



HAL
open science

A Haar component for quantum limits on locally symmetric spaces

Nalini Anantharaman, Lior Silberman

► **To cite this version:**

Nalini Anantharaman, Lior Silberman. A Haar component for quantum limits on locally symmetric spaces. 2010. hal-00521050

HAL Id: hal-00521050

<https://hal.science/hal-00521050>

Preprint submitted on 24 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A HAAR COMPONENT FOR QUANTUM LIMITS ON LOCALLY SYMMETRIC SPACES

NALINI ANANTHARAMAN AND LIOR SILBERMAN

ABSTRACT. We prove lower bounds for the entropy of limit measures associated to non-degenerate sequences of eigenfunctions on locally symmetric spaces of non-positive curvature. In the case of certain compact quotients of the space of positive definite $n \times n$ matrices (any quotient for $n = 3$, quotients associated to inner forms in general), measure classification results then show that the limit measures must have a Lebesgue component. This is consistent with the conjecture that the limit measures are absolutely continuous.

CONTENTS

1. Introduction	1
2. Background and notation regarding semisimple Lie groups	10
3. Quantization and pseudodifferential operators	14
4. The WKB Ansatz	19
5. The Cotlar–Stein argument.	24
6. Measure Rigidity	35
References	38

1. INTRODUCTION

1.1. Background and motivations. The study of high-energy Laplacian eigenfunctions on negatively curved manifolds has progressed considerably in recent years. In the so-called “arithmetic” case, Elon Lindenstrauss has proved the Quantum Unique Ergodicity conjecture for Hecke eigenfunctions on congruence quotients of the hyperbolic plane [16]. In the “general case” (variable negative curvature, with no arithmetic structure), the first author has proved that semiclassical limits of eigenfunctions have positive Kolmogorov–Sinai entropy, in a joint work with Stéphane Nonnenmacher [1, 3, 4].

The two approaches are very different, but have in common the central role of the notion of entropy. In Lindenstrauss’ work, an entropy bound is obtained from arithmetic considerations [5], and then combined with the measure rigidity phenomenon to prove Quantum Unique Ergodicity.

It is very natural to ask about a possible generalization of these results to locally symmetric spaces of higher rank and nonpositive curvature. In this case the Laplacian will be

N. Anantharaman wishes to acknowledge the support of Agence Nationale de la Recherche, under the grants ANR-09-JCJC-0099-01 and ANR-07-BLAN-0361.

replaced by the entire algebra of translation-invariant differential operators, as proposed by Silberman and Venkatesh in [23]. A generalization of the entropic bound of [5] has been worked out by these authors in the adelic case, and as a result they could prove a form of Arithmetic Quantum Unique Ergodicity in the case of the locally symmetric space $\Gamma \backslash SL_n(\mathbb{R})$, when n is prime and Γ is derived from a division algebra over \mathbb{Q} [24]. The goal of this paper is to generalize the “non-arithmetic” approach of [3, 4] in this context – that is to say, prove an entropy bound without using the Hecke operators or other arithmetic techniques. Doing so, we will not require some of the assumptions used in [24]: we will work with an arbitrary connected semisimple Lie group with finite center G , Γ will be any *cocompact* lattice in G , and we will not use the Hecke operators. Combining the entropy bound with the measure classification results of [8, 9, 17], in the case of $G = SL_3(\mathbb{R})$, Γ arbitrary, or $G = SL_n(\mathbb{R})$, n arbitrary but Γ derived from a division algebra over \mathbb{Q} , we will prove a weakened form of Quantum Unique Ergodicity : any semiclassical measure has the Haar measure as an ergodic component¹.

In addition to the intrinsic interest of locally symmetric spaces, there is yet another motivation to study these models. So far, the entropic bound of [3, 4] is not satisfactory for manifolds of variable negative curvature ([1] proves that the entropy is positive, but without giving an explicit bound). Gabriel Rivière has been able to treat the case of surfaces [19, 20]; he is even able to work in nonpositive curvature, but the case of higher dimensions remains open. The problem comes from the existence of several distinct Lyapunov exponents at each point. Locally symmetric spaces are an attempt to make some progress in this direction : we will deal with flows that have distinct Lyapunov exponents, some of which may even vanish. Still, considerable simplifications arise from the fact that they are homogeneous spaces, and that the stable and unstable foliations are smooth. It would be extremely interesting to extend the techniques of [3, 4, 19, 20] to systems that are not uniformly hyperbolic (euclidean billiards would be the ultimate goal).

Let G be a connected semisimple Lie group with finite center, $K < G$ be a maximal compact subgroup, $\Gamma < G$ a uniform lattice. We will work on the symmetric space $\mathbf{S} = G/K$, the compact quotient² $\mathbf{Y} = \Gamma \backslash G/K$, and the homogeneous space $\mathbf{X} = \Gamma \backslash G$. We will endow G with its Killing metric, yielding a G -invariant Riemannian metric on G/K , with nonpositive curvature.

Call \mathcal{D} the algebra of G -invariant differential operators on \mathbf{S} ; it follows from the structure of semisimple Lie algebras that this algebra is commutative and finitely generated [11, Ch. II §4.1, §5.2]. The number of generators, to be denoted r , coincides with the real rank of \mathbf{S} (that is the dimension of a maximal flat totally geodesic submanifold), and, in a more algebraic fashion, with the dimension of \mathfrak{a} , a maximal abelian semisimple subalgebra³ of

¹Unfortunately, we are not able to extend the method to the case of $\Gamma = SL_n(\mathbb{Z})$, which is not cocompact – unless we input the extra assumption that there is no escape of mass to infinity, or that the mass escapes very fast.

²We do not assume that Γ is torsion free. When speaking of smooth functions on \mathbf{Y} , we have in mind smooth functions on \mathbf{S} that are Γ -invariant.

³We shall denote \mathfrak{g} the Lie algebra of G , \mathfrak{k} the Lie algebra of K , and so on.

\mathfrak{g} orthogonal to \mathfrak{k} . More background and notations concerning Lie groups are given in Section 2.

Remark 1.1. The algebra \mathcal{D} always contains the Laplacian. If the symmetric space \mathbf{S} has rank $r = 1$, then \mathcal{D} is generated by the Laplacian.

Example 1.2. The case $G = SO_o(d, 1)$ yields the d -dimensional hyperbolic space $S = \mathbb{H}^d$ (of rank 1), already dealt with in [3, 4].

We will focus on the example of $G = SL_n(\mathbb{R})$, $K = SO(n, \mathbb{R})$. In that case, \mathfrak{g} is the set of matrices with trace 0, \mathfrak{k} the antisymmetric matrices, and one can take \mathfrak{a} to be the set of diagonal matrices with trace 0. The connected group generated by \mathfrak{a} is denoted A , in this example it is the set of diagonal matrices of determinant 1 and with positive entries. The rank is $r = n - 1$.

We will be interested in Γ -invariant joint eigenfunctions of \mathcal{D} ; in other words, eigenfunctions of \mathcal{D} that go to the quotient $\Gamma \backslash G/K$. If we choose a set of generators of \mathcal{D} , the collection of eigenvalues can be represented as an element of \mathbb{R}^r . We will recall in Section 2.2 that it is more natural to parametrize the eigenvalue by an element $\nu \in \mathfrak{a}_{\mathbb{C}}^*$, the complexified dual of \mathfrak{a} . More precisely, $\nu \in \mathfrak{a}_{\mathbb{C}}^*/W$ where W is the Weyl group of G , a finite group given by M'/M where M' is the normalizer of A , and M the centralizer of A , in K .

1.2. Semiclassical limit. Silberman and Venkatesh suggested to study the L^2 -normalized eigenfunctions (ψ) in the limit $\|\nu\| \rightarrow +\infty$, as a variant of the very popular question of understanding high-energy eigenfunctions of the Laplacian. The question of “quantum ergodicity” is to understand the asymptotic behaviour of the family of probability measures $d\bar{\mu}_{\psi}(y) = |\psi(y)|^2 dy$ on $\mathbf{Y} = \Gamma \backslash G/K$. They considered the case where $\frac{\nu}{\|\nu\|}$ has a limit $\nu_{\infty} \in \mathfrak{a}_{\mathbb{C}}^*/W$, with the sequence ν satisfying a certain number of additional assumptions that we shall recall later. For the moment, we just note that the real parts $\Re(\nu)$ are uniformly bounded, so that $\Re(\nu_{\infty}) = 0$ ([23, Thm. 2.7 (3)]). We will denote $\Lambda_{\infty} = \Im m(\nu_{\infty}) = -i\nu_{\infty}$.

1.3. Symplectic lift vs. representation-theoretic lift. The locally symmetric space \mathbf{Y} should be thought of as the *configuration space* of our dynamical system. To properly analyze the dynamics it is necessary to move to an appropriate *phase space*. Once we *lift* the eigenfunctions there, the measures become approximately invariant under the dynamics and we can apply the tools of ergodic theory. Two different kinds of lifts have been considered thus far: the *microlocal lift* (we also call it the *symplectic lift*) lifts the measure $\bar{\mu}_{\psi}$ to a distribution $\tilde{\mu}_{\psi}$ on the cotangent bundle $T^*\mathbf{Y} = \Gamma \backslash T^*(G/K)$, taking advantage of its symplectic structure. This construction applies in great generality, for example when \mathbf{Y} is any compact Riemannian manifold. The *representation theoretic lift* used in [26, 16, 23, 24, 6], specific to locally symmetric spaces, lifts the measure $\bar{\mu}_{\psi}$ to a measure μ_{ψ} defined on $\mathbf{X} = \Gamma \backslash G$, taking advantage of the homogeneous space structure of G/K .

The two lifts are very natural, and closely related. In our proofs we will use a lot the symplectic point of view, as we will use the Helgason-Fourier transform of L^2 functions, and interpret it geometrically as a decomposition into lagrangian states. But we will also

need to translate our results in terms of the representation theoretic lift, in order to apply some measure classification results from [8, 9].

In the symplectic point of view, the dynamics is defined as follows. On $T^*(G/K)$, consider the algebra \mathcal{H} of smooth G -invariant Hamiltonians, that are polynomial in the fibers of the projection $T^*(G/K) \rightarrow G/K$. This algebra is isomorphic to the algebra of W -invariant polynomials on \mathfrak{a}^* (consider the restriction on $\mathfrak{a}^* \subset T_o^*(G/K)$). The structure theory of semisimple Lie algebras shows that \mathcal{H} is isomorphic to a polynomial ring in r generators. Moreover, the elements of \mathcal{H} commute under the Poisson bracket. Thus, we have on $T^*(G/K)$ a family of r independent commuting Hamiltonian flows H_1, \dots, H_r . The Killing metric, seen as a function on $T^*(G/K)$, always belongs to \mathcal{H} , and its symplectic gradient generates the geodesic flow. Of course, since all these flows are G -equivariant, they descend to the quotient $T^*\mathbf{Y}$.

Joint energy layers of \mathcal{H} are naturally parametrized by elements $\Lambda \in \mathfrak{a}^*/W$. This is easy to explain geometrically: fix a point in G/K , say the origin $o = eK$. Consider the flat totally geodesic submanifold $A.o \subset G/K$ going through o . It is isometric to \mathbb{R}^r , and the cotangent space $T_o^*(A.o)$ is naturally isomorphic to \mathfrak{a}^* . If $\mathcal{E} \subset T^*(G/K)$ is a joint energy layer of \mathcal{H} (or equivalently a G -orbit in $T^*(G/K)$), then there exists $\Lambda \in \mathfrak{a}^*$ such that $\mathcal{E} \cap T_o^*(A.o) = W.\Lambda$. See [13] for details. We will denote \mathcal{E}_Λ the energy layer of parameter Λ .

In Section 3 we will use a quantization procedure to associate to every Γ -invariant eigenfunction ψ a distribution $\tilde{\mu}_\psi$ on $T^*\mathbf{Y}$, called its microlocal lift. This distribution projects to $\bar{\mu}_\psi$ on \mathbf{Y} . This is a very standard construction, and so is the following theorem, which is an avatar of propagation of singularities for solutions of partial differential equations:

Theorem 1.3. *Assume that $\|\nu\| \rightarrow +\infty$, and that $\frac{\nu}{\|\nu\|}$ has a limit ν_∞ . Denote $\Lambda_\infty = -i\nu_\infty \in \mathfrak{a}^*/W$. Any limit (in the distribution sense) of the sequence $\tilde{\mu}_\psi$ is a probability measure on $T^*\mathbf{Y}$, carried by the energy layer $\mathcal{E}_{\Lambda_\infty}$, and invariant under the family of Hamiltonian flows generated by \mathcal{H} .*

In order to transport this statement to get an A -invariant measure on $\Gamma \backslash G$, we must now make some assumptions on Λ_∞ . Silberman and Venkatesh assume that ν_∞ is a regular element of $\mathfrak{a}_\mathbb{C}^*$, in the sense that it is not fixed by any non-trivial element of W , and they show that it implies $\Re(\nu_n) = 0$ for all but a finite number of ν_n s in the sequence. The element ν_∞ being regular is, of course, equivalent to Λ_∞ being regular; and this is also equivalent to the energy layer Λ_∞ being regular, in the sense that the differentials dH_1, \dots, dH_r are independent there [13].

There is a surjective map

$$(1.1) \quad \pi : G/M \times \mathfrak{a}^* \longrightarrow T^*(G/K)$$

$$(1.2) \quad (gM, \lambda) \longmapsto (gK, g.\lambda).$$

Remember that M is the centralizer of A in K . The image of $G/M \times \{\lambda\}$ under π is the energy layer \mathcal{E}_λ . The map $\pi_\lambda : G/M \times \{\lambda\} \rightarrow \mathcal{E}_\lambda$ is a diffeomorphism if and only if λ is regular (otherwise π_λ is not injective). Under π_λ^{-1} , the action of the Hamiltonian flow Φ_H^t

generated by $H \in \mathcal{H}$ on \mathcal{E}_λ is conjugate to

$$gM \mapsto g \exp(t dH(\lambda))M.$$

The same statements hold after quotienting on the left by Γ . Since H is a function on \mathfrak{a}^* , the differential $dH(\lambda)$ is an element of \mathfrak{a} . Denoting $R(e^{tX})$ the one-parameter flow on G/M generated by $X \in \mathfrak{a}$ (acting by multiplication on the right), we can rephrase this by writing

$$\pi \circ R(e^{tdH(\lambda)}) = \Phi_H^t \circ \pi \quad \text{on } \mathcal{E}_\lambda.$$

If λ is regular, the elements $dH(\lambda)$ can be shown to span \mathfrak{a} as H varies over \mathcal{H} . Otherwise, we have [13]

$$(1.3) \quad \{dH(\lambda), H \in \mathcal{H}\} = \{X \in \mathfrak{a}, \forall \alpha \in \Delta, (\langle \alpha, \lambda \rangle = 0 \implies \alpha(X) = 0)\},$$

where $\Delta \subset \mathfrak{a}^*$ is the set of roots.

Thus, Theorem 1.3 may be rephrased as follows:

Theorem 1.4. *Assume Λ_∞ is regular. Then any limit (in the distribution sense) of the sequence $\tilde{\mu}_\psi$ yields a probability measure on $\Gamma \backslash G/M$, invariant under the right action of A by multiplication.*

This theorem was proved in [23, Thm. 1.6 (3)] using the representation-theoretic lift; the equivariance of that lift shows that the construction is compatible with the Hecke operators on $\Gamma \backslash G$. It is also shown there that the symplectic lift $\tilde{\mu}_\psi$ and the representation theoretic lift μ_ψ have the same asymptotic behaviour as ν tends to infinity, and if we identify $\mathcal{E}_{\Lambda_\infty} \subset \Gamma \backslash T^*(G/K)$ with $\Gamma \backslash G/M$.

Definition 1.5. We will call any limit point of the sequence $\tilde{\mu}_\psi$ (or μ_ψ) a *semiclassical measure* in the direction Λ_∞ .

Semiclassical measures in a regular direction are, equivalently, positive measures on $T^*(\Gamma \backslash G/K)$ (carried by a regular energy layer), positive measures on $\Gamma \backslash G/M$, or positive measures on $\Gamma \backslash G$ (which are M -invariant).

1.4. Entropy bounds. Our main result is a non-trivial lower bound on the entropy of semiclassical measures. We fix $H \in \mathcal{H}$, and we consider the corresponding Hamiltonian flow Φ_H^t on $\mathcal{E}_{\Lambda_\infty}$, which has Lyapunov exponents

$$-\chi_J(H) \leq \dots \leq -\chi_1(H) \leq 0 \leq \chi_1(H) \leq \dots \leq \chi_J(H).$$

In addition, the Lyapunov exponent 0 appears trivially with multiplicity r , as a consequence of the existence of r integrals of motion. The dimension of $\mathcal{E}_{\Lambda_\infty}$ is $r + 2J$. The integer J , the rank r and the dimension d of G/K are related by $d = J + r$. In general, the Lyapunov exponents are measurable functions on the phase space, but here, because of the homogeneous structure, the Lyapunov exponents are constants.

In the following theorem we will denote $\chi_{\max}(H) = \chi_J(H)$, the largest Lyapunov exponent. We denote $h_{KS}(\mu, H)$ the Kolmogorov-Sinai entropy of a (Φ_H^t) -invariant probability

measure μ . We recall the Ruelle-Pesin inequality,

$$h_{KS}(\mu, H) \leq \sum_j \chi_j(H),$$

which holds for any (Φ_H^t) -invariant probability measure μ .

Theorem 1.6. (*Symplectic version*) *Let μ be a semiclassical measure in the direction Λ_∞ . Assume that Λ_∞ is regular.*

For $H \in \mathcal{H}$, we consider the corresponding Hamiltonian flow Φ_H^t on $\mathcal{E}_{\Lambda_\infty}$. Then

$$(1.4) \quad h_{KS}(\mu, H) \geq \sum_{j: \chi_j(H) \geq \frac{\chi_{\max}(H)}{2}} \left(\chi_j(H) - \frac{\chi_{\max}(H)}{2} \right).$$

Continuing with the assumption that Λ_∞ is regular, we can transport the theorem to $\Gamma \backslash G/M$. If we fix a 1-parameter subgroup (e^{tX}) of A (with $X \in \mathfrak{a}$), it is well known that the (non trivial) Lyapunov exponents of the flow (e^{tX}) acting on X/M are the real numbers $(\alpha(X))$, where $\alpha \in \mathfrak{a}^*$ run over the set of roots Δ (see Section 2 for background related to Lie groups). If α is a root then so is $-\alpha$ (one of the two will be called *positive*, the other *negative*). The notion of positivity is explained in detail later. For now it suffices to note that we may assume that $\alpha(X) \geq 0$ for positive roots α . We write $\alpha_{\max}(X)$ for $\max_\alpha \alpha(X)$ (this is the largest Lyapunov exponent of the associated Hamiltonian flow). Each root occurs with *multiplicity* m_α , which must be taken into account in the statements below (the corresponding Lyapunov exponent $\alpha(X)$ would be counted repeatedly, m_α times).

Theorem 1.7. (*Group-theoretic version*) *Let μ be a semiclassical measure in the direction Λ_∞ . Assume that Λ_∞ is regular.*

Let (e^{tX}) ($X \in \mathfrak{a}$) be a one parameter subgroup of A such that $\alpha(X) \geq 0$ for all positive roots α .

Let $h_{KS}(\mu, X)$ be the entropy of μ with respect to the flow (e^{tX}) . Then

$$(1.5) \quad h_{KS}(\mu, X) \geq \sum_{\alpha: \alpha(X) \geq \frac{\alpha_{\max}(X)}{2}} m_\alpha \left(\alpha(X) - \frac{\alpha_{\max}(X)}{2} \right).$$

Our lower bound is positive for all non-zero X , in fact greater than $\frac{\alpha_{\max}(X)}{2}$. In [1, 3], the first author and S. Nonnenmacher had conjectured the following stronger bound

$$h_{KS}(\mu, H) \geq \frac{1}{2} \sum_j \chi_j(H)$$

or equivalently

$$(1.6) \quad h_{KS}(\mu, X) \geq \frac{1}{2} \sum_{\alpha > 0} m_\alpha \cdot \alpha(X).$$

We are still unable to prove it, except in one case: when all the positive Lyapunov exponents are equal to each other, so that formula (1.5) reduces to (1.6). One case is that of hyperbolic d -space ($G = SO(d, 1)$) alluded to above. Another, the main focus of the present paper, is

of the “extremely irregular” elements of the torus in $G = SL_n(\mathbb{R})$. These are the elements conjugate under the Weyl group to

$$X = \text{diag}(n - 1, -1, \dots, -1).$$

1.5. Application: towards Quantum Unique Ergodicity on locally symmetric spaces. In Section 6 we combine our entropy bounds with measure classification results. Let $n \geq 3$, $G = SL_n(\mathbb{R})$, $\Gamma < G$ a cocompact lattice. Let μ be a semiclassical measure on $\Gamma \backslash G$ in the regular direction Λ_∞ .

