounds themselves (through a narrow-to-coarse mesh interpolation) !

Those drawbacks kept us from trying it, but further investigation may allow simplifications and more efficiency.

Conclusion

Two schools confront with one another in the field of stochastic processes. One of them is to search for analytical solutions of the differential equations satisfied by the distributions of probabilities. The other one is to evaluate these distributions by Monte-Carlo simulations. The advantage of the second approach is a high efficiency.

Under technical hypotheses, this work shows it is easy to find the equations of evolution of the probability density function of a *transiting* process. We give the procedure to construct the equations practically. In the discrete case, these equations are standard, coupled partial differential equations. In the continuous case, they involve differential terms and principal value integrals. In both cases, they are globally linear, and stable under reasonnable circumstances. From a practical point of view, one can approximate a continuous system by a discrete one, by the discretization of the Poisson parameter. At the crossing of the probability theory and analysis, we find properties such as the conservation law, which help simplify the equations and get less complex, but equivalent problems.

Well-known computational techniques can be used to find numerical solutions to the equations, and the increasing power of computation gives a concrete hope to obtaining accurate solutions in real time. Here, we have used finite difference and finite element methods. But many other methods such as finite recurrences are still to explore. From a practical point of view, strict upper bounds for the statitics computed are sometimes sought, but there is no efficient method to find such bounds yet. Monte-Carlo simulations do not offer the possibility of finding upper bounds.

Appendix A

MATLAB V modules for the finite element method

Module init_meq3.m

```
N=51; % Number of subsystems
Psimin=-150*pi/180;
Psimax=150*pi/180;

% Load geometry/boundary condition description matrices b g
load meq_geom.mat

% Initialize/refine mesh
[p,e,t]=initmesh(g);
[p,e,t]=refinemesh(g,p,e,t);
[p,e,t]=refinemesh(g,p,e,t);
% p=jigglemesh(p,e,t,'Opt','minimum','Iter',10);
np=size(p,2);

% Precomputation of reduction matrices
[QQ,GG,HH,RR]=assemb(b,p,e);
[kerH,kerHorth]=pdenullorth(HH);
```

```
% Parameters
% dt=10/3600
lambda=4
V=300
sigm=30

% Precomputation of the coupling matrix KK
Psi=linspace(Psimin, Psimax,N);
setK
q=length(K);
KK=toeplitz(K(ceil(q/2):ceil(q/2)+N-1));
sKK=sum(KK);

% Preassembling of invariant matrices
mastereq3
```

Module iter_meq3.m

```
% Iterates of the finite element methods in time
%
% Convention :  N components k=1...N denoted u(:,k)
% Starting point u0(:,1:N)  -- iter_meq3 --> u1(:,1:N)
%
% Pre-assembling of the problem is assumed. Assemble only the coupling terms
% (which depend on u0)
%
% Coupling matrix (Toeplitz) : KK

epsilon= .1  % Artificial stabilization term
f_tot=lambda*(u0*KK)*dPsi; % Note : KK is a symmetric matrix

for k=1:N,
  k,
  f=f_tot(:,k)-lambda*u0(:,k)*sKK(k)*dPsi;

   % Assemble matrix F
```

```
% Interpolated values of f, primarily given at node points
f_intrp=pdeintrp(p,t,f);

f_intrp=f_intrp.*ar/3;
F=  sparse(it1,1,f_intrp,np,1);
F=F+sparse(it2,1,f_intrp,np,1);
F=F+sparse(it3,1,f_intrp,np,1);



% Assemble psi-depending matrices
psi_curr=Psi(k);
G=Gx*(V*(cos(psi_curr)-1))+Gy*(V*sin(psi_curr)); % gamma = V [cos(psi)-1, sin(psi)]
E=(Ex-Hx)*(V*(cos(psi_curr)-1))+(Ey-Hy)*(V*sin(psi_curr));

% Be aware : formula true only for host flying from W to E (psiH=0)
% L=(1+cos(psi_curr)^2)*Lxx+sin(psi_curr)*cos(psi_curr)*(Lxy+Lyx)+sin(psi_curr)^2*Lyy;
% R=(1+cos(psi_curr)^2)*Rxx+sin(psi_curr)*cos(psi_curr)*(Rxy+Ryx)+sin(psi_curr)^2*Ryy;

L=L+epsilon*(Lxx+Lyy); % Stabilization term
R=R+epsilon*(Rxx+Ryy);

MM=M-sigm^2/2*dt*(L-R);
FF=(M-E*dt)*u0(:,k)+(F+G)*dt;

% Reduction of the system   MM*u1(:,k)=FF
% to force satisfaction of boundary conditions by u1

ud=full(kerHorth*((HH*kerHorth)\RR));
FFF=kerH'*(FF-MM*ud);
MMM=kerH'*MM*kerH;
u1(:,k)=kerH*(MMM\FFF)+ud; % Solve next time step


end;
```

