Perturbation Theory and Control in Classical or Quantum Mechanics by an Inversion Formula

Michel VITTOT *

Centre de Physique Théorique †
CNRS Luminy, case 907 – F-13288 Marseille cedex 9 – France

November 7, 2003

Abstract

We consider a perturbation of an "integrable" Hamiltonian and give an expression for the canonical or unitary transformation which "simplifies" this perturbed system. The problem is to invert a functional defined on the Lie-algebra of observables. We give a bound for the perturbation in order to solve this inversion. And apply this result to a particular case of the control theory, as a first example, and to the "quantum adiabatic transformation", as another example.

Key-words: perturbation series, algebra of observables, resonances, inversion formula, control theory

http://www.cpt.univ-mrs.fr

* vittot@cpt.univ-mrs.fr
† Unité Propre de Recherche 7061

Contents

1 - Algebraic Framework

We start with a vector space $\mathcal{A}$ which we call the space of observables.
We will apply this theory to 2 main examples: the classical mechanics, in which $\mathcal{A}$ is an (abelian) algebra of functions defined on the phase space, and the quantum mechanics, in which $\mathcal{A}$ is an algebra of operators on some Hilbert space. By taking a basis of projectors in $\mathcal{A}$ we will consider these operators as a collection of "matrix elements", i.e. also as functions.
We assume that $\mathcal{A}$ is endowed with a Lie stucture, i.e. there exists a mapping (called the "bracket") from $\mathcal{A}$ into the space $\mathfrak{L}(\mathcal{A})$ of linear operators of $\mathcal{A}$:

$$\begin{aligned} \{\ldots\}: \quad & \mathcal{A} \to \mathfrak{L}(\mathcal{A}) \\ & V \mapsto \{V\} \end{aligned} \tag{1.1}$$

which satisfies:

$$\forall V, W \in \mathcal{A} \quad \{V\}W = -\{W\}V \tag{1.2}$$

$$\text{and} \quad \{\{V\}W\} = \{V\}\{W\} - \{W\}\{V\} \tag{1.3}$$

so that the "bracket" mapping is linear in its argument. It is also antisymmetric, by (1.2) and satisfies the Jacobi identity (1.3). We consider a fixed element H of $\mathcal{A}$ which we call a Hamiltonian. The "motion" or the "flow" generated by H is the 1-parameter group:

$$\forall t \in \mathbb{R} \quad e^{t\{H\}} : \mathcal{A} \to \mathcal{A} \tag{1.4}$$

The exponential of $\{H\}$ is defined by the usual power series:

$$e^{t\{H\}} := \sum_{n=0}^{\infty} \frac{\{t.H\}^n}{n!} \tag{1.5}$$

and is an automorphism of the Lie structure, as is proven below in (5.4).
Actually what is only needed is that $\{H\}$ is a linear operator "affiliated" to $\mathcal{A}$ which means that (1.4) is still valid: $\{H\}$ is assumed to be the generator of a 1-parameter group of automorphisms, even if H is not an element of $\mathcal{A}$.

Our problem is to find a relation (for instance a "conjugation") between the flow generated by a perturbation of H, denoted $H+V$ for some $V \in \mathcal{A}$, and the flow of H. Poincaré called this problem "the main problem of the dynamics". Of course this relation would be mainly useful if we have some information on the flow of the unperturbed hamiltonian H.

We will study 2 types of problems:

Problem 1 - *If we are permitted to modify the perturbed hamiltonian $H+V$ by adding a "small" term $f(V)$, we will try to find a relation between the flow of $H+V+f(V)$ and the flow of H. Then the term $f(V)$ will be called a "control" term.*

Of course we want to exclude the trivial solution $f(V) = -V$ so we have added the supplementary condition on this term to be "small", for instance quadratic in V. So we stabilize the perturbed hamiltonian by a small control term, well adapted to the problem.

Problem 2 - *If we are not permitted to modify the perturbed hamiltonian (for instance if we investigate the motion of some planet), or we don't want to modify it, then we will try to find a good "change of coordinates", i.e. an automorphism of $\mathcal{A}$, which connects the 2 flows: perturbed and unperturbed.*

We will search this automorphism under the form of some exponential $e^{\{\Gamma W\}}$ where we define Γ as follows.

Let us make an important assumption on H, which will be satisfied when H is "integrable", as seen below:

Hypothesis 1 - *We assume that there exists a linear operator $\mathcal{N} : \mathcal{A} \to \mathcal{A}$ such that the following product is (formally) well defined from $\mathcal{A}$ into $\mathcal{A}$:*

$$\Gamma := \{H\}^{-1}\mathcal{N} \tag{1.6}$$

and such that:

$$\{H\}\mathcal{N} = \{H\} \tag{1.7}$$

So we have:

$$\mathcal{N} = \{H\}\Gamma \tag{1.8}$$

Hence Γ is a pseudo-inverse of $\{H\}$:

$$\{H\}^2\,\Gamma = \{H\} \tag{1.9}$$

Let us remind that it is impossible to find a strict inverse of $\{H\}$ since it always has a non-trivial kernel (for instance $\{H\}H^\alpha = 0$ for any $\alpha \in \mathbb{R}$ such that $H^\alpha \in \mathcal{A}$). Hence $\{H\}^{-1}$ is not defined everywhere, but the product (1.6) is (formally) defined.
Any element V of $\mathcal{A}$ such that $\{H\}V = 0$, is constant under the flow of H:

$$e^{t\{H\}}V = V \tag{1.10}$$

So the vector space $\mathrm{Ker}\,\{H\}$ is called the set of "constants of motion". It is a sub-Lie-algebra of $\mathcal{A}$ since (using (1.3)):

$$\text{If} \quad \{H\}V = \{H\}W = 0 \quad \text{Then} \quad \{H\}\{V\}W = \{V\}\{H\}W + \{\{H\}V\}W = 0 + 0 \tag{1.11}$$

Let us also note that (1.7) can be rewritten as:

$$\{H\}\mathcal{R} = 0 \tag{1.12}$$

$$\text{where} \qquad \mathcal{R} := \mathbf{1} - \mathcal{N} \tag{1.13}$$

which means that the range $\mathrm{Rg}\,\mathcal{R}$ of the operator $\mathcal{R}$ is included in $\mathrm{Ker}\,\{H\}$. The $\mathbf{1}$ is the identity in the algebra $\mathcal{L}(\mathcal{A})$ of endomorphisms of $\mathcal{A}$. The notation $\mathcal{R}$ designates the "resonant part" and the notation $\mathcal{N}$ the "non-resonant part". Hence $\{H\}^{-1}$ is well defined on the space $\mathrm{Rg}\,\mathcal{N}$. Let us remind that $\{H\}, \mathcal{R}, \mathcal{N}, \Gamma$ are elements of $\mathcal{L}(\mathcal{A})$.

Application 1 - *Classical Mechanics:*

$\mathcal{A}$ is the algebra of $\mathcal{C}^\infty$ real-valued functions of (p, q) in some domain of $\mathbb{R}^{2L}$ or functions of $(A, \theta) \in \mathcal{H} \times \mathbb{T}^L$ for some domain $\mathcal{H}$ of $\mathbb{R}^L$ (L is the number of degree of freedoms). It is more convenient to make a Fourier transformation in θ, and so $\mathcal{A}$ may be taken as the algebra of functions V of $(A, \Delta) \in \mathcal{H} \times \mathbb{Z}^L$ into $\mathbb{C}$, such that $V(A, -\Delta) = V(A, \Delta)^*$. The Lie structure is the Poisson bracket. We can compute the operators $\mathcal{R}$ and Γ when H is "integrable", i.e. when H is a function of the "actions variables" only, and does not depend on the "angles variables":

$$H(A, \theta) = h(A) \tag{1.14}$$

or after a Fourier transformation in θ:

$$H(A, \Delta) = h(A).\delta_0(\Delta) \tag{1.15}$$

with δ the Kronecker symbol. Indeed, let us denote by $\omega(A)$ the derivative (gradient) of h with respect to A. Then the bracket $\{H\} = \omega(A).\partial_\theta$ and so, after a Fourier transformation, it is given by:

$$(\{H\}V)(A, \Delta) = \mathbf{i}.(\omega(A) \cdot \Delta).V(A, \Delta) \tag{1.16}$$

on any element V of $\mathcal{A}$. We have used a scalar product $\omega(A) \cdot \Delta$. So that the operator $\mathcal{R}$ is:

$$(\mathcal{R}V)(A, \Delta) = V(A, \Delta).\chi(\omega(A) \cdot \Delta = 0) \tag{1.17}$$

where we introduce a characteristic function $\chi(\omega(A)\cdot\Delta = 0)$ which is 1 when $\omega(A)\cdot\Delta = 0$ and is 0 otherwise. We could have written it as $\chi(\Delta \in \omega(A)^{\perp} \cap \mathbb{Z}^L)$. Similarly:

$$(\mathcal{N}V)(A,\Delta) = V(A,\Delta).\chi(\omega(A)\cdot\Delta \neq 0) \tag{1.18}$$

Then the action of the operator Γ is given by:

$$(\Gamma V)(A,\Delta) := \frac{\chi(\omega(A)\cdot\Delta \neq 0)}{\mathbf{i}.(\omega(A)\cdot\Delta)} \cdot V(A,\Delta) \tag{1.19}$$

See section 6 for more details and section 7 where we introduce a norm.