The measure μ can be written uniquely as a sum of an absolutely continuous measure and a singular measure (with respect to Lebesgue or Haar measure). Since μ is invariant under the action of A , the same holds for both components. Because the Haar measure is known to be ergodic for the action of A , the absolutely continuous part of μ is, in fact, proportional to Haar measure. We call this the *Haar component* of μ . Its total mass is the *weight* of this component.

Theorem 1.8. *Let $n = 3$. Then μ has a Haar component of weight $\geq \frac{1}{4}$.*

Theorem 1.9. *Let $n = 4$. Then either μ has a Haar component, or each ergodic component*

*is the Haar measure on a closed orbit of the group $\begin{pmatrix} * & * & * & 0 \\ * & * & * & 0 \\ * & * & * & 0 \\ 0 & 0 & 0 & * \end{pmatrix}$ (or one of its 4 images under the Weyl group), and the components invariant by each of these 4 subgroups have total weight $\frac{1}{4}$.*

In fact, the result is slightly stronger: if some “extremely irregular” element acts on μ with entropy strictly larger than half of its entropy w.r.t. Haar measure, then there is a Haar component.

It does not seem to be possible to push this technique beyond SL_4 . The problem is that there are large subgroups (in the style of those occurring in Theorem 1.9) whose closed orbits support measures of large entropy. For particular lattices, however, these large subgroups do not have closed orbits, so the only possible non-Haar components have small entropy and cannot account for all the entropy. For co-compact lattices this occurs, for example, when Γ is the set of elements of reduced norm 1 of an order in a central division algebra over \mathbb{Q} , or more generally for any lattice commensurable with one obtained this way (we say that Γ is *associated* to the division algebra). Such lattices are said to be of “inner type” since they correspond to inner forms of SL_n over \mathbb{Q} (there also exist non-uniform lattices of inner type, corresponding to central simple \mathbb{Q} -algebras which are not division algebras). For a brief description of the construction and references see Section 6.

Theorem 1.10. *For $n \geq 3$ let $\Gamma < SL_n(\mathbb{R})$ be a lattice associated to a division algebra over \mathbb{Q} , and let μ be a semiclassical measure on $\Gamma \backslash SL_n(\mathbb{R})$ in a regular direction. Then μ has a Haar component of weight $\geq \frac{\frac{n+1}{2}-t}{n-t} > 0$ where t is the largest proper divisor of n .*

It is not surprising that strongest implication is for n prime (so that there are few intermediate algebraic measures). Indeed, setting $t = 1$ we find $w_\Delta \geq \frac{1}{2}$ in that case. However for n prime Silberman-Venkatesh [24] show that the semiclassical measures associated to *Hecke eigenfunctions* are equal to Haar measure. The main impact of Theorem 1.10 is thus when the n is composite, where previous methods only showed that semiclassical measures are convex combinations of algebraic measures but could not establish that Haar measure occurs in the combination.

Remark 1.11. We compare here our result with that of [24]. That paper studies the case of lattices in $G = PGL_n(\mathbb{R})$ associated to division algebras of prime degree n and joint eigenfunctions of \mathcal{D} and of the Hecke operators. It is then shown that *any ergodic component of a semiclassical measure* μ has positive entropy; it follows that μ must be the Haar measure. Our result is neither stronger nor weaker:

- We cannot prove that all ergodic components of μ have positive entropy, only that the total entropy of μ is positive. Hence, we are not able to exclude components of zero entropy;
- On the other hand, our lower bound on the total entropy (1/2 of the maximal entropy) is explicit and quite strong. This allows to detect the presence of a Haar component in a variety of cases;
- In particular, for $n = 3$ we do not need any assumption on the cocompact lattice Γ ; and for Γ associated to a division algebra, our result holds for all n .
- The Hecke-operator method applies more naturally to adelic quotients $\mathbb{G}(\mathbb{Q}) \backslash \mathbb{G}(\mathbb{A}) / K_\infty K_f$. When G is a form of SL_n there is no distinction, but when $G = PGL_n$ the adelic quotients are typically *disjoint unions* of quotients $\Gamma \backslash G$. Even when the quotient is compact, G -invariance of the limit measure does not show that all components have the same proportion of the mass. Our result applies to each connected component separately.
- We do not assume that our eigenfunctions are also eigenfunctions of the Hecke operators: this means that multiplicity of eigenvalues is not an issue in this work.
- The methods of Silberman-Venkatesh apply to non-cocompact lattices as well.

1.6. Hyperbolic dispersive estimate. The proof of Theorem 1.6 (and 1.7) follows the main ideas of [3], with a major difference which lies in an improvement of the “hyperbolic dispersive estimate” : [1, Thm. 1.3.3] and [3, Thm. 2.7]. If we applied directly the result of [3], we would get

$$h_{KS}(\mu, H) \geq \sum_k \left(\chi_k(H) - \frac{\chi_{\max}(H)}{2} \right).$$

This inequality is often trivial (the right-hand term being negative) whereas in (1.4) we managed to get rid of the negative terms $\left(\chi_k(H) - \frac{\chi_{\max}(H)}{2} \right)$.

Since the “hyperbolic dispersive estimate” has an intrinsic interest, and is the core of this paper, we state it here as one of our main results. We fix a quantization procedure, set at scale $\hbar = \|\nu\|^{-1}$, that associates to any reasonable function a on $T^*\mathbf{Y}$ an operator $\text{Op}_\hbar(a)$ on $L^2(\mathbf{Y})$. An explicit construction is given in Section 3. In particular, it is useful to know

that Op_{\hbar} can be defined so that, if $H \in \mathcal{H}$ is real valued, $\text{Op}_{\hbar}(H)$ is a self-adjoint operator belonging to \mathcal{D} . More explicitly, $\text{Op}_{\hbar}(H)$ is defined so that $\text{Op}_{\hbar}(H)\psi_{\nu} = H(-i\hbar\nu)\psi_{\nu}$ for any \mathcal{D} -eigenfunction ψ_{ν} , with spectral parameter ν (hence the choice of the normalisation $\hbar = \|\nu\|^{-1}$).

Let $(P_k)_{k=1,\dots,K}$ be a family of smooth real functions on \mathbf{Y} , such that

$$(1.7) \quad \forall x \in \mathbf{Y}, \quad \sum_{k=1}^K P_k^2(x) = 1 .$$

We assume that the diameter of the supports of the functions P_k is small enough. We will also denote P_k the operator of multiplication by $P_k(x)$ on the Hilbert space $L^2(\mathbf{Y})$.

We denote $U^t = \exp(i\hbar^{-1}t \text{Op}_{\hbar}(H))$ the propagator of the ‘‘Schrödinger equation’’ generated by the Hamiltonian H . This is a unitary Fourier Integral Operator associated with the classical Hamiltonian flow Φ_H^{-t} . The \hbar -dependence of U will be implicit in our notations. We fix a small discrete time step η .

Throughout the paper we will use the notation $\widehat{A}(t) = U^{-t\eta} \widehat{A} U^{t\eta}$ for the quantum evolution at time $t\eta$ of an operator \widehat{A} . For each integer $T \in \mathbb{N}$ and any sequence of labels $\underline{\omega} = (\omega_{-T}, \dots, \omega_{-1}, \omega_0, \dots, \omega_{T-1})$, $\omega_i \in [1, K]$ (we say that the sequence $\underline{\omega}$ is of length $|\underline{\omega}| = 2T$), we define the operators

$$(1.8) \quad P_{\underline{\omega}} = P_{\omega_{T-1}}(T-1) P_{\omega_{T-2}}(T-2) \dots P_{\omega_0} P_{\omega_{-1}}(-1) \dots P_{\omega_{-T}}(-T) .$$

We fix a smooth, compactly supported function χ on $T^*\mathbf{Y}$, supported in a tubular neighbourhood of size ϵ of the energy layer $\mathcal{E}_{\Lambda_{\infty}}$ (which is assumed to be regular); and we define

$$(1.9) \quad P_{\underline{\omega}}^{\chi} = P_{\omega_{T-1}}(T-1) P_{\omega_{T-2}}(T-2) \dots P_{\omega_0}^{1/2} \text{Op}(\chi) P_{\omega_0}^{1/2} P_{\omega_{-1}}(-1) \dots P_{\omega_{-T}}(-T) .$$

The operator $P_{\underline{\omega}}^{\chi}$ should be thought of as $P_{\underline{\omega}}$ restricted to a spectral window around the energy layer $\mathcal{E}_{\Lambda_{\infty}}$.

Theorem 1.12. *Fix $H \in \mathcal{H}$, and a time step η , small enough. Let $\mathcal{K} > 0$ be fixed, arbitrary. Let $\chi \in \mathcal{C}^{\infty}(T^*\mathbf{Y})$, supported in a tubular neighbourhood of size ϵ of the regular energy layer $\mathcal{E}_{\Lambda_{\infty}}$. Assume that ϵ , as well as the diameters of the supports of each P_k , are small enough.*

Then, there exists $\hbar_{\mathcal{K}} > 0$ such that, for all $\hbar \in (0, \hbar_{\mathcal{K}})$, for $T = \lfloor \frac{\mathcal{K} |\log \hbar|}{\eta} \rfloor$, and for every sequence $\underline{\omega}$ of length T ,

$$(1.10) \quad \|P_{\underline{\omega}}^{\chi}\| \leq C \hbar^{-c\epsilon} \prod_{k, \chi_k(H) \geq \frac{1}{2\mathcal{K}}} \frac{e^{-T\eta \chi_k(H)}}{\hbar^{1/2}}$$

where the $\chi_k(H)$ denote the Lyapunov exponents of Φ_H^t on the energy layer $\mathcal{E}_{\Lambda_{\infty}}$. The constant C does not depend on \mathcal{K} nor on H , whereas c does.

The method used in [3] only yielded the upper bound:

$$(1.11) \quad \|P_{\underline{\omega}}^{\chi}\| \leq C \hbar^{-c\epsilon} \prod_k \frac{e^{-T\eta \chi_k(H)}}{\hbar^{1/2}}$$

This is clearly not optimal when Φ_H^t has some neutral, or slowly expanding directions. For instance, if $H = 0$ then $\Phi_H^t = I$ has only neutral directions. In this case, (1.11) reads

$$(1.12) \quad \|P_{\underline{\omega}}^{\chi}\| \leq C \hbar^{-\frac{d}{2}-c\epsilon},$$

where d is the dimension of \mathbf{Y} , which is obviously much worse (for any T) than the trivial bound

$$(1.13) \quad \|P_{\underline{\omega}}^{\chi}\| \leq 1.$$

On the other hand, if some of the $\chi_k(H)$ are (strictly) positive, then (1.11) is much better than the trivial bound (1.13), for very large $T\eta$. The bound given by Theorem 1.12 interpolates between the two, for $T\eta \sim \mathcal{K} |\log \hbar|$.

The proof of the hyperbolic dispersion estimates is quite technical, and occupies Sections 3, 4, 5. It uses a version of the pseudodifferential calculus adapted to the geometry of locally symmetric spaces, based on Helgason's version of the Fourier transform for this spaces, and inspired by the work of Zelditch in the case of $G = SL(2, \mathbb{R})$ [27]. We point out the fact that an alternative proof of Theorem 1.12 is given in [2], based on more conventional Fourier analysis. The reader might prefer to read [2] instead of Sections 3, 4, 5, however we feel that the two techniques have an interest of their own.

We will not repeat here the argument that leads from Theorem 1.12 to the entropy bound Theorem 1.6; it would be an exact repetition of the argument given in [3, §2]. Let us just make one comment : in this argument, we are limited to $\mathcal{K} = \frac{1}{\chi_{\max}(H)}$ (the time $T_E = \frac{|\log \hbar|}{\chi_{\max}(H)}$ is sometimes called the Ehrenfest time for the Hamiltonian H , and corresponds to the time where the approximation of the quantum flow U^t by the classical flow Φ_H^t breaks down). This means that we eventually keep the Lyapunov exponents such that $\chi_k(H) \geq \frac{\chi_{\max}(H)}{2}$, and explains why this restriction appears in (1.4).

2. BACKGROUND AND NOTATION REGARDING SEMISIMPLE LIE GROUPS

Our terminology follows Knapp [15].

2.1. Structure. Let G denote a non-compact connected simple Lie group with finite center⁴. We choose a Cartan involution Θ for G , and let $K < G$ be the Θ -fixed maximal

⁴If G is semisimple our discussion remains valid, but one can even do something finer, as remarked in [23, §5.1]. After decomposing \mathfrak{g} into simple factors $\oplus \mathfrak{g}^{(j)}$, and assuming that the Cartan involution, the subalgebra \mathfrak{a} , etc. are compatible with this decomposition, one can decompose the spectral parameter ν into its components $\nu^{(j)} \in \mathfrak{a}^{(j)*}$. Instead of assuming that $\|\nu\| \rightarrow +\infty$ and $\frac{\nu}{\|\nu\|}$ has a regular limit ν_{∞} , one can assume the same independently for each component $\nu^{(j)}$. This means that we do not have to assume that all the norms $\|\nu^{(j)}\|$ go to infinity at the same speed.

compact subgroup. Let $\mathfrak{g} = \text{Lie}(G)$, and let θ denote the differential of Θ , giving the Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ with $\mathfrak{k} = \text{Lie}(K)$. Let $\mathbf{S} = G/K$ be the symmetric space, with $o = eK \in \mathbf{S}$ the point with stabilizer K . We fix a G -invariant metric on G/K : observe that the tangent space at the point o is naturally identified with \mathfrak{p} , and endow it with the Killing form. For a lattice $\Gamma < G$ we write $\mathbf{X} = \Gamma \backslash G$ and $\mathbf{Y} = \Gamma \backslash G/K$, the latter being a locally symmetric space of non-positive curvature. In this paper, we shall always assume that \mathbf{X} and \mathbf{Y} are compact.

Fix now a maximal abelian subalgebra $\mathfrak{a} \subset \mathfrak{p}$.

We denote by $\mathfrak{a}_{\mathbb{C}}$ the complexification $\mathfrak{a} \otimes \mathbb{C}$. We denote by \mathfrak{a}^* (resp. $\mathfrak{a}_{\mathbb{C}}^*$) the real dual (resp. the complex dual) of \mathfrak{a} . For $\nu \in \mathfrak{a}_{\mathbb{C}}^*$, we define $\Re(\nu), \Im(\nu) \in \mathfrak{a}^*$ to be the real and imaginary parts of ν , respectively. For $\alpha \in \mathfrak{a}^*$, set $\mathfrak{g}_{\alpha} = \{X \in \mathfrak{g}, \forall H \in \mathfrak{a} : \text{ad}(H)X = \alpha(H)X\}$, $\Delta = \Delta(\mathfrak{a} : \mathfrak{g}) = \{\alpha \in \mathfrak{a}^* \setminus \{0\}, \mathfrak{g}_{\alpha} \neq \{0\}\}$ and call the latter the (restricted) roots of \mathfrak{g} with respect to \mathfrak{a} . The subalgebra \mathfrak{g}_0 is θ -invariant, and hence $\mathfrak{g}_0 = (\mathfrak{g}_0 \cap \mathfrak{p}) \oplus (\mathfrak{g}_0 \cap \mathfrak{k})$. By the maximality of \mathfrak{a} in \mathfrak{p} , we must then have $\mathfrak{g}_0 = \mathfrak{a} \oplus \mathfrak{m}$ where $\mathfrak{m} = Z_{\mathfrak{k}}(\mathfrak{a})$, the centralizer of \mathfrak{a} in \mathfrak{k} .

The Killing form of \mathfrak{g} induces a standard inner product $\langle \cdot, \cdot \rangle$ on \mathfrak{p} , and by duality on \mathfrak{p}^* . By restriction we get an inner product on \mathfrak{a}^* with respect to which $\Delta(\mathfrak{a} : \mathfrak{g}) \subset \mathfrak{a}^*$ is a root system. The associated Weyl group, generated by the root reflections s_{α} , will be denoted $W = W(\mathfrak{a} : \mathfrak{g})$. This group is also canonically isomorphic to $N_K(\mathfrak{a})/Z_K(\mathfrak{a})$. In what follows we will represent any element w of the Weyl group by a representative in $N_K(\mathfrak{a}) \subset K$ (taking care to only make statements that do not depend on the choice of a representative), and the action of $w \in W(\mathfrak{a} : \mathfrak{g})$ on \mathfrak{a} or \mathfrak{a}^* will be given by the adjoint representation $\text{Ad}(w)$. The fixed-point set of any s_{α} is a hyperplane in \mathfrak{a}^* , called a wall. The connected components of the complement of the union of the walls are cones, called the (open) Weyl chambers. A subset $\Pi \subset \Delta(\mathfrak{a} : \mathfrak{g})$ will be called a system of simple roots if every root can be uniquely expressed as an integral combination of elements of Π with either all coefficients non-negative or all coefficients non-positive. For a simple system Π , the open cone $C_{\Pi} = \{\nu \in \mathfrak{a}^*, \forall \alpha \in \Pi : \langle \nu, \alpha \rangle > 0\}$ is an (open) Weyl chamber. The closure of an open chamber will be called a closed chamber; we will denote in particular $\overline{C_{\Pi}} = \{\nu \in \mathfrak{a}^*, \forall \alpha \in \Pi : \langle \nu, \alpha \rangle \geq 0\}$. The Weyl group acts simply transitively on the chambers and simple systems. The action of $W(\mathfrak{a} : \mathfrak{g})$ on \mathfrak{a}^* extends in the complex-linear way to an action on $\mathfrak{a}_{\mathbb{C}}^*$ preserving $i\mathfrak{a}^* \subset \mathfrak{a}_{\mathbb{C}}^*$, and we call an element $\nu \in \mathfrak{a}_{\mathbb{C}}^*$ regular if it is fixed by no non-trivial element of $W(\mathfrak{a} : \mathfrak{g})$. Since $-C_{\Pi} \subset \mathfrak{a}^*$ is a chamber, there is a unique $w_{\ell} \in W(\mathfrak{a} : \mathfrak{g})$, called the “long element”, such that $\text{Ad}(w_{\ell}).C_{\Pi} = -C_{\Pi}$. Note that $w_{\ell}^2.C_{\Pi} = C_{\Pi}$ and hence $w_{\ell}^2 = e$. Also, w_{ℓ} depends on the choice of Π but we suppress this from the notation.

Fixing a simple system Π we get a notion of positivity. We will denote by Δ^+ the set of positive roots, by $\Delta^- = -\Delta^+$ the set of negative roots. We use $\rho = \frac{1}{2} \sum_{\alpha > 0} (\dim \mathfrak{g}_{\alpha}) \alpha \in \mathfrak{a}^*$ to denote half the sum of the positive roots. For $\mathfrak{n} = \bigoplus_{\alpha > 0} \mathfrak{g}_{\alpha}$ and $\bar{\mathfrak{n}} = \Theta \mathfrak{n} = \bigoplus_{\alpha < 0} \mathfrak{g}_{\alpha}$ we have $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{a} \oplus \mathfrak{m} \oplus \bar{\mathfrak{n}}$. Note that $\bar{\mathfrak{n}} = \text{Ad}(w_{\ell}).\mathfrak{n}$. We also have (“Iwasawa decomposition”) $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{a} \oplus \mathfrak{k}$. We can therefore uniquely write every $X \in \mathfrak{g}$ in the form $X = X_{\mathfrak{n}} + X_{\mathfrak{a}} + X_{\mathfrak{k}}$. We also write $H_0(X)$ for $X_{\mathfrak{a}}$.

Let $N, A, \bar{N} < G$ be the connected subgroups corresponding to the subalgebras $\mathfrak{n}, \mathfrak{a}, \bar{\mathfrak{n}} \subset \mathfrak{g}$ respectively, and let $M = Z_K(\mathfrak{a})$. Then $\mathfrak{m} = \text{Lie}(M)$, though M is not necessarily connected. Moreover $P_0 = NAM$ is a minimal parabolic subgroup of G , with the map $N \times A \times M \rightarrow P_0$ being a diffeomorphism. The map $N \times A \times K \rightarrow G$ is a (surjective) diffeomorphism (Iwasawa decomposition), so for $g \in G$ there exists a unique $H_0(g) \in \mathfrak{a}$ such that $g = n \exp(H_0(g))k$ for some $n \in N, k \in K$. The map $H_0 : G \rightarrow \mathfrak{a}$ is continuous; restricted to A , it is the inverse of the exponential map.

We will use the G -equivariant identification between G/M and $G/K \times G/P_0$, given by $gM \mapsto (gK, gP_0)$. The quotient G/P_0 can also be identified with K/M .

Starting from H_0 we define a ‘‘Busemann function’’ B on $G/K \times G/P_0 \sim G/M$:

$$(2.1) \quad B(gK, g_1P_0) = H_0(k^{-1}g),$$

where k is the K -part in the KAN decomposition of g_1 (if g_1 is defined modulo P_0 , then k is defined modulo M). Equivalently, if $gM \in G/M$, we have $B(gM) = a$, where $g = kna$ is the KNA decomposition of g (if g is defined modulo M , then a is uniquely defined and k is defined modulo M).

