Reprinted from:

Computation and Control II (Proceedings, Bozeman, 1990, K. L. Bowers & J. Lund eds.) Progress in Systems and Control Theory, vol. 11, pp. 325-350, Birkhäuser, 1991

ACUITY OF OBSERVATION FOR INVARIANT EVOLUTION EQUATIONS

DOROTHY I. WALLACE AND JOSEPH A. WOLF

Mathematics and Computer Science
Dartmouth College
Hanover, New Hampshire 03755

Department of Mathematics
University of California
Berkeley, California 94720

Abstract

In a recent paper [4] we studied discrete observability for invariant evolution equations on compact homogeneous spaces, e.g. for the heat equation on the sphere. The observations there were given by simultaneous measurements, corresponding to function evaluations. The initial data was observed as a limit of truncations deduced from a finite number of measurements. That procedure naturally involves two types of errors. First, observations *qua* evaluation functionals are restricted to a finite part of the Fourier Peter Weyl expansion; that restriction implicitly involves a convolution. See (1.4) and (1.6) below. We think of the resulting error as the error in the head of the approximation. Second, the actual initial data minus the truncations are the usual type of error terms; we think of them as the error in the tail of the approximation. In this paper we show that the error in the head depends linearly on the error in the tail. We then investigate the extent to which smoothness of the initial data function controls the tail error through a set of Sobolev inequalities. We also investigate consequences of polynomial spectral growth conditions on the rate of vanishing of the tail error. Finally, we specialize these results to riemannian symmetric spaces.

Research partially supported by N.S.A. Grant MDA 904 89 H 2021.
Research partially supported by N.S.F. Grant DMS 88 05816.

1 Introduction

We recently [4] studied discrete observability for evolution equations

$$D_x f(x:t) + \frac{\partial}{\partial t} f(x:t) = 0 \text{ and } f(x:0) = b(x) \tag{1.1}$$

on compact homogeneous spaces $X = G/K$, where D is a closed densely defined G-invariant operator of $L^2(X)$. The initial data is given by the function $b \in L^2(X)$. Our result (see Theorem 2.14 below for a precise statement) was that the evolution equation is discretely observable at any time t_0 and near any point $x_0 \in X$. In other words, given a neighborhood U of x_0 that meets every component of X, there is a sequence of locations $\{x_1, x_2, \cdots, x_n \cdots\} \subset U$ such that the solution matching the observed values

$$\{f(x_1:t_0),\ f(x_2:t_0),\ \cdots,\ f(x_n:t_0),\ \cdots\} \tag{1.2}$$

is unique.

In practice one proceeds by decomposing the set of all representations of G as an increasing union of finite subsets. This gives us $L^2(X)$ as a G-invariant increasing union of finite dimensional invariant subspaces $E_r(X)$, and we use the first $n_r = dim\ E_r(X)$ observations

$$\{f(x_1:t_0),\ f(x_2:t_0),\ \cdots,\ f(x_{n_r}:t_0)\} \tag{1.3}$$

as follows. For each t we have the orthogonal projection $f_r(\cdot:t)$ of $f(\cdot:t)$ to $E_r(X)$. It is determined by $f_r(\cdot:t_0)$, and $f_r(\cdot:t_0)$ is uniquely defined by the values $f(x_k:t_0)$, $1 \leqq k \leqq n_r$, of (1.3). This determination can be viewed as solving a linear system defined by restriction of the evaluation functionals $\psi_k(\phi) = \phi(x_k)$ from $L^2(X)$ to $E_r(X)$. That restriction is convolution with the sum of the characters of the irreducible summands (ignoring multiplicity) of the representation of G on $E_r(X)$. We do *not* want to have to compute this convolution. So instead we use the approximations $\widetilde{f}_r$ to f_r, $\widetilde{f}_r(\cdot:t) \in E_r(X)$ for all t, defined by the evaluation functionals without any restriction,

$$\widetilde{f}_r(x_k:t_0) = f(x_k:t_0) \text{ for } 1 \leqq k \leqq n_r\ . \tag{1.4}$$

Now we have two approximation errors, for the total approximation error

$$f(x:t) - \widetilde{f}_r(x:t) \tag{1.5}$$

splits as the sum of the error in the head of the approximation, the "head error",

$$f_r(x:t) - \widetilde{f}_r(x:t) \tag{1.6}$$

and the more standard "tail error"

$$f(x:t) - f_r(x:t). \tag{1.7}$$

In this paper we show how the head error depends linearly on the tail error, and we discuss the decay rates of those errors.