Application 2 - *Quantum Mechanics:*

$\mathcal{A}$ is the algebra of operators on some separable Hilbert space. More precisely we assume that $\mathcal{A}$ has a unity (denoted by $\mathbf{1}$), and we consider a maximal family of mutually orthogonal projectors $P_A \in \mathcal{A}$ where A varies in some countable set $\mathcal{H}$. That means:

$$\begin{aligned} \sum_{A\in\mathcal{H}} P_A &= \mathbf{1} && (1.20)\\ \forall A, A' \in \mathcal{H} \quad P_A.P_{A'} &= P_A.\delta_{A,A'} && (1.21) \end{aligned}$$

with δ the Kronecker symbol. $\mathcal{R}$ and Γ are easily written when H can be "diagonalized". We choose H as follows:

$$H = \sum_{A\in\mathcal{H}} h(A).P_A \tag{1.22}$$

for some function $h : \mathcal{H} \to \mathbb{R}$. So $h(\mathcal{H})$ is the spectrum of H and the projectors are the spectral projectors of H. We can choose h to be the identity function by taking $\mathcal{H}$ as the spectrum of H.

Remark 1 - *Reciprocally, if $\mathcal{A}$ were endowed with an involution ("*", named "adjonction") we could have started by giving a self-adjoint operator H with pure point spectrum, and then we could have taken the spectral projectors of H, as the family of projectors P_A.*

Then any element V of $\mathcal{A}$ can be written as:

$$V = \sum_{A,A'\in\mathcal{H}} V_{A,A'} \quad \text{where} \quad V_{A,A'} := P_A V P_{A'} \tag{1.23}$$

It is convenient to introduce the set:

$$\mathcal{G}(A) := \mathcal{H} - A := \{A' - A\}_{A'\in\mathcal{H}} \tag{1.24}$$

and to define:

$$\forall A \in \mathcal{H} \quad \forall \Delta \in \mathcal{G}(A) \quad V(A,\Delta) := P_{A+\Delta} V P_A \tag{1.25}$$

i.e. $V(A,\Delta) = V_{A+\Delta,A}$. For instance:

$$H(A,\Delta) = h(A).\delta_0(\Delta).P_A \tag{1.26}$$

which is similar to (1.15).

The Lie structure is given by the commutator. Actually we multiply it by $\mathbf{i}/\hbar$ for some constant $\hbar$ which has the correct dimensionality, so that $t\{H\}$ has no dimension:

$$\forall V, W \in \mathcal{A} \quad \{V\}(W) := \mathbf{i}(V.W - W.V)/\hbar \tag{1.27}$$

Let us now turn to the action of the operators $\{H\}$ and $\mathcal{R}$:

$$(\{H\}V)(A, \Delta) = \mathbf{i}.\Big(h(A+\Delta) - h(A)\Big).V(A, \Delta)/\hbar \tag{1.28}$$

Hence:

$$(\mathcal{R}V)(A, \Delta) = V(A, \Delta).\ \chi\Big(h(A+\Delta) = h(A)\Big) \tag{1.29}$$

and:

$$(\mathcal{N}V)(A, \Delta) = V(A, \Delta).\ \chi\Big(h(A+\Delta) \neq h(A)\Big) \tag{1.30}$$

Then the operator Γ is given by:

$$(\Gamma V)(A, \Delta) = \frac{\hbar.\chi\Big(h(A+\Delta) \neq h(A)\Big)}{\mathbf{i}.\Big(h(A+\Delta) - h(A)\Big)} \cdot V(A, \Delta) \tag{1.31}$$

See section 6 for more details and section 7 where we introduce a norm.

Remark: we have introduced an arbitrary constant $\hbar$ in (1.27). But we will notice that the operator Γ always appears with a bracket around it. And the pair $\{\Gamma \ldots\}$ is independant of $\hbar$, since this constant is in the numerator and the denominator.

2 - The Main Theorems

Theorem 1 - *The control problem (problem 1) is solved by an explicit formula. Let us first define the functions* F and $f : \mathcal{A} \to \mathcal{A}$ *by:*

$$F(V) := e^{-\{\Gamma V\}}\mathcal{R}V + \frac{1 - e^{-\{\Gamma V\}}}{\{\Gamma V\}}\mathcal{N}V \tag{2.1}$$

$$f(V) := F(V) - V \tag{2.2}$$

Then we have:

$$\forall t \in \mathbb{R} \qquad e^{t\{H+V+f(V)\}} = e^{-\{\Gamma V\}}.e^{t\{H\}}.e^{t\{\mathcal{R}V\}}.e^{\{\Gamma V\}} \tag{2.3}$$

The meaning of the second term in (2.1) is the following:

$$\frac{1 - e^{-\{\Gamma V\}}}{\{\Gamma V\}} = \sum_{n\in\mathbb{N}} \frac{\{-\Gamma V\}^n}{n+1!} = \int_0^1 ds.e^{-s\{\Gamma V\}} \tag{2.4}$$

We will prove this theorem in section 5. This will solve the control problem if we can check that $f(V)$ is indeed smaller than V. This will be proved in section 7. To see this, it is sufficient to notice that in the expression (2.1) of $F(W)$, the terms $e^{-\{\Gamma W\}}$ and $(1 - e^{-\{\Gamma W\}})/\{\Gamma W\}$ are near $\mathbf{1}$ when $W \approx 0$. So:

$$F(W) \approx \mathcal{R}W + \mathcal{N}W := W \quad \text{so that} \quad f(W) = \mathcal{O}(W^2) \tag{2.5}$$

as is seen in (7.10).
The formula (2.3) connects the perturbed flow, modified by a control term, with the unperturbed flow.
The new factor, the flow of $\mathcal{R}V$ will turn out to commute with the flow of H: cf. (5.9).

The second problem (the "change of coordinates") is solved by an inversion formula. Let us rewrite (2.3) with $F(V)$ instead of $V+f(V)$, and with W instead of V:

$$\forall t \in \mathbb{R} \qquad e^{t\{H+F(W)\}} = e^{-\{\Gamma W\}}.e^{t\{H\}}.e^{t\{\mathcal{R}W\}}.e^{\{\Gamma W\}} \tag{2.6}$$

Theorem 2 - *If we can find W such that $F(W)=V$,* i.e. *$W=F^{-1}(V)$, then:*

$$\forall t \in \mathbb{R} \qquad e^{t\{H+V\}} = e^{-\{\Gamma W\}}.e^{t\{H\}}.e^{t\{\mathcal{R}W\}}.e^{\{\Gamma W\}} \tag{2.7}$$

Here again, the flow of $\mathcal{R}W$ commute with the flow of H.
Hence we need to invert the function F. But F is near the identity function around 0. More precisely we will find a ball around 0 in $\mathcal{A}$, for some norm, such that the difference between F and the identity function (what we called f in (2.2)) is Lipschitz and contractant: cf. (2.5). So F can be inverted, at least around 0. See section 7.

So to summarize this introduction: the first problem is solved explicitly for any perturbation, but we need to assume some smallness on the size of the perturbation to ensure that the "control" term is smaller than the original perturbation. And the second problem is solved by an inversion formula, which also needs some smallness on the size of the perturbation to ensure that F is invertible. To be more precise, we need some norm on the algebra $\mathcal{A}$: this is done in section 7.

The equation in the unknown W (i.e. the inversion $W=F^{-1}(V)$) may be named the "Hamilton-Jacobi equation". Indeed it yields the automorphism $e^{\{\Gamma W\}}$.

Remark: if we also assume that $\{H\}\Gamma = \Gamma\{H\}$ (as in examples (1.19) or (1.31)) then $\mathcal{N}^2 = \mathcal{N}$ i.e. the range $\operatorname{Rg}\mathcal{R}$ of the operator $\mathcal{R}$ is exactly the set of constants of motion $\operatorname{Ker}\{H\}$.

3 - Localization of the Action Variable

Let us study the case of the classical mechanics, as above, with action-angles variables. Usually, we localize the action variable A near some point A_0 i.e. we make the canonical change of variables, from (A,θ) to (A_1,θ_1) with:

$$A = A_0 + \varepsilon A_1 \qquad \theta = \theta_1 \tag{3.1}$$

along with a rescaling of the hamiltonian:

$$H_1(A) := H(A_0 + \varepsilon.A)/\varepsilon \tag{3.2}$$

where we choose some positive constant ε. This transformation is called a "canonical similarity" since it preserves the symplectic form, up to the multiplicative constant ε.

