# *Dirichlet forms methods*

# *for*

# *error calculus and sensitivity analysis*

Nicolas BOULEAU, Osaka university, november 2004

These lectures propose tools for studying sensitivity of models to scalar or functional parameters. A Dirichlet forms based language is developed in order to manage the propagation of the variances and the biases of errors through mathematical models. Examples will be given in physics, numerical analysis and finance.

In the two first lectures, the intuitive calculus of Gauss for the propagation of errors will be connected with the rigorous mathematical framework of error structures. Then the main features of error structures will be studied until infinite products in order to construct error structures on functional spaces and on the Monte Carlo space.

The third lecture will be devoted to error calculus on the Wiener space with the classical Ornstein-Uhlenbeck structure or generalized Mehler-type structures with applications to SDE and finance.

In the fourth lecture will be tackled the question of identifying an error structure. The main role of the Fisher information will be exposed and also other methods based on asymptotic Hopf-type theorems or based on Donsker invariance theorem.

Reference : N. Bouleau, *Error Calculus for Finance and Physics,* De Gruyter 2003.

# First lecture **Propagation of errors : from Gauss to Dirichlet forms**

A) Propagation of errors

- 1 Historical outlook (or landscape)
  - the propagation calculus of Gauss
  - its coherence property, non coherence of other formulae
- 2 The propagation of errors
  - by non linear maps : variances and bias
  - bias by quadratic map
  - intuitive error calculus
- 3 Examples of propagation calculations
  - Gaussian variables
  - triangle
  - oscillographe
- 4 Examples of dynamical systems sensitivity analysis
  - Feigenbaum transition
  - piecewise linear system
  - Lorenz attractor

B) Error structures

- Languages with or without extension tool : the example of probability theory
- Adding an extension tool to the language of Gauss : the idea
- Definition of error structures
- Examples
- Comparison of approaches

# Second lecture **Error structures and sensitivity analysis**

C) Properties of error structures

- recalling the definition
- Lipschitzian calculus
- images
- densities and Dloc
- finite and infinite products
- the gradient and the sharp
- Integration by part formulae

D) Application in simulation : error calculus on the Monte Carlo space

- structures with or without border terms
- Sensitivity analysis of a Markov chain

E) Application in numerical analysis

- sensitivity of an ODE to a functional coefficient
- comments on finite elements methods

## Third lecture **New tools for finance**

F) Error structures on the Wiener space

G) error structures on the Poisson space

H) sensitivity analysis of an SDE, application to finance

I) A non classical approach to finance

## Fourth lecture **Links with statistics and empirical data**

J) Identifying an error structure

- The link with the Fisher information and its stability
- Natural error structures for some dynamical systems, extension of the "arbitrary functions method"

K) Convergence in Dirichlet law and transition from finite to infinite dimension

- Central limit theorem
- Convergence of an erroneous random walk to the erroneous Brownian motion :

extension of Donsker theorem and applications.

# First lecture
## **Propagation of errors : from Gauss to Dirichlet forms**

A) Propagation of errors

- 1/ Historical outlook (or landscape)
    - the propagation calculus of Gauss
    - its coherence property, non coherence of other formulae
- 2/ The propagation of errors
    - by non linear maps : variances and bias
    - bias by quadratic map
    - intuitive error calculus
- 3/ Examples of propagation calculations
    - simulation of Gaussian variables
    - triangle
- 4/ Examples of dynamical systems sensitivity analysis
    - Feigenbaum transition
    - piecewise linear system
    - Lorenz attractor

B) Error structures

- Languages with or without extension tool : the example of probability theory
- Adding an extension tool to the language of Gauss : the idea
- Definition of error structures
- Examples
- Comparison of approaches