In G/K , a ‘‘flat’’ is a maximal flat totally geodesic submanifold. Every flat is of the form $\{gaK, a \in A\}$ for some $g \in G$. The space of flats can be naturally identified with G/MA , or with an open dense subset of $G/P_0 \times G/\bar{P}_0$, via the G -equivariant map

$$gMA \mapsto (gP_0, g\bar{P}_0)$$

where $\bar{P}_0 = MAN\bar{N} = w_\ell P_0 w_\ell^{-1}$. We will also use the following injective map from G/MA into $G/P_0 \times G/P_0$,

$$gMA \mapsto (gP_0, gw_\ell P_0).$$

Its image is an open dense subset of $G/P_0 \times G/P_0$, namely $\{(g_1P_0, g_2P_0), g_2^{-1}g_1 \in P_0 w_\ell P_0\}$. Finally we recall the Bruhat decomposition $G = \sqcup_{w \in W(\mathfrak{a}; \mathfrak{g})} P_0 w P_0$, with $P_0 w_\ell P_0$ being an open dense subset (the ‘‘big cell’’).

2.2. The universal enveloping algebra; Harish-Chandra isomorphisms. We analyze the structure of \mathcal{D} by comparing it with other algebras of differential operators. For a Lie algebra \mathfrak{s} we write $\mathfrak{s}_\mathbb{C}$ for its complexification $\mathfrak{s} \otimes_{\mathbb{R}} \mathbb{C}$. In particular, $\mathfrak{g}_\mathbb{C}$ is a complex semisimple Lie algebra. We fix a maximal abelian subalgebra $\mathfrak{b} \subset \mathfrak{m}$ and let $\mathfrak{h} = \mathfrak{a} \oplus \mathfrak{b}$. Then $\mathfrak{h}_\mathbb{C}$ is a Cartan subalgebra of $\mathfrak{g}_\mathbb{C}$, with an associated root system $\Delta(\mathfrak{h}_\mathbb{C} : \mathfrak{g}_\mathbb{C})$ satisfying $\Delta(\mathfrak{a} : \mathfrak{g}) = \{\alpha|_{\mathfrak{a}}\}_{\alpha \in \Delta(\mathfrak{h}_\mathbb{C}; \mathfrak{g}_\mathbb{C})} \setminus \{0\}$.

If $\mathfrak{s}_\mathbb{C}$ is a complex Lie algebra, we denote by $U(\mathfrak{s}_\mathbb{C})$ its universal enveloping algebra; $U(\mathfrak{g}_\mathbb{C})$ is isomorphic to the algebra of left- G -invariant differential operators on G with complex coefficients [10].

There is an isomorphism, called the Harish-Chandra isomorphism, between the algebra \mathcal{D} of G -invariant differential operators on G/K and the algebra $\mathcal{D}_W(A)$ of A - and W -invariant differential operators on $A \sim \mathbb{R}^r$. The latter is obviously isomorphic to $U(\mathfrak{a}_\mathbb{C})^W$, the subalgebra of $U(\mathfrak{a}_\mathbb{C})$ formed of W -invariant elements. Since $\mathfrak{a}_\mathbb{C}$ is abelian, $U(\mathfrak{a}_\mathbb{C})$ can be identified to the space of polynomial functions on \mathfrak{a}^* with complex coefficients.

The Harish-Chandra isomorphism $\Gamma : \mathcal{D} \rightarrow \mathcal{D}_W(A)$ can be realized in a geometric way as follows [11, Cor. II.5.19]. Consider the flat subspace $A.o \subset G/K$, naturally identified with A . Fixing $D \in \mathcal{D}$, let $\Delta_N(D)$ be the translation-invariant differential operator on A (that is, an element of $U(\mathfrak{a})$) given by

$$[\Delta_N(D)f](a) = D\tilde{f}(a.o),$$

for $a \in A$, $f \in C^\infty(A.o)$, and where \tilde{f} stands with the unique N -invariant function on G/K that coincides with f on $A.o$. Then, we define

$$\Gamma : D \mapsto e^{-\rho} \circ \Delta_N(D) \circ e^\rho,$$

remembering that ρ is half the sum of positive roots and thus can be seen as a function on A . Note that

$$e^{-\rho} \circ \Delta_N(D) \circ e^\rho = \tau_\rho \cdot \Delta_N(D),$$

where τ_ρ is the automorphism of $U(\mathfrak{a})$ defined by putting $\tau_\rho(X) = X + \rho(X)$ for every $X \in \mathfrak{a}$.

In what follows, we denote by $\mathcal{Z}(\mathfrak{g}_\mathbb{C})$ the center of $U(\mathfrak{g}_\mathbb{C})$. Thus, $\mathcal{Z}(\mathfrak{g}_\mathbb{C})$ is the algebra of G -bi-invariant operators. Differentiating the action of G on \mathbf{S} gives a map $\mathcal{Z}(\mathfrak{g}_\mathbb{C}) \rightarrow \mathcal{D}$. For the next lemma we shall compare the isomorphism Γ with an isomorphism $\omega_{HC} : \mathcal{Z}(\mathfrak{g}_\mathbb{C}) \rightarrow U(\mathfrak{h}_\mathbb{C})^{W(\mathfrak{h}_\mathbb{C}:\mathfrak{g}_\mathbb{C})}$, also called the Harish-Chandra isomorphism⁵.

Lemma 2.1. *Assume that the restriction from $\mathfrak{h}_\mathbb{C}$ to \mathfrak{a} induces a surjection from $U(\mathfrak{h}_\mathbb{C})^{W(\mathfrak{h}_\mathbb{C}:\mathfrak{g}_\mathbb{C})}$ to $U(\mathfrak{a}_\mathbb{C})^W$ (thought of as functions on the respective linear spaces).*

Let $D \in \mathcal{D}$, of degree \bar{d} . Then there exists $Z \in \mathcal{Z}(\mathfrak{g}_\mathbb{C})$ such that Z and D coincide on (right-) K -invariant functions, and such that

$$Z - \tau_{-\rho}\Gamma(D) \in U(\mathfrak{n}_\mathbb{C})U(\mathfrak{a}_\mathbb{C})^{\bar{d}-2} + U(\mathfrak{g}_\mathbb{C})\mathfrak{k}_\mathbb{C}.$$

Remark 2.2. The assumption is automatically satisfied when G is split. It is also satisfied when G/K is a *classical* symmetric space, that is when G is a classical group [11, p. 341]. In fact the lemma itself is Proposition II.5.32 of [11], with the difference of degree between Z and $\tau_{-\rho}\Gamma(D)$ made precise.

Proof. Let $D \in \mathcal{D}$ be of degree \bar{d} , so that $\Gamma(D) \in U(\mathfrak{a}_\mathbb{C})^W$ is a polynomial of degree $\leq \bar{d}$. By assumption, we can extend $\Gamma(D)$ to an element of $U(\mathfrak{h}_\mathbb{C})^{W(\mathfrak{h}_\mathbb{C}:\mathfrak{g}_\mathbb{C})}$. Consider $Z_1 = \omega_{HC}^{-1}\Gamma(D)$. It is shown in [23, Cor. 4.4] that

$$Z_1 - \tau_{-\rho}\Gamma(D) \in U(\mathfrak{n}_\mathbb{C})U(\mathfrak{a}_\mathbb{C})^{\bar{d}-2} + U(\mathfrak{g}_\mathbb{C})\mathfrak{k}_\mathbb{C}.$$

It is not completely clear that Z_1 and D coincide on K -invariant functions, but the above formula shows that $\Gamma(Z_1) - \Gamma(D)$ is of degree $\leq \bar{d} - 2$, and hence that $Z_1 - D$ has degree at most $\bar{d} - 2$.

By descending induction on the degree of $\Gamma(Z) - \Gamma(D)$, we see that we can thus construct $Z \in \mathcal{Z}(\mathfrak{g}_\mathbb{C})$ such that

$$Z - \tau_{-\rho}\Gamma(D) \in U(\mathfrak{n}_\mathbb{C})U(\mathfrak{a}_\mathbb{C})^{\bar{d}-2} + U(\mathfrak{g}_\mathbb{C})\mathfrak{k}_\mathbb{C}$$

⁵This is the isomorphism denoted by γ_{HC} in [23], and defined by $\gamma_{HC}(z) = \tau_{\rho_b}\text{pr}(z)$, where $\text{pr}(z) \in U(\mathfrak{h}_\mathbb{C})$ is such that $z - \text{pr}(z) \in U(\mathfrak{n}_\mathbb{C})U(\mathfrak{a}_\mathbb{C}) + U(\mathfrak{g}_\mathbb{C})\mathfrak{k}_\mathbb{C}$.

and such that $\Gamma(Z) - \Gamma(D) = 0$ (which precisely means that Z and D coincide on right- K -invariant functions). \square

2.3. The Helgason-Fourier transform. For any $\theta \in G/P_0$, $\nu \in \mathfrak{a}_{\mathbb{C}}^*$, the function

$$e_{\nu, \theta} : x \in G/K \mapsto e^{(\rho + \nu)B(x, \theta)}$$

is a joint eigenfunction of \mathcal{D} , and one can verify easily (for instance in the case $\theta = eM$) that

$$De_{\nu, \theta} = [\Gamma(D)](\nu)e_{\nu, \theta},$$

for every $D \in \mathcal{D}$. Here we have seen $\Gamma(D)$ as a W -invariant polynomial on $\mathfrak{a}_{\mathbb{C}}^*$.

In fact for any joint eigenfunction ψ of \mathcal{D} there exists $\nu \in \mathfrak{a}_{\mathbb{C}}^*$ such that

$$D\psi = [\Gamma(D)](\nu)\psi$$

for every $D \in \mathcal{D}$ [11, Ch. II Thm. 5.18, Ch. III Lem. 3.11]. The parameter ν is called the “spectral parameter” of ψ ; it is uniquely determined up to the action of W .

The Helgason–Fourier transform gives the spectral decomposition of a function $u \in C_c^\infty(\mathbf{S})$ on the “basis” $(e_{\nu, \theta})$ of eigenfunctions of \mathcal{D} . It is defined as

$$(2.2) \quad \tilde{u}(\lambda, \theta) = \int_{\mathbf{S}} u(x)e_{-i\lambda, \theta}(x)dx,$$

($\lambda \in \mathfrak{a}^*$, $\theta \in G/P_0$). It has an inversion formula:

$$u(x) = \int_{\theta \in G/P_0, \lambda \in \overline{C_{\Pi}}} \tilde{u}(\lambda, \theta)e_{i\lambda, \theta}(x)d\theta|c(\lambda)|^{-2}d\lambda.$$

Here $d\theta$ denotes the normalized K -invariant measure on $G/P_0 \sim K/M$. The function c is the so-called Harish-Chandra function, given by the Gindikin-Karpelevic formula [11, Thm. 6.14, p. 447].

The Plancherel formula reads

$$\|u\|_{L^2(\mathbf{S})}^2 = \int_{\theta \in G/P_0, \lambda \in \overline{C_{\Pi}}} |\tilde{u}(\lambda, \theta)|^2 d\theta |c(\lambda)|^{-2} d\lambda.$$

Remark 2.3. For $D \in \mathcal{D}$, D acts on u by

$$Du(x) = \int_{\theta \in G/P_0, \lambda \in \overline{C_{\Pi}}} [\Gamma(D)](i\lambda)\tilde{u}(\lambda, \theta)e_{i\lambda, \theta}(x)d\theta|c(\lambda)|^{-2}d\lambda$$

3. QUANTIZATION AND PSEUDODIFFERENTIAL OPERATORS

In this section we develop a pseudodifferential calculus for \mathbf{S} , inspired by the work of Zelditch [27]. We do not push the analysis as far as in [27] (a more detailed analysis is done in Michael Schröder’s thesis [22]). For us, the most important feature of this quantization is that it is based on the Helgason-Fourier transform, in other words, on the spectral decomposition of the algebra \mathcal{D} .

3.1. Semiclassical Helgason transform. We now introduce a parameter \hbar . In the sequel it will tend to 0 at the same speed as $\|\nu\|^{-1}$; the reader may identify the two. The parameter will be assumed to go to infinity in the conditions of §1.2, the limit ν_∞ assumed to be regular.

From now on we rescale the parameter space \mathfrak{a}^* of the Helgason–Fourier transform by \hbar . We define the semiclassical Fourier transform, $\widehat{u}_\hbar(\lambda, \theta) = \widetilde{u}(\hbar^{-1}\lambda, \theta)$. Thus, for $u \in C_c^\infty(\mathbf{S})$, we rewrite equation (2.2) as:

$$\widehat{u}_\hbar(\lambda, \theta) = \int_{\mathbf{S}} u(x) e_{-i\hbar^{-1}\lambda, \theta}(x) dx$$

($\lambda \in \overline{C_\Pi}, \theta \in G/P_0$). The inversion formula now reads

$$u(x) = \int_{\theta \in G/P_0, \lambda \in \overline{C_\Pi}} \widehat{u}_\hbar(\lambda, \theta) e_{i\hbar^{-1}\lambda, \theta}(x) d\theta |c_\hbar(\lambda)|^{-2} d\lambda,$$

with the “semiclassical Harish-Chandra c -function”,

$$|c_\hbar(\lambda)|^{-2} = \hbar^{-r} |c(\hbar^{-1}\lambda)|^{-2}.$$

Remark 3.1. By the Gindikin-Karpelevic formula, we have

$$|c(\hbar^{-1}\lambda)|^{-2} \asymp \hbar^{-\dim \mathfrak{n}}$$

uniformly for λ in a compact subset of C_Π , and thus

$$|c_\hbar(\lambda)|^{-2} \asymp \hbar^{-d}$$

where $d = \dim \mathfrak{a} + \dim \mathfrak{n} = \dim(G/K)$.

We also adjust the Plancherel formula to

$$\|u\|_{L^2(\mathbf{S})}^2 = \int |\widehat{u}_\hbar(\lambda, \theta)|^2 d\theta |c_\hbar(\lambda)|^{-2} d\lambda.$$

In the sequel we will always use the semiclassical Fourier transform, and will in general denote \widehat{u} instead of \widehat{u}_\hbar .

3.2. Pseudodifferential calculus on \mathbf{Y} . We identify the functions on the quotient $\mathbf{Y} = \Gamma \backslash G/K$ (respectively $T^*\mathbf{Y}$) with the Γ -invariant functions on $\mathbf{S} = G/K$ (resp. $T^*(G/K)$). If Γ has torsion, we shall use “smooth function on \mathbf{Y} ” to mean a Γ -invariant smooth function on \mathbf{S} . For a compactly supported function χ on \mathbf{S} , we denote $\Pi_\Gamma \chi(x) = \sum_\gamma \chi(\gamma \cdot x)$. This sum is finite for any $x \in \mathbf{S}$, and hence defines a function on \mathbf{Y} .

On \mathbf{S} , we fix once and for all a positive, smooth and compactly supported function ϕ such that $\sum_{\gamma \in \Gamma} \phi(\gamma \cdot x) \equiv 1$. We call such a function a “smooth fundamental cutoff” or a “smooth fundamental domain”. Here we have used the assumption that \mathbf{Y} is compact. We also introduce $\tilde{\phi} \in C_c^\infty(\mathbf{S})$ which is identically 1 on the support of ϕ . We note that for any $D \in \mathcal{D}$ and for any smooth Γ -invariant u on \mathbf{S} we have

$$(3.1) \quad \Pi_\Gamma \left(\tilde{\phi} D(\phi u) \right) = \Pi_\Gamma D(\phi u) = D \Pi_\Gamma \phi u = Du.$$

The analogue of left-quantization on \mathbb{R}^n in our setting associates to a function a on $G/K \times G/P_0 \times C_\Pi$ the operator which acts on $u \in C_c^\infty(G/K)$ by

$$(3.2) \quad \text{Op}_\hbar^L(a) u(x) = \int_{\theta \in G/P_0, \lambda \in \overline{C_\Pi}} a(x, \theta, \lambda) \widehat{u}(\lambda, \theta) e_{i\hbar^{-1}\lambda, \theta}(x) d\theta |c_\hbar(\lambda)|^{-2} d\lambda .$$

A similar formula was introduced by Zelditch in [27] (with $\hbar = 1$) in the case $G = SL(2, \mathbb{R})$; it is shown there that $a \mapsto \text{Op}_\hbar^L(a)$ is G -equivariant. The operator $\text{Op}_\hbar^L(a)$ can be defined if a belongs to a nice class of functions (possibly depending on \hbar). If a is smooth enough and has reasonable growth, it will be a pseudodifferential operator. We give the regularity assumptions on a below. In any case, we shall always require a to be of the form $b \circ \pi$, where b is a symbol on $T^*(G/K)$ and π was defined in (1.1); besides, we will assume that b is supported away from the singular G -orbits in $T^*(G/K)$ (which means that a is supported away from the walls in C_Π). This allows to identify a in a natural way with a function defined on (a subset of) $T^*(G/K)$.

Let us define *symbols of order m* on $T^*(G/K)$ (independent of \hbar) in the usual fashion :

$$S^m(G/K) := \{a \in C^\infty(T^*(G/K))\}$$

for every compact $F \subset G/K$, for every α, β , there exists C such that

$$|D_z^\alpha D_\xi^\beta a(x, \xi)| \leq C(1 + |\xi|)^{m-|\beta|} \text{ for all } (x, \xi) \in T^*(G/K), x \in F\}.$$

We also define *semiclassical symbols of order m and degree l* — thus called because they depend on a parameter \hbar :

$$(3.3) \quad S^{m,l}(G/K) = \{a_\hbar(x, \xi) = \hbar^l \sum_{j=0}^{\infty} \hbar^j a_j(x, \xi), a_j \in S^{m-j}\}.$$

This means that $a_\hbar(x, \xi)$ has an asymptotic expansion in powers of \hbar , in the sense that

$$a - \hbar^l \sum_{j=0}^{N-1} \hbar^j a_j \in \hbar^{l+N} S^{m-N}$$

for all N , uniformly in \hbar . In this context, we denote $S^{-\infty, +\infty} = \bigcap_{m \geq 0} S^{-m, m}$.

Remark 3.2. As indicated above, we define symbols on $G/K \times G/P_0 \times C_\Pi$ by transporting the standard definition on $T^*(G/K)$ through the map π (1.1). We will exclusively consider the case where a vanishes outside a fixed neighbourhood of the singular G -orbits in $T^*(G/K)$. In other words, a can be identified (through (1.1)) with a function on $G/K \times G/P_0 \times C_\Pi$, that vanishes in a neighbourhood of $G/K \times G/P_0 \times \partial C_\Pi$. Defining a good pseudodifferential calculus using formula (3.2) for symbols supported near the walls of C_Π raises delicate issues about the behaviour of the c -function near the walls, and we do not address this problem here. This is one among several reasons why we assume that Λ_∞ is regular in our main theorem.

We now project this construction down to functions on \mathbf{Y} , which we identify with Γ -invariant functions on \mathbf{S} . Here we do not follow Zelditch, who defined the action of $\text{Op}_\hbar(a)$

on Γ -invariant functions in a global manner, using the Helgason-Fourier decomposition of such functions. We continue to work locally, which is sufficient for our purposes.

For us, the quantization of $a \in S^{m,k} \cap C^\infty(T^*\mathbf{Y})$ (supported away from singular G -orbits) is defined to act on $u \in C^\infty(\mathbf{Y})$ by:

$$(3.4) \quad \text{Op}_\hbar(a) u = \Pi_\Gamma \tilde{\phi} \text{Op}_\hbar^L(a) \phi u \in C^\infty(\mathbf{Y}).$$

Note that (3.1) and Remark 2.3 imply that $\text{Op}_\hbar(H) = \Gamma^{-1}[H(-i\hbar\bullet)]$ for $H \in \mathcal{H}$.

The image of $S^{m,k}$ by this quantization will be denoted $\Psi^{m,k}(\mathbf{Y})$. This quantization procedure depends on the fundamental cutoff ϕ and on $\tilde{\phi}$. However, this dependence only appears at second order in \hbar . The space $\Psi^{m,k}(\mathbf{Y})$ itself is perfectly well defined modulo $\Psi^{-\infty,+\infty}(\mathbf{Y}) = \cap_{k',m'} \Psi^{m',k'}(\mathbf{Y})$. Moreover, it coincides with the more usual definition of pseudodifferential operators, defined using the euclidean Fourier transform in local coordinates⁶.

3.3. Action of $\text{Op}_\hbar(H)$ on WKB states. Fix a Hamiltonian $H \in \mathcal{H}$.

The letter H will stand for several different objects which are canonically related: a function H on $T^*(G/K)$, a W -invariant polynomial function on \mathfrak{a}^* , and an element of $U(\mathfrak{a})^W$. As such, we can also let H act as a left- G -invariant differential operator⁷ on G or G/M .

In the following lemma, all functions on G/K and G/M are lifted to functions on G , and in that sense we can apply to them any differential operator on G . If b is a function defined on $G/M = G/K \times G/P_0$, and θ is an element of G/P_0 , we denote b_θ the function defined on G/K by $b_\theta(x) = b(x, \theta)$.