When H is integrable in the usual sense, with actions-angles variables, we can expand the hamiltonian $H = H(A)$ around A_0:

$$H(A_0+\varepsilon A_1) = c+\varepsilon.\omega\cdot A_1+q(\varepsilon.A_1) \quad \text{where} \quad \omega := H'(A_0) \quad \text{and} \quad q(0) = q'(0) = 0 \tag{3.3}$$

(i.e. q is quadratic in A_1) and where the additive constant $c := H(A_0)$ is not relevant and will be forgotten. Hence:

$$H_1(A) = \omega \cdot A + \varepsilon.\tilde{q}(A) \tag{3.4}$$

where $\tilde{q}(A) := q(\varepsilon.A)/\varepsilon^2$ is of order ε^0.
Let us introduce a perturbation V, as above, and choose $\varepsilon := ||V||^{1/2}$ for some norm. Under the rescaling (3.2) of the hamiltonian, the perturbation is also divided by ε and so becomes:

$$V_1(A,\theta) := V(A_0 + \varepsilon.A, \theta)/\varepsilon \tag{3.5}$$

which is of order ε, so that the perturbed hamiltonian is now:

$$H_1(A) + V_1(A,\theta) = \omega \cdot A + \varepsilon.V_2(A,\theta) \tag{3.6}$$

$$\text{with} \quad V_2(A,\theta) = \tilde{q}(A) + V_1(A,\theta)/\varepsilon \tag{3.7}$$

which is of order ε^0. So to summarize, it is always possible to assume that the integrable part is an harmonic oscillator, at least locally in the variable A, i.e. in a region (around any fixed A_0) where A (in the hamiltonian (3.6)) is of order ε^0, but it may be wrong when A is of order $1/\varepsilon$. Hence in this case of classical mechanics, localized in action variable, the operator Γ is:

$$\Gamma = \frac{1}{\omega\cdot\partial}\cdot\mathcal{N} \quad \text{with} \quad \partial := \partial_\theta \tag{3.8}$$

and the operator $\mathcal{N}$ given by (1.18) becomes here, after a Fourier transformation from the angles θ to the integer vector Δ:

$$(\mathcal{N}V)(A,\Delta) = V(A,\Delta).\chi(\omega\cdot\Delta \neq 0) \tag{3.9}$$

4 - Non-Resonant Hamiltonians

We can also define the notion of “non-resonance” as follows.

Definition 1 - *In classical mechanics, H is “non-resonant” iff:*

$$\forall\ V, W \ \in \mathrm{Ker}\,\{H\} \quad \text{we have} \quad \{V\}W = 0 \tag{4.1}$$

In quantum mechanics, H is “non-resonant” iff:

$$\forall\ V, P \ \in \mathrm{Ker}\,\{H\} \ \text{ s.t. } \ P^2 = P \quad \text{we have} \quad \{V\}P = 0 \tag{4.2}$$

Hence in the classical case:

$$\forall W, V \in \mathcal{A} \quad \{\mathcal{R}W\}\mathcal{R}V = 0 \quad \text{ i.e. } \quad \{\mathcal{R}\mathcal{A}\}\mathcal{R} = 0 \tag{4.3}$$

The above definition means that any 2 constants of motion of H do commute together. After the localisation of the action variable, as in the preceding section 3, the above non-resonant condition (4.1) is equivalent to the "usual" non-resonance condition, which is in classical mechanics:

$$\omega^{\perp} \cap \mathbb{Z}^{L} = \{0\} \tag{4.4}$$

i.e. there are no integer vector Δ orthogonal to ω, except for $\Delta = 0$. And the operator $\mathcal{R}$ may be written as the multiplication by the characteristic function $\chi(\Delta = 0)$.
We will give an example of a resonant hamiltonian in (9.1), for which we can still apply our control theory.

In quantum mechanics, the non-resonance condition (4.2) means that any spectral projector of H commutes with any constant of motion. The non-resonance condition is "usually" defined for the Floquet case:

$$H = \sum_{k\in\mathbb{Z}, A\in\mathbb{N}} (k + h(A)).P_{k,A} \tag{4.5}$$

for some projectors $P_{k,A}$ which are mutually orthogonal. For instance the function $h(A)$ may be taken as $\omega.A$ or $\omega.A^2$. The condition (4.2) is satisfied exactly when the set of the spectral gaps intersects $\mathbb{Z}$ in the only point 0:

$$\mathcal{G} \cap \mathbb{Z} = \{0\} \quad \text{where} \quad \mathcal{G} := \{h(A) - h(B) \ \text{ s.t. } \ A, B \in \mathbb{N}\} \tag{4.6}$$

For the case where $h(A) = \omega.A^a$ for some positive integer a, the condition (4.6) exactly means that the "frequency" ω is "non-resonant" in the usual sense. We see that the "resonance" condition is a different notion than the degeneracy property, i.e. the dimension of the spectral projectors.

Let us note that the r.h.s. of (2.7) is called the "normal form" of the (perturbed) flow on the l.h.s. When H is "resonant", then (2.7) is called the "resonant normal form".

5 - Proof of Theorems 1 & 2

Actually the theorems (2.3) & (2.7) are 2 different interpretations of the same formula (2.3) or (2.6). Let us first prove:

Proposition 1

$$\forall W \in \mathcal{A} \qquad H + F(W) = e^{-\{\Gamma W\}}(H + \mathcal{R}W) \tag{5.1}$$

Proof: indeed from the definition (2.1) and (1.8), $F(W)$ can be rewritten as:

$$F(W) - e^{-\{\Gamma W\}}\mathcal{R}W = \frac{1 - e^{-\{\Gamma W\}}}{\{\Gamma W\}}\{H\}\Gamma W = -\frac{1 - e^{-\{\Gamma W\}}}{\{\Gamma W\}}\{\Gamma W\}H = e^{-\{\Gamma W\}}H - H \tag{5.2}$$

where we used the antisymmetry (1.2). *Q.E.D.* ∎

Proof of (2.6): Let us now take the brackets of the 2 sides of (5.1):

$$\{H + F(W)\} = \{e^{-\{\Gamma W\}}(H + \mathcal{R}W)\} \tag{5.3}$$

But:

$$\forall V, W \in \mathcal{A} \quad \{e^{\{V\}}W\} = e^{\{V\}}.\{W\}.e^{-\{V\}} \tag{5.4}$$

Indeed:

$$\forall V, W \in \mathcal{A}, \ \ \forall n \in \mathbb{N} \quad \{\{V\}^n W\} = \sum_{k=0}^{n} \binom{n}{k}.\{V\}^{n-k}.\{W\}.\{-V\}^k \tag{5.5}$$

The proof of (5.5) is an easy recurrence from the case $n = 1$: cf (1.3). Hence:

$$\{H + F(W)\} = e^{-\{\Gamma W\}}.\{H + \mathcal{R}W\}.e^{\{\Gamma W\}} \tag{5.6}$$

Let us now exponentiate the 2 sides of (5.6) (multiplied by any $t \in \mathbb{R}$):

$$e^{t\{H+F(W)\}} = \exp\Big[t.e^{-\{\Gamma W\}}.\{H + \mathcal{R}W\}.e^{\{\Gamma W\}}\Big] = e^{-\{\Gamma W\}}.e^{t\{H+\mathcal{R}W\}}.e^{\{\Gamma W\}} \tag{5.7}$$

where we have used:

$$\forall A, B \quad e^{A^{-1}.B.A} = A^{-1}.e^B.A \tag{5.8}$$

To finish the proof of (2.6), there remains to show that $\{H\}$ and $\{\mathcal{R}W\}$ commute, in the algebra of linear operators on $\mathcal{A}$. But from (1.3), (1.12):

$$\{H\}.\{\mathcal{R}W\} - \{\mathcal{R}W\}.\{H\} = \{\{H\}\mathcal{R}W\} = 0 \tag{5.9}$$

The formula (2.6) is proven. And also the theorems 1 and 2 which are only some rewrittings of it. *Q.E.D.* ■

6 - The 2 Main Applications

Application 1 - *Classical Mechanics:*

Here, an automorphism is called a "canonical transformation".
First we have to explicit the Poisson bracket:

$$\forall (A, \Delta) \in \mathcal{H} \times \mathbb{Z}^L \tag{6.1}$$

$$(\{W\}V)(A, \Delta) = \mathbf{i} \sum_{\Delta' \in \mathbb{Z}^L} W'(A, \Delta - \Delta') \cdot \Delta'.V(A, \Delta') - V'(A, \Delta - \Delta') \cdot \Delta'.W(A, \Delta')$$

where W' is the derivative (gradient) of W with respect to A. And we have used a scalar product $W'(A, \Delta - \Delta') \cdot \Delta'$. If H is "non-resonant", the action of the operator $\mathcal{R}$ on any observable W consists in keeping (in the Fourier coefficients of W) the only term with $\Delta = 0$, i.e. the average over the angles variables θ. So $\mathcal{R}W$ is a function of A only (i.e. also "integrable"), and any 2 functions of A only, do commute mutually. Then the flow $e^{t\{\mathcal{R}W\}}$ is of the same type as the flow $e^{t\{H\}}$, which is:

$$e^{t\{H\}} = e^{t.\omega(A)\partial_\theta} \tag{6.2}$$

This is the operator translating the variable θ by $t.\omega(A)$:

$$(e^{t\{H\}}V)(A, \theta) = V(A, \theta + t.\omega(A)) \tag{6.3}$$

or after a Fourier transformation in θ:

$$(e^{t\{H\}}V)(A, \Delta) = e^{\mathbf{i}t.\omega(A)\cdot\Delta}.V(A, \Delta) \tag{6.4}$$

And similarly for $e^{t\{\mathcal{R}W\}}$, with $\omega(A)$ replaced by $(\mathcal{R}W)'(A) = (\mathcal{R}W')(A)$.