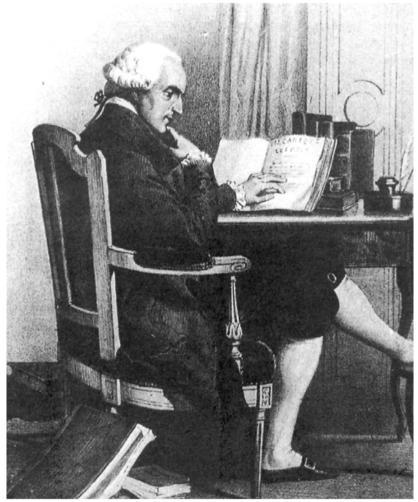
# Historical insights, error theory and least squares

Legendre

Nouvelles méthodes pour la détermination des orbites des planètes (1805)

Gauss

Theoria motus corporum coelestium (1809)



Laplace

Théorie analytique des probabilités (1811)

Least squares method for chosing the “best” value

Famous argument leading to the normal law for the errors

Least squares method for solving linear systems

Twelve years after his argument showing the importance of the normal law as probability law for the errors (*Theoria motus corporum coelestium* 1809), Gauss was interested in the propagation of errors (*Theoria Combinationis* 1821).

Given a quantity $U = F(V_1, V_2, \ldots)$ function of other erroneous quantities $V_1, V_2, \ldots$ he states the problem of computing the quadratic error to fear on $U$ knowing the quadratic errors $\sigma_1^2, \sigma_2^2, \ldots$ on $V_1, V_2, \ldots$, these errors being supposed small and independent. His answer is the following formula

$$\sigma_U^2 = (\frac{\partial F}{\partial V_1})^2 \sigma_1^2 + (\frac{\partial F}{\partial V_2})^2 \sigma_2^2 + \tag{1}$$

he gives also the covariance between the error on $F$ and the error of an other function of the $V_i$'s.

$$(1) \qquad \sigma_U^2 = (\frac{\partial F}{\partial V_1})^2 \sigma_1^2 + (\frac{\partial F}{\partial V_2})^2 \sigma_2^2 +$$

Formula (1) possesses a property which makes it highly better, in several questions, than other formulas used here and there in textbooks during the 19th and 20th centuries. It is a *coherence* property. With a formula such that

$$(2) \qquad \sigma_U = |\frac{\partial F}{\partial V_1}|\sigma_1 + |\frac{\partial F}{\partial V_2}|\sigma_2 + \ldots$$

errors can depend on the manner the function $F$ is written : in dimension 2 already composing an injective linear map with its inverse leads with formula (2) to the fact that the identity map increases the errors what is hardly acceptable.

This doesn't happen in Gauss' calculus. Introducing the operator

$$L = \frac{1}{2}\sigma_1^2 \frac{\partial^2}{\partial V_1^2} + \frac{1}{2}\sigma_2^2 \frac{\partial^2}{\partial V_2^2} + \dots$$

and supposing the functions smooth, we remark that formula (1) can be written

$$\sigma_U^2 = LF^2 - 2FLF$$

and the coherence of this calculus comes from the coherence of the transport of a differential operator by a function : if $L$ is such an operator, $u$ and $v$ injective regular maps, denoting the operator $\varphi \to L(\varphi \circ u) \circ u^{-1}$ by $\theta_u L$ we have $\theta_{v \quad u} L = \theta_v(\theta_u L)$.

**Now, it is clear that if the errors are correlated, Gauss formula becomes**

$$\sigma_U^2 = \sum_{ij} \frac{\partial F}{\partial V_i}\frac{\partial F}{\partial V_j}\sigma_{ij}$$