2 Invariant Evolution Equations

In order to be more specific we must recall some of the general results of [4]. Let X be a homogeneous space G/K where G is a compact Lie group, and let

$$D : L^2(X) \to L^2(X) \tag{2.1}$$

be a closed densely defined operator that commutes with the action

$$[L(g)f](x) = f(g^{-1}x). \tag{2.2}$$

The action (2.2) is the left regular representation G on $L^2(X)$. The corresponding evolution equation (analog of the heat equation) for initial data $b(x)$ is

$$D_x f(x:t) + \frac{\partial}{\partial t} f(x:t) = 0 \text{ and } f(x:0) = b(x) \tag{2.3}$$

on $X \times \mathbb{R}$. The usual heat equation is the case where D is the Laplace-Beltrami operator for a G-invariant riemannian metric on X.

In this context, observability is the study of just which types of data on (samples of) the values $f(x:t)$ allow us to reconstruct the function $b(x)$ accurately. Here, as in [4], we consider a restricted version of this question. Suppose that we have a sequence of points $\{x_1, x_2, \cdots\} \subset X$ and we are allowed to sample (observe) the $f(x_i:t)$ at some time t_0. In [4] we saw just when it is possible to deduce $b(x) = f(x:0)$ for all $x \in X$ from this data. In particular we imposed conditions on the x_i that make this deduction possible. The conditions were such that, given $x_0 \in X$ and a neighborhood of x_0 in X, the sequence $\{x_i\}$ will be contained in the neighborhood. We need a little more structure theory in order to describe those conditions.

The following, well known in the case of the heat equation, is one of the basic results of [4].

Proposition 2.4. *If D is a normal operator on $L^2(X)$, then there is a complete orthonormal set $\{\phi_j\}$ in $L^2(X)$ of eigenfunctions of D. If $\{\phi_j\}$ is any such orthonormal set, $D\phi_j = \lambda_j\phi_j$, then the $L^2(X)$ solutions to (2.3) are just the functions of the form*

$$f(x:t) = \sum_j a_j e^{-t\lambda_j}\phi_j(x), \quad a_j \in \mathbb{C} \tag{2.5}$$

for $x \in X$ and for $t \in \mathbb{R}$ in the range such that $\sum_j |a_j e^{-t\lambda_j}|^2 < \infty$.

This is well known in the case of the heat equation. The idea of our proof is to use G-invariance and the compactness implicit in the Peter-Weyl Theorem to replace compactness in the argument which shows that the Laplacian Δ has discrete spectrum. The Peter-Weyl Theorem gives a decomposition

$$L^2(G) = \sum_{\pi\in\widehat{G}} V_\pi \otimes V_\pi^* \text{ and } L^2(X) = \sum_{\pi\in\widehat{G}} V_\pi \otimes (V_\pi^*)^K . \tag{2.6a}$$

$\widehat{G}$ is the set of (equivalence classes of) irreducible unitary representations of G and V_π is the (finite dimensional) vector space on which G is represented by π. We identify $V_\pi \otimes V_\pi^*$ with the complex span of the matrix coefficient functions for π,

$$v \otimes w^* \text{ corresponds to the function } x \mapsto \langle v\ ,\ \pi(x)w\rangle$$

where $w^* \in V^*$ corresponds to inner product with $w \in V$. Thus $V_\pi \otimes (V_\pi^*)^K$ is the subspace consisting of functions $f \in V_\pi \otimes V_\pi^*$ such that $f(gk) = f(g)$ for all $g \in G$ and $k \in K$. Those are the functions in $V_\pi \otimes V_\pi^*$ that can be (and will be) viewed as functions on G/K.