Application 2 - *Quantum Mechanics:*

When $\mathcal{A}$ has an involution ("*", named "adjonction": cf. Remark 1), and when $H = H^*$ the associated automorphism is called a "unitary transformation".
The bracket is defined as:

$$\forall A \in \mathcal{H} \quad \forall \Delta \in \mathcal{G}(A) \tag{6.5}$$

$$(\{W\}V)(A,\Delta) = \frac{\mathbf{i}}{\hbar} \sum_{\Delta' \in \mathcal{G}(A)} W(A+\Delta', \Delta-\Delta').V(A,\Delta') - V(A+\Delta',\Delta-\Delta').W(A,\Delta')$$

We can put this bracket in a form similar to the Poisson bracket (6.1), by adding and substracting 2 terms:

$$\forall A \in \mathcal{H} \;\; \forall \Delta \in \mathcal{G}(A) \tag{6.6}$$

$$(\{W\}V)(A,\Delta) = \sum_{\Delta' \in \mathcal{G}(A)} \mathbf{i}\left[\frac{W(A+\Delta',\Delta-\Delta') - W(A,\Delta-\Delta')}{\hbar}\right].V(A,\Delta') -$$

$$\mathbf{i}\left[\frac{V(A+\Delta',\Delta-\Delta') - V(A,\Delta-\Delta')}{\hbar}\right].W(A,\Delta') + \{W(A,\Delta')\}V(A,\Delta-\Delta')$$

where the last term is a short notation for:

$$\mathbf{i}\left[\frac{W(A,\Delta').V(A,\Delta-\Delta') - V(A,\Delta-\Delta').W(A,\Delta')}{\hbar}\right] \tag{6.7}$$

The formula (6.6) is reminiscent of the famous "correspondance principle" between the classical mechanics and the quantum mechanics: if the set $\mathcal{H}$ becomes more and more dense in $\mathbb{R}$ (for instance if it is $\hbar\mathbb{Z}$) with $\hbar \to 0$, in such a way that A remains constant, then the last term (6.7) tends to 0 and the first term in bracket tends to the derivative of W with respect to A, multiplied by Δ'. So that the quantum bracket becomes the classical one. Cf. [1] for a precise formulation of this fact, in some particular cases.
From now on, we will choose $\hbar = 1$, since we won't use this semi-classical limit.
Here again, if H is "non-resonant", $\mathcal{R}W$ is a diagonal matrix, and so its flow is of the same type as the flow of H, since:

$$(e^{t\{H\}}V)(A,\Delta) = e^{\mathbf{i}t.[h(A+\Delta)-h(A)]}.V(A,\Delta) \tag{6.8}$$

and similarly for the other flow.

7 - Quantitative Estimates

We start by chosing an arbitrary norm on $\mathcal{A}$. And we replace $\mathcal{A}$ by its closure with respect to this norm. We deduce a canonical norm for the operator Γ which acts bilinearly on $\mathcal{A}$:

$$|||\Gamma||| := \sup_{V,W \in \mathcal{A} \text{ s.t. } ||V||=||W||=1} ||\{\Gamma W\}V|| \tag{7.1}$$

We make an important assumption:

Hypothesis 2

$$|||\Gamma||| < \infty \tag{7.2}$$

This hypothesis is necessary to be able to apply the so-called "local bijection theorem" to invert the function F. When (7.2) is not true, it can be replaced by a weaker one, but then we need to use the more complicated theorem of Nash-Moser, cf [10], [9], [7], [5], or the Newton iterative method, as in the KAM theory. They are based on a Frechet structure on $\mathcal{A}$ i.e. an infinite sequence of norms $||\ldots||_{s-1} \geq ||\ldots||_s$ instead of only 1 norm. The hypothesis (7.2) is still required but with a weaker norm:

$$|||\Gamma|||_\alpha := \sup_{s\in\mathbb{N}} \sup_{V,W\in\mathcal{A} \text{ s.t. } ||V||_{s-1}=||W||_s=1} ||\{\Gamma V\}W||_s \,.\, \alpha(s) \tag{7.3}$$

where $\alpha : \mathbb{N} \to \mathbb{R}_+^*$ and $\lim_{s\to\infty} \alpha(s) = 0$. Hence Γ may be bounded if we admit some loss of "regularity".
Let us also define a norm:

$$||||\mathcal{R}|||| := \sup_{W\in\mathcal{A} \text{ s.t. } ||W||=1} ||\mathcal{R}W|| \tag{7.4}$$

Another assumption, is on the operator $\mathcal{R}$:

Hypothesis 3

$$||||\mathcal{R}|||| \leq 1 \tag{7.5}$$

It is fulfiled for many norms, for instance those given in (10.1): cf (10.2).

Then F is invertible when $||V||$ is small enough:

Theorem 3 - *Let V be an element of $\mathcal{A}$ and Γ be defined by (1.6), or explicitly by (1.19) or (1.31). Under the hypothesis 1, 2 and 3:*

$$\text{If} \quad ||V|| \leq \frac{1}{5|||\Gamma|||} \tag{7.6}$$

$$\text{Then} \quad \frac{24}{35} \leq \frac{||F^{-1}(V)||}{||V||} \leq \frac{24}{13} \tag{7.7}$$

Proof: Let us start by expanding $f(W)$ as given by (2.2) in series:

$$f(W) = \sum_{n=1}^{\infty} \{-\Gamma W\}^n \cdot \frac{n\mathcal{R}+1}{n+1!} \cdot W \tag{7.8}$$

and use the definition (7.1):

$$||\{\Gamma W\}V|| \leq |||\Gamma|||.||W||.||V|| \tag{7.9}$$

and the hypothesis (7.5), so that:

$$||f(W)|| \leq \sum_{n=1}^{\infty} (|||\Gamma|||.||W||)^n \cdot \frac{n||||\mathcal{R}||||+1}{n+1!} \cdot ||W|| \leq (e^{|||\Gamma|||.||W||} - 1).||W|| \tag{7.10}$$

This proves that $f(W) = \mathcal{O}(W^2)$. To compute the derivative of $f(W)$ with respect to W we need the following formula, valid for any derivation $\partial \in \mathrm{Der}(\mathcal{A})$ of some algebra $\mathcal{A}$:

$$\partial e^V = \int_0^1 dt.e^{t.V}.\partial V.e^{(1-t).V} \tag{7.11}$$

which may be rewritten as:

$$\partial e^V = \left(\frac{e^{\{V\}} - 1}{\{V\}} \partial V \right) . e^V \tag{7.12}$$

where the Lie-bracket $\{\ldots\}$ is given by the commutator. Indeed:

$$e^V . W . e^{-V} = e^{\{V\}} W \tag{7.13}$$

The proof of (7.13) is obtained by power expanding the exponentials:

$$\forall n \in \mathbb{N} \quad \partial(V^n) = \sum_{k=1}^{n} \binom{n}{k} . \{V\}^{k-1} . (\partial V) . V^{n-k} \tag{7.14}$$

which is proven by a simple recurrence. Similarly for (7.11):

$$\partial e^V = \sum_{N=0}^{\infty} \frac{\partial(V^N)}{N!} = \sum_{n,k=0}^{\infty} \frac{V^n . \partial V . V^k}{(n+k+1)!} \tag{7.15}$$

whereas:

$$\int_0^1 dt . e^{t.V} . \partial V . e^{(1-t).V} = \sum_{n,k=0}^{\infty} \int_0^1 dt . \frac{t^n . (1-t)^k}{n! . k!} V^n . \partial V . V^k \tag{7.16}$$

These 2 expressions coincide after using:

$$\int_0^1 dt . \frac{t^n . (1-t)^k}{n! . k!} = \frac{1}{(n+k+1)!} \tag{7.17}$$

Let us note that (7.11) is a generalization of the following formula (valid when t varies in a finite set) to the case where t is a continuous variable:

$$\partial \left(\prod_t V_t \right) = \sum_t \left(\prod_{\tau < t} V_\tau \right) (\partial V_t) \left(\prod_{\tau > t} V_\tau \right) \tag{7.18}$$

When ∂V commute with V we retrieve the usual formula:

$$\partial e^V = (\partial V) . e^V = e^V . (\partial V) \tag{7.19}$$

Let us apply (7.11) to the Lie algebra $\mathfrak{L}(\mathcal{A})$ of the endomorphisms of $\mathcal{A}$ (the space of observables), for which the bracket is indeed the commutator. And we take for ∂ the derivation with respect to W:

$$\partial e^{\{\Gamma W\}} = \int_0^1 dt . e^{t.\{\Gamma W\}} . \{\Gamma \ldots\} . e^{(1-t).\{\Gamma W\}} \tag{7.20}$$

which can be rewritten as:

$$\partial e^{\{\Gamma W\}} = \left(\frac{e^{\{\{\Gamma W\}\}} - 1}{\{\{\Gamma W\}\}} . \{\Gamma \ldots\} \right) . e^{\{\Gamma W\}} \tag{7.21}$$

where the double bracket is the bracket in $\mathfrak{L}(\mathcal{A})$ i.e. the commutator:

$$\forall X \in \mathfrak{L}(\mathcal{A}) \quad \{\{\Gamma W\}\} X := \{\Gamma W\} . X - X . \{\Gamma W\} \tag{7.22}$$

Therefore:

$$F'(W)V = \Big(\psi\big(\{\{\Gamma W\}\}\big)\{\Gamma V\}\Big)e^{-\{\Gamma W\}}\mathcal{R}W + e^{-\{\Gamma W\}}\mathcal{R}V + \left(\frac{1-e^{-\{\Gamma W\}}}{\{\Gamma W\}}\right)\mathcal{N}V + \Big(\varphi\big(\{\{\Gamma W\}\}\big)\{\Gamma V\}\Big)e^{-\{\Gamma W\}}\mathcal{N}W \tag{7.23}$$

where:

$$\psi(X) := \frac{1-e^{-X}}{X} \qquad\qquad \varphi(X) := \int_0^1 ds.\psi(sX) \tag{7.24}$$