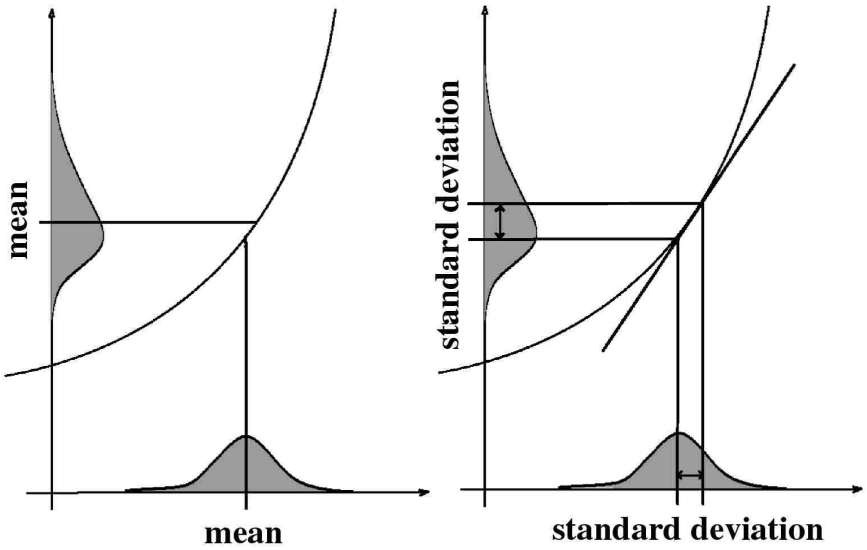
**and in general $\sigma_{ij}$ depend on the values of the $V_i$'s, so that we obtain the general formula**

$$\sigma_U^2 = \sum_{ij} \frac{\partial F}{\partial V_i}\frac{\partial F}{\partial V_j}\sigma_{ij}(V_1, V_2, \ldots)$$

# 2. The propagation of errors

**Let us consider an erroneous quantity with a centered error and let us apply successively non-linear applications**

quadratic bias




**We observe the following properties**

- **The error doesn't remain centered : a bias appears**

- **The variances transmit with a first order differential calculus**

$$\sigma_{n+1}^2 = f_{n+1}'^2(x_n)\ \sigma_n^2$$

- **The biases and the variances keep (except special case) the same order of magnitude**
- **The biases follow a second order differential calculus involving the variances**

$$\mathrm{bias}_{n+1} = f_{n+1}'(x_n)\mathrm{bias}_n + \frac{1}{2} f_{n+1}''(x_n)\sigma_n^2$$

## Intuitive notion of error structure

The preceding example shows that the quadratic error operator $\Gamma$ naturally polarizes into a bilinear operator (as the covariance operator in probability theory), which is a first-order differential operator.

**1.** We thus adopt the following temporary definition of an *error structure*:

An error structure is a space equipped with an operator $\Gamma$ acting upon real functions

$$(\Omega, \Gamma)$$

and satisfying the following properties:

a) *Symmetry*

$$\Gamma[F, G] = \Gamma[G, F];$$

b) *Bilinearity*

$$\Gamma\left[\sum_i \lambda_i F_i, \sum_j \mu_j G_j\right] = \sum_{ij} \lambda_i \mu_j \Gamma[F_i, G_j];$$

c) *Positivity*

$$\Gamma[F] = \Gamma[F, F] \geq 0$$

d) *Functional calculus on regular functions*

$$\begin{aligned}\Gamma[\Phi(F_1, \ldots, F_p), \Psi(G_1, \ldots, G_q)] \\ = \sum_{i,j} \Phi_i'(F_1, \ldots, F_p)\Psi_j'(G_1, \ldots, G_q)\Gamma[F_i, G_j].\end{aligned}$$

**2.** In order to take in account the biases, we also have to introduce a bias operator $A$, a linear operator acting on regular functions through a second order functional calculus involving $\Gamma$ :

$$\begin{aligned}A[\Phi(F_1, \ldots, F_p)] = \sum_i \Phi_i'(F_1, \ldots, F_p)A[F_i] \\ + \frac{1}{2} \sum_{ij} \Phi_{ij}''(F_1, \ldots, F_p)\Gamma[F_i, F_j]\end{aligned}$$

**Example**

Let us consider the usual way of simulating two normal random variables

$$\begin{cases} X &= R\cos 2\pi V \\ Y &= R\sin 2\pi V \end{cases}$$

with $R = \sqrt{-2\log U}$ in other words $U = e^{-\frac{R^2}{2}}$.