Lemma 3.3. *Let $H \in \mathcal{H}$ be of degree \bar{d} , and let b be a smooth function on G/M . Fix $\lambda \in \mathfrak{a}^*$. Then, there exist $D_k \in U(\mathfrak{n}_\mathbb{C})U(\mathfrak{a}_\mathbb{C})$ of degree $\leq k$ (depending on λ and on H) such that for any $\theta \in G/P_0$, for any $x \in G/K$,*

$$\text{Op}_\hbar(H)[b_\theta \cdot e_{i\hbar^{-1}\lambda, \theta}](x) = \left(H(\lambda)b(x, \theta) - i\hbar[dH(\lambda).b](x, \theta) + \sum_{k=2}^{\bar{d}} \hbar^k D_k b(x, \theta) \right) e_{i\hbar^{-1}\lambda, \theta}(x).$$

On the right H is seen as a function on \mathfrak{a}^* , so its differential $dH(\lambda)$ is an element of \mathfrak{a} , and it acts as a differential operator of order 1 on G/M . Each operator D_k actually defines a differential operator on G/M .

Proof. By linearity, it is enough to treat the case where $H \in U(\mathfrak{a})^W$ is homogeneous of degree \bar{d} . In this case, we have

$$\text{Op}_\hbar(H) = \hbar^{\bar{d}} \text{Op}_1(H) = \hbar^{\bar{d}} \Gamma^{-1}[H(-i\bullet)].$$

⁶This could be checked by testing the action of $\text{Op}_\hbar(a)$ on a local plane wave of the form $\phi(x)e^{\frac{i\xi \cdot x}{\hbar}}$ in local euclidean coordinates. One then uses the stationary phase method and the facts that the complex phase of $e_{\hbar^{-1}\lambda, \theta}$ is $\hbar^{-1}\lambda B(x, \theta)$, and that the covector $(x, d_x \lambda B(x, \theta)) \in T_x^*(G/K)$ corresponds precisely to (x, θ, λ) under the identification (1.1).

⁷We have also introduced the differential operator $\text{Op}_1(H) = \Gamma^{-1}[H(-i\bullet)]$ acting on G/K . These are not the same objects, but [23, Cor. 4.4] relates the two.

Consider the operator Z related to $D = \text{Op}_1(H)$ by Lemma 2.1. We have

$$\text{Op}_1(H)[b_\theta \cdot e_{i\hbar^{-1}\lambda, \theta}](x) = Z[b_\theta \cdot e_{i\hbar^{-1}\lambda, \theta}](x).$$

In what follows we consider the point $(x, \theta) \in G/K \times G/P_0$. We choose a representative of θ in K (θ is then defined modulo M , but the calculations do not depend on the choice of this representative). We write $x = \theta naK$. This means that (x, θ) represents the point $\theta naM \in G/M$. All functions on G/K and G/M are lifted to functions on G , and in that sense we can apply to them any differential operator on G .

By Lemma 2.1, we have

$$Z[b_\theta \cdot e_{i\hbar^{-1}\lambda, \theta}](x) = Z[b_\theta \cdot e_{i\hbar^{-1}\lambda, \theta}](\theta na) = \tau_{-\rho}H(-i\bullet) \cdot [b_\theta \cdot e_{i\hbar^{-1}\lambda, \theta}](\theta na) + D[b_\theta \cdot e_{i\hbar^{-1}\lambda, \theta}](\theta na)$$

where $D \in U(\mathfrak{n}_{\mathbb{C}})U(\mathfrak{a}_{\mathbb{C}})^{\bar{d}-2}$.

Because of the identity

$$e_{i\hbar^{-1}\lambda, \theta}(\theta nag) = e^{(\rho+i\hbar^{-1}\lambda)B(\theta na)} e^{(\rho+i\hbar^{-1}\lambda)H_0(g)},$$

(valid for any $g \in NA$) we see that, for any $D \in U(\mathfrak{n}_{\mathbb{C}})U(\mathfrak{a}_{\mathbb{C}})$, the term $D[e_{i\hbar^{-1}\lambda, \theta}](\theta na)$ is of the form $C e_{i\hbar^{-1}\lambda, \theta}(\theta na)$, where the constant C depends on D and $\hbar^{-1}\lambda$. This constant C is in fact polynomial in $\hbar^{-1}\lambda$.

This results in an expression :

$$\begin{aligned} Z[b_\theta \cdot e_{i\hbar^{-1}\lambda, \theta}](x) &= Z[b_\theta \cdot e_{i\hbar^{-1}\lambda, \theta}](\theta na) = \tau_{-\rho}H(-i\bullet) \cdot [b_\theta \cdot e_{i\hbar^{-1}\lambda, \theta}](\theta na) \\ &\quad + \left[\sum_{k=0}^{\bar{d}-2} \hbar^{-k} D_{\bar{d}-k} b(\theta na) \right] e_{i\hbar^{-1}\lambda, \theta}(\theta na) \end{aligned}$$

where $D_{\bar{d}-k} \in U(\mathfrak{n}_{\mathbb{C}})U(\mathfrak{a}_{\mathbb{C}})$ depends only on λ and H .

A term in \hbar^{-k} can only arise if $e_{i\hbar^{-1}\lambda, \theta}$ is differentiated k times; but Z being of degree \bar{d} , we see then that $D_{\bar{d}-k}$ can be of order $\bar{d} - k$ at most. The last term, when multiplied by $\hbar^{\bar{d}}$, becomes $\sum_{k=2}^{\bar{d}} \hbar^k D_k b$. We do not know a priori if the function $D_{\bar{d}-k} b$ (defined on G) is M -invariant, but the sum $\sum_{k=0}^{\bar{d}-2} \hbar^{-k} D_{\bar{d}-k} b$ necessarily defines an M -invariant function on G , since all the other terms do. Since \hbar is arbitrary, we see that each D_k must necessarily send an M -invariant function to an M -invariant function.

Finally, we write

$$\begin{aligned} \tau_{-\rho}H(-i\bullet) \cdot [b_\theta \cdot e_{i\hbar^{-1}\lambda, \theta}](\theta naM) &= H(-i\bullet) [b_\theta \cdot e_{i\hbar^{-1}\lambda-\rho, \theta}] \cdot e_{0, \theta}(\theta naM) \\ &= [\tau_{i\hbar^{-1}\lambda} H(-i\bullet) \cdot b_\theta] \cdot e_{i\hbar^{-1}\lambda, \theta}(\theta naM). \end{aligned}$$

When multiplying by $\hbar^{\bar{d}}$, and using the Taylor expansion of H at λ , we have

$$\hbar^{\bar{d}} \tau_{i\hbar^{-1}\lambda} H(-i\bullet) = H(\lambda) - i\hbar dH(\lambda) + \sum_{k=2}^{\bar{d}} \frac{(-i\hbar)^k}{k!} d^{(k)} H(\lambda).$$

□

We will refer to a function of the form $x \mapsto b_\theta(x) e_{i\hbar^{-1}\lambda, \theta}(x)$ as a *WKB state*, using the language of semiclassical analysis.

3.4. Symplectic lift. Let ψ be a \mathcal{D} -eigenfunction, of spectral parameter ν . We let $\hbar = \|\nu\|^{-1}$ (the choice of the norm here is arbitrary, one can take the Killing norm for instance). We sometimes write $\psi = \psi_\nu$ to indicate the spectral parameter, but this notation is imprecise in that ψ may not be uniquely determined by ν .

To ψ_ν we attach a distribution $\tilde{\mu}_\psi$ (sometimes denoted $\tilde{\mu}_\nu$) on $T^*\mathbf{Y}$: for $a \in C_c^\infty(T^*\mathbf{Y})$ set

$$\tilde{\mu}_\psi(a) = \langle \psi, \text{Op}_\hbar(a)\psi \rangle_{L^2(\mathbf{Y})}$$

As described in Section 1 we are trying to classify weak-* limits of the distributions $\tilde{\mu}_\nu$ in the limit $\nu \rightarrow \infty$. We fix such a limit (“semiclassical measure”) μ and a sequence $(\psi_j)_{j \in \mathbb{N}} = (\psi_{\nu_j})_{j \in \mathbb{N}}$ of eigenfunctions such that the corresponding sequence $(\tilde{\mu}_{\nu_j})$ converges weak-* to μ . In the sequel we write ν for ν_j . We assume that ν goes to infinity in the conditions of paragraph 1.2, the limit ν_∞ assumed to be regular. We let $\hbar = \|\nu\|^{-1}$. Writing $\Lambda = \Lambda_\nu = \hbar \mathfrak{S}m(\nu)$ we have $\Lambda \rightarrow \Lambda_\infty = \mathfrak{S}m(\nu_\infty) = -i\nu_\infty$. Note that $\Re(\nu)$ is bounded [15, §16.5(7) & Thm. 16.6]), so that $\hbar\nu = i\lambda_\nu + O(\hbar)$. Necessarily ν_∞ is purely imaginary.

With the notations of Section 2.2, the state ψ_ν satisfies

$$(3.5) \quad \text{Op}_\hbar(H).\psi_\nu = H(-i\hbar\nu)\psi_\nu$$

for all $H \in \mathcal{H}$. From now on, we fix a Hamiltonian $H \in \mathcal{H}$. The letter H will stand for two different objects that are canonically related: a function H on $T^*(G/K)$ (G -invariant and polynomial in the fibers of the projection $T^*(G/K) \rightarrow G/K$), a W -invariant polynomial function on \mathfrak{a}^* , an element of $U(\mathfrak{a})^W$.

We denote $X_\Lambda = dH(\Lambda) \in \mathfrak{a}$. Since Λ is only defined up to an element of W , so is X_Λ . One can assume that $\alpha(X_{\Lambda_\infty}) \geq 0$ for all $\alpha \in \Delta^+$. For simplicity (and without loss of generality), we will also assume that Λ_∞ belongs to the Weyl chamber C_Π .

Other miscellaneous notations: d is the dimension of G/K , r is the rank, and J the dimension of N (so that $d = r + J$). We call \tilde{J} the number of roots. We index the positive roots $\alpha_1, \dots, \alpha_{\tilde{J}}$ in such a way that $\alpha_1(X_{\Lambda_\infty}) \leq \alpha_2(X_{\Lambda_\infty}) \leq \dots \leq \alpha_{\tilde{J}}(X_{\Lambda_\infty})$ (with our previous notations, we have $\alpha_{\tilde{J}}(X_{\Lambda_\infty}) = \chi_{\max}(H)$). We fix \mathcal{K} as in Theorem 1.12, and we denote $j_0 = j_0(X_{\Lambda_\infty})$ the largest index j such that $\alpha_j(X_{\Lambda_\infty}) < \frac{1}{2\mathcal{K}}$.

With $w_\ell \in W$ the long element, we set: $\mathfrak{n}_{\text{fast}} = \bigoplus_{j > j_0} \mathfrak{g}_{\alpha_j}$, $\mathfrak{n}_{\text{slow}} = \bigoplus_{j \leq j_0} \mathfrak{g}_{\alpha_j}$, $\bar{\mathfrak{n}}_{\text{fast}} = \bigoplus_{j > j_0} \mathfrak{g}_{w_\ell.\alpha_j}$, $\bar{\mathfrak{n}}_{\text{slow}} = \bigoplus_{j \leq j_0} \mathfrak{g}_{w_\ell.\alpha_j}$. $J_0 = \dim \mathfrak{n}_{\text{slow}} = \sum_{j \leq j_0} m_{\alpha_j}$. The spaces $\mathfrak{n}_{\text{fast}}$ and $\bar{\mathfrak{n}}_{\text{fast}}$ are subalgebras, in fact ideals, in \mathfrak{n} , $\bar{\mathfrak{n}}$ respectively; they generate subgroups $N_{\text{fast}}, \bar{N}_{\text{fast}}$ that are normal in N, \bar{N} respectively.

4. THE WKB ANSATZ

We now start the proof of Theorem 1.12. We first describe how the operator P_ω^x acts on WKB states. In Section 5, we will use the fact that these states form a kind of basis to estimate the norm of the operator.

4.1. Goal of this section. Fix a sequence $\underline{\omega} = (\omega_{-T}, \dots, \omega_{-1}, \omega_0, \dots, \omega_{T-1})$, of length $2T$ chosen so that $T\eta \leq \mathcal{K}|\log \hbar|$. Theorem 1.12 requires us to estimate the norm of the

operator $P_{\underline{\omega}}^X$ acting on $L^2(\mathbf{Y})$ (for a suitable choice of the time step η). This operator is the same as $U^{-(T-1)\eta}\mathcal{P}$ where

$$\mathcal{P} = P_{\omega_{T-1}} U^\eta \dots U^\eta P_{\omega_0}^{1/2} \text{Op}_h(\chi) P_{\omega_0}^{1/2} U^\eta \dots P_{\omega_{-T+1}} U^\eta P_{\omega_{-T}},$$

where we recall that

$$U^t = \exp(i\hbar^{-1}t \text{Op}_h(H)).$$

On the ‘‘energy layer’’ \mathcal{E}_λ , U^t quantizes the action of e^{-tX_λ} , in other words the time $-t$ of the Hamiltonian flow generated by H . Under the action of e^{-tX_λ} for $t \geq 0$, elements of \mathfrak{n} are expanded and elements of $\bar{\mathfrak{n}}$ are contracted (the vector X_Λ may be singular, so that these stable or unstable spaces can also contain neutral directions).

In what follows we estimate the norm of \mathcal{P} . To do so, we will first describe how \mathcal{P} acts on our Fourier basis $e_{i\hbar^{-1}\lambda,\theta}$, using the technique of WKB expansion (§4.2). Then, we will use the Cotlar-Stein lemma (§5) to estimate as precisely as possible the norm of \mathcal{P} .

The sequence $\omega_{-T}, \dots, \omega_{T-1}$ is fixed throughout this section. Instead of working with functions on \mathbf{Y} we work with functions on G/K that are Γ -invariant. For instance, P_ω is the multiplication operator by the Γ -invariant function P_ω . We assume that each connected component of the support of P_ω has very small diameter (say ϵ). We will fix Q_ω , a function in $C_c^\infty(\mathbf{S})$ such that $\Pi_\Gamma Q_\omega = P_\omega$ and such that the support of Q_ω has diameter ϵ . We also denote Q_ω the corresponding multiplication operator. Finally we need to introduce Q'_ω in $C_c^\infty(\mathbf{S})$ which is identically 1 on the support of Q_ω and supported in a set of diameter 2ϵ .

We decompose

$$(4.1) \quad \mathcal{P} = \mathcal{S}^* \mathcal{U}_\chi$$

where

$$\mathcal{U}_\chi = \text{Op}(\chi) P_{\omega_0}^{1/2} U^\eta P_{\omega_{-1}} \dots U^\eta \widehat{P}_{\omega_{-T+1}} U^\eta P_{\omega_{-T}}$$

and

$$\mathcal{S} = P_{\omega_0}^{1/2} \dots U^{-\eta} P_{\omega_{T-2}} U^{-\eta} P_{\omega_{T-1}}.$$

4.2. The WKB Ansatz for the Schrödinger propagator . We recall some standard calculations, already done in [3], with some additional simplifications coming from the fact that the functions $e_{i\hbar^{-1}\lambda,\theta}$ are eigenfunctions of $\text{Op}_h(H)$.

On \mathbf{S} , let us try to solve

$$-i\hbar \frac{\partial \tilde{u}}{\partial t} = \text{Op}_h(H) \tilde{u},$$

in other words

$$\tilde{u}(t) = U^t \tilde{u}(0),$$

with initial condition the WKB state $\tilde{u}(0, x) = a_h(0, x) e_{i\hbar^{-1}\lambda,\theta}(x)$. We only consider $t \geq 0$. We assume that a_h is compactly supported and has an asymptotic expansion in all C^l norms as $a_h \sim \sum_{k \geq 0} \hbar^k a_k$. We look for approximate solution up to order \hbar^M , in the form

$$u(t, x) = e^{\frac{itH(\lambda)}{\hbar}} e_{i\hbar^{-1}\lambda,\theta}(x) a_h(t, x) = e^{\frac{itH(\lambda)}{\hbar}} e_{i\hbar^{-1}\lambda,\theta}(x) \sum_{k=0}^{M-1} \hbar^k a_k(t, x).$$

Let us denote

$$(4.2) \quad u(t, x) = e^{\frac{itH(\lambda)}{\hbar}} e_{i\hbar^{-1}\lambda, \theta}(x) a_{\hbar}(t, x, \theta, \lambda) = e^{\frac{itH(\lambda)}{\hbar}} e_{i\hbar^{-1}\lambda, \theta}(x) \sum_{k=0}^{M-1} \hbar^k a_k(t, x, \theta, \lambda)$$

to keep track of the dependence on θ and λ ; the pair (x, θ) then represents an element of $G/K \times G/P_0 = G/M$. Identifying powers of \hbar , and using Lemma 3.3, we find the conditions:

$$(4.3) \quad \begin{cases} \frac{\partial a_0}{\partial t}(x, \theta) = [dH(\lambda).a_0](x, \theta) & \text{(0-th transport equation)} \\ \frac{\partial a_k}{\partial t}(x, \theta) = [dH(\lambda).a_k](x, \theta) + i \sum_{l=2}^{\bar{d}} \sum_{l+m=k+1} D_l a_m(x, \theta) & \text{(k-th transport equation)} \end{cases} .$$

The equations (4.3) can be solved explicitly by

$$a_0(t, (x, \theta), \lambda) = a_0(0, (x, \theta)) e^{tX_\lambda},$$

in other words

$$a_0(t) = R(e^{tX_\lambda}) a_0(0),$$

where R here denotes the action of A on functions on G/M by right translation; and

$$a_k(t) = R(e^{tX_\lambda}) a_k(0) + \int_0^t R(e^{(t-s)X_\lambda}) \left(i \sum_{l=2}^{\bar{d}} \sum_{l+m=k+1} D_l a_m(s, x, \theta) \right) ds.$$

If we now define u by (4.2), u solves

$$-i\hbar \frac{\partial \tilde{u}}{\partial t} = \text{Op}_{\hbar}(H) \tilde{u} - e^{\frac{itH(\lambda)}{\hbar}} e_{i\hbar^{-1}\lambda, \theta} \left[\sum_{l=2}^{\bar{d}} \sum_{k=M+1-l}^{M-1} \hbar^{k+l} D_l a_k \right]$$

and thus

$$\begin{aligned} \|u(t) - U^t u(0)\|_{L^2(\mathfrak{S})} &\leq \int_0^t \left[\sum_{l=2}^{\bar{d}} \sum_{k=M+1-l}^{M-1} \hbar^{k+l-1} \|D_l a_k(s)\|_{L^2(\mathfrak{S})} \right] ds \\ &\leq t e^{(2M+\bar{d}-2)t \max_{\alpha \in \Delta^+} \alpha(X_\lambda)^-} \left[\sum_{l=2}^{\bar{d}} \sum_{k=M+1-l}^{M-1} \hbar^{k+l-1} \sum_{j=0}^k \|a_{k-j}(0)\|_{C^{2j+l}} \right] \\ &\leq Ct \hbar^M e^{(2M+\bar{d}-2)t \max_{\alpha \in \Delta^+} \alpha(X_\lambda)^-} \left[\sum_{k=0}^{M-1} \|a_k(0)\|_{C^{2(M-k)+\bar{d}-2}} \right]. \end{aligned}$$

Since D_k belongs to $U(\mathfrak{n}_{\mathbb{C}})U(\mathfrak{a}_{\mathbb{C}})$, in the co-ordinates (x, θ) it only involves differentiation with respect to x . We also recall that D_k is of order k . We have used the following estimate on the flow $R(e^{tX_\lambda})$ (for $t \geq 0$) :

$$\left\| \frac{d^N}{dx^N} a((x, \theta) e^{tX_\lambda}) \right\| \leq e^{-tN \min_{\alpha \in \Delta^+} \alpha(X_\lambda)} \left\| \frac{d^N}{dx^N} a((x, \theta)) \right\|$$

and we have denoted $x^- = \max(-x, 0)$.

Remark 4.1. In what follows we will always have $\lambda \in \text{supp}(\chi)$, where by assumption χ is supported on a tubular neighbourhood of size ϵ of $\mathcal{E}_{\Lambda_\infty}$, and $\alpha(\Lambda_\infty) \geq 0$ for $\alpha \in \Delta^+$. For such λ we have $\alpha(X_\lambda) \geq -\epsilon$ for all $\alpha \in \Delta^+$. We see that our approximation method makes sense if t is restricted by $\hbar^M e^{(2M+\bar{d}-2)t\epsilon} \ll 1$. Since ϵ can be chosen arbitrarily small, we can assume that the WKB approximation is good for $t \leq 3\mathcal{K}|\log \hbar|$.