Let us now use the formula (7.9):

$$||e^{\{\Gamma W\}}.V|| \leq e^{|||\Gamma|||.||W||}.||V|| \tag{7.25}$$

so that:

$$||\partial e^{\{\Gamma W\}}.V|| \leq \int_0^1 dt.e^{t.|||\Gamma|||.||W||}.|||\Gamma|||.e^{(1-t).|||\Gamma|||.||W||}.||V|| \;=\; e^{|||\Gamma|||.||W||}.|||\Gamma|||.||V|| \tag{7.26}$$

And we don't need the formula (7.23). Hence $f'(W)$ has a norm bounded by:

$$||f'(W)|| \;\leq\; \sum_{n=1}^{\infty} n.(|||\Gamma|||.||W||)^{n-1}.|||\Gamma|||\cdot\frac{n||||\mathcal{R}||||+1}{n+1!}\cdot||W|| + (|||\Gamma|||.||W||)^{n}\cdot\frac{n||||\mathcal{R}||||+1}{n+1!} \tag{7.27}$$

so that, using (7.5):

$$||f'(W)|| \leq e^{|||\Gamma|||.||W||}.(|||\Gamma|||.||W||+1)-1 \tag{7.28}$$

Let us call γ the solution of the transcendental equation:

$$e^{\gamma}(\gamma+1)-1 = 1 \quad \text{i.e.} \quad \gamma = 0.3748225258118948\ldots > 13/35 \tag{7.29}$$

Hence:

$$\text{If} \quad |||\Gamma|||.||W|| < \gamma \quad \text{Then} \quad ||f'(W)|| < 1 \tag{7.30}$$

and so:

$$||F'(W)|| = ||1+f'(W)|| \geq 1 - ||f'(W)|| > 0 \tag{7.31}$$

Then F is invertible and:

$$\begin{aligned} ||F(W)|| = ||W+f(W)|| &\leq ||W||.\left(1+\frac{||f(W)||}{||W||}\right) \\ &\leq ||W||.(1+e^{\gamma}-1) = e^{\gamma}.||W|| \end{aligned} \tag{7.32}$$

Likewise:

$$||F(W)|| \geq ||W||.\left(1-\frac{||f(W)||}{||W||}\right) \geq ||W||.(2-e^{\gamma}) \tag{7.33}$$

So if we replace W by $F^{-1}(V)$ we get:

$$e^{-\gamma} \leq \frac{||F^{-1}(V)||}{||V||} \leq \frac{1}{2-e^{\gamma}} \tag{7.34}$$

under the condition (7.30), i.e. if:

$$||V|| < \frac{(2 - e^\gamma).\gamma}{|||\Gamma|||} \tag{7.35}$$

Indeed we will have in that case:

$$||W|| = ||F^{-1}(V)|| \leq \frac{||V||}{2 - e^\gamma} < \frac{\gamma}{|||\Gamma|||} \tag{7.36}$$

To prove the theorem 3, there remains to use the value of γ: $e^\gamma = \frac{2}{\gamma+1} < \frac{35}{24}$. *Q.E.D.* ∎
Let us note that under the condition (7.6), the new term (the "control term", $f(V)$) will be smaller than V (cf. (7.10)):

$$\text{If} \quad ||V|| \leq \frac{1}{5|||\Gamma|||} \qquad \text{Then} \quad \frac{||f(V)||}{||V||} \leq e^{1/5} - 1 \quad \left(< \frac{2}{9}\right) \tag{7.37}$$

8 - A Formal Series for the Inverse of F

We want an expansion of $W = F^{-1}(V)$ in powers of V, using the expansion of $F(W) = W + f(W)$ given in (7.8). For that purpose we can use an extension of the Lagrange inversion formula which was established to invert a function from $\mathbb{C} \to \mathbb{C}$ in power series. An extension to the case where the argument is not a number but a function (like here), was given in [12], and latter in [2]. We first rewrite the definition of W as a fixed-point problem:

$$F(W) = V \iff W = G(W) := V - f(W) \tag{8.1}$$

A Taylor expansion of G is given by:

$$G(W) = \sum_{n \geq 0} \hat{G}(n) W^n \tag{8.2}$$

with (cf. 7.8):

$$\hat{G}(0) := G(0) = V \quad \text{and} \quad \hat{G}(1) := 0 \tag{8.3}$$

$$\forall n \geq 2 \quad \hat{G}(n) := \frac{G^{(n)}(0)}{n!} = \mathcal{S}(-1)^n \{\Gamma \ldots\}^{n-1} \cdot \frac{(n-1)\mathcal{R} + 1}{n!} \ldots \tag{8.4}$$

is an n-linear completely symmetric application from $\mathcal{A}^n$ into $\mathcal{A}$. Its n arguments are symbolized by the n "slots". In (8.2), this n-linear mapping is applied to identical arguments: n times W. When applied to n general arguments $W_1, \ldots, W_n$, it would give:

$$\hat{G}(n)(W_1, \ldots, W_n) = \mathcal{S}(-1)^n \{\Gamma W_1\} \ldots \{\Gamma W_{n-1}\} \cdot \frac{(n-1)\mathcal{R} + 1}{n!} W_n \tag{8.5}$$

The operator $\mathcal{S}$ is the symmetrization of the arguments which yields, when applied to an n-linear application T:

$$(\mathcal{S}T)(W_1, \ldots, W_n) := \frac{1}{n!} \sum_{\sigma \in \text{Permutations}} T(W_{\sigma(1)}, \ldots, W_{\sigma(n)}) \tag{8.6}$$

We will omit the parenthesis and the commas in using such tensors. So $W^M = (W, W, \ldots, W)$ (M times). For instance:

$$G'(0)(W_1) := \lim_{\lambda\to 0} \frac{G(\lambda W_1) - G(0)}{\lambda} \tag{8.7}$$

which is 0, for the function defined in (8.1), since it is quadratic in its argument. And $G''(0)$ is a tensor of order 2.

$$G''(0)(W_1, W_2) := \lim_{\lambda\to 0}\lim_{\mu\to 0} \frac{G(\lambda W_1 + \mu W_2) - G(\lambda W_1) - G(\mu W_2) + G(0)}{\lambda.\mu} \tag{8.8}$$

which can be easily computed to be:

$$\hat{G}(2)W_1.W_2 = \{\Gamma W_1\}\frac{\mathcal{R}+1}{4}W_2 + \{\Gamma W_2\}\frac{\mathcal{R}+1}{4}W_1 \tag{8.9}$$

Theorem 4 - *The above mentionned extension of the Lagrange inversion formula says that the solution of any fixed-point equation $W = G(W)$, with a function G given by a series (8.2), is formally:*

$$W = \sum_{N\geq 1}\sum_{\nu\in\mathcal{T}(N)} \hat{G}(\nu_{N-1})\hat{G}(\nu_{N-2})\ldots\hat{G}(\nu_1)\hat{G}(\nu_0) \tag{8.10}$$

where:

$$\begin{aligned}\mathcal{T}(N) := \{\nu = (\nu_0, \nu_1, \ldots, \nu_{N-1}) \in \{0, 1, \ldots, N-1\}^N \quad \text{s.t.} \\ \forall k \in \{0, 1, \ldots, N-1\} : \ |\nu|_k \leq k \quad \text{and} \quad |\nu|_{N-1} = N-1\}\end{aligned} \tag{8.11}$$

with:

$$|\nu|_k := \nu_0 + \nu_1 + \ldots + \nu_k \tag{8.12}$$

The expansion (8.10) is only useful if $\hat{G}(n)$ are "small", since we expand in powers of them. In our case (8.3), (8.4), only $\hat{G}(0)$ is small, and $\hat{G}(1) = 0$. But $\hat{G}(n)$ is of order 1, when $n \geq 2$. So we have to rearrange the series (8.10):

$$W = V + \sum_{M\geq 2} W_M \tag{8.13}$$

where:

$$W_M := \sum_{N=M+1}^{2M-1} \quad \sum_{\nu\in\mathcal{T}(N) \text{ s.t. } \nu^{-1}(0)=M} \hat{G}(\nu_{N-1})\hat{G}(\nu_{N-2})\ldots\hat{G}(\nu_1)\hat{G}(\nu_0) \tag{8.14}$$

with the restriction "$\nu^{-1}(0) = M$" in the sum over ν, meaning that ν must take exactly M times the value 0. Indeed this is exactly the case "$\nu_k = 0$" which produces a factor V. We can also notice that in the definition (8.11) of $\mathcal{T}(N)$ we can restrict $\nu \in \{0, 2, 3, \ldots, N-1\}$ since in our case $\hat{G}(1) = 0$. Before we explicit the first terms of the solution (8.13), let us understand heuristically the series (8.10) or (8.13). A first approximation of the solution of $W = G(W)$ is $W \approx G(0) := V$. Then a second approximation is $W \approx$