Then $(X, Y)$ is a reduced normal pair when $U$ and $V$ are independent uniformly distributed on $[0, 1]$. Let us suppose that the variances of the errors on $U$ and $V$ are such that

$$\begin{array}{ll} i) & \Gamma[V, V] = 1 \\ ii) & \Gamma[R, R] = 4\pi^2 R^2 \\ iii) & \Gamma[V, R] = 0. \end{array}$$

hypothesis i) means that the error on $V$ does'nt depend on $V$,
hypothesis ii) means that *the proportional error* on $R$ is constant, it is equivalent, by the functional calculus, to the hypothesis $\Gamma[U, U] = 16\pi^2 U^2 \log^2 U$,
and hypothesis iii) means that the errors on $V$ and $R$ are uncorrelated.

Then supposing $\Gamma$ satisfies the functional calculus, we get

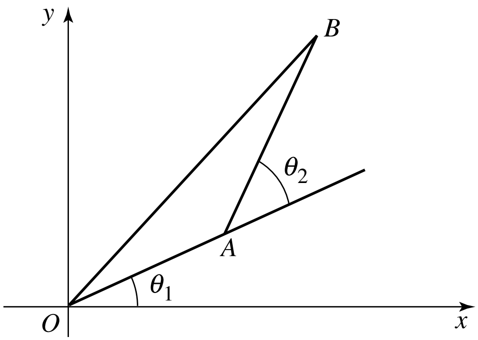
$$\begin{aligned} \Gamma[X, X] &= \cos^2 2\pi V.\Gamma[R, R] + R^2 4\pi^2 \sin^2 2\pi V = 4\pi^2 R^2 \\ \Gamma[Y, Y] &= 4\pi^2 R^2 \\ \Gamma[X, Y] &= \cos 2\pi V \sin 2\pi V.\Gamma[R, R] + 4\pi^2 R^2(-\sin 2\pi V \cos 2\pi V) = 0 \end{aligned}$$

With these hypotheses on $U$ and $V$ and their errors, we can conclude that the variances of the errors on $X$ and $Y$ depend only on the radius $R$ and the co-error on $X$ and $Y$ vanishes.

If we think $(U, V)$ or $(R, V)$ as the data of the model and $(X, Y)$ as the output, we see that this error calculus may be seen as a *sensitivity analysis* but dealing with more information than simple derivation since we obtain co-sensitivity as well.

## error calculation for a triangle

. Suppose we are drawing a triangle with a graduated rule and protractor: we take the polar angle of $OA$, say $\theta_1$, and set $OA = \ell_1$; next we take the angle $(OA, AB)$, say $\theta_2$, and set $AB = \ell_2$.




**1) Select hypotheses on errors**

$\ell_1, \ell_2$ and $\theta_1, \theta_2$ and their errors can be modeled as follows:

$$\left((0, L)^2 \times (0, \pi)^2, \mathcal{B}\big((0, L)^2 \times (0, \pi)^2\big), \frac{d\ell_1}{L}\frac{d\ell_2}{L}\frac{d\theta_1}{\pi}\frac{d\theta_2}{\pi}, \mathbb{D}, \Gamma\right)$$

where

$$\mathbb{D} = \left\{ f \in L^2\left(\frac{d\ell_1}{L}\frac{d\ell_2}{L}\frac{d\theta_1}{\pi}\frac{d\theta_2}{\pi}\right) : \frac{\partial f}{\partial \ell_1}, \frac{\partial f}{\partial \ell_2}, \frac{\partial f}{\partial \theta_1}, \frac{\partial f}{\partial \theta_2} \in L^2\left(\frac{d\ell_1}{L}\frac{d\ell_2}{L}\frac{d\theta_1}{\pi}\frac{d\theta_2}{\pi}\right)\right\}$$

and

$$\Gamma[f] = \ell_1^2\left(\frac{\partial f}{\partial \ell_1}\right)^2 + \ell_1\ell_2\frac{\partial f}{\partial \ell_1}\frac{\partial f}{\partial \ell_2} + \ell_2^2\left(\frac{\partial f}{\partial \ell_2}\right)^2 + \left(\frac{\partial f}{\partial \theta_1}\right)^2 + \frac{\partial f}{\partial \theta_1}\frac{\partial f}{\partial \theta_2} + \left(\frac{\partial f}{\partial \theta_2}\right)^2,$$