Module mastereq3.m

```
% Master equation solved by finite element method

it1=t(1,:);
it2=t(2,:);
it3=t(3,:);

% All couples with elements in {1,2,3}
c123=[ 1 2 ; 2 1 ; 1 3 ; 3 1 ; 2 3 ; 3 2 ];

%-------------- Area contributions --------------
% Assemble x-component of matrix E
[ar,g1x,g1y,g2x,g2y,g3x,g3y]=pdetrg(p,t);
f1x=g1x.*ar/3;
f2x=g2x.*ar/3;
f3x=g3x.*ar/3;
% Non-diagonal elements
Ex=   sparse(it1,it2,f1x,np,np);
Ex=Ex+sparse(it1,it3,f1x,np,np);
Ex=Ex+sparse(it2,it1,f2x,np,np);
Ex=Ex+sparse(it2,it3,f2x,np,np);
Ex=Ex+sparse(it3,it2,f3x,np,np);
Ex=Ex+sparse(it3,it1,f3x,np,np);

% Diagonal elements are provably zero - except for boundary node points !
Ex=Ex+sparse(it1,it1,f1x,np,np);
Ex=Ex+sparse(it2,it2,f2x,np,np);
Ex=Ex+sparse(it3,it3,f3x,np,np);



% Assemble y-component of matrix E
f1y=g1y.*ar/3;
f2y=g2y.*ar/3;
f3y=g3y.*ar/3;
```

```
% Non-diagonal elements
Ey=   sparse(it1,it2,f1y,np,np);
Ey=Ey+sparse(it1,it3,f1y,np,np);
Ey=Ey+sparse(it2,it1,f2y,np,np);
Ey=Ey+sparse(it2,it3,f2y,np,np);
Ey=Ey+sparse(it3,it2,f3y,np,np);
Ey=Ey+sparse(it3,it1,f3y,np,np);

% Diagonal elements are provably zero, except for boundary node points
Ey=Ey+sparse(it1,it1,f1y,np,np);
Ey=Ey+sparse(it2,it2,f2y,np,np);
Ey=Ey+sparse(it3,it3,f3y,np,np);


% Assemble matrix Rxx
r12x=ar.*(g1x.*g2x);
r23x=ar.*(g2x.*g3x);
r31x=ar.*(g3x.*g1x);
Rxx=  sparse(it1,it2,r12x,np,np); % Off-diagonal elements
Rxx=Rxx+sparse(it2,it3,r23x,np,np);
Rxx=Rxx+sparse(it3,it1,r31x,np,np);
Rxx=Rxx+Rxx.';

r1x=ar.*(g1x.^2);
r2x=ar.*(g2x.^2);
r3x=ar.*(g3x.^2);
Rxx=Rxx+sparse(it1,it1,r1x,np,np);
Rxx=Rxx+sparse(it2,it2,r2x,np,np);
Rxx=Rxx+sparse(it3,it3,r3x,np,np);


% Assemble matrix Ryy
r12y=ar.*(g1y.*g2y);
r23y=ar.*(g2y.*g3y);
r31y=ar.*(g3y.*g1y);
```

```
Ryy=    sparse(it1,it2,r12y,np,np); % Off-diagonal elements
Ryy=Ryy+sparse(it2,it3,r23y,np,np);
Ryy=Ryy+sparse(it3,it1,r31y,np,np);
Ryy=Ryy+Ryy.';

r1y=ar.*(g1y.^2);
r2y=ar.*(g2y.^2);
r3y=ar.*(g3y.^2);
Ryy=Ryy+sparse(it1,it1,r1y,np,np);
Ryy=Ryy+sparse(it2,it2,r2y,np,np);
Ryy=Ryy+sparse(it3,it3,r3y,np,np);

% Assemble matrix Rxy
r12xy=ar.*(g1x.*g2y);
r23xy=ar.*(g2x.*g3y);
r31xy=ar.*(g3x.*g1y);
r21xy=ar.*(g2x.*g1y);
r32xy=ar.*(g3x.*g2y);
r13xy=ar.*(g1x.*g3y);

Rxy=    sparse(it1,it2,r12xy,np,np); % Off-diagonal elements
Rxy=Rxy+sparse(it2,it3,r23xy,np,np);
Rxy=Rxy+sparse(it3,it1,r31xy,np,np);

Rxy=Rxy+sparse(it2,it1,r21xy,np,np);