Remark 4.2. On the quotient $\mathbf{Y} = \Gamma \backslash \mathbf{S}$, the same method applies to find an approximate solution of $U^t \Pi_\Gamma u(0)$ in the form $\Pi_\Gamma u(t)$, with the same bound

$$(4.4) \quad \|\Pi_\Gamma u(t) - U^t \Pi_\Gamma u(0)\|_{L^2(\mathbf{Y})} \leq Ct\hbar^M e^{ct(2M+\bar{d}-2)} \left[\sum_{k=0}^{M-1} \|a_k(0)\|_{C^{2(M-k)+\bar{d}-2}} \right],$$

provided that the projection $\mathbf{S} \rightarrow \mathbf{Y}$ is bijective when restricted to the support of $a_h(t)$. If λ stays in a compact set and if the support of $a_h(0)$ has small enough diameter ϵ , this condition will be satisfied in a time interval $t \in [0, T_0]$. In the applications below, we may and will always assume that $\eta < T_0$.

We can iterate the previous WKB construction T times to get the following description of the action of \mathcal{U}_χ on $\Pi_\Gamma Q'_{\omega_{-T}} e_{i\hbar^{-1}\lambda, \theta}$ (the induction argument to control the remainders at each step is the same as in [3] and we won't repeat it here):

Proposition 4.3.

$$(4.5) \quad \mathcal{U}_\chi(\Pi_\Gamma Q'_{\omega_{-T}} e_{i\hbar^{-1}\lambda, \theta}) = \Pi_\Gamma \left[e^{\frac{iT\eta H(\lambda)}{\hbar}} e_{i\hbar^{-1}\lambda, \theta} A_M^{(T)}(\bullet, \theta, \lambda) \right] + \mathcal{O}_{L^2(\mathbf{Y})}(\hbar^M) \|Q'_{\omega_{-T}} e_{i\hbar^{-1}\lambda, \theta}\|_{L^2(\mathbf{S})}$$

where

$$A_M^{(T)}(x, \theta, \lambda) = \sum_{k=0}^{M-1} \hbar^k a_k^{(T)}(x, \theta, \lambda).$$

The function $a_0^{(T)}(x, \theta, \lambda)$ is equal to

$$a_0^{(T)}(x, \theta, \lambda) = \chi(\lambda) P_{\omega_0}^{1/2}(x) P_{\omega_{-1}}((x, \theta) e^{\eta X_\lambda}) P_{\omega_{-2}}((x, \theta) e^{2\eta X_\lambda}) \dots Q_{\omega_{-T}}((x, \theta) e^{T\eta X_\lambda}),$$

where we have lifted the functions P_ω (originally defined on G/K) to $G/M = G/K \times G/P_0$. The functions $a_k^{(T)}$ have the same support as $a_0^{(T)}$. Moreover, if we consider $a_k^{(T)}$ as a function of (x, θ) , that is, as a function on G/M , we have the following bound

$$\|Z_\alpha^m a_k^{(T)}\| \leq P_{k,m,Z_\alpha}(T) \sup_{j=0,\dots,T} \{e^{-(m+2k)j\eta \alpha(X_\lambda)}\}$$

if Z_α belongs to \mathfrak{g}_α ($P_{k,m,Z_\alpha}(T)$ is polynomial in T). In particular, for $\alpha \in \Delta^+$,

$$\|Z_\alpha^m a_k^{(T)}\| \leq P_{k,m,Z_\alpha}(T) e^{(m+2k)T\eta \epsilon}$$

The energy parameter λ will always stay ϵ -close to Λ_∞ . Recall that we denote by the same letter ϵ the diameter of the support of each Q_ω . We choose ϵ and η (the time step)

small enough to ensure the following: there exists $\gamma = \gamma_{\omega_{-T}, \dots, \omega_0} \in \Gamma$ (independent of θ or λ) such that

$$(4.6) \quad a_0^{(T)}(x, \theta, \lambda) = \chi(\lambda) Q_{\omega_0}^{1/2} \circ \gamma^{-1}(x) P_{\omega_{-1}}((x, \theta) e^{\eta X \lambda}) P_{\omega_{-2}}((x, \theta) e^{2\eta X \lambda}) \dots Q_{\omega_{-T}}((x, \theta) e^{T\eta X \lambda}).$$

This means that the function $a_0^{(T)}(\bullet, \theta, \lambda)$ is supported in a single connected component of the support of $P_{\omega_0}^{1/2}$.

We will also use the following variant:

Proposition 4.4. *Let $\gamma = \gamma_{\omega_{-T}, \dots, \omega_0}$.*

$$\mathcal{U}_\chi(Q'_{\omega_{-T}} \circ \gamma \quad e_{i\hbar^{-1}\lambda, \theta}) = e^{\frac{iT\eta H(\lambda)}{\hbar}} e_{i\hbar^{-1}\lambda, \theta}(x) A_M^{(T)} \circ \gamma(x, \theta, \lambda) + \mathcal{O}(\hbar^M) \|Q'_{\omega_{-T}} \circ \gamma \quad e_{i\hbar^{-1}\lambda, \theta}\|$$

where

$$A_M^{(T)}(x, \theta, \lambda) = \sum_{k=0}^{M-1} \hbar^k a_k^{(T)}(x, \theta, \lambda).$$

Remark 4.5. For the operator \mathcal{S} , analogous results can be obtained if we replace everywhere λ by $w_\ell \lambda$, $-t$ by $+t$, and the label ω_{-j} by ω_{+j} .

Remark 4.6. Let $u, v \in L^2(\mathbf{Y})$. We explain how the previous Ansatz can be used to estimate the scalar product $\langle v, \mathcal{U}_\chi u \rangle_{L^2(\mathbf{Y})}$ (up to a small error). This is done by decomposing u and v , locally, into a combination of the functions $e_{i\hbar^{-1}\lambda, \theta}$ (using the Helgason-Fourier transform), and inputting our Ansatz into this decomposition.

In more detail, we note that $P_{\omega_{-T}} = P_{\omega_{-T}} \Pi_\Gamma Q_{\omega_{-T}}'^2$, so that $\mathcal{U}_\chi u = \mathcal{U}_\chi \Pi_\Gamma Q_{\omega_{-T}}'^2 u$. We use the Fourier decomposition to write

$$Q_{\omega_{-T}}'^2 u(x) = Q'_{\omega_{-T}}(x) \int_{\theta \in G/P_0, \lambda \in \overline{C_\Pi}} \widehat{Q'_{\omega_{-T}} u}(\lambda, \theta) e_{i\hbar^{-1}\lambda, \theta}(x) d\theta |c_\hbar(\lambda)|^{-2} d\lambda.$$

By Cauchy-Schwarz and the asymptotics of the c -function (Remark 3.1), we note that

$$\int_{\chi(\lambda) \neq 0} |\widehat{Q'_{\omega_{-T}} u}(\lambda, \theta)| d\theta |c_\hbar(\lambda)|^{-2} d\lambda = \mathcal{O}(\hbar^{-d/2}) \|u\|_{L^2(\mathbf{Y})},$$

and write

$$(4.7) \quad \begin{aligned} \langle v, \mathcal{U}_\chi u \rangle_{L^2(\mathbf{Y})} &= \left\langle v, \mathcal{U}_\chi \Pi_\Gamma Q_{\omega_{-T}}'^2 u \right\rangle_{L^2(\mathbf{Y})} \\ &= \int_{\chi(\lambda) \neq 0} \widehat{Q'_{\omega_{-T}} u}(\lambda, \theta) \left\langle v, \mathcal{U}_\chi \Pi_\Gamma Q'_{\omega_{-T}} e_{i\hbar^{-1}\lambda, \theta} \right\rangle d\theta |c_\hbar(\lambda)|^{-2} d\lambda + \mathcal{O}(\hbar^\infty) \|u\|_{L^2(\mathbf{Y})} \|v\|_{L^2(\mathbf{Y})}. \end{aligned}$$

We now use Proposition 4.3 to replace \mathcal{U}_χ by the Ansatz,

$$\begin{aligned} \left\langle v, \mathcal{U}_\chi \Pi_\Gamma Q'_{\omega_{-T}} e_{i\hbar^{-1}\lambda, \theta} \right\rangle_{L^2(\mathbf{Y})} &= \left\langle v, e^{\frac{iT\eta H(\lambda)}{\hbar}} e_{i\hbar^{-1}\lambda, \theta} A_M^{(T)}(\bullet, \theta, \lambda) \right\rangle_{L^2(\mathbf{S})} + \mathcal{O}(\hbar^M) \|v\|_{L^2(\mathbf{Y})} \\ &= \left\langle Q'_{\omega_0} \circ \gamma^{-1} \cdot v, e^{\frac{iT\eta H(\lambda)}{\hbar}} e_{i\hbar^{-1}\lambda, \theta} A_M^{(T)}(\bullet, \theta, \lambda) \right\rangle_{L^2(\mathbf{S})} + \mathcal{O}(\hbar^M) \|v\|_{L^2(\mathbf{Y})} \\ &= \left\langle Q'_{\omega_0} v, e^{\frac{iT\eta H(\lambda)}{\hbar}} e_{i\hbar^{-1}\lambda, \theta} \circ \gamma A_M^{(T)}(\gamma \bullet, \theta, \lambda) \right\rangle_{L^2(\mathbf{S})} + \mathcal{O}(\hbar^M) \|v\|_{L^2(\mathbf{Y})} \end{aligned}$$

for $\gamma = \gamma_{\omega_{-T}, \dots, \omega_0}$ defined above. Thus,

$$(4.8) \quad \langle v, \mathcal{U}_\chi u \rangle_{L^2(\mathbf{Y})} = \int_{\chi(\lambda) \neq 0} \widehat{Q'_{\omega_{-T}} u}(\lambda, \theta) \left\langle Q'_{\omega_0} v, e^{\frac{iT\eta H(\lambda)}{\hbar}} e_{i\hbar^{-1}\lambda, \theta} \circ \gamma \quad A_M^{(T)}(\gamma_\bullet, \theta, \lambda) \right\rangle_{L^2(\mathbf{S})} d\theta |c_\hbar(\lambda)|^{-2} d\lambda \\ + \mathcal{O}(\hbar^{M-d/2}) \|v\|_{L^2(\mathbf{Y})} \|u\|_{L^2(\mathbf{Y})}.$$

In this last line we see that replacing the exact expression of \mathcal{U}_χ by the Ansatz induces an error of $\mathcal{O}(\hbar^{M-d/2}) \|v\|_{L^2(\mathbf{Y})} \|u\|_{L^2(\mathbf{Y})}$. We will take M very large, depending on the constant \mathcal{K} in Theorem 1.12, so that the error $\mathcal{O}(\hbar^{M-d/2})$ is negligible compared to the bound announced in the theorem.

5. THE COTLAR–STEIN ARGUMENT.

We now use the previous approximations of \mathcal{U}_χ and \mathcal{S} to estimate the norm of \mathcal{P} . This is done in a much finer, and more technical manner, than in [1, 3], because we want to eliminate the slowly expanding/contracting directions.

5.1. The Cotlar-Stein lemma.

Lemma 5.1. *Let E, F be two Hilbert spaces. Let $(A_\alpha) \in \mathcal{L}(E, F)$ be a countable family of bounded linear operators from E to F . Assume that for some $R > 0$ we have*

$$\sup_\alpha \sum_\beta \|A_\alpha^* A_\beta\|^{\frac{1}{2}} \leq R$$

and

$$\sup_\alpha \sum_\beta \|A_\alpha A_\beta^*\|^{\frac{1}{2}} \leq R$$

Then $A = \sum_\alpha A_\alpha$ converges strongly and A is a bounded operator with $\|A\| \leq R$.

We refer for instance to [7] for the proof.

5.2. A non-stationary phase lemma. The following lemma is just a version of integration by parts.

Lemma 5.2. *Let Ω be an open set in a smooth manifold. Let Z be a vector field on Ω and μ be a measure on Ω with the property that $\int (Zf) d\mu = \int f J d\mu$ for every smooth function f and for some smooth J .*

Let $S \in C^\infty(\Omega, \mathbb{R})$ and $a \in C_c^\infty(\Omega)$. Assume that ZS does not vanish. Consider the integral

$$(5.1) \quad I_\hbar = \int e^{\frac{iS(x)}{\hbar}} a(x) d\mu(x).$$

Then we have $I_\hbar = i\hbar \int e^{\frac{iS(x)}{\hbar}} D_Z a(x) d\mu(x)$, where the operator D_Z is defined by

$$D_Z a = Z \left(\frac{a}{ZS} \right) - \frac{aJ}{ZS}.$$

If we iterate this formula n times we get

$$I_{\hbar} = (i\hbar)^n \int e^{\frac{iS(x)}{\hbar}} D_Z^n a(x) d\mu(x)$$

and D_Z^n has the form

$$D_Z^n a = \sum_{m \geq n, k+m \leq 2n, \sum l_j \leq n} f_{k, (l_j), m} \frac{Z^k a Z^{l_1} S \dots Z^{l_r} S}{(ZS)^m}$$

where the $f_{k, (l_j), m}(x)$ are smooth functions that do not depend on a nor S .

5.3. Study of several phase functions.

5.3.1. Sum of two Helgason phase functions.

Proposition 5.3. (i) Let $g_1 P_0, g_2 P_0 \in G/P_0$ be two points on the boundary. Let $\lambda, \mu \in \overline{C_{\Pi}}$ be two elements of the closed nonnegative Weyl chamber. Consider the function on G/K ,

$$(5.2) \quad gK \mapsto \lambda.H_0(g_1^{-1}gK) + \mu.H_0(g_2^{-1}gK).$$

Then, this map has critical points if and only if $\mu = -\text{Ad}(w_{\ell}).\lambda$.

(ii) Let $\lambda, \mu \in C_{\Pi}$ be two (regular) elements of the positive Weyl chamber. Let $g_1 P_0, g_2 P_0 \in G/P_0$ be two points on the boundary, and assume that $g_1^{-1}g_2 \in P_0 w_{\ell} P_0$ (we don't assume here that the conclusion of (i) is satisfied). Write $g_1^{-1}g_2 = b_1 w_{\ell} b_2$ with $b_1, b_2 \in P_0$.

Then, the set of critical points for variations of the form

$$t \mapsto \lambda.H_0(e^{tX} g_1^{-1}gK) + \mu.H_0(e^{tX} g_2^{-1}gK),$$

with $X \in \mathfrak{n}$ is precisely $\{gK, g \in g_1 b_1 A\}$. Moreover, these critical points are non-degenerate.

Remark 5.4. The set of critical points is $\{gK, g \in g_1 P_0, g w_{\ell} \in g_2 P_0\}$, that is, the flat in G/K determined by the two boundary points $g_1 P_0, g_2 P_0$.

Proof. (i) It is enough to consider the case $g_1 = e$. By the Bruhat decomposition, we know that there exists a unique $w \in W$ such that $g_2 \in BwB$, that is, $g_2 = b_1 w b_2$ for some $b_1, b_2 \in B$. The map (5.2) has the same critical points as the map

$$(5.3) \quad gK \mapsto \lambda.H_0(gK) + \mu.H_0(w^{-1}b_1^{-1}gK),$$

and those are the image under $gK \mapsto b_1 gK$ of the critical points of

$$(5.4) \quad gK \mapsto \lambda.H_0(gK) + \mu.H_0(w^{-1}gK).$$

For $X \in \mathfrak{a}$ the derivative at $t = 0$ of

$$(5.5) \quad t \mapsto \lambda.H_0(e^{tX} gK) + \mu.H_0(w^{-1}e^{tX} gK)$$

is $\lambda(X) + \mu(\text{Ad}(w^{-1})X)$. Thus, for the map (5.4) to have critical points, we must have

$$\lambda(X) + \mu(\text{Ad}(w^{-1})X) = 0$$

for every $X \in \mathfrak{a}$. Letting X vary over the dual basis to a positive basis of \mathfrak{a}^* , we see that $\mu = -\text{Ad}(w).\lambda$ is nonnegative, and this is only possible if $\mu = -\text{Ad}(w_{\ell}).\lambda$ (this does not necessarily mean that $w = w_{\ell}$ if λ is not regular).

(ii) Here we assume that μ and λ are regular, and that we are in the “generic” case where $g_1^{-1}g_2 \in P_0w_\ell P_0$. Starting from (5.4), we now consider variations of the form

$$(5.6) \quad t \mapsto \lambda.H_0(e^{tX}gK) + \mu.H_0(w_\ell^{-1}e^{tX}gK)$$

for $X \in \mathfrak{n}$. The term $\lambda.H_0(e^{tX}gK)$ is constant, and it remains to deal with $\mu.H_0(w_\ell^{-1}e^{tX}gK)$. Write $g = w_\ell anK$, $n \in N, a \in A$, and denote $Y = \text{Ad}(w_\ell).X \in \bar{\mathfrak{n}}$, $Y' = \text{Ad}(a^{-1})Y$. We have

$$\mu.H_0(w_\ell^{-1}e^{tX}gK) = \mu.H_0(e^{tY}anK) = \mu(a) + \mu.H_0(e^{tY'}nK) = \mu(a) + \mu.H_0(n^{-1}e^{tY'}nK).$$

Hence

$$\frac{d}{dt}\mu.H_0(e^{tY}anK) = \mu.H_0(\text{Ad}(n^{-1})Y').$$

We see that the set of critical points of (5.6) is the set of those points gK , with $g = w_\ell anK$ such that n satisfies $\mu.H_0(\text{Ad}(n^{-1})Y') = 0$ for all $Y' \in \bar{\mathfrak{n}}$. Since μ is regular, one can check that this implies $n = e$. This proves the first assertion of (ii).

Finally, assume that we are at a critical point, that is, $gK = aK$ in (5.6). We calculate the second derivative at $t = 0$ of $t \mapsto \mu.H_0(w_\ell^{-1}e^{tX}aK)$ when $X \in \mathfrak{n}$. We keep the same notation as above for Y and Y' .

Let $U = Y' - \theta(Y') \in \mathfrak{k}$. By the Baker-Campbell-Hausdorff formula, we have

$$(5.7) \quad e^{tY'} = e^{t\theta(Y') + \frac{t^2}{2}[Y', \theta(Y')] + \mathcal{O}(t^3)} e^{tU} = e^{t\theta(Y')} e^{\frac{t^2}{2}[Y', \theta(Y')] + \mathcal{O}(t^3)} e^{tU}.$$

Remember that $\theta(Y') \in \mathfrak{n}$, and that H_0 is left- N -invariant. This calculation shows that the second derivative of $t \mapsto \mu.H_0(w_\ell^{-1}e^{tX}aK)$ is the quadratic form

$$X \mapsto \mu([Y', \theta(Y')]),$$

where $Y' = \text{Ad}(a^{-1})\text{Ad}(w_\ell).X$. This is a non-degenerate quadratic form if μ is regular. \square

5.3.2. *Variations with respect to \bar{N} .* In this section we need the decomposition $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{a} \oplus \mathfrak{m} \oplus \bar{\mathfrak{n}}$. We will denote $\pi_{\mathfrak{n}}, \pi_{\mathfrak{a}}, \pi_{\bar{\mathfrak{n}}}$ the corresponding projections. We note that $\pi_{\mathfrak{a}} = H_0$, since $\bar{\mathfrak{n}} \subset \mathfrak{n} + \mathfrak{k}$.

Lemma 5.5. *Fix $n \in N$ and $a \in A$. Then there exist two neighbourhoods V_1, V_2 of 0 in $\bar{\mathfrak{n}}$, and a diffeomorphism $\Psi = \Psi_{na}: V_1 \rightarrow V_2$ such that*

$$e^{-Y_1}nae^{Y_2} \in NA, Y_1 \in V_1, Y_2 \in V_2 \iff Y_2 = \Psi(Y_1).$$

Moreover, the differential at 0 of Ψ (denoted Ψ'_0) preserves the subalgebra $\bar{\mathfrak{n}}_{\text{slow}}$. Finally, if we write $e^{-Y}nae^{\Psi(Y)} = n(Y)a(Y)$, we have

$$a'_0.Y = \pi_{\mathfrak{a}}[\text{Ad}(na)\Psi'_0(Y)].$$

Proof. We apply the implicit function theorem. For $Y_1 = 0$, the differential of $Y_2 \mapsto nae^{Y_2}(na)^{-1}$ at $Y_2 = 0$ is $Y_2 \mapsto \text{Ad}(na).Y_2$. What we need to check is the equivalence of

$\pi_{\bar{n}}[\text{Ad}(na).Y_2] = 0$ and $Y_2 = 0$, which is the case since $\text{Ad}(na)$ preserves $\mathfrak{n} \oplus \mathfrak{a} \oplus \mathfrak{m}$. So the existence of Ψ is proved, in addition the differential Ψ'_0 is defined by

$$Y = \pi_{\bar{n}}[\text{Ad}(na).\Psi'_0.Y]$$

for $Y \in \bar{\mathfrak{n}}$. Since $\text{Ad}(na)$ preserves the space $\mathfrak{n} \oplus \mathfrak{a} \oplus \mathfrak{m} \oplus \bar{\mathfrak{n}}_{\text{slow}}$ (without preserving the decomposition, of course), $\Psi'_0.Y$ must belong to $\bar{\mathfrak{n}}_{\text{slow}}$ if Y does.