This quadratic error operator indicates that the errors on lengths $\ell_1, \ell_2$ are uncorrelated with those on angles $\theta_1, \theta_2$ (i.e. no term in $\frac{\partial f}{\partial \ell_i}\frac{\partial f}{\partial \theta_j}$). Such a hypothesis proves natural when measurements are conducted using different instruments. The bilinear operator associated with $\Gamma$ is

$$\Gamma[f, g] = \ell_1^2\frac{\partial f}{\partial \ell_1}\frac{\partial g}{\partial \ell_1} + \frac{1}{2}\ell_1\ell_2\left(\frac{\partial f}{\partial \ell_1}\frac{\partial g}{\partial \ell_2} + \frac{\partial f}{\partial \ell_2}\frac{\partial g}{\partial \ell_1}\right) + \ell_2^2\frac{\partial f}{\partial \ell_2}\frac{\partial g}{\partial \ell_2}$$
$$+ \frac{\partial f}{\partial \theta_1}\frac{\partial g}{\partial \theta_1} + \frac{1}{2}\left(\frac{\partial f}{\partial \theta_1}\frac{\partial g}{\partial \theta_2} + \frac{\partial f}{\partial \theta_2}\frac{\partial g}{\partial \theta_1}\right) + \frac{\partial f}{\partial \theta_2}\frac{\partial g}{\partial \theta_2}.$$

**2) Compute the errors on significant quantities using functional calculus on $\Gamma$**

**Take point $B$ for instance:**

$$\begin{aligned}
X_B &= \ell_1 \cos\theta_1 + \ell_2 \cos(\theta_1+\theta_2), \quad Y_B = \ell_1 \sin\theta_1 + \ell_2 \sin(\theta_1+\theta_2) \\
\Gamma[X_B] &= \ell_1^2 + \ell_1\ell_2(\cos\theta_2 + 2\sin\theta_1 \sin(\theta_1+\theta_2)) \\
&\quad + \ell_2^2(1 + 2\sin^2(\theta_1+\theta_2)) \\
\Gamma[Y_B] &= \ell_1^2 + \ell_1\ell_2(\cos\theta_2 + 2\cos\theta_1 \cos(\theta_1+\theta_2)) \\
&\quad + \ell_2^2(1 + 2\cos^2(\theta_1+\theta_2)) \\
\Gamma[X_B, Y_B] &= -\ell_1\ell_2 \sin(2\theta_1+\theta_2) - \ell_2^2 \sin(2\theta_1 + 2\theta_2).
\end{aligned}$$

**For the area of the triangle, the formula area$(OAB) = \frac{1}{2}\ell_1\ell_2\sin\theta_2$ yields:**

$$\Gamma[\mathrm{area}(OAB)] = \frac{1}{4}\ell_1^2\ell_2^2(1 + 2\sin^2\theta_2).$$

**The proportional error on the triangle area**

$$\frac{(\Gamma[\mathrm{area}(OAB)])^{1/2}}{\mathrm{area}(OAB)} = \left(\frac{1}{\sin^2\theta_2} + 2\right)^{1/2} \geq \sqrt{3}$$

**reaches a minimum at $\theta_2 = \frac{\pi}{2}$ when the triangle is rectangular. From $OB^2 = \ell_1^2 + 2\ell_1\ell_2\cos\theta_2 + \ell_2^2$, we obtain**

$$\begin{aligned}
\Gamma[OB^2] &= 4\big[(\ell_1^2+\ell_2^2)^2 + 3(\ell_1^2+\ell_2^2)\ell_1\ell_2\cos\theta_2 + 2\ell_1^2\ell_2^2\cos^2\theta_2\big] \\
&= 4OB^2(OB^2 - \ell_1\ell_2\cos\theta_2)
\end{aligned}$$