In the left regular representation (2.2) of G, the G-module structure implicit in (2.6a) is

$$L^2(G) = \sum_{\pi\in\widehat{G}} deg(\pi)V_\pi \text{ and } L^2(X) = \sum_{\pi\in\widehat{G}} mult(1_K, \pi|_K)V_\pi \tag{2.6b}$$

where $deg(\pi)$ is the degree of the representation π and $mult(1_K, \pi|_K)$ is the multiplicity of the trivial representation 1_K in the restriction $\pi|_K$.

The key observation in the proof of (2.4) is that linear operator D and its adjoint D^* preserve each of the finite dimensional summands

$$A(\pi) = V_\pi \otimes (V_\pi^*)^K \subset L^2(X). \tag{2.7}$$

Express $\widehat{G}$ as an increasing union of finite subsets $\widehat{G}_r$:

$$\widehat{G}_r = \{\pi_\nu \in \widehat{G} \mid \|\nu\| < r\} \tag{2.8}$$

where $\pi_\nu \in \widehat{G}$ has highest weight ν. This filtration of $\widehat{G}$ defines a filtration of $L^2(X)$, expressing it as an increasing union of finite dimensional subspaces

$$L^2(X) = \bigcup_{r>0} E_r(X) \tag{2.9a}$$

where

$$E_r(X) = \sum_{\|\nu\|<r} V_{\pi_\nu} \otimes (V_{\pi_\nu}^*)^K = \sum_{\|\nu\|<r} mult(1_K, \pi_\nu|_K)V_\pi . \tag{2.9b}$$

The complete orthonormal set $\{\phi_j\}$ in $L^2(X)$ of eigenfunctions of D in Proposition 2.4 can be constructed as an increasing union of orthonormal bases of the finite dimensional subspaces $E_r(X) \subset L^2(X)$ of (2.9). So we can fix one such complete orthonormal set $\Phi = \bigcup_{r>0} \Phi_r$ where

$$\Phi_r = \{\phi_1, \cdots, \phi_{n_r}\} \text{ is an orthonormal basis of } E_r(X) \text{ with } D\phi_j = \lambda_j\phi_j. \tag{2.10}$$

Definition 2.11. *The evolution equation (2.3) is* **discretely observable at** $x_0 \in X$ *if, for every time t_0 and every neighborhood V of x_0 in X, there is a countable subset $\{x_1, x_2, \cdots\} \subset V$ with the following property. If f is the solution to the heat equation (2.3) for initial data $b \in L^2(S^n)$, and if $f_r(\cdot : t_0)$ denotes the orthogonal projection of $f(\cdot : t_0)$ to $E_r(X)$, then the $f_r(x_i : t_0)$, $1 \leqq i \leqq n_r$, determine f_r.*

The finite dimensional spaces $E_r(X)$ are discretely observable in the sense of [7] because the action of G on $E_r(X)$ is a subrepresentation of the left regular representation of G on $L^2(G)$. More precisely, in [4] we prove

Proposition 2.12. *Fix $x_0 \in X$. Then G has a countable subset $S = \bigcup_{r>0} S_r$ with $S_r = \{s_1, \cdots, s_{n_r}\}$ such that the function evaluations $\psi_j : \phi \mapsto \phi(s_j^{-1}x_0)$, $1 \leqq j \leqq n_r$, form a basis of the linear dual space of $E_r(X)$. If U is an open subset of G that meets every connected component, then we can find $S \subset U$.*

To prove Proposition 2.12, one views the point evaluations $\psi_{(x)}(\phi) = \phi(x)$ as linear functionals on $C^\infty(X)$ and argues that $\{\psi_{(u^{-1}x_0)}|_{E_r(X)} \mid u \in U\}$ spans the dual space of $E_r(X)$. That gives

$$S_r = \{s_1, \cdots, s_{n_r}\} \subset U \tag{2.13a}$$

such that

$$\text{the } \psi_{(s_j^{-1}x_0)}|_{E_r(X)}, 1 \leqq j \leqq n_r, \text{ form a basis of } E_r(X). \tag{2.13b}$$

Then if $r < v$ we choose $S'_{r,v} = \{s_{n_r+1}, \cdots, s_{n_v}\} \subset U$ so that

$$\{\psi_{(s_j^{-1}x_0)}|_{E_r(X)^\perp \cap E_v(X)}\}_{n_r<j\leqq n_v} \text{ is a basis of } E_r(X)^\perp \cap E_v(X) \quad (2.13c)$$

and we take $S_v = S_r \cup S'_{r,v}$.