The last formula is simply obtained by differentiating $e^{-Y}nae^{\Psi(Y)} = n(Y)a(Y)$. \square

For the next lemma we need to recall our two decompositions $\bar{\mathfrak{n}} = \sum_{k \leq j_0} \mathfrak{g}_{w_\ell.\alpha_k} \oplus \sum_{j > j_0} \mathfrak{g}_{w_\ell.\alpha_k} = \bar{\mathfrak{n}}_{\text{slow}} \oplus \bar{\mathfrak{n}}_{\text{fast}}$ and $\mathfrak{n} = \sum_{k \leq j_0} \mathfrak{g}_{\alpha_k} \oplus \sum_{j > j_0} \mathfrak{g}_{\alpha_k} = \mathfrak{n}_{\text{slow}} \oplus \mathfrak{n}_{\text{fast}}$. The space $\mathfrak{n}_{\text{fast}}$ is an ideal of \mathfrak{n} , and we call N_{fast} the associated (normal) subgroup.

Lemma 5.6. (i) *The set*

$$\{n \in N, H_0(\text{Ad}(n)Y) = 0 \quad \forall Y \in \bar{\mathfrak{n}}_{\text{slow}}\}$$

is, near identity, a submanifold of N , tangent to $\mathfrak{n}_{\text{fast}}$.

(ii) *If n is close enough to identity, if we write $n = e^{\sum \alpha T_\alpha}$ with $T_\alpha \in \mathfrak{g}_\alpha$ ($\alpha \in \Delta^+$), and if $\mu \in \mathfrak{a}^*$ is a regular element, we have*

$$\left| \mu.H_0 \left(\text{Ad}(n) \frac{\theta(T_\beta)}{\|\theta(T_\beta)\|} \right) \right| \geq C_\mu \|T_\beta\|$$

with $C_\mu > 0$.

Proof. The differential of $n \mapsto H_0(\text{Ad}(n)Y)$ is $Z \mapsto H_0([Z, Y])$ ($Z \in \mathfrak{n}$). Write $Z = Z_{\text{slow}} + Z_{\text{fast}}$, $Z_{\text{slow}} = \sum Z_\alpha$, with $Z_\alpha \in \mathfrak{g}_\alpha$, and take $Y = \theta(Z_\beta)$ for some β . We have $H_0([Z, Y]) = -\langle Z_\beta, Z_\beta \rangle H_\beta$ where $H_\beta \in \mathfrak{a}$ is the coroot [14, Ch. VI §5, Prop. 6.52]. Note that $\mu(H_\beta) = \langle \mu, \beta \rangle$, hence $\mu(H_\beta) \neq 0$ if μ is regular. This proves the lemma. \square

5.4. First decomposition of \mathcal{P} . We want to use the Cotlar-Stein lemma to estimate the norm of the operator \mathcal{P} , defined in (4.1). To do so, we will decompose \mathcal{P} into many pieces. Our first decomposition of \mathcal{P} is obtained by covering the boundary G/P_0 by a finite number of small sets $\Omega_1, \dots, \Omega_M$ described below. We use the fact that there is a neighbourhood Ω of eP_0 in G/P_0 that is diffeomorphic to a neighbourhood of e in \bar{N} , via the map

$$\begin{aligned} \bar{N} &\longrightarrow G/P_0 \\ \bar{n} &\mapsto \bar{n}P_0. \end{aligned}$$

Using compactness, we can find an open cover of G/P_0 by a finite number of open sets $\Omega_1, \dots, \Omega_M$ such that, for every m , there exists $g_m \in G$ with $\Omega_m \subset g_m\Omega \subset g_m\bar{N}P_0$. Introduce a family of smooth functions χ_{Ω_m} on G/P_0 such that χ_{Ω_m} is supported inside Ω_m and $\sum_m \chi_{\Omega_m} \equiv 1$. We then define the pseudodifferential operators

$$Q_m u(x) = \int \widehat{u}(w_\ell.\lambda, k) Q'_{\omega_0}(x) \chi_{\Omega_m}(k) e^{i\hbar^{-1}w_\ell.\lambda.k} dk |c_\hbar(\lambda)|^{-2} d\lambda,$$

and

$$\mathcal{P}_m u = \Pi_\Gamma \mathcal{S}^* Q_m^* \mathcal{U}_\chi u$$

Obviously, $\mathcal{P} = \sum_m \mathcal{P}_m$. The sum over m is finite, and we now fix m . The variable k stays in $g_m \overline{N} P_0$.

Remark 5.7. Let $\gamma = \gamma_{\omega_{T-1}, \dots, \omega_0}$ defined as in (4.6). Proposition (4.4) (and Remark 4.5) can be generalized to

$$(5.8) \quad Q_m \mathcal{S} \left(Q'_{\omega_{T-1}} \circ \gamma \ e_{ih^{-1}w_\ell, \mu, k} \right) = e^{\frac{-iT\eta H(\mu)}{\hbar}} e_{ih^{-1}w_\ell, \mu, k} B_M^{(T)} \circ \gamma (x, k, w_\ell, \mu) + \mathcal{O}_{L^2(\mathbf{S})}(\hbar^M) \|Q'_{\omega_T} \circ \gamma \ e_{ih^{-1}w_\ell, \mu, k}\|_{L^2(\mathbf{S})},$$

where now

$$(5.9) \quad B_M^{(T)}(x, k, w_\ell, \mu) = \sum_{k=0}^{M-1} \hbar^k b_k^{(T)}(x, k, w_\ell, \mu),$$

$$b_0^{(T)}(x, k, w_\ell, \mu) = \chi_{\Omega_m}(k) P_{\omega_0}^{1/2}(x) P_{\omega_1}((x, k) e^{-\eta X_{w_\ell, \mu}}) P_{\omega_2}((x, k) e^{-2\eta X_{w_\ell, \mu}}) \dots Q_{\omega_{T-1}}((x, k) e^{-(T-1)\eta X_{w_\ell, \mu}}) \\ = \chi_{\Omega_m}(k) Q_{\omega_0}^{1/2} \circ \gamma^{-1}(x) P_{\omega_1}((x, \theta) e^{-\eta X_{w_\ell, \mu}}) P_{\omega_2}((x, k) e^{-2\eta X_{w_\ell, \mu}}) \dots Q_{\omega_{T-1}}((x, k) e^{-(T-1)\eta X_{w_\ell, \mu}})$$

and the next terms have the same support as the leading one (their derivatives are bounded the same way as in Proposition 4.3).

In the next paragraphs we will concentrate our attention on brackets of the form:

$$\left\langle Q'_{\omega_{T-1}} \circ \gamma_2 \ e_{ih^{-1}w_\ell, \mu, k}, \mathcal{S}^* Q_m^* \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 \ e_{ih^{-1}\lambda, \theta} \right\rangle_{L^2(\mathbf{S})},$$

for $\lambda, \mu \in C_\Pi$, $\theta, k \in G/P_0$. We take $\gamma_1 = \gamma_{\omega_{-T}, \dots, \omega_0}$ and $\gamma_2 = \gamma_{\omega_{T-1}, \dots, \omega_0}$ as defined in (4.6). These are none other than the matrix elements of the operator \mathcal{P}_m in the Fourier basis $e_{ih^{-1}\lambda, \theta}$.

5.5. Second decomposition of \mathcal{P} . The index m being fixed, we will apply the Cotlar-Stein lemma to bound the norm of \mathcal{P}_m . We decompose \mathcal{P}_m as a sum of countably many operators, and this decomposition is much more involved.

We have assumed that we have a diffeomorphism from a relatively compact subset of \overline{N} to Ω_m : $\bar{n}_1 \mapsto g_m \bar{n}_1 P_0$. We can write the Haar measure on Ω_m as $dk = \text{Jac}(\bar{n}_1) d\bar{n}_1$, where Jac is a smooth function on \overline{N} (we suppress from the notation its dependence on g_m). An element $(x, k) \in G/K \times \Omega_m$ corresponding to the point $g_m \bar{n}_1 n_1 a_1 M \in G/M$ can also be represented as $(g_m \bar{n}_1 n_1 a_1 K, g_m \bar{n}_1) \in G/K \times g_m \overline{N}$. Accordingly we now write denote $e_{ih^{-1}w_\ell, \mu, g_m \bar{n}_1}$ for $e_{ih^{-1}w_\ell, \mu, k}$.

Let us look at a scalar product $\left\langle Q_m \mathcal{S} Q'_{\omega_{T-1}} \circ \gamma_2 \ e_{ih^{-1}w_\ell, \mu, g_m \bar{n}_1 P_0}, \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 \ e_{ih^{-1}\lambda, \theta} \right\rangle$. We only need to consider the generic case where $\theta \in g_m \bar{n}_1 P_0 w_\ell P_0$, that is, θ is of the form $g_m \bar{n}_1 n_1 w_\ell P_0$ (with $n_1 \in N$). In addition, we always assume that λ and μ are regular. Proposition 5.3 (ii) tells us that the stationary points of the phase function

$$gK \mapsto \lambda.H_0(\theta^{-1}gK) - (w_\ell, \mu).H_0(k^{-1}gK), \quad k = g_m \bar{n}_1 P_0$$

with respect to variations

$$(g_m \bar{n}_1 n_1) e^{tX} (g_m \bar{n}_1 n_1)^{-1} gK, \quad X \in \mathfrak{n},$$

are the points of the form $gK = g_m \bar{n}_1 n_1 a_1 K$ with $a_1 \in A$. Thus the set of critical points is of codimension J . The stationary phase method then gives:

$$(5.10) \quad \left\langle Q_m \mathcal{S} Q'_{\omega_{T-1}} \circ \gamma_2 e_{i\hbar^{-1} w_\ell \cdot \mu, g_m \bar{n}_1 P_0}, \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 e_{i\hbar^{-1} \lambda, \theta} \right\rangle \\ = \hbar^{J/2} \int_{a_1 \in A} d(\lambda, a_1) C_\hbar(g_m \bar{n}_1 n_1 a_1 M, \lambda, w_\ell \cdot \mu) \bar{e}_{i\hbar^{-1} w_\ell \cdot \mu, g_m \bar{n}_1 P_0}(g_m \bar{n}_1 n_1 a_1 K) \\ e_{i\hbar^{-1} \lambda, g_m \bar{n}_1 n_1 w_\ell P_0}(g_m \bar{n}_1 n_1 a_1 K) da_1$$

where $C_\hbar(g_m \bar{n}_1 n_1 a_1 M, \lambda, w_\ell \cdot \mu) \sim \sum \hbar^k c_k(g_m \bar{n}_1 n_1 a_1 M, \lambda, w_\ell \cdot \mu)$ and

$$c_0(g_m \bar{n}_1 n_1 a_1 M, \lambda, w_\ell \cdot \mu) = \\ \left(A_M^{(T)} \circ \gamma_1(g_m \bar{n}_1 n_1 a_1 w_\ell M, \lambda) \right) \left(\bar{B}_M^{(T)} \circ \gamma_2(g_m \bar{n}_1 n_1 a_1 M, w_\ell \cdot \mu) \right).$$

(and the next terms have the same support as the leading one). The term $d(\lambda, a_1)$ is the prefactor involving the hessian of the phase function in the application of the method of stationary phase, it is a smooth function. So the asymptotics of our scalar product only takes into account the elements $g_m \bar{n}_1 n_1 a_1 M$ with

$$A_M^{(T)} \circ \gamma_1(g_m \bar{n}_1 n_1 a_1 w_\ell M, \lambda) \bar{B}_M^{(T)} \circ \gamma_2(g_m \bar{n}_1 n_1 a_1 M, w_\ell \cdot \mu) \neq 0.$$

Lemma 5.8. *Assume that the diameter of Ω and of $\text{supp } Q_{\omega_0}$ is smaller than ϵ . Then there exist $n_0 \in N$ and a_0 in A such that*

$$B_M^{(T)} \circ \gamma_2(g_m \bar{n}_1 n_1 a_1 M, w_\ell \cdot \mu) \neq 0$$

implies $n_1 a_1 = n_0 a_0 g$, where $g \in NA$ is ϵ -close to identity.

Proof. Just note from the expression of $B_M^{(T)} \circ \gamma_2$ that, if it is not 0, we must have

$$g_m \bar{n}_1 n_1 a_1 \in \text{supp } Q_{\omega_0}.$$

The element g_m varies in a finite set and \bar{n}_1 varies over Ω which is of diameter $\leq \epsilon$. We also assume that $\text{supp } Q_{\omega_0}$ is of diameter $\leq \epsilon$, so that n_1 (and a_1) must both vary in sets of diameter $\leq \epsilon$. \square

It follows that $\bar{n}_1 n_1 a_1 M$ itself is ϵ -close to $n_0 a_0 M$ in G/M . From now on we write $g_m \bar{n}_1 n_1 a_1 M = g_m n_0 a_0 g M$, where $gM \in G/M$ varies in a neighbourhood of eM of diameter $\leq \epsilon$. We will always choose a representative $g \in \exp(\mathfrak{n} \oplus \mathfrak{a} \oplus \bar{\mathfrak{n}})$. By G -equivariance we may assume $g_m n_0 a_0 = 1$, which we do from now on.

Proposition 5.9. *(Contracting and expanding foliations)*

- (1) *Let μ be such that $\alpha_k(X_\mu) > 0$ for all $\alpha_k \in \Delta^+$ with $k > j_0$ (this is of course the case if μ is close enough to Λ_∞). Suppose we have gM and $g'M$ both ϵ -close to eM such that $B_M^{(T)} \circ \gamma_2(gM, w_\ell \cdot \mu) \neq 0$ and $B_M^{(T)} \circ \gamma_2(g'M, w_\ell \cdot \mu) \neq 0$, then $g'^{-1}g = \exp(X + \sum_{\alpha \in \Delta^+} Y_\alpha + \sum_{\alpha \in \Delta^+} Y_{w_\ell \cdot \alpha})$ with $X \in \mathfrak{a}, Y_\alpha \in \mathfrak{g}_\alpha, \|X\|, \|Y_\alpha\| \leq \epsilon$, and $\|Y_{w_\ell \cdot \alpha_k}\| \leq \epsilon e^{-T\eta(w_\ell \cdot \alpha_k)(X_{w_\ell \cdot \mu})} = \epsilon e^{-T\eta\alpha_k(X_\mu)}$ for $k > j_0$.*

- (2) Similarly, assume that $\alpha_k(X_\lambda) > 0$ for all $\alpha_k \in \Delta^+$ with $k > j_0$. Suppose we have gM and $g'M$ both ϵ -close to eM such that $A_M^{(T)} \circ \gamma_1(gw_\ell M, \lambda) \neq 0$ and $A_M^{(T)} \circ \gamma_1(g'w_\ell M, \lambda) \neq 0$. Then $g'^{-1}g = \exp(X + \sum_\alpha Y_\alpha)$ with $X \in \mathfrak{a}, Y_\alpha \in \mathfrak{g}_\alpha, \|X\|, \|Y_\alpha\| \leq \epsilon, \|Y_{\alpha_k}\| \leq \epsilon e^{-T\eta\alpha_k(X_\lambda)}$ for $k > j_0$.

Actually, the claim holds for all k (not only for $k > j_0$), but we will only use it for $k > j_0$. For the other indices, there is something more optimal to do.

Proof. Assume that the term $B_M^{(T)} \circ \gamma_2(gM, w_\ell \cdot \mu)$ does not vanish. The evolution equation (5.9) shows that we must have⁸

- $ge^{-(T-1)\eta X_{w_\ell \cdot \mu}} M \in \gamma_2^{-1} \cdot \text{supp } Q_{\omega_{T-1}};$
- $gM \in \text{supp } Q_{\omega_0}.$

If gM and $g'M$ both satisfy the two conditions above, then we see that $g'^{-1}g$ must be ϵ -close to identity. For ϵ small enough we can write this element using the co-ordinates described in part (1) of the claim. Also, $e^{(T-1)\eta X_{w_\ell \cdot \mu}} g'^{-1}g e^{-(T-1)\eta X_{w_\ell \cdot \mu}}$ must stay in the fixed compact set

$$M[\text{supp } Q_{\omega_{T-1}}]^{-1} \text{supp } Q_{\omega_{T-1}} M \subset G.$$

Writing the action of A in the co-ordinate system gives the claim. The proof of the second part is similar. \square

Finally we write $gM = \bar{n}naM$ with $\bar{n} \in \bar{N}, n \in N, a \in A$ all ϵ -close to 1. We decompose $n = e^Y n_{\text{fast}}$, and $\bar{n} = e^{\bar{Y}} \bar{n}_{\text{fast}}$, $Y \in \mathfrak{n}_{\text{slow}} \simeq \mathbb{R}^{J_0}, \bar{Y} \in \bar{\mathfrak{n}}_{\text{slow}} \simeq \mathbb{R}^{J_0}$ both ϵ -close to 0 (we fix a vector space isomorphism that sends the root spaces to the coordinate axes of \mathbb{R}^{J_0}); and $n_{\text{fast}} \in N_{\text{fast}}, \bar{n}_{\text{fast}} \in \bar{N}_{\text{fast}}$ both ϵ -close to 1. The quantity ϵ is fixed, but can be chosen as small as we wish. Note that the previous Proposition restricts n_{fast} and \bar{n}_{fast} to sets of measure $\prod_{k>j_0} \epsilon e^{-T\eta m_{\alpha_k} \alpha_k(X_\lambda)}$ and $\prod_{k>j_0} \epsilon e^{-T\eta m_{\alpha_k} \alpha_k(X_\mu)}$, respectively.

We will now break \mathcal{P}_m into countably many pieces,

$$\mathcal{P}_m = \sum_{(\bar{y}, y, t, \lambda_0) \in \mathbb{Z}^{2J_0} \times \mathbb{Z}^{2r}} \mathcal{P}_{m, (\bar{y}, y, t, \lambda_0)}$$

to which we will apply the Cotlar-Stein lemma.

For $j = J_0$ and $j = r$ choose a smooth nonnegative compactly supported function χ^j on \mathbb{R}^j such that

$$(5.11) \quad \sum_{y \in \mathbb{Z}^j} \chi^j(Y - y) \equiv 1$$

and such that $\chi^j(Y) \cdot \chi^j(Y + 2y) = 0$ for all $Y \in \mathbb{R}^j$ and $y \in \mathbb{Z}^j \setminus \{0\}$.

Let $(\bar{y}, y) \in \mathbb{Z}^{2J_0}$ and let $(t, \lambda_0) \in \mathbb{Z}^{2r}$. Denote 2^+ a fixed real number > 2 . Define $\chi_{(\bar{y}, y)}^h(\bar{Y}, Y) = \chi^{J_0}(\hbar^{-1/2^+} \bar{Y} - \bar{y}) \chi^{J_0}(\hbar^{-1/2^+} Y - y)$; and $\chi_{\lambda_0}^h(\lambda) = \chi^r(\hbar^{-1/2^+} \lambda - \lambda_0)$ and $\chi_t^h(a) = \chi^r(\hbar^{-1/2^+} a - t)$. Also define $\chi_{(\bar{y}, y, t)}^h(gM) = \chi_{(\bar{y}, y, t)}^h(\bar{Y}, Y) \chi_t^h(a)$ if gM is an element of G/M that can be decomposed as $gM = e^{\bar{Y}} \bar{n}_{\text{fast}} e^Y n_{\text{fast}} aM$, as described above.

⁸Here the Q_ω are treated as functions on G/M that factor through G/K .

We define a bounded operator $\mathcal{S}_{m,(\bar{y},y,t,\lambda_0)} : L^2(G/K) \longrightarrow L^2(G/K)$ by

$$\mathcal{S}_{m,(\bar{y},y,t,\lambda_0)} [e^{i\hbar^{-1}w_\ell \cdot \mu, k}](x) \stackrel{\text{def}}{=} e^{\frac{-iT\eta H(\mu)}{\hbar}} e^{i\hbar^{-1}w_\ell \cdot \mu, k}(x) \chi_{(\bar{y},y,t)}^\hbar(x, k) \chi_{\lambda_0}^\hbar(\mu) B_M^{(T)} \circ \gamma_2(x, k, w_\ell \cdot \mu).$$

We then define

$$\mathcal{P}_{m,(\bar{y},y,t,\lambda_0)} \stackrel{\text{def}}{=} \Pi_\Gamma Q'_{\omega_{T-1}} \circ \gamma_2 \mathcal{S}_{m,(\bar{y},y,t,\lambda_0)}^* \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1.$$

It can be checked that

$$\|\mathcal{P}_m - \sum_{(\bar{y},y,t,\lambda_0) \in \mathbb{Z}^{2k_0+2r}} \mathcal{P}_{m,(\bar{y},y,t,\lambda_0)}\|_{L^2(\mathbf{Y}) \rightarrow L^2(\mathbf{Y})} = \mathcal{O}(\hbar^{M-d/2}),$$

by noting that the sum $\sum_{(\bar{y},y,t,\lambda_0) \in \mathbb{Z}^{2k_0+2r}} \mathcal{S}_{m,(\bar{y},y,t,\lambda_0)}$ gives back our Ansatz (5.8) for $Q_m \mathcal{S}$, and by arguing as in (4.7) that the difference between $Q_m \mathcal{S}$ and the Ansatz is of order $\mathcal{O}(\hbar^{M-d/2})$. Again we choose M large enough so that the error $\mathcal{O}(\hbar^{M-d/2})$ is negligible compared to the bound announced in Theorem 1.12.