Rxy=Rxy+sparse(it3,it2,r32xy,np,np);
Rxy=Rxy+sparse(it1,it3,r13xy,np,np);

r1xy=ar.*(g1x.*g1y); % Diagonal elements
r2xy=ar.*(g2x.*g2y);
r3xy=ar.*(g3x.*g3y);
Rxy=Rxy+sparse(it1,it1,r1xy,np,np);
Rxy=Rxy+sparse(it2,it2,r2xy,np,np);
Rxy=Rxy+sparse(it3,it3,r3xy,np,np);

% Assemble matrix Rxy
```

```
r12yx=ar.*(g1y.*g2x);
r23yx=ar.*(g2y.*g3x);
r31yx=ar.*(g3y.*g1x);
r21yx=ar.*(g2y.*g1x);
r32yx=ar.*(g3y.*g2x);
r13yx=ar.*(g1y.*g3x);

Ryx=    sparse(it1,it2,r12yx,np,np); % Off-diagonal elements
Ryx=Ryx+sparse(it2,it3,r23yx,np,np);
Ryx=Ryx+sparse(it3,it1,r31yx,np,np);

Ryx=Ryx+sparse(it2,it1,r21yx,np,np);
Ryx=Ryx+sparse(it3,it2,r32yx,np,np);
Ryx=Ryx+sparse(it1,it3,r13yx,np,np);

r1yx=ar.*(g1y.*g1x); % Diagonal elements
r2yx=ar.*(g2y.*g2x);
r3yx=ar.*(g3y.*g3x);
Ryx=Ryx+sparse(it1,it1,r1yx,np,np);
Ryx=Ryx+sparse(it2,it2,r2yx,np,np);
Ryx=Ryx+sparse(it3,it3,r3yx,np,np);



% Assemble matrix M
aod=ar/12;
ad=2*aod;
M=  sparse(it1,it2,aod,np,np); % Off-diagonal elements
M=M+sparse(it2,it3,aod,np,np);
M=M+sparse(it3,it1,aod,np,np);
M=M+M.';
M=M+sparse(it1,it1,ad,np,np); % Diagonal elements
M=M+sparse(it2,it2,ad,np,np);
M=M+sparse(it3,it3,ad,np,np);
```

```
%------------ Boundary contributions ------------

% Indices of edges with domain on the right. 'ie' assumed nonzero
ie= find(e(6,:)==0 & e(7,:)>0);
e_r=e(:,ie);
nie=length(ie);
perm=[2:nie,1];

% Assemble vectors Gx and Gy (boundary condition : u=g=1)
ie1=e_r(1,:);
ie2=e_r(2,:);

dx=p(1,ie2)-p(1,ie1);
dy=p(2,ie2)-p(2,ie1);
d=sqrt(dx.^2+dy.^2); % Side length

nxe=-dy./d; % Outward norm vector
nye=dx./d;

% In the following, the index ie2 is a trick to designate node P(j+1) when
% dealing with segments [P(j),P(j+1)] and [P(j+1),P(j+2)]
Gx=sparse(ie2,1,.5*nxe.*d,np,1);
Gx=Gx+sparse(ie2,1,.5*nxe(perm).*d(perm),np,1);

Gy=sparse(ie2,1,.5*nye.*d,np,1);
Gy=Gy+sparse(ie2,1,.5*nye(perm).*d(perm),np,1);

% Preparation of data for further computation
gx1=[];
gy1=[];
gx2=[];
gy2=[];

% Index of the triangles with edges [e_r(1,:),e_r(2,:)], in this order
for k=1:nie,
  i=1;
```

```
  while (isempty(find(t(c123(i,1),:)==e_r(1,k) & t(c123(i,2),:)==e_r(2,k)) & (i<7)))
    i=i+1;
  end;
  if i==7, error('Impossible to find edge in triangles'); return; end;
  gx1=[gx1,eval(['g',num2str(c123(i,1)),'x(',num2str(k),')'])];
  gy1=[gy1,eval(['g',num2str(c123(i,1)),'y(',num2str(k),')'])];
  gx2=[gx2,eval(['g',num2str(c123(i,2)),'x(',num2str(k),')'])];