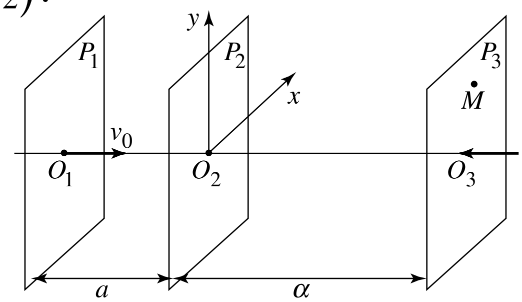
**and by $\Gamma[OB] = \frac{\Gamma[OB^2]}{4OB^2}$, we have:**

$$\frac{\Gamma[OB]}{OB^2} = 1 - \frac{\ell_1\ell_2\cos\theta_2}{OB^2}$$

**thereby providing the result that the proportional error on $OB$ is minimal when $\ell_1 = \ell_2$ and $\theta_2 = 0$ and in this case $\frac{(\Gamma[OB])^{1/2}}{OB} = \frac{\sqrt{3}}{2}$.**

## • Cathodic tube

An oscillograph is modeled in the following way. After acceleration by an electric field, electrons arrive at point $O_1$ at a speed $v_0 > 0$ orthogonal to plane $P_1$. Between parallel planes $P_1$ and $P_2$, a magnetic field $\vec{B}$ orthogonal to $O_1O_2$ is acting; its components on $O_2x$ and $O_2y$ are $\left(B_1, B_2\right)$.




**Equation of the model.** The physics of the problem is classical. The gravity force is negligible, the Lorenz force $q\vec{v}\wedge\vec{B}$ is orthogonal to $\vec{v}$ such that the modulus $|\vec{v}|$ remains constant and equal to $v_0$, and the electrons describe a circle of radius $R = \frac{mv_0}{e|\vec{B}|}$.

If $\theta$ is the angle of the trajectory with $O_1O_2$ as it passes through $P_2$, we then have:

$$\theta = \arcsin\frac{a}{R}$$
$$|O_2A| = R(1-\cos\theta)$$

and

$$A = \left(|O_2A|\frac{B_2}{|\vec{B}|}, -|O_2A|\frac{B_1}{|\vec{B}|}\right).$$

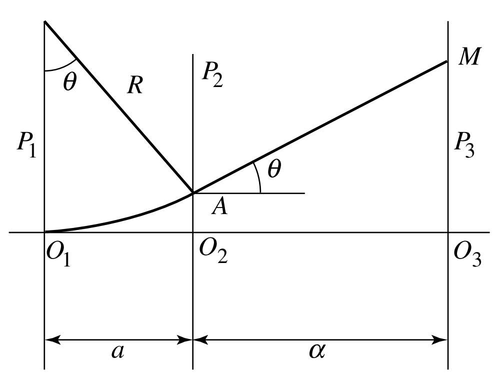



Figure in the plane of the trajectory

The position of $M$ is thus given by:

$$\begin{cases} M = (X, Y) \\ X = \left(\dfrac{mv_0}{e|\vec{B}|}(1-\cos\theta) + d\tan\theta\right)\dfrac{B_2}{|\vec{B}|} \\ Y = -\left(\dfrac{mv_0}{e|\vec{B}|}(1-\cos\theta) + d\tan\theta\right)\dfrac{B_1}{|\vec{B}|} \\ \theta = \arcsin \dfrac{ae}{mv_0}|\vec{B}| \end{cases} \tag{1}$$