Let us now look at a scalar product $\langle \mathcal{S}_{m,(\bar{y},y,t,\lambda_0)} e^{i\hbar^{-1}w_\ell \cdot \mu, \bar{n}P_0}, \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 e^{i\hbar^{-1}\lambda, \theta} \rangle$. We need only consider the generic case where $\theta \in \bar{n}P_0 w_\ell P_0$, that is, θ is of the form $\theta = \bar{n}n w_\ell P_0$ (with $n \in N$). From the previous discussions, it follows that this scalar product is non-negligible only if \bar{n} and n stay in some sets of diameters $\leq \epsilon$; and, without loss of generality, we have assumed they are both ϵ -close to 1. As in (5.10), we have by the stationary phase method

$$\begin{aligned} (5.12) \quad & \left\langle \mathcal{S}_{m,(\bar{y},y,t,\lambda_0)} e^{i\hbar^{-1}w_\ell \cdot \mu, \bar{n}P_0}, \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 e^{i\hbar^{-1}\lambda, \theta} \right\rangle \\ &= \hbar^{J/2} \int_{a \in A} d(\lambda, a) C_h^{(\bar{y},y,t,\lambda_0)}(\bar{n}naM, \lambda, w_\ell \cdot \mu) \bar{e}_{i\hbar^{-1}w_\ell \cdot \mu, \bar{n}P_0}(\bar{n}naK) e_{i\hbar^{-1}\lambda, \bar{n}nw_\ell P_0}(\bar{n}naK) da \\ &= \hbar^{J/2} \int_{a \in A} d(\lambda, a) C_h^{(\bar{y},y,t,\lambda_0)}(\bar{n}naM, \lambda, w_\ell \cdot \mu) \bar{e}_{i\hbar^{-1}w_\ell \cdot \mu, \bar{n}P_0}(\bar{n}naK) e_{i\hbar^{-1}\lambda, \bar{n}nw_\ell P_0}(\bar{n}naK) da. \end{aligned}$$

where $C_h^{(\bar{y},y,t,\lambda_0)}(\bar{n}naM, \lambda, w_\ell \cdot \mu) = \sum \hbar^k c_k(\bar{n}naM, \lambda, w_\ell \cdot \mu)$ and

$$c_0(\bar{n}naK, \bar{n}P_0, \bar{n}nw_\ell P_0, \lambda, w_\ell \cdot \mu) = A_M^{(T)} \circ \gamma_1(\bar{n}naw_\ell M, \lambda) \bar{B}_M^{(T)} \circ \gamma_2(\bar{n}naM, w_\ell \cdot \mu) \times \chi_{(\bar{y},y,t)}^\hbar(\bar{n}naM) \chi_{\lambda_0}^\hbar(\mu)$$

(the next terms have the same support as the leading one). Remember the notation $\bar{n} = e^{\bar{Y}} \bar{n}_{fast}$, $n = e^Y n_{fast}$. By Proposition 5.9, and by definition of the cut-off functions χ^{J_0}, χ^r , our scalar product is non-negligible only if \bar{Y}, Y stay in a set of measure $\hbar^{J_0/2^+}$, and n_{fast}, \bar{n}_{fast} stay in a set of measure $\prod_{k>j_0} e^{-T\eta m_{\alpha_k} \alpha_k(X_\lambda)}$ and $\prod_{k>j_0} e^{-T\eta m_{\alpha_k} \alpha_k(X_\mu)}$, respectively.

5.6. Norm of $\mathcal{P}_{m,(\bar{x},x,s,\mu_0)}^* \mathcal{P}_{m,(\bar{y},y,t,\lambda_0)}$. We are now ready to check the first assumption of the Cotlar-Stein lemma, that is, to bound from above the norm of $\mathcal{P}_{m,(\bar{x},x,s,\mu_0)}^* \mathcal{P}_{m,(\bar{y},y,t,\lambda_0)}$.

Let $u, v \in L^2(\Gamma \backslash G/K)$. We write

$$\begin{aligned} & \langle \mathcal{P}_{m,(\bar{x},x,s,\mu_0)} v, \mathcal{P}_{m,(\bar{y},y,t,\lambda_0)} u \rangle_{\Gamma \backslash G/K} \\ &= \left\langle Q'_{\omega_{T-1}} \circ \gamma_2 \mathcal{S}_{m,(\bar{x},x,s,\mu_0)}^* \mathcal{U}_\chi Q'^2_{\omega_{-T}} \circ \gamma_1 v, Q'_{\omega_{T-1}} \circ \gamma_2 \mathcal{S}_{m,(\bar{y},y,t,\lambda_0)}^* \mathcal{U}_\chi Q'^2_{\omega_{-T}} \circ \gamma_1 u \right\rangle_{G/K} \\ &= \left\langle \mathcal{S}_{m,(\bar{x},x,s,\mu_0)}^* \mathcal{U}_\chi Q'^2_{\omega_{-T}} \circ \gamma_1 v, \mathcal{S}_{m,(\bar{y},y,t,\lambda_0)}^* \mathcal{U}_\chi Q'^2_{\omega_{-T}} \circ \gamma_1 u \right\rangle_{G/K} + \mathcal{O}(\hbar^\infty) \|u\| \|v\|. \end{aligned}$$

We develop fully this scalar product using the Fourier transform.

$$\begin{aligned} (5.13) \quad & \left\langle \mathcal{S}_{m,(\bar{x},x,s,\mu_0)}^* \mathcal{U}_\chi Q'^2_{\omega_{-T}} \circ \gamma_1 v, \mathcal{S}_{m,(\bar{y},y,t,\lambda_0)}^* \mathcal{U}_\chi Q'^2_{\omega_{-T}} \circ \gamma_1 u \right\rangle_{G/K} \\ &= \int d\theta d\theta' |c_\hbar(\lambda)|^{-2} d\lambda |c_\hbar(\lambda')|^{-2} d\lambda' \widehat{Q'_{\omega_{-T}} \circ \gamma_1 u(\lambda, \theta)} \overline{\widehat{Q'_{\omega_{-T}} \circ \gamma_1 v(\lambda', \theta')}} \\ & \left\langle \mathcal{S}_{m,(\bar{x},x,s,\mu_0)}^* \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 e^{i\hbar^{-1}\lambda', \theta'}, \mathcal{S}_{m,(\bar{y},y,t,\lambda_0)}^* \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 e^{i\hbar^{-1}\lambda, \theta} \right\rangle_{G/K} \\ &= \int d\theta d\theta' dk |c_\hbar(\lambda)|^{-2} d\lambda |c_\hbar(\lambda')|^{-2} d\lambda' |c_\hbar(\mu)|^{-2} d\mu \widehat{Q'_{\omega_{-T}} \circ \gamma_1 u(\lambda, \theta)} \overline{\widehat{Q'_{\omega_{-T}} \circ \gamma_1 v(\lambda', \theta')}} \\ & \left\langle \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 e^{i\hbar^{-1}\lambda', \theta'}, \mathcal{S}_{m,(\bar{x},x,s,\mu_0)} e^{i\hbar^{-1}w_\ell, \mu, k} \right\rangle \left\langle \mathcal{S}_{m,(\bar{y},y,t,\lambda_0)} e^{i\hbar^{-1}w_\ell, \mu, k} \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 e^{i\hbar^{-1}\lambda, \theta} \right\rangle_{G/K} \\ &= \int d\theta d\theta' \text{Jac}(\bar{n}) d\bar{n} |c_\hbar(\lambda)|^{-2} d\lambda |c_\hbar(\lambda')|^{-2} d\lambda' |c_\hbar(\mu)|^{-2} d\mu \widehat{Q'_{\omega_{-T}} \circ \gamma_1 u(\lambda, \theta)} \overline{\widehat{Q'_{\omega_{-T}} \circ \gamma_1 v(\lambda', \theta')}} \\ & \left\langle \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 e^{i\hbar^{-1}\lambda', \theta'}, \mathcal{S}_{m,(\bar{x},x,s,\mu_0)} e^{i\hbar^{-1}w_\ell, \mu, \bar{n}P_0} \right\rangle \left\langle \mathcal{S}_{m,(\bar{y},y,t,\lambda_0)} e^{i\hbar^{-1}w_\ell, \mu, \bar{n}P_0}, \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 e^{i\hbar^{-1}\lambda, \theta} \right\rangle_{G/K} \end{aligned}$$

Finally, in equation (5.13), we write $\theta = \bar{n}n w_\ell P_0$ and $\theta' = \bar{n}n' w_\ell P_0$ (we can do so on a set of full measure). We have shown in (5.12) that

$$\begin{aligned} (5.14) \quad & \left\langle \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 e^{i\hbar^{-1}\lambda', \theta'}, \mathcal{S}_{m,(\bar{x},x,s,\mu_0)} e^{i\hbar^{-1}w_\ell, \mu, \bar{n}P_0} \right\rangle \left\langle \mathcal{S}_{m,(\bar{y},y,t,\lambda_0)} e^{i\hbar^{-1}w_\ell, \mu, \bar{n}P_0}, \mathcal{U}_\chi Q'_{\omega_{-T}} \circ \gamma_1 e^{i\hbar^{-1}\lambda, \theta} \right\rangle_{G/K} \\ &= \hbar^J \int_{a \in A} d(\lambda, a) C_h^{(\bar{y},y,t,\lambda_0)}(\bar{n}naM, \lambda, w_\ell, \mu) \bar{e}_{i\hbar^{-1}w_\ell, \mu, \bar{n}}(\bar{n}naK) e_{i\hbar^{-1}\lambda, \bar{n}nw}(\bar{n}naK) da \\ & \int_{a' \in A} d(\lambda', a') \bar{C}_h^{(\bar{x},x,s,\mu_0)}(\bar{n}n'a'M, \lambda', w_\ell, \mu) e_{i\hbar^{-1}w_\ell, \mu, \bar{n}}(\bar{n}n'a'K) \bar{e}_{i\hbar^{-1}\lambda', \bar{n}n'w}(\bar{n}n'a'K) da'. \end{aligned}$$

Already we can note that $C_h^{(\bar{y},y,t,\lambda_0)}(\bar{n}naM, \lambda, w_\ell, \mu) \bar{C}_h^{(\bar{x},x,s,\mu_0)}(\bar{n}n'a'M, \lambda', w_\ell, \mu)$ can only be non zero if $\chi_{(\bar{y},y,t)}^{\hbar}(\bar{n}naM) \chi_{(\bar{x},x,s)}^{\hbar}(\bar{n}n'a'M) \neq 0$, and from the way we chose χ^{J_0} this can happen only for $\|\bar{x} - \bar{y}\| \leq 2$. For the same reason, it can only be non zero if $\|\mu_0 - \lambda_0\| \leq 2$.

Now we try to show that (5.13) decays fast when $\|x - y\|$ gets large. Under the last integral in (5.13) we have a function of the pair $(\bar{n}na, \bar{n}n'a')$. We have an oscillatory integral of the form (5.1), with a phase

$$\begin{aligned} S(\bar{n}na, \bar{n}n'a') &= \lambda.B(\bar{n}naw_\ell M) + (w_\ell, \mu)[B(\bar{n}n'a') - B(\bar{n}na)] - \lambda'.B(\bar{n}n'a'w_\ell M) \\ &= \lambda.B(\bar{n}naw_\ell M) + (w_\ell, \mu)[a' - a] - \lambda'.B(\bar{n}n'a'w_\ell M), \end{aligned}$$

where B is the function defined in (2.1). We want to do “integration by parts with respect to \bar{n} ” (as in Lemma 5.2). However, because the derivatives of S with respect to \bar{n} are tricky to compute, it is preferable to use a vector field Z whose definition is a bit delicate but with the property that $Z.B(\bar{n}naw_\ell M) = 0$ and $Z.B(\bar{n}n'a'w_\ell M) = 0$.

Consider a variation of the form

$$\Psi^\tau : (\bar{n}na, \bar{n}n'a') \mapsto (\bar{n}ne^{\tau Y}a, \bar{n}n'a'a^{-1}e^{\Psi(\tau Y)}a) = \bar{n}n(e^{\tau Y}a, n^{-1}n'a'a^{-1}e^{\Psi(\tau Y)}a),$$

for $Y \in \bar{\mathfrak{n}}$, and $\Psi = \Psi_{n^{-1}n'a'a^{-1}}$ defined in lemma 5.5. By definition of Ψ , the two elements $\bar{n}ne^{\tau Y}a$ and $\bar{n}n'a'a^{-1}e^{\Psi(\tau Y)}a$ are in the same NA orbit, for all τ . Such a variation preserves the terms $B(\bar{n}n'a'w_\ell M)$ and $B(\bar{n}naw_\ell M)$. We call Z the vector field $\frac{d\Psi^\tau}{d\tau}|_{\tau=0}$. We take $Y \in \mathfrak{g}_{w_\ell \cdot \alpha_k}$ with $1 \leq k \leq j_0$. We note that each term of the product

$$Q'_{\omega_{-T}} \widehat{\circ \gamma_1} u(\lambda, \bar{n}nw_\ell P_0) \overline{Q'_{\omega_{-T}} \widehat{\circ \gamma_1} v(\lambda', \bar{n}n'w_\ell P_0)} e_{i\hbar^{-1}\lambda, \bar{n}nw_\ell}(\bar{n}naK) \bar{e}_{i\hbar^{-1}\lambda', \bar{n}n'w_\ell}(\bar{n}n'a'K)$$

is invariant under Ψ^τ . The function $C_h^{(\bar{y}, y, t, \lambda_0)}(\bar{n}naM) \bar{C}_h^{(\bar{x}, x, s, \mu_0)}(\bar{n}n'a'M)$ satisfies

$$\|Z^m C_h^{(\bar{y}, y, t, \lambda_0)}(\bar{n}naM) \bar{C}_h^{(\bar{x}, x, s, \mu_0)}(\bar{n}n'a'M)\| \leq C(m) \hbar^{-m/2^+}$$

just by the definition of $C_h^{(\bar{y}, y, t, \lambda_0)}$ and $C_h^{(\bar{x}, x, s, \mu_0)}$. Now we want to apply the nonstationary phase lemma 5.2, so we need to understand $ZS = Z[(w_\ell \cdot \mu)[B(\bar{n}n'a') - B(\bar{n}na)]]$.

Lemmas 5.5 and 5.6 tell us that if we write $n^{-1}n' = \exp(T)$ with $T = \sum_\alpha T_\alpha$ ϵ -close to 0, choose β among the slow exponents so that the norm of T_β is comparable to the norm of $\log(n^{-1}n')_{\text{slow}}$ and take $Y \in \bar{\mathfrak{n}}_{\text{slow}}$ of norm 1 such that $\Psi'_0(Y) = \theta(T_\beta)$ then

$$(5.15) \quad |ZS(\bar{n}na, \bar{n}n'a')| \geq C \|\log(n^{-1}n')_{\text{slow}}\|.$$

Note that we have $\|\log(n^{-1}n')_{\text{slow}}\| \geq \hbar^{1/2^+} (\|x-y\| - 4)$ if $C_h^{(\bar{y}, y, t, \lambda_0)}(\bar{n}naM) \bar{C}_h^{(\bar{x}, x, s, \mu_0)}(\bar{n}n'a'M) \neq 0$.

We now apply Lemma 5.2 to the last expression of integral (5.13), integrating by parts \tilde{M} -times using the vector field Z . This yields that $\langle \mathcal{P}_{m, (\bar{x}, x, s, \mu_0)} v, \mathcal{P}_{m, (\bar{y}, y, t, \lambda_0)} u \rangle_{\mathbf{Y}}$ is bounded from above by

$$\frac{C(M) \hbar^{\tilde{M}(1-2/2^+)}}{\max(16, \|x-y\|)^{\tilde{M}}} \hbar^J \int \text{Jac}(n) dn \text{Jac}(n') dn' \text{Jac}(\bar{n}) d\bar{n} da da' \chi_{(\bar{y}, y, t)}^\hbar(\bar{n}naM) \chi_{(\bar{x}, x, s)}^\hbar(\bar{n}n'a'M) |c_\hbar(\lambda)|^{-2} d\lambda |c_\hbar(\lambda')|^{-2} d\lambda' |c_\hbar(\mu)|^{-2} d\mu \chi_{\lambda_0}^\hbar(\mu) \chi_{\mu_0}^\hbar(\mu) |Q'_{\omega_{-T}} \widehat{\circ \gamma_1} u(\lambda, \bar{n}nw_\ell P_0) \overline{Q'_{\omega_{-T}} \widehat{\circ \gamma_1} v(\lambda', \bar{n}n'w_\ell P_0)}|$$

for an arbitrarily large integer \tilde{M} . For any \bar{n}, n, n' , we have

$$\int da da' \chi_{(\bar{y}, y, t)}^\hbar(\bar{n}naM) \chi_{(\bar{x}, x, s)}^\hbar(\bar{n}n'a'M) = \mathcal{O}(\hbar^{2r/2^+}),$$

so the previous bound becomes

$$\frac{\hbar^{\tilde{M}(1-2/2^+)}}{\max(16, \|x-y\|)^{\tilde{M}}} \hbar^J \hbar^{2r/2^+} \int \text{Jac}(n) dn \text{Jac}(n') dn' \text{Jac}(\bar{n}) d\bar{n} |c_{\hbar}(\lambda)|^{-2} d\lambda |c_{\hbar}(\lambda')|^{-2} d\lambda' |c_{\hbar}(\mu)|^{-2} d\mu \\ \chi_{\lambda_0}^{\hbar}(\mu) \chi_{\mu_0}^{\hbar}(\mu) |Q'_{\omega_{-T}} \widehat{\circ} \gamma_1 u(\lambda, \bar{n} n w_{\ell} P_0) Q'_{\omega_{-T}} \widehat{\circ} \gamma_1 v(\lambda', \bar{n} n' w_{\ell} P_0)|$$

Similarly, \tilde{M} integrations by parts in (5.13) with respect to the variable μ allows to gain a factor $\frac{\hbar^{\tilde{M}(1-1/2^+)}}{\|a-a'\|^{\tilde{M}}} \leq \frac{\hbar^{\tilde{M}(1-2/2^+)}}{\|t-s\|^{\tilde{M}}}$ if $\|t-s\|$ is large enough. Integrations by parts with respect to a allow to gain a factor $\frac{\hbar^{\tilde{M}(1-1/2^+)}}{\|\lambda-\mu\|^{\tilde{M}}}$; and integrations by parts with respect to a' allow to gain a factor $\frac{\hbar^{\tilde{M}(1-1/2^+)}}{\|\lambda'-\mu\|^{\tilde{M}}}$. In particular, the contribution to (5.13) of those λ, λ', μ with $\|\lambda' - \mu\| \geq \hbar^{1/2}$ or $\|\lambda - \mu\| \geq \hbar^{1/2}$ is $\mathcal{O}(\hbar^{\infty})$. In these cases the application of the non-stationary phase lemma 5.2 is made simpler by the fact that the phase S is linear in μ, a and a' .

We find that $\langle \mathcal{P}_{m,(\bar{x},x,s,\mu_0)} v, \mathcal{P}_{m,(\bar{y},y,t,\lambda_0)} u \rangle_{\mathbf{Y}}$ is bounded from above by

$$(5.16) \quad \frac{1}{\max(16, \|x-y\|)^{\tilde{M}}} \frac{1}{\max(16, \|t-s\|)^{\tilde{M}}} \hbar^J \hbar^{2r/2^+} \int \text{Jac}(n) dn \text{Jac}(n') dn' \text{Jac}(\bar{n}) d\bar{n} |c_{\hbar}(\lambda)|^{-2} d\lambda |c_{\hbar}(\lambda')|^{-2} d\lambda' |c_{\hbar}(\mu)|^{-2} d\mu \\ \chi_{\lambda_0}^{\hbar}(\mu) \chi_{\mu_0}^{\hbar}(\mu) |Q'_{\omega_{-T}} \widehat{\circ} \gamma_1 u(\lambda, \bar{n} n w_{\ell} P_0) Q'_{\omega_{-T}} \widehat{\circ} \gamma_1 v(\lambda', \bar{n} n' w_{\ell} P_0)|.$$

In this integral, λ', λ, μ are all ϵ -close to Λ_{∞} , and each of them runs over a set of volume $\hbar^{r/2^+}$; \bar{n} runs over a set of measure $\hbar^{J_0/2^+} \prod_{k>j_0} e^{-T\eta m_{\alpha_k} \alpha_k(X_{\mu})}$, n runs over a set of measure $\hbar^{J_0/2^+} \prod_{k>j_0} e^{-T\eta m_{\alpha_k} \alpha_k(X_{\lambda})}$, and n' runs over a set of measure $\hbar^{J_0/2^+} \prod_{k>j_0} e^{-T\eta m_{\alpha_k} \alpha_k(X_{\lambda'})}$. Using Cauchy-Schwarz and the Plancherel formula we find that the integral

$$\int \text{Jac}(n) dn \text{Jac}(n') dn' |c_{\hbar}(\lambda)|^{-2} d\lambda |c_{\hbar}(\lambda')|^{-2} d\lambda' |Q'_{\omega_{-T}} \widehat{\circ} \gamma_1 u(\lambda, \bar{n} n w_{\ell} P_0) Q'_{\omega_{-T}} \widehat{\circ} \gamma_1 v(\lambda', \bar{n} n' w_{\ell} P_0)|$$

is bounded by $\hbar^{-d} \hbar^{J_0/2^+} \hbar^{r/2^+} \prod_{k>j_0} e^{-T\eta m_{\alpha_k} \alpha_k(X_{\Lambda_{\infty}})} \hbar^{-JK\epsilon} \|u\|_{L^2(\mathbf{Y})} \|v\|_{L^2(\mathbf{Y})}$.