  gy2=[gy2,eval(['g',num2str(c123(i,2)),'y(',num2str(k),')'])];
end;

% Results of the last computation :
%  * gx1 contains the x-gradient of the element associated
% with vertice number 1 (ie1) in the boundary edge
% * gx2 contains the x-gradient of the element associated
% with vertice number 2 (ie2) in the boundary edge
%  etc...

% Assemble matrix Lxx

% Off-diagonal terms
l12xx=.5*d.*(nxe.*gx2);
l21xx=.5*d.*(nxe.*gx1);

Lxx=    sparse(ie1,ie2,l12xx,np,np);
Lxx=Lxx+sparse(ie2,ie1,l21xx,np,np);

% Diagonal terms
l1xx1=.5*d.*(nxe.*gx1);
l1xx2=.5*d.*(nxe.*gx2);
Lxx=Lxx+sparse(ie1,ie1,l1xx1,np,np);
Lxx=Lxx+sparse(ie2,ie2,l1xx2,np,np);



% Off-diagonal terms
l12yx=.5*d.*(nye.*gx2);
l21yx=.5*d.*(nye.*gx1);
```

```
Lyx=    sparse(ie1,ie2,l12yx,np,np);
Lyx=Lyx+sparse(ie2,ie1,l21yx,np,np);

% Diagonal terms
l1yx1=.5*d.*(nye.*gx1);
l1yx2=.5*d.*(nye.*gx2);
Lyx=Lyx+sparse(ie1,ie1,l1yx1,np,np);
Lyx=Lyx+sparse(ie2,ie2,l1yx2,np,np);


% Off-diagonal terms
l12xy=.5*d.*(nxe.*gy2);
l21xy=.5*d.*(nxe.*gy1);

Lxy=    sparse(ie1,ie2,l12xy,np,np);
Lxy=Lxy+sparse(ie2,ie1,l21xy,np,np);

% Diagonal terms
l1xy1=.5*d.*(nxe.*gy1);
l1xy2=.5*d.*(nxe.*gy2);
Lxy=Lxy+sparse(ie1,ie1,l1xy1,np,np);
Lxy=Lxy+sparse(ie2,ie2,l1xy2,np,np);

% Assemble matrix Lyy

% Off-diagonal terms
l12yy=.5*d.*(nye.*gy2);
l21yy=.5*d.*(nye.*gy1);

Lyy=    sparse(ie1,ie2,l12yy,np,np);
Lyy=Lyy+sparse(ie2,ie1,l21yy,np,np);

% Diagonal terms
l1yy1=.5*d.*(nye.*gy1);
l1yy2=.5*d.*(nye.*gy2);
```

```
Lyy=Lyy+sparse(ie1,ie1,l1yy1,np,np);
Lyy=Lyy+sparse(ie2,ie2,l1yy2,np,np);


%------------ Boundary contributions 2 ------------

% Indices of edges with domain on the left. 'ie' assumed nonzero
ie_l= find(e(6,:)>0 & e(7,:)==0);
e_l=e(:,ie_l);
nie_l=length(ie_l);
perm=[2:nie_l,1];


% Assemble vectors Gx and Gy (boundary condition : u=g=1)
ie1_l=e_l(1,:);
ie2_l=e_l(2,:);


dx_l=p(1,ie2_l)-p(1,ie1_l);
dy_l=p(2,ie2_l)-p(2,ie1_l);
d_l=sqrt(dx_l.^2+dy_l.^2); % Side length

nxe_l=dy_l./d_l; % Outward norm vector
nye_l=-dx_l./d_l;


% Assemble matrix Hx

% Off-diagonal terms
h12x=d_l/6.*nxe_l;

Hx=    sparse(ie1_l,ie2_l,h12x,np,np);
Hx=Hx+Hx.';

% Diagonal terms
h1x=d_l/3.*nxe_l;
```

```
Hx=Hx+sparse(ie1_l,ie1_l,h1x,np,np);


% Assemble matrix Hy

% Off-diagonal terms
h12y=d_l/6.*nye_l;

Hy=    sparse(ie1_l,ie2_l,h12y,np,np);
Hy=Hy+Hy.';

% Diagonal terms
h1y=d_l/3.*nye_l;
Hy=Hy+sparse(ie1_l,ie1_l,h1y,np,np);
```