$G(G(0)) \approx G(0) + G'(0).G(0)$. And a third one, at the order V^3 , is $W \approx G(G(G(0))) \approx G[G(0) + G'(0).G(0)]$ i.e.:

$$W \approx G(0) + G'(0).G(0) + G'(0).G'(0).G(0) + \frac{1}{2}.G''(0).G(0).G(0) \tag{8.15}$$

and so on. We are building the series (8.10). Now we have to check that the general term of (8.10) is indeed an element of $\mathcal{A}$ and not an arbitrary m-linear mapping for some m. Each term is a composition of high-order tensors, and this product is actually a "vector" i.e. an element of $\mathcal{A}$. This is due to the definition of $\mathfrak{T}(N)$: we note that $\hat{G}(n)$ is of type $V.\bar{V}^n$ i.e. it is a (sum of) tensorial product(s) of a vector V and of n covectors $\bar{V}$. More generally we say that a tensor is of type $V^m.\bar{V}^n$ when it is n times covariant and m times contravariant. This is a convenient way to consider such tensors, in order to keep track of all subsequent contractions. Of course the ordering is crucial: $V.\bar{V}$ is a matrix (a tensor of order 2) whereas $\bar{V}.V$ is a scalar (a scalar product). Each time a covector follows (on the left) a vector, we make the contraction, i.e. we reduce the tensorial product to a scalar. Actually we just want to give a meaning to the composition of multilinear applications. To this purpose we can define:

$$\forall B \in \mathcal{L}(\mathcal{A}^n, \mathcal{A}) \quad \forall C \in \mathcal{L}(\mathcal{A}^m, \mathcal{A}) \quad B.C \in \mathcal{L}(\mathcal{A}^{m+n-1}, \mathcal{A}) \tag{8.16}$$

by:

$$(B.C)(V_1, \ldots, V_m, V_{m+1}, \ldots, V_{m+n-1}) := B\Big(C(V_1, \ldots, V_m), V_{m+1}, \ldots, V_{m+n-1}\Big) \in \mathcal{A} \tag{8.17}$$

Now we note that for any $\nu \in \mathfrak{T}(N) : \nu_0 = 0$ since $|\nu|_0 = 0$. So the right-most factor in (8.10) is a vector: $\hat{G}(0) = V \in \mathcal{A}$.
Then the preceding one is $\hat{G}(\nu_1)$ where $\nu_1 \leq 1$ since $|\nu|_1 = \nu_0 + \nu_1$ has to be less than 1. Hence $\hat{G}(\nu_1)$ is either a vector (if $\nu_1 = 0$) or a matrix (if $\nu_1 = 1$). In the latter case the product $\hat{G}(1)\hat{G}(0)$ is a vector, and in the former case we have a "bi-vector" which will be made into a vector by the higher-order terms: indeed if $\nu_1 = 0$ then there exists a $k \geq 2$ such that $\nu_k \geq 2$ since $|\nu|_{N-1}$ has to be $N-1$. More precisely we note that the product of a tensor $V^a.\bar{V}^b$ by a tensor $V^c.\bar{V}^d$ is a tensor of type:

$$V^a.\bar{V}^b.V^c.\bar{V}^d = V^{a+c-\min(b,c)}.\bar{V}^{b+d-\min(b,c)} \tag{8.18}$$

So the product $\hat{G}(\nu_k)\ldots\hat{G}(\nu_0)$ is a tensor of type:

$$\hat{G}(\nu_k)\ldots\hat{G}(\nu_0) = V^{\max(k-|\nu|_k,1)}.\bar{V}^0 \tag{8.19}$$

i.e. without any covector, since we assume $|\nu|_k \leq k \quad \forall k$. Finally for $k = N-1$, $|\nu|_{N-1}$ has to be $N-1$, so (8.10) is indeed a sum of vectors.
The idea of the proof of (8.10) is to plug the r.h.s. of (8.10) into the expression (8.2) of $G(W)$ and to recognise that we get the same expansion than (8.10), i.e. this is also W.

A simple illustration (which is not of our type) of the theorem 4 is when:

$$G(W) := V.e^{\bar{Y}\cdot W} \tag{8.20}$$

for some $V \in \mathcal{A}$ and $\bar{Y} \in \mathcal{A}^* := \mathcal{L}(\mathcal{A}, \mathbb{R})$. In that simple case, the theorem 4 gives the solution of the "transcendental" equation in W:

$$W = V.e^{\bar{Y}\cdot W} \tag{8.21}$$

as:

$$W = V.\lambda(\bar{Y} \cdot V) \quad \text{where} \quad \lambda(x) := \sum_{n \geq 0} \frac{(n+1)^{n-1}}{n!} x^n \tag{8.22}$$

So that $\lambda(x)$ converges if $|x| \leq e^{-1}$ (but may be extended if $x < -e^{-1}$).

We can rearrange the series (8.13) as follows:

$$W_M := \sum_{N=1}^{M-1} \sum_{\lambda \in \mathcal{B}(N,M)} \sum_{\mu \in \mathcal{C}(N,\lambda)} \hat{G}(\lambda_1^\mu) \dots \hat{G}(\lambda_{N+M}^\mu) \tag{8.23}$$

where $\hat{G}$ is defined in (8.3) and (8.4), and:

$$\mathcal{B}(N, M) := \{\lambda \in \mathbb{N}_*^N \ \text{ s.t. } \quad |\lambda|_N = M - 1\} \tag{8.24}$$

and:

$$\mathcal{C}(N, \lambda) := \{\mu \in \mathbb{N}_*^N \ \text{ s.t. } \quad \forall\, 1 \leq n \leq N \quad |\mu|_n \leq n + |\lambda|_{n-1}\} \tag{8.25}$$

with:

$$|\mu|_n := \mu_1 + \ldots + \mu_n \tag{8.26}$$

ans similarly for $|\lambda|_n$ with $|\lambda|_0 := 0$. Finally in (8.23) the integer λ_l^μ is defined for any $1 \leq l \leq N + M$ by:

$$\begin{aligned} \lambda_l^\mu &= \lambda_n + 1 \quad \text{if} \quad l = |\mu|_n \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{8.27}$$

This means that $\lambda_l^\mu = 0$ when l is not an element of the set of the values $|\mu|_n$, when n varies in $\{1, \ldots, N\}$. This case occurs M times when l varies in $\{1, \ldots, N + M\}$. Otherwise when $l = |\mu|_n$ for some n, then $\lambda_l^\mu = \lambda_n + 1$, for this n.

Finally we can also rearrange the Lagrange series in our specific case (8.3) & (8.4) into:

$$W_M := \sum_{N=1}^{M-1} \sum_{\lambda \in \mathcal{B}(N,M)} \sum_{\nu \in \mathcal{D}(N,\lambda)} (-1)^{N+M+1} \{\Gamma \ldots\}^{\lambda_N} \frac{\lambda_N \mathcal{R} + 1}{\lambda_N + 1!} V^{\nu_N} \cdot \ldots \cdot \{\Gamma \ldots\}^{\lambda_1} \frac{\lambda_1 \mathcal{R} + 1}{\lambda_1 + 1!} V^{\nu_1} \tag{8.28}$$

where:

$$\mathcal{D}(N, \lambda) := \{\nu \in \mathbb{N}^N \ \text{ s.t. } \quad |\nu|_N = M \quad \& \quad \forall\, 1 \leq n \leq N \quad |\nu|_n > |\lambda|_n\} \tag{8.29}$$

Let us note that:

$$||\hat{G}(n)|| \leq \frac{|||\Gamma|||^{n-1}}{(n-1)!} \tag{8.30}$$

and so:

$$||W_M|| \leq c_M.|||\Gamma|||^{M-1}.||V||^M \tag{8.31}$$

for some positive constant c_M. Indeed: $\lambda_1 + \ldots + \lambda_N = M - 1$.

Let us explicit the first orders of the expansion of the solution W of our problem (8.1)-(8.4), as given by (8.13) or (8.23):

$$F^{-1}(V) = W = \sum_{M \geq 1} W_M \qquad \text{where} \quad W_1 = V \qquad W_2 = \{\Gamma V\} \frac{\mathcal{R} + 1}{2} V \tag{8.32}$$

$$W_3 = \left(\{\Gamma V\}\frac{\mathcal{R}+1}{2}\right)^2 V - \{\Gamma V\}^2\frac{2\mathcal{R}+1}{6}V + \left\{\Gamma\left(\{\Gamma V\}\frac{\mathcal{R}+1}{2}V\right)\right\}\frac{\mathcal{R}+1}{2}V$$

$$W_4 = \left(\{\Gamma V\}\frac{\mathcal{R}+1}{2}\right)^3 V + \{\Gamma V\}^3\frac{3\mathcal{R}+1}{24}V - \{\Gamma V\}^2\frac{2\mathcal{R}+1}{6}\{\Gamma V\}\frac{\mathcal{R}+1}{2}V -$$
$$\left\{\Gamma\left(\{\Gamma V\}\frac{\mathcal{R}+1}{2}V\right)\right\}\frac{\mathcal{R}+1}{2}\{\Gamma V\}\frac{\mathcal{R}+1}{2}V - \left\{\Gamma\left(\{\Gamma V\}^2\frac{2\mathcal{R}+1}{6}V\right)\right\}\frac{\mathcal{R}+1}{2}V -$$
$$\{\Gamma V\}\frac{\mathcal{R}+1}{2}\left\{\{\Gamma V\}^2\frac{2\mathcal{R}+1}{6}V\right\} + \{\Gamma V\}\frac{\mathcal{R}+1}{2}\left\{\Gamma\left(\{\Gamma V\}\frac{\mathcal{R}+1}{2}V\right)\right\}\frac{\mathcal{R}+1}{2}V +$$
$$\left\{\Gamma\left(\left\{\Gamma\left(\{\Gamma V\}\frac{\mathcal{R}+1}{2}V\right)\right\}\frac{\mathcal{R}+1}{2}V\right)\right\}\frac{\mathcal{R}+1}{2}V + \left\{\Gamma\left(\left(\{\Gamma V\}\frac{\mathcal{R}+1}{2}\right)^2 V\right)\right\}\frac{\mathcal{R}+1}{2}V -$$
$$\left\{\{\Gamma V\}\Gamma\left(\{\Gamma V\}\frac{\mathcal{R}+1}{2}V\right)\right\}\frac{2\mathcal{R}+1}{6}V$$

and so on for W_5... We have proven in Theorem 3 that this series converges at least when:

$$||V|| \ \leq \ \frac{1}{5|||\Gamma|||} \tag{8.33}$$

The paper [6] proves that a series similar to this one (the Lindstedt series, which is local in the variable A) converges. The key ingredient is the compensation between the terms of different signs.