The integral $\int \text{Jac}(\bar{n}) d\bar{n} |c_{\hbar}(\mu)|^{-2} d\mu$ adds another factor $\hbar^{-d} \hbar^{J_0/2^+} \hbar^{r/2^+} \prod_{k>j_0} e^{-T\eta m_{\alpha_k} \alpha_k(X_{\Lambda_{\infty}})} \hbar^{-JK\epsilon}$. Overall we find that

$$\|\mathcal{P}_{m,(\bar{x},x,s,\mu_0)}^* \mathcal{P}_{m,(\bar{y},y,t,\lambda_0)}\| \\ \leq \frac{1}{\max(16, \|x-y\|)^{\tilde{M}}} \frac{1}{\max(16, \|t-s\|)^{\tilde{M}}} \hbar^{J+4r/2^+-2d+2J_0/2^+} \prod_{k>j_0} e^{-2T\eta m_{\alpha_k} \alpha_k(X_{\Lambda_{\infty}})} \hbar^{-2JK\epsilon}$$

and vanishes for $\|\bar{x} - \bar{y}\| > 2$ or $\|\mu_0 - \lambda_0\| > 2$.

Choosing \tilde{M} large enough, we can sum over all $(\bar{y}, y, t, \lambda_0)$, and we find

$$\sum_{(\bar{y}, y, t, \lambda_0) \in \mathbb{Z}^{2J_0+2r}} \|\mathcal{P}_{m,(\bar{x}, x, s, \mu_0)}^* \mathcal{P}_{m,(\bar{y}, y, t, \lambda_0)}\|^{1/2} \leq \hbar^{J/2+2r/2^+-d+J_0/2^+} \prod_{k>j_0} e^{-T\eta m_{\alpha_k} \alpha_k(X_{\Lambda_\infty})} \hbar^{-JK\epsilon}.$$

Remembering that $J = d - r$ and that 2^+ could be chosen arbitrarily close to 2, we get

$$\sum_{(\bar{y}, y, t, \lambda_0) \in \mathbb{Z}^{2J_0+2r}} \|\mathcal{P}_{m,(\bar{x}, x, s, \mu_0)}^* \mathcal{P}_{m,(\bar{y}, y, t, \lambda_0)}\|^{1/2} \leq \hbar^{\frac{J_0-J}{2}-c\epsilon} \prod_{k>j_0} e^{-T\eta m_{\alpha_k} \alpha_k(X_{\Lambda_\infty})}$$

with a constant c that depends on \mathcal{K} .

5.7. Norm of $\mathcal{P}_{m,(\bar{x}, x, s, \mu_0)}^* \mathcal{P}_{m,(\bar{y}, y, t, \lambda_0)}$. Using a similar calculation reversing the roles of \bar{N} and N , we get the same bound,

$$\sum_{(\bar{y}, y, t, \lambda_0) \in \mathbb{Z}^{2J_0+2r}} \|\mathcal{P}_{m,(\bar{x}, x, s, \mu_0)} \mathcal{P}_{m,(\bar{y}, y, t, \lambda_0)}^*\|^{1/2} \leq \hbar^{\frac{J_0-J}{2}-c\epsilon} \prod_{k>j_0} e^{-T\eta m_{\alpha_k} \alpha_k(X_{\Lambda_\infty})}.$$

Using the Cotlar-Stein lemma and the fact that the $\alpha(X_{\Lambda_\infty})$ coincide with the Lyapunov exponents $\chi(H)$ on the energy layer $\mathcal{E}_{\Lambda_\infty}$, we get Theorem 1.12.

6. MEASURE RIGIDITY

In this section we prove Theorems 1.8, 1.9 and 1.10. The proofs combine our entropy bounds with the measure classification results of [8, 9] and the orbit classification results of [17, 25] which give information about A -invariant and ergodic measures that have a large entropy.

Proposition 6.1. (*Measure rigidity theory*) *Let G be a split group, and let μ be an ergodic A -invariant measure on $X = \Gamma \backslash G$.*

(1) [8, Lem. 6.2] *there exist constants $s_\alpha(\mu) \in [0, 1]$ associated to the roots $\alpha \in \Delta$, such that for any $a \in A$,*

$$h_{KS}(\mu, a) = \sum_{\alpha \in \Delta} s_\alpha(\mu) (\log \alpha(a))^+.$$

Here $t^+ = \max\{0, t\}$ for $t \in \mathbb{R}$. Furthermore, $s_\alpha(\mu) = 1$ if and only if μ is invariant by the root subgroup U_α .

(2) [8, Prop. 7.1] *Assume that $s_\alpha(\mu), s_\beta(\mu) > 0$ for two roots $\alpha, \beta \in \Delta$ such that $\alpha + \beta \in \Delta$. Then $s_{\alpha+\beta}(\mu) = 1$.*

(3) [8, Thm. 4.1(iv)] *If G is locally isomorphic to SL_n and $s_\alpha(\mu) > 0$ for all α , then μ is G -invariant.*

(4) [9, Cor. 3.4] *In the case $G = SL_n$, we have $s_\alpha(\mu) = s_{-\alpha}(\mu)$ for all roots α .*

We do not know if (4) holds in general.

Now let μ be an A -invariant probability measure with ergodic decomposition $\mu = \int_X \mu_x d\mu(x)$. For each subset $R \subset \Delta$ let X_R be the set of ergodic components μ_x such

that $\{\alpha, s_\alpha(\mu_x) > 0\} = R$. Write $w_R = \mu(X_R)$ and if $w_R > 0$, let $\mu_R = \frac{1}{w_R} \int_{X_R} \mu_x d\mu(x)$, so that $\mu = \sum_R w_R \mu_R$. From Proposition 6.1(1) we have for $a \in A$

$$h_{KS}(\mu_R, a) \leq \sum_{\alpha \in R} (\log \alpha(a))^+,$$

(this is in fact an avatar of the Ruelle-Pesin inequality) and hence

$$h_{KS}(\mu, a) \leq \sum_R w_R \sum_{\alpha \in R} (\log \alpha(a))^+.$$

By Proposition 6.1(2) it is enough to consider those R that are closed under the addition of roots. In the case $G = SL_n$, parts (3) and (4) show, respectively, that it is enough to consider those R which are symmetric and that also $\mu_\Delta = \mu_{Haar}$.

Proposition 6.2. *Let $G = SL_3(\mathbb{R})$, Γ a lattice in G , and μ an A -invariant probability measure on $\Gamma \backslash G$, such that $h_{KS}(\mu, a) \geq \frac{1}{2} h_{KS}(\mu_{Haar}, a)$ for $a = e^X$, $X = \text{diag}(2, -1, -1)$, $\text{diag}(-1, 2, -1)$, and $\text{diag}(-1, -1, 2)$. Then $w_\Delta \geq \frac{1}{4}$, that is, the Haar component has weight at least $\frac{1}{4}$.*

Proof. The possible sets R are $\Delta, \emptyset, \{\alpha, -\alpha\}$. In the case of SL_n the roots are indexed by $\{ij, 1 \leq i, j \leq n, i \neq j\}$: α_{ij} is defined by $\alpha_{ij}(X) = X_{ii} - X_{jj}$. Consider $a = \text{diag}(e^1, e^1, e^{-2})$. Then $h_{KS}(\mu_{Haar}, a) = 6$ (since $s_\alpha = 1$ for all α), $h_{KS}(\mu_\emptyset, a) = 0$, $h_{KS}(\mu_{12}, a) = 0$, $h_{KS}(\mu_{13}, a) \leq 3$, $h_{KS}(\mu_{23}, a) \leq 3$. Thus,

$$(6.1) \quad 3 \leq h_{KS}(\mu, a) \leq 3w_{13} + 3w_{23} + 6w_\Delta.$$

This implies

$$w_\Delta - w_{12} \geq 1 - (w_\Delta + w_{12} + w_{13} + w_{23}) \geq 0.$$

By symmetry it follows that $w_\Delta \geq w_{13}$ and $w_\Delta \geq w_{23}$. Returning to (6.1), it follows that $3 \leq 12w_\Delta$.

In fact, if $h_{KS}(\mu, a) \geq (\frac{1}{3} + \epsilon) h_{KS}(\mu_{Haar}, a)$ for $a = e^X$, $X = \text{diag}(2, -1, -1)$, $\text{diag}(-1, 2, -1)$, or $\text{diag}(-1, -1, 2)$, then $w_\Delta \geq \frac{3}{2}\epsilon$. \square

Putting together Theorem 1.7 and Proposition 6.2 gives Theorem 1.8.

For SL_4 the analogue of Proposition 6.2 is given below. Theorem 1.9 is an immediate corollary.

Proposition 6.3. *Let $G = SL_4(\mathbb{R})$, μ an A -invariant probability measure on $\Gamma \backslash G$, such that $h_{KS}(\mu, a) \geq (\frac{1}{2} + \epsilon) h_{KS}(\mu_{Haar}, a)$ for $a = e^X$, X in the Weyl orbit of $\text{diag}(3, -1, -1, -1)$. Then $w_\Delta \geq 2\epsilon$. If $\epsilon = 0$ and there is no Haar component, then each ergodic component is*

the Haar measure on a closed orbit of the group $\begin{pmatrix} * & * & * & 0 \\ * & * & * & 0 \\ * & * & * & 0 \\ 0 & 0 & 0 & * \end{pmatrix}$ *(or one of its 4 images*

under the Weyl group), and the components invariant by any of these 4 subgroups have total weight $\frac{1}{4}$.

Theorem 1.8 and its analogue for $G = SL_4$ apply to any lattice Γ . On the other hand for $G = SL_n$ with n large some quotients $\Gamma \backslash G$ support ergodic invariant measures of large entropy other than Haar measure, so our entropy bound is not strong enough to obtain a Haar component. However, for some lattices Γ there are further restrictions on the set of ergodic components, so that non-Haar measures have much smaller entropy. This is the case where Γ is a lattice associated to a division algebra.

We give here a quick outline of the construction, referring the reader to [25] and its references (or [18]) for a detailed discussion. Let F be a central simple algebra of degree n over \mathbb{Q} and assume that F splits over \mathbb{R} , that is that $F \otimes_{\mathbb{Q}} \mathbb{R} \simeq M_n(\mathbb{R})$. Next, let $\mathcal{O} \subset F$ be an *order*, that is a subring whose additive group is generated by a basis for F over \mathbb{Q} . Finally, let $\mathcal{O}^1 \subset SL_n(\mathbb{R})$ denote the subgroup of elements of \mathcal{O} with determinant 1 (“reduced norm 1”). Such \mathcal{O}^1 are in fact lattices; any lattice $\Gamma < SL_n(\mathbb{R})$ commensurable with some \mathcal{O}^1 is said to be of *inner type*. We simply say that they are *associated* to the algebra F . Our Theorem 1.7 applies when the lattice is co-compact, which is the case if and only if F is a division algebra.

We shall need the fact that those measure rigidity results of [9] which are stated specifically for $SL_n(\mathbb{Z})$ apply, in fact, to any lattice of inner type, since the proof of Lemma 5.2 of that paper carries over to the more general situation. We give the easy argument here:

Lemma 6.4. *Let $\Gamma < SL_n(\mathbb{R})$ be a lattice of inner type. Then there is no $\gamma \in \Gamma$, diagonalizable in $SL_n(\mathbb{R})$, such that ± 1 are not eigenvalues of γ and all eigenvalues of γ are simple except for precisely one which occurs with multiplicity two.*

Proof. Say that Γ is associated to the central simple algebra F , and let \mathcal{O} be an order in F such that $\Gamma \cap \mathcal{O}^1$ has finite index in Γ .

Assume by contradiction that there exists γ as in the statement, and choose r so that $\gamma^r \in \mathcal{O}$. Since \mathcal{O} is a ring with a finitely generated additive group, the Cayley-Hamilton Theorem shows that γ^r is integral over \mathbb{Z} . It follows that every eigenvalue of γ^r , hence of γ , is an algebraic integer. The fact that $\det(\gamma) = 1$ now shows that the rational eigenvalues of γ must be integral divisors of 1, so by assumption all eigenvalues of γ are irrational. Let $f(x) \in \mathbb{R}[x]$ be the characteristic polynomial of γ , when γ is thought of as an element of $SL_n(\mathbb{R})$. We will show $f(x) \in \mathbb{Q}[x]$. Then the multiplicity the eigenvalues of f would be Galois invariant giving the desired contradiction. For the last claim extend scalars to \mathbb{C} and note that the usual proof that the reduced trace and norm belong to \mathbb{Q} applies to the entire characteristic polynomial. \square

Proposition 6.5. *Let $n \geq 3$ and let t be the largest proper divisor of n . Let $G = SL_n(\mathbb{R})$ and let $\Gamma < G$ be a lattice of inner type. Let μ be an A -invariant probability measure on $\Gamma \backslash G$ such that $h_{KS}(\mu, a) \geq \frac{1}{2} h_{KS}(\mu_{Haar}, a)$ for $a = e^X$, X a Weyl conjugate of $\text{diag}(n - 1, -1, \dots, -1)$. Then $w_{\Delta} \geq \frac{\frac{(n+1)}{2} - t}{n-t} > 0$. In other words, μ must contain an ergodic component proportional to Haar measure.*

Theorem 1.10 follows.

Proof. As above, let μ_x be an ergodic component of μ that has positive entropy with respect to e^X . By [9, Thm. 1.3] (replacing Lemma 5.2 of that paper with Lemma 6.4 above) μ_x must be *algebraic*: there exists a closed subgroup H containing A , and a closed orbit zH in $\Gamma \backslash G$, such that μ_x is the H -invariant measure on zH . By [17] (the arguments are contained in the proof of Lemma 4.1 and Lemma 6.2) and [25] (see Thm 1.2 and §4.2), H must be reductive, and conjugate to the connected component of $GL_k(\mathbb{R})^l \cap SL_n(\mathbb{R})$; where $n = kl$ and $GL_k(\mathbb{R})^l$ denotes the block-diagonal embedding of l copies of $GL_k(\mathbb{R})$ into $GL_n(\mathbb{R})$.

By the discussion following Proposition 6.1 we see that for such lattices Γ the possible sets R are obtained by partitioning n into l subsets B_1, B_2, \dots, B_l of equal size k , and letting

$$R = \{\alpha_{ij}, 1 \leq i, j \leq n, \exists u \text{ such that } i \in B_u \text{ and } j \in B_u\}.$$

Consider $a = \text{diag}(e^{n-1}, e^{-1}, \dots, e^{-1})$. Then $h_{KS}(\mu_{Haar}, a) = n(n-1)$, and for every subset R defined as above by a non-trivial partition, we have $h_{KS}(\mu_R) \leq n(t-1)$. The inequality $h_{KS}(\mu, a) \geq \frac{1}{2}h_{KS}(\mu_{Haar}, a)$ now shows that

$$w_\Delta(n-1) + \sum_{R \neq \Delta} w_R(t-1) \geq \frac{n-1}{2}.$$

In other words we have

$$w_\Delta(n-1) + (1 - w_\Delta)(t-1) \geq \frac{n-1}{2},$$

which is equivalent to the statement of the theorem. \square

REFERENCES

- [1] N. Anantharaman, *Entropy and the localization of eigenfunctions*, Ann. of Math. (2) 168 (2008), no. 2, 435–475.
- [2] N. Anantharaman, *Exponential decay for products of Fourier integral operators*, arXiv:1007.2066 (2010)
- [3] N. Anantharaman, S. Nonnenmacher, *Entropy of semiclassical measures of the Walsh-quantized baker's map*, Ann. H. Poincaré 8, 37–74 (2007)
- [4] N. Anantharaman, H. Koch and S. Nonnenmacher, *Entropy of eigenfunctions*, New Trends of Mathematical Physics, selected contributions of the 15th International Congress on Mathematical Physics, Springer (2009), 1–22.
- [5] J. Bourgain, E. Lindenstrauss, *Entropy of quantum limits*, Comm. Math. Phys. 233, 153–171 (2003).
- [6] U. Bunke, M. Olbrich, *On Quantum Ergodicity for Vector Bundles*, Acta Appl. Math. 90, 19–41 (2006)
- [7] M. Dimassi, J. Sjöstrand, *Spectral asymptotics in the semi-classical limit*, London Mathematical Society Lecture Note Series, 268. Cambridge University Press, Cambridge, 1999.
- [8] M. Einsiedler, A. Katok, *Invariant measures on G/Γ for split simple Lie groups G* . Dedicated to the memory of Jürgen K. Moser. Comm. Pure Appl. Math. 56 (2003), no. 8, 1184–1221.
- [9] M. Einsiedler, A. Katok, E. Lindenstrauss, *Invariant measures and the set of exceptions to Littlewood's conjecture*. Ann. of Math. (2) 164 (2006), no. 2, 513–560.
- [10] R. Godement, *Introduction à la théorie des groupes de Lie*, Tome 2. Publications Mathématiques de l'Université Paris VII 12 Université de Paris VII, U.E.R. de Mathématiques, Paris, 1982.

- [11] S. Helgason, *Groups and geometric analysis. Integral geometry, invariant differential operators, and spherical functions*. Pure and Applied Mathematics, 113. Academic Press Inc., Orlando, 1984.
- [12] S. Helgason, *Geometric analysis on symmetric spaces*. Second edition. Mathematical Surveys and Monographs, 39. American Mathematical Society, Providence, RI, 2008. .
- [13] J. Hilgert *An ergodic Arnold-Liouville theorem for locally symmetric spaces*. Twenty years of Białowieża: a mathematical anthology, 163–184, World Sci. Monogr. Ser. Math., 8, World Sci. Publ., Hackensack, NJ, 2005.
- [14] A. Knapp, *Lie groups beyond an introduction*. Progress in Mathematics, 140. Birkhäuser Boston, Inc., Boston, MA, 1996.
- [15] A. Knapp, *Representation theory of semisimple groups*. An overview based on examples. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1986.
- [16] E. Lindenstrauss, *Invariant measures and arithmetic quantum unique ergodicity*, Annals of Math. **163**, 165–219 (2006)
- [17] E. Lindenstrauss, B. Weiss, *On sets invariant under the action of the diagonal group*. Ergodic Theory Dynam. Systems 21 (2001), no. 5, 1481–1500.
- [18] V. Platonov and A. Rapinchuk, *Algebraic groups and number theory*. Translated from the 1991 Russian original by Rachel Rowen. Pure and Applied Mathematics, 139. Academic Press, Inc., Boston, MA, 1994. xii+614 pp.
- [19] G. Rivière, *Entropy of semiclassical measures in dimension 2*, to appear in Duke Math. J.
- [20] G. Rivière, *Entropy of semiclassical measures for surfaces of nonpositive curvature*, preprint (2009).
- [21] Z. Rudnick and P. Sarnak, *The behaviour of eigenstates of arithmetic hyperbolic manifolds*, Commun. Math. Phys. **161**, 195–213 (1994).
- [22] M. Schröder, Patterson-Sullivan distributions for symmetric spaces of the noncompact type, PhD dissertation, Universität Paderborn, 2010.
- [23] L. Silberman, A. Venkatesh, *On quantum unique ergodicity for locally symmetric spaces*, Geom. Funct. Anal. **17** no. 3, 960–998 (2007).
- [24] L. Silberman, A. Venkatesh, *Entropy bounds for Hecke eigenfunctions on division algebras*, to appear in Geom. Funct. Anal.
- [25] G. Tomanov, *Actions of maximal tori on homogeneous spaces*. Rigidity in dynamics and geometry (Cambridge, 2000), 407–424, Springer, Berlin, 2002.
- [26] S.A. Wolpert, *The modulus of continuity for $\Gamma_0(m)/\mathbb{H}$ semi-classical limits*, Commun. Math. Phys. **216**, 313–323 (2001)
- [27] S. Zelditch, *Pseudodifferential analysis on hyperbolic surfaces*, J. Funct. Anal. **68**, no. 1, 72105 (1986).

LABORATOIRE DE MATHÉMATIQUE, UNIVERSITÉ D'ORSAY PARIS XI, 91405 ORSAY CEDEX, FRANCE
E-mail address: Nalini.Anantharaman@math.u-psud.fr

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER BC V6T 1Z2,
 CANADA
E-mail address: lior@math.ubc.ca