Bibliography

[1] Russel A.Paielli and Heinz Erzberger. Conflict probability estimation for free flight. *Journal of Guidance, Control and Dynamics*, pages 588–596, May 1997.

[2] J.M. Bony. *Cours d'Analyse.* Ecole Polytechnique, "Edition du Bicentenaire", 1994.

[3] G.E. Forsythe. *Finite Difference-methods for Partial Differential Equations.* Wiley, 1960.

[4] C.W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences.* Springer-Verlag, 1985 (2nd ed.).

[5] I.I. Gikhman and A.V. Skorohod. *Stochastic Differential Equations.* Springer-Verlag, 1972. Translated from Russian by Kenneth Wickwire.

[6] G.J. and H.Blom. Air traffic risk modelling. In *Proceedings of the 32nd Conference on Decision and Control*, pages 1464–1469.

[7] G.Rigal. *Filtrage non linéaire, résolution particulaire et application au traitement du signal.* PhD thesis, Université Paul Sabatier, 1993.

[8] G.Strang. *Introduction to applied mathematics.* Wellesley-Cambridge Press, 1986.

[9] H.Vailong. A posteriori upper bounds for linear outputs. Master's thesis, Massachusetts Institute of Technology, 1997.

[10] K. Itô. *Lectures on Stochastic processes.* 1961. Reproduced from typewritten copy.

[11] J.L. Lions and P.G. Ciarlet. *Handbook of Numerical Analysis.* North Holland, 1989.

[12] Mathworks. *Partial Differential Equation Toolbox Reference Guide.* MATLAB V guide - Not available in libraries.

[13] A. Pazy. *Semigroups of Linear Operators and Their Applications to Partial Differential Equations.* Springer, Applied Mathematical Sciences, 1983 (2nd printing).

[14] H. Risken. *The Fokker-Planck Equation : Methods of Solution and Applications.* Springer-Verlag, 1989 (2nd ed.).

[15] S. Shreve and I. Karatzas. *Brownian Motion and Stochastic Calculus.* 1991 (2nd ed.).

[16] S.Karlin and H.M.Taylor. *Second Course in Stochastic Processes.* Academic Press. 1981.

[17] Lee C. Yang and James K. Kuchar. Prototype conflict alerting system for free flight. *Journal of Guidance, Control and Dynamics,* pages 768–773, July 1997.