Remark: as we have noticed at the end of section 1, the solution W is independant of the constant $\hbar$ since the operator Γ always appears with a bracket around it. And the pair $\{\Gamma\ldots\}$ is independant of $\hbar$.

9 - Example 1: Classical Control Theory

We will study a simple model, introduced in [11] and described in [3], of a charged particle in a plasma, in a Tokamak, which is a reactor for the controlled thermonuclear fusion. We consider a section of the Tokamak as a Poincaré section of a dynamical system with 1 degree of freedom. So that the particle has 1 degree of freedom (p, q), but it is embedded in a complicated electric field, depending on time. The fast motion given by the strong magnetic field has been averaged out. The hamiltonian is:

$$H(p,q,E,\tau) + V(p,q,E,\tau) \quad \text{where} \quad H(p,q,E,\tau) = E \tag{9.1}$$
$$\text{and} \quad V(p,q,E,\tau) = \sum_{n,m,k\neq 0} \varepsilon^{(1)}_{n,m,k}.\sin(nq+mp+k\tau) \tag{9.2}$$

The extended canonical coordinates are (E, τ) so that the motion of the new dynamical variable τ is trivial:

$$e^{t.\{H+V\}}\tau = \tau + t \tag{9.3}$$

Here we have applied the flow to the observable $\tau : (p,q,E,\tau) \mapsto \tau$. And of course V is independant of the variable E, which is the variable canonically conjugate to τ. Let

us note that in (9.2), the variable k must be integer (or at least away from 0) but the variables m, n may be integers or reals: in that case the sum over them should be replaced by an integral.
The hamiltonian H is resonant since $\{H\} = \partial_\tau$ so that:

$$(\mathcal{R}V)(p,q,\tau) = \oint d\tau . V(p,q,\tau) \tag{9.4}$$

is independent of τ. Indeed a Fourier transformation in τ yields:

$$(\hat{\mathcal{R}V})(p,q,E,k) = \hat{V}(p,q,E,0).\ \chi(k=0) \tag{9.5}$$

For the perturbation (9.2) we have: $\mathcal{R}V = 0$. The action of the operator Γ is defined by:

$$(\Gamma V)(p,q,\tau) = -\sum_{n,m,k\neq 0} \frac{\varepsilon^{(1)}_{n,m,k}}{k} \cdot \cos(nq + mp + k\tau) \tag{9.6}$$

Then the "control term" $f(V)$ can be explicitly computed, cf (7.8) with $\mathcal{R}V = 0$:

$$f(V) = \sum_{s\geq 2} f_s \quad \text{where} \quad f_s := \frac{\{-\Gamma V\}^{s-1}}{s!} V \tag{9.7}$$

Hence:

$$f(V)(p,q,\tau) = \sum_{n,m,k\neq 0} \varepsilon_{n,m,k} . \sin(nq + mp + k\tau) \tag{9.8}$$

where:

$$\varepsilon_{n,m,k} := \sum_{s\geq 2} (-1)^{s-1} \cdot \frac{\varepsilon^{(s)}_{n,m,k}}{s!} \tag{9.9}$$

with:

$$\varepsilon^{(s)}_{n,m,k} := \sum_{N,M,K\neq 0} \frac{\varepsilon^{(1)}_{N,M,K}}{K} \cdot \left(\varepsilon^{(s-1)}_{N+n,M+m,K+k} - \varepsilon^{(s-1)}_{N-n,M-m,K-k}\right) .(M.n - N.m) \tag{9.10}$$

The proof of (9.10) is based on:

$$\begin{aligned} &\Big\{\cos(Nq + Mp + K\tau)\Big\} \sin(nq + mp + k\tau) = \\ &\qquad \frac{1}{2} \cdot (N.m - M.n) \cdot \Big[\sin[(N+n)q + (M+m)p + (K+k)\tau] + \\ &\qquad\qquad \sin[(N-n)q + (M-m)p + (K-k)\tau]\Big] \end{aligned} \tag{9.11}$$

so that we can iterate and compute $\{\Gamma V\}^s\ V$.

A simple case where the control term can be computed explicitly is the following. Let us choose:

$$\varepsilon \in \mathbb{R}^*, \quad b > 1/\sqrt{2}, \quad \sigma = \pm 1, \quad m, n \in \mathbb{R} \quad \text{s.t.} \quad m \neq n \tag{9.12}$$

And we take V as a particular case of (9.2), a sum of 2 waves:

$$V(p,q,\tau) = \varepsilon.\sigma.\sqrt{2b^2 - 1}. \sin(nq + mp + \tau) - \varepsilon. \sin(q + p + \tau) \tag{9.13}$$

We have taken $m \neq n$ to avoid that V depends only on a single variable $q+p$. The role of b (and σ) is to permit 2 different coupling constants. Then the control term is a sum of only 5 waves:

$$\begin{aligned} f(V) = \hat{\varepsilon}^2.\sigma.\sin\Big((n-1)q+(m-1)p\Big) + \\ \tilde{\varepsilon}^3.\Bigg[\sqrt{2b^2-1}.\Big(\sin(q+p+\tau)+\sin\Big((2n-1)q+(2m-1)p+\tau\Big)\Big) - \\ \sigma.\Big(\sin(nq+mp+\tau)+\sin\Big((2-n)q+(2-m)p+\tau\Big)\Big)\Bigg] \end{aligned} \tag{9.14}$$

where:

$$\hat{\varepsilon} := \frac{(2b^2-1)^{1/4}}{b} \cdot \left(\frac{1-\cos(\varepsilon.b.(m-n))}{|m-n|}\right)^{1/2} \approx |\varepsilon|.|m-n|^{1/2}.\Big(\frac{2b^2-1}{4}\Big)^{1/4} \tag{9.15}$$

$$\tilde{\varepsilon} := \frac{(4b^2-2)^{1/6}}{b} \cdot \left(\varepsilon.b - \frac{\sin(\varepsilon.b.(m-n))}{m-n}\right)^{1/3} \approx \varepsilon.|m-n|^{2/3}.\Big(\frac{2b^2-1}{18}\Big)^{1/6} \tag{9.16}$$

We have indicated the first order of the expansion of $\hat{\varepsilon}$ or $\tilde{\varepsilon}$ when ε is small. So that $f(V) = \mathcal{O}(V^2)$.
Cf also [3] for some numerical experiments that prove the effectiveness of this method, when the coefficients $\varepsilon_{n,m,k}$ are taken to reflect some properties of a realistic field, in a Tokamak:

$$\varepsilon_{n,m,k} = \frac{\varepsilon}{(n^2+m^2)^{\frac{3}{2}}} \cdot \chi(1 \leq n^2+m^2 \leq N^2).\ \chi(k=1) \tag{9.17}$$

for some constant ε proportional to the inverse of the (strong) magnetic field and for some "cut-off" N. In that case the first term of the control, $\varepsilon^{(2)}_{n,m,k}$ is vanishing when $k \neq 0$.

Let us summarize the method of control, in this case, where $\mathcal{R}V = 0$ and $\{H\} = \partial_\tau$:

$$\forall t \in \mathbb{R} \qquad e^{t\{H+V+f(V)\}} = e^{-\{\Gamma V\}}.e^{t\partial_\tau}.e^{\{\Gamma V\}} \tag{9.18}$$

so that the distance between the dynamical variable p (or q) at the time t and at the initial time is:

$$p_t - p_0 = \Big(e^{t\{H+V+f(V)\}} - 1\Big)p_0 \tag{9.19}$$

Let us write p instead of p_0, and use the "telescopic" formula:

$$\forall a,b,c: \quad a.b.c - 1 = (a-1) + a.(b-1) + a.b.(c-1) \tag{9.20}$$

And we replace the flow in (9.19) by its decomposition (9.18):

$$p_t - p = \Big[(e^{-\{\Gamma V\}}-1) + e^{-\{\Gamma V\}}.(e^{t\partial_\tau}-1) + e^{-\{\Gamma V\}}.e^{t\partial_\tau}.(e^{\{\Gamma V\}}-1)\Big]p \tag{9.21}$$

But the middle term vanishes since $\partial_\tau p = 0$, so that:

$$p_t - p = (e^{-\{\Gamma V\}}-1)p + e^{-\{\Gamma V\}}.e^{t\partial_\tau}.(e^{\{\Gamma V\}}-1)p \tag{9.22}$$

And we can divide and multiply by $\{\Gamma V\}$, and use the antisymmetry (1.2):

$$p_t - p = \Big(\frac{1 - e^{-\{\Gamma V\}}}{\{\Gamma V\}}\Big)\{p\}\Gamma V - e^{-\{\Gamma V\}}.e^{t\partial_\tau}.\Big(\frac{e^{\{\Gamma V\}} - 1}{\{\Gamma V\}}\Big)\{p\}\Gamma V \tag{9.23}$$

Finally, let us note that $\{p\} = -\partial_q$, i.e. the formula (9.23) can be explicitly computed.