To study the sensitivity of point $M$ to the magnetic field $\vec{B}$, we assume that $\vec{B}$ varies in $[-\lambda, \lambda]^2$ , equipped with the error structure

$$\left([-\lambda,\lambda]^2, \mathcal{B}\left([-\lambda,\lambda]^2\right), \mathbb{P}, \mathbb{D}, u \to u_1'^2 + u_2'^2\right).$$

We can now compute the quadratic error on $M$, i.e. the matrix

$$\underset{=}{\Gamma}\left[M, M^t\right] = \begin{pmatrix} \Gamma[X] & \Gamma[X,Y] \\ \Gamma[X,Y] & \Gamma[Y] \end{pmatrix} = \begin{pmatrix} (\frac{\partial X}{\partial B_1})^2 + (\frac{\partial X}{\partial B_2})^2 & \frac{\partial X}{\partial B_1}\frac{\partial Y}{\partial B_1} + \frac{\partial X}{\partial B_2}\frac{\partial Y}{\partial B_2} \\ \frac{\partial X}{\partial B_1}\frac{\partial Y}{\partial B_1} + \frac{\partial X}{\partial B_2}\frac{\partial Y}{\partial B_2} & (\frac{\partial Y}{\partial B_1})^2 + (\frac{\partial Y}{\partial B_2})^2 \end{pmatrix}$$

Using some approximations to simplfy the calculations, we obtain

$$\begin{aligned} &\underset{=}{\Gamma}\left[M, M^t\right] = \\ &= \begin{pmatrix} 4\beta^2 B_1^2 B_2^2 + \left(\alpha + \beta B_1^2 + 3\beta B_2^2\right)^2 & -2\beta B_1 B_2\left(\alpha + 3\beta\left(B_1^2 + B_2^2\right)\right) \\ -2\beta B_1 B_2\left(\alpha + 3\beta\left(B_1^2 + B_2^2\right)\right) & 4\beta^2 B_1^2 B_2^2 + \left(\alpha + \beta B_2^2 + 3\beta B_1^2\right)^2 \end{pmatrix} \end{aligned} \tag{2}$$

It follows in particular, by computing the determinant, that the law of $M$ is absolutely continuous.

If we now suppose that the inaccuracy on the magnetic field stems from a noise in the electric circuit responsible for generating $\vec{B}$ and that this noise is *centered*: $A\big[B_1\big] = A\big[B_2\big] = 0$, we can compute *the bias* of the errors on $M = (X, Y)$

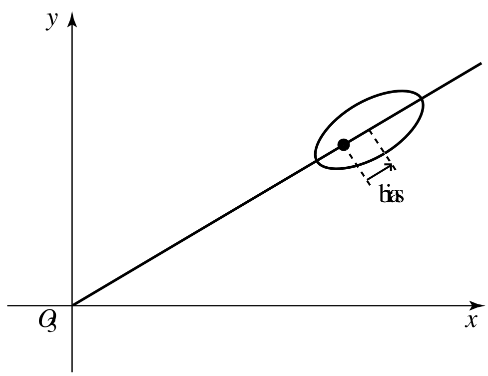
$$A[X] = \frac{1}{2}\frac{\partial^2 X}{\partial B_1^2}\Gamma\big[B_1\big] + \frac{1}{2}\frac{\partial^2 X}{\partial B_2^2}\Gamma\big[B_2\big]$$
$$A[Y] = \frac{1}{2}\frac{\partial^2 Y}{\partial B_1^2}\Gamma\big[B_1\big] + \frac{1}{2}\frac{\partial^2 Y}{\partial B_2^2}\Gamma\big[B_2\big]$$

which yields

$$A[X] = 4\beta B_2$$
$$A[Y] = -4\beta B_1.$$

By comparison with (2), we can observe that:

$$A\big[\overrightarrow{O_3 M}\big] = \frac{4\beta}{\alpha + \beta|\vec{B}|^2}\,\overrightarrow{O_3 M}.$$




The appearance of biases in the absence of bias on the hypotheses is specifically due to the fact that the method considers the errors, although infinitesimal, to be *random* quantities. This feature will be highlighted in the following table.