Propositions 2.4 and 2.12 combine with (2.9) and (2.10) to give a particular type of observability for the evolution equation (2.3).

Theorem 2.14. *Let X be a homogeneous space G/K where G is a compact Lie group. Let D be a closed densely defined G-invariant normal operator on $L^2(X)$. Choose a complete orthonormal set $\Phi = \bigcup_{r>0} \Phi_r$ in $L^2(X)$ as in (2.10). Then the $L^2(X)$ solutions to the evolution equation (2.3) are just the functions*

$$f(x:t) = lim_{r\to\infty} \sum_{j=1}^{n_r} a_j e^{-t\lambda_j}\phi_j(x) \quad (2.15)$$

for $x \in X$ and for $t \in \mathbb{R}$ in the range such that $\sum_j |a_j e^{-t\lambda_j}|^2 < \infty$. The solution $f(x:t)$ of the evolution equation is discretely observable at any time t_0 on any open subset of X that meets every topological component. In other words, let U be an open subset of G that meets every topological component. Then there is an increasing union $S = \bigcup_{r>0} S_r \subset G$ with each $S = \{s_1, \cdots, s_{n_r}\} \subset U$ such that each partial sum

$$f_r(x:t) = \sum_{j=1}^{n_r} a_j e^{-t\lambda_j}\phi_j(x) \quad (2.16)$$

is determined by the "observations" $f_r(s_j^{-1}x_0 : t_0)$, $1 \leqq j \leqq n_r$.

The point of this paper is to work out estimates for the head and tail errors at each finite stage in the observability process of Theorem 2.14.

3 The Head-Tail Estimate

We now derive the head-tail estimate mentioned at the end of §1, comparing the errors (1.6) and (1.7).

Define $d_k = f(x_k : t_0)$ for $1 \leqq k < \infty$. Then we have

$$d_k = f(x_k:t_0) = \sum_{j=1}^{\infty} a_j e^{-t_0\lambda_j}\phi_j(x_k) = {}^tC \cdot V(x_k:t_0) \quad (3.1)$$

where

$$C = \begin{pmatrix} 1 \\ 1 \\ \vdots \end{pmatrix} \quad \text{and} \quad V(x:t) = \begin{pmatrix} a_1 e^{-t\lambda_1}\phi_1(x) \\ a_2 e^{-t\lambda_2}\phi_2(x) \\ \vdots \end{pmatrix} \quad (3.2)$$

are infinite column vectors. Note that $x_k = s_k^{-1}x_0$ implies $\phi_j(x_k) = [L(s_k)\phi_j](x_0)$, so

$$V(x_k:t) = L(s_k) \cdot V(x_0:t). \quad (3.3)$$

Define

$$D = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} {}^tC \cdot L(s_1) \\ {}^tC \cdot L(s_2) \\ \vdots \end{pmatrix}. \quad (3.4)$$

Then (3.1), (3.3) and (3.4) combine as

$$D = M \cdot V(x_0:t_0). \quad (3.5)$$

In other words, we "solve" the system (3.1) in the form $V(x_0 : t_0) = M^{-1} \cdot D$.

The approximate truncation $\widetilde{f}_r$ of (1.4) is similarly determined by the $d_k = \widetilde{f}_r(x_k : t_0)$, this time for $1 \leqq k \leqq n_r$,

$$d_k = \widetilde{f}_r(x_k:t_0) = \sum_{j=1}^{n_r} \widetilde{a}_j e^{-t_0\lambda_j}\phi_j(x_k) = {}^tC_r \cdot \widetilde{V}_r(x_k:t_0) \quad (3.6)$$

where

$$C_r = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad \text{and} \quad \widetilde{V}_r(x:t) = \begin{pmatrix} \widetilde{a}_1 e^{-t\lambda_1}\phi_1(x) \\ \widetilde{a}_2 e^{-t\lambda_2}\phi_2(x) \\ \vdots \\ \widetilde{a}_{n_r} e^{-t\lambda_{n_r}}\phi_{n_r}(x) \end{pmatrix} \quad (3.7)$$

are $n_r \times 1$ column vectors. As before,

$$\widetilde{V}_r(x_k:t) = L_r(s_k) \cdot \widetilde{V}_r(x_0:t). \quad (3.8)$$

where $L_r(s) = L(s)|_{E_r(X)}$. Let

$$D_r = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n_r} \end{pmatrix} \quad \text{and} \quad M_r = \begin{pmatrix} {}^tC \cdot L(s_1) \\ {}^tC \cdot L(s_2) \\ \vdots \\ {}^tC \cdot L(s_{n_r}) \end{pmatrix}. \quad (3.9)$$

Then (3.6), (3.8) and (3.9) combine as

$$D_r = M_r \cdot \widetilde{V}_r(x_0 : t_0). \tag{3.10}$$

So, similarly, we "solve" the system (3.6) in the form $\widetilde{V}_r(x_0 : t_0) = M_r^{-1} \cdot D_r$.

Now let us compare (3.5) and (3.10). Separating off the part that corresponds to $E_r(X)$ we split

$$D = \begin{pmatrix} D_r \\ D_r' \end{pmatrix}, \quad C = \begin{pmatrix} C_r \\ C_r' \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} V_r \\ V_r' \end{pmatrix} \tag{3.11}$$

into pieces of sizes $n_r \times 1$ and $\infty \times 1$, and

$$L(s) = \begin{pmatrix} L_r(s) & 0 \\ 0 & L_r'(s) \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} M_r & N_r \\ M_r' & N_r' \end{pmatrix}. \tag{3.12}$$

Then (3.5) and (3.10) say

$$M_r \cdot \widetilde{V}_r(x_0 : t_0) = D_r = M_r \cdot V_r(x_0 : t_0) + N_r \cdot V_r'(x_0 : t_0). \tag{3.13}$$

In other words,

$$M_r \cdot \{V_r(x_0 : t_0) - \widetilde{V}_r(x_0 : t_0)\} + N_r \cdot V_r'(x_0 : t_0) = 0. \tag{3.14}$$

That gives the head-tail estimate

Proposition 3.15. *The head error $V_r(x_0 : t_0) - \widetilde{V}_r(x_0 : t_0)$ and the tail error $V_r'(x_0 : t_0)$ are related by*

$$V_r(x_0 : t_0) - \widetilde{V}_r(x_0 : t_0) = M_r^{-1} \cdot N_r \cdot V_r'(x_0 : t_0). \tag{3.16}$$

4 Bounds Derived From Unitarity of the Regular Representation

In order to use the head-tail estimate (3.16) one needs decay information on the matrices M_r^{-1} and N_r. In this section we describe that part of the needed information that comes out of unitarity of the regular representation of G on $L^2(X)$.

We start by enumerating the irreducible constituents of the regular representation,

$$L(s) = \begin{pmatrix} \pi_{\nu_1}(s) & & \\ & \pi_{\nu_2}(s) & \\ & & \ddots \end{pmatrix} \tag{4.1}$$

where π_ν denotes the irreducible representation of highest weight ν, where for each i, ν_i occurs at most $deg\ \pi_{\nu_i}$ times, and where $||\nu_1|| \leqq ||\nu_2|| \leqq \cdots$. The rows of M are infinite row vectors,

$${}^tC \cdot L(s_k) = (\, m_1(s_k) \quad m_2(s_k) \quad \cdots) \tag{4.2}$$

where $m_i(s)$ is a row vector of length $d(\nu_i) = deg\ \pi_{\nu_i}$,

$$m_i(s) = (\, m_{i,1}(s) \quad m_{i,2}(s) \quad \cdots \quad m_{i,d(\nu_i)}(s)\,). \tag{4.3}$$

In view of (4.1) we can express

$$m_i(s) = {}^tc_i \cdot \pi_{\nu_1}(s) \text{ where } c_i = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \text{ is } d(\nu_i) \times 1. \tag{4.4}$$

So $m_i(s)$ is the coordinate expression of ${}^tc_i$ in the orthonormal frame composed of the rows of the unitary matrix $\pi_{\nu_i}(s)$. That says

$$||m_i(s)||^2 = \sum_{j=1}^{d(\nu_i)} |m_{i,j}(s)|^2 = d(\nu_i). \tag{4.5}$$

We view (4.5) as a bound on the growth of any row ${}^tC \cdot L(s_k)$ of M — and thus also of its submatrices M_r and N_r — in terms of the degree polynomial $d(\cdot)$. Lemma 4.10 below uses this bound to give us an estimate on the column vectors

$$N_r \cdot V_r'(x_0 : t) = \begin{pmatrix} \sum_{||\nu_i|| \geqq r} \sum_{1 \leqq j \leqq d(\nu_i)} m_{i,j}(s_1) a_{i,j} e^{-t\lambda_{i,j}} \phi_{i,j}(x_0) \\ \sum_{||\nu_i|| \geqq r} \sum_{1 \leqq j \leqq d(\nu_i)} m_{i,j}(s_2) a_{i,j} e^{-t\lambda_{i,j}} \phi_{i,j}(x_0) \\ \vdots \\ \sum_{||\nu_i|| \geqq r} \sum_{1 \leqq j \leqq d(\nu_i)} m_{i,j}(s_{n_r}) a_{i,j} e^{-t\lambda_{i,j}} \phi_{i,j}(x_0) \end{pmatrix}. \tag{4.6}$$

Here we converted to multi-indices following the pattern of the $m_{i,j}$.

The Schur orthogonality relations give us $|\phi_{i,j}(x)| \leqq d(\nu_i)$; compare [4, 3.8]. Combine this with (4.5) for

$$\sum_{1 \leqq j \leqq d(\nu_i)} |m_{i,j}(s) \phi_{i,j}(x)|^2 \leqq d(\nu_i)^3. \tag{4.7}$$

We have the initial condition function $b = \sum b_\nu \in L^2(X)$ with $b_\nu \in A(\pi_\nu)$. Write

$$\lambda(\nu_i) = min_{1\leqq j\leqq d(\nu_i)}\mathrm{Re}\ \lambda_{i,j} \quad (4.8a)$$

so that

$$\sum\nolimits_{1\leqq j\leqq d(\nu_i)}|a_{i,j}e^{-t\lambda_{i,j}}|^2 \leqq ||b_{\nu_i}||^2 e^{-t\lambda(\nu_i)}. \quad (4.8b)$$

Combine (4.7) and (4.8) for

$$\sum\nolimits_{1\leqq j\leqq d(\nu_i)}|m_{i,j}(s_k)a_{i,j}e^{-t\lambda_{i,j}}\phi_{i,j}(x_0)|^2 \leqq d(\nu_i)^3||b_{\nu_i}||^2 e^{-t\lambda(\nu_i)}. \quad (4.9)$$

We summarize as follows.

Proposition 4.10. *The typical entry in the column vector $N_r \cdot V_r'(x_0 : t)$ of (4.6) has upper bound*

$$\sum_{||\nu_i||\geqq r}\ \sum_{1\leqq j\leqq d(\nu_i)} |m_{i,j}(s_k)a_{i,j}e^{-t\lambda_{i,j}}\phi_{i,j}(x_0)|^2 \leqq \sum_{||\nu_i||\geqq r} d(\nu_i)^3||b_{\nu_i}||^2 e^{-t\lambda(\nu_i)}. \quad (4.11)$$

Propositions 3.15 and 4.10 show just how the rate of decay of the head error depends on the decay rate of the $||b_{\nu_i}||$, the growth rate of the $\lambda(\nu_i)$, and the lowest eigenvalue of M_r as a function of r.

5 Example: The Heat Equation on the Sphere

The heat equation on the sphere $S^n = SO(n+1)/SO(n)$, for initial data $b(x)$, is given by

$$\Delta_x f(x:t) + \frac{\partial}{\partial t}f(x:t) = 0 \text{ and } f(x:0) = b(x), \quad x \in S^n \text{ and } t \geqq 0 \quad (5.1)$$

where Δ is the (positive) Laplace-Beltrami operator[1] on the sphere. The function $f(x:t)$ represents temperature distribution at time t on S^n evolving from temperature distribution $b(x)$ at time 0.

[1]We use the sign of the Laplace-Beltrami operator corresponding to the Laplacian $\Delta = -\sum \frac{\partial^2}{\partial x_i^2}$ on euclidean space, because this both usual and natural in differential geometry and in group representation theory.

Suppose $n > 1$. Then the negative of the Killing form of $SO(n+1)$ induces a riemannian metric of constant positive curvature on S^n. This is the multiple of the standard curvature $+1$ metric for which Δ has eigenvalue $||\nu+\rho||^2 - ||\rho||^2$ on $A(\pi_\nu)$, as follows. Setting aside those ν for which $A(\pi_\nu) = 0$ we are left with a 1-parameter family of highest weights ν_h, $h \geqq 0$, for $SO(n+1)$. This one parameter family satisfies $mult(1_{SO(n)}, \pi_\nu|_{SO(n)}) = 1$. In other words[2] each π_ν occurs exactly once on $L^2(S^n)$. It follows that Δ has spectrum

eigenvalues	multiplicities
$\lambda_h = \|\|\nu_h + \rho\|\|^2 - \|\|\rho\|\|^2$	$d_h = dimA(\pi_{\nu_h})$
$\frac{(n-1)h+h^2}{2n-2}$	$\frac{n-1+2h}{n-1}\prod_{k=1}^{n-2}\frac{k+h}{k}$

(5.2)

for $h \geqq 0$. See [1] or compare [4].

Here $\nu_h = h\nu_1$ and $||\nu_1||^2 = \frac{1}{2n-2}$, so $||\nu_h||^2 = \frac{h^2}{2n-2}$. Thus $||\nu_h|| \geqq r$ as in (4.11) if and only if $h \geqq r\sqrt{2n-2}$. So we denote

$$h(r) = r\sqrt{2n-2} \text{ for } r \geqq 0 \quad (5.3)$$

in order that summation over the range $||\nu_h|| \geqq r$ be the same as summation over $h \geqq h(r)$.

The initial data function $b \in L^2(S^n)$, so $\sum_{h=0}^{\infty}||b_{\nu_h}||^2 < \infty$. This forces $||b_{\nu_h}||^2 \to 0$, that is[3], $||b_{\nu_h}||^2 = o(1)$ as $h \to \infty$. Along with (5.2), now Proposition 4.10 reduces in the present context to

Lemma 5.4. *In the case of the heat equation (5.1) on the sphere S^n, $n > 1$, the typical entry of the column vector $N_r \cdot V_r'(x_0 : t)$ of (4.6) satisfies*

$$\sum_{h\geqq h(r)}\ \sum_{1\leqq j\leqq d_h} |m_{h,j}(s_k)a_{h,j}e^{-t\lambda_h}\phi_{h,j}(x_0)|^2 = o\left(\sum_{h\geqq h(r)}\left\{d_h^3 \exp\left(-t\frac{(n-1)h+h^2}{2n-2}\right)\right\}\right) \quad (5.5)$$

[2]The connection is given by the Frobenius Reciprocity Theorem.

[3]We use the standard definition: $p = o(q)$ as $s \to s_0$ if, for every $\epsilon > 0$, there is a neighborhood U of s_0 such that $|p(s)| < \epsilon q(s)$ for $s \in U \setminus \{s_0\}$.

as $r \to \infty$, where $d_h = \left(\frac{n-1+2h}{n-1} \prod_{k=1}^{n-2} \frac{k+h}{k}\right)$.

The column vector $N_r \cdot V_r'(x_0 : t)$ of (4.6) has height $n_r = d_0 + d_1 + \cdots + d_{h(r)-1}$. Since $h(r)$ is a multiple of r and d_h is polynomial of degree $n-1$ in h this says that

$$n_r \text{ is a polynomial of degree } n \text{ as a function of } r. \tag{5.6}$$

Now combine Proposition 3.15, (5.3), Lemma 5.4 and (5.6) to see the first assertion (5.8) of

Proposition 5.7. *In the case of the heat equation* (5.1) *on the sphere* $S^n, n > 1$, *the head error*

$$|V_r(x_0 : t) - \widetilde{V}_r(x_0 : t)| = o\left(||M_r^{-1}|| r^n \sum_{h \geqq h(r)} \left\{ d_h^3 \exp\left(-t\frac{(n-1)h + h^2}{2n-2}\right)\right\}\right) \tag{5.8}$$

as $r \to \infty$, *where* $d_h = \left(\frac{n-1+2h}{n-1} \prod_{k=1}^{n-2} \frac{k+h}{k}\right)$. *Fix* $t > 0$. *Then the* $h(r)$ *summand eventually dominates,*

$$|V_r(x_0 : t) - \widetilde{V}_r(x_0 : t)| = o\left(||M_r^{-1}|| \cdot r^{4n-3} \cdot e^{-t(r^2 + r\sqrt{(n-1)/2})}\right) \tag{5.9}$$

as $r \to \infty$. *In particular, if*

$$\limsup_{r \to \infty} ||M_r^{-1}|| \cdot r^{4n-3} \cdot e^{-t(r^2 + r\sqrt{(n-1)/2})} < \infty \tag{5.10a}$$

then

$$\lim_{r \to \infty} |V_r(x_0 : t) - \widetilde{V}_r(x_0 : t)| = 0. \tag{5.10b}$$

Proof. We noted that (5.8) follows from Proposition 3.15, (5.3), Lemma 5.4 and (5.6). Evidently, (5.9) and (5.10a) imply (5.10b). So we need only check that (5.8) implies (5.9).

Fix $t > 0$. Notice $\lim_{h \to \infty} d_{h+1}/d_h = 1$. Also note that, for $h \gg 0$,

$$\frac{\exp\left(-t\frac{(n-1)(h+1)+(h+1)^2}{2n-2}\right)}{\exp\left(-t\frac{(n-1)h+h^2}{2n-2}\right)} = \exp\left(-\frac{t}{2}\right) \exp\left(-t\frac{2h+1}{2n-2}\right).$$

Now for $r \gg 0$ we have

$$\left(\frac{n-1+2h(r)}{n-1} \prod_{k=1}^{n-2} \frac{k+h(r)}{k}\right)^3 \exp\left(-t\frac{(n-1)h(r) + h(r)^2}{2n-2}\right)$$
$$\geqq \sum_{h > h(r)} \left\{\left(\frac{n-1+2h}{n-1} \prod_{k=1}^{n-2} \frac{k+h}{k}\right)^3 \exp\left(-t\frac{(n-1)h+h^2}{2n-2}\right)\right\}.$$

Thus (5.9) follows from (5.8). □

A glance at [1] will convince the reader that the story is essentially the same for the heat equation on any compact symmetric space of rank 1. In fact it is essentially the same for the heat equation on any symmetric space of compact type. In this connection see §9 below.

6 General Bounds on the Head Error

In this section we work out the general results that correspond to the specific results of §5. Our main result is Theorem 6.16, which exhibits the delicate interplay between the spectral properties of D and the M_r and smoothness properties of the initial data function $b \in L^2(X)$.

We will need to apply the Sobolev Inequalities to the initial data function $b(\cdot) = f(\cdot : 0)$ in order to control decay of the norms $||b_{\nu_h}||$ and the terms $d(\nu_h)^3 ||b_{\nu_h}||^2$ that occur (with i instead of h) in Proposition 4.10. This enhances the role of the term $h^{-\epsilon}$ that occurs in (5.5) and (5.8), and the term $r^{-\epsilon}$ in (5.9) and (5.10). In Theorem 6.16 this shows how increased smoothness for b implies faster convergence for certain bounds on the head error $|V_r(x_0 : t) - \widetilde{V}_r(x_0 : t)|$, allowing us the describe some general conditions under which $\lim_{r \to \infty} |V_r(x_0 : t) - \widetilde{V}_r(x_0 : t)| = 0$.

The Sobolev Inequalities are essentially the same for homogeneous spaces $X = G/K$ where G is a compact Lie group, as for euclidean spaces. We recall the basic facts from Wallach's