When we apply an approximate control term φ instead of the exact one $f(V)$, the formula (9.18) becomes:

$$e^{t\{H+V+\varphi\}} = e^{-\{\Gamma V\}}.e^{-\{\Gamma W\}}.e^{t\partial_\tau}.e^{t\{\mathcal{R}W\}}.e^{\{\Gamma W\}}.e^{\{\Gamma V\}} \tag{9.24}$$

where:

$$W := F^{-1}\Big(e^{\{\Gamma V\}}.\big(\varphi - f(V)\big)\Big) \tag{9.25}$$

When we take $\varphi = 0$, (9.25) is the begining of the KAM recursive method. See [5] for more details, in the case of quantum mechanics.

10 - Example 2: Quantum Adiabatic Transformation

An example of typical norm on $\mathcal{A}$ is given by an arbitrary "weight" function $g(A, \Delta) > 0$. Let us define a norm on the algebra $\mathcal{A}$ by:

$$||V|| := \sup_A \sum_\Delta |V(A, \Delta)|/g(A, \Delta) \tag{10.1}$$

In the quantum mechanical case, we will take the usual L^2-operator norm on each "block" $V(A, \Delta)$. This choice is irrelevant in the case where all the projectors P_A are finite-dimensional. We could have chosen $\sup_{A,\Delta}$ but this would have just been (approximately) a modification of the "weight" function g.

Lemma 1 *We have:*

$$||\mathcal{R}V|| \leq ||V|| \quad \& \quad ||\mathcal{N}V|| \leq ||V|| \tag{10.2}$$

Proof: The operators $\mathcal{R}$ and $\mathcal{N}$ are implemented by some characteristic functions: cf (1.17), (1.18), (1.29), (1.30). So the sup & sum in (10.1) are restricted by some conditions: therefore the norm decreases. *Q.E.D.* ∎

Hence the hypothesis 3 is fulfiled. Let us consider as before a hamiltonian H but where now the perturbation depends on time. We need to extend the algebra and the full hamiltonian is now: $D + H$ where D is the "derivative with respect to time". We can apply the proposition 1 (5.1) but we still use the same operator Γ i.e. the pseudo-inverse of the bracket with H, and also the same function F defined in (2.1). Let us rewrite (5.1):

$$\forall W \in \mathcal{A} \quad H + F(W) = e^{-\{\Gamma W\}}(H + \mathcal{R}W) \tag{10.3}$$

Now we want some information on $D + H + V$:

$$D + H + F(W) = D + e^{-\{\Gamma W\}}(H + \mathcal{R}W) = e^{-\{\Gamma W\}}(e^{\{\Gamma W\}}D + H + \mathcal{R}W) \tag{10.4}$$

where:

$$W := F^{-1}(V) \tag{10.5}$$

Hence:
$$D + H + F(W) = e^{-\{\Gamma W\}}\Big(D + \frac{e^{\{\Gamma W\}}-1}{\{\Gamma W\}}.\{\Gamma W\}D + H + \mathcal{R}W\Big) \tag{10.6}$$
so, using (1.2), and the notation $\dot W := \{D\}W$:
$$D + H + V = e^{-\{\Gamma W\}}\big(D + H_1 + V_1\big) \tag{10.7}$$
with:
$$H_1 := H + \mathcal{R}W \qquad V_1 := -\frac{e^{\{\Gamma W\}}-1}{\{\Gamma W\}}.\Gamma\dot W \tag{10.8}$$
Indeed $\{D\}\Gamma W = \Gamma\{D\}W = \Gamma\dot W$. The formula (10.7) is useful if the derivative $\dot V$ of the perturbation with respect to time is "smaller" than the perturbation V itself. This is the "adiabatic hypothesis". In that case we can also show that $\dot W$ is small, since $W \approx V$. So the new perturbation is "smaller" than the original perturbation, or more qualitatively it is approximately: $\Gamma\dot W$ i.e. $\{H\}^{-1}\{D\}V$. This means that it has losed some regularity in time, but gained some regularity in its "spatial" dependance, since generally $\{H\}^{-1}$ is a regularizing operator. The difference between H and the new H_1 is of the same size as the original perturbation, but "diagonalized" with respect to H since it commutes with it. Cf also [4] for an iterative proof of this result.
Let us finally note that it is possible to iterate this procedure:
$$D + H + V = e^{-\{\Gamma W\}}e^{-\{\Gamma_1 W_1\}}\big(D + H_2 + V_2\big) \qquad \text{where} \qquad W_1 := F_1^{-1}(V_1) \tag{10.9}$$
and:
$$H_2 := H_1 + \mathcal{R}_1 W_1 \qquad V_2 := -\frac{e^{\{\Gamma_1 W_1\}}-1}{\{\Gamma_1 W_1\}}.\Gamma_1\dot W_1 \tag{10.10}$$
and with $\mathcal{R}_1, \mathcal{N}_1, \Gamma_1, F_1$ defined as before, but with H replaced by H_1. And so on. Generally this iteration doesn't converge: so we must stop it at an optimized order. This has been done in [8], with a different method and framework.
Finally let us explicit the norm (7.1) of the operator Γ for the norm chosen in (10.1):
$$|||\Gamma||| \leq \sup_{A,A' \text{ s.t. } A\neq A'} \frac{\psi(A,A')}{|h(A)-h(A')|} \tag{10.11}$$
where:
$$\psi(A,A') := g_{A,A'}.\ \sup_{A''}\max\Big(\frac{g_{A,A''}}{g_{A',A''}}\ ,\ \frac{g_{A',A''}}{g_{A,A''}}\Big) \tag{10.12}$$
and $g_{A,A'} = g(A, A'-A)$. The proof is an easy estimation of the norm (7.1). Hence the hypothesis 2 (7.2) requires that:
$$\forall A, A' \text{ s.t. } A \neq A' \qquad |h(A)-h(A')| \geq \gamma.\psi(A,A') \tag{10.13}$$
for some constant $\gamma > 0$. This is a "Diophantine" condition. It depends on the choice of the weight function g, i.e. on the regularity of the perturbations V we want to consider. For instance, when $\mathcal{H} = \mathbb{N}^*$:
$$\text{If} \quad g_{A,A'} = \varphi_A \cdot \varphi_{A'} \quad \text{then} \quad \psi(A,A') = \max(\varphi_A^2, \varphi_{A'}^2) \tag{10.14}$$
As a particular case we can take, for some $\alpha \geq 1$:
$$h(A) = A^\alpha \quad \text{and} \quad \varphi_A = A^{\frac{\alpha-1}{2}} \tag{10.15}$$
for which the condition (10.13) is satisfied.
See [5] for more details.

11 - References

[1] J. BELLISSARD, M. VITTOT: "Heisenberg's picture and non commutative geometry of the semi classical limit in quantum mechanics", Annales de l'Institut Henri Poincaré, A, vol 52, 3, (1990) p 175-235.

[2] T. CARLETTI: "The Lagrange inversion formula on non-Archimedean fields. Non-Analytical Form of Differential and Finite Difference Equations", Discrete and Continuous Dynamical Systems, Series A, Vol 9, 4, (July 2003), p 835-858. Archived in http://arxiv.org/pdf/math/0110135

[3] G. CIRAOLO, C. CHANDRE, R. LIMA, M. PETTINI, M. VITTOT, Ch. FIGARELLA, Ph. GHENDRIH: "Controlling chaotic transport in an Hamiltonian model of interest to magnetized plasmas", http://arxiv.org/pdf/nlin.CD/0304040 (2003).

[4] P. DUCLOS, O. LEV, P. STOVICEK, M. VITTOT: "Progressive diagonalization and applications", Proceedings of the Conference "Operator Algebras & Mathematical Physics", Constanţa (Roumanie, 2001), R. Purice Ed., Theta Foundation, Bucarest (2003).

[5] P. DUCLOS, O. LEV, P. STOVICEK, M. VITTOT: "Weakly regular Floquet Hamiltonians with pure point spectrum", http://arxiv.org/pdf/math-ph/0103041, Reviews in Mathematical Physics, 14 6 (2002), p 1-38

[6] L.H. ELIASSON: "Hamiltonian systems with linear normal form near an invariant torus", in "Non-Linear Dynamics" (Bologne, 1988), G. Turchetti Ed., World Scientific, Singapore (1989), p 11-29.

[7] R.S. HAMILTON: "The inverse function theorem of Nash and Moser", Bull. Amer. Math. Soc. (N.S.), 7 (1), (1982), p 65-222.

[8] A. JOYE, Ch.-Ed. PFISTER: "Exponential Estimates in Adiabatic Quantum Evolution", International Congress of Mathematical Physics (ICMP 1997), De Wit, Bracken, Gould & Pearce Eds, International Press, Cambridge. 6 (1999), p.309-315

[9] J. MOSER: "A new technique for the construction of solutions of nonlinear differential equations", Proc. Nat. Acad. Sci. U.S.A., 47, (1961), p 1824-1831.

[10] J. NASH: "The imbedding problem for Riemannian manifolds", Ann. of Math. (2), 63, (1956), p 20-63.

[11] M. PETTINI: "Low Dimensional Hamiltonian Models for Non-Collisional Diffusion of Charged Particles", in "Non-Linear Dynamics" (Bologne, 1988), G. Turchetti Ed., World Scientific, Singapore (1989), p 287-296.

[12] M. VITTOT: "Lindstedt perturbation series in Hamiltonian mechanics: explicit formulation via a multidimensional Bürmann-Lagrange formula", Preprint CNRS Luminy CPT-91/P.2603, Marseille (1991). Archived in http://ccdb3fs.kek.jp/cgi-bin/img/allpdf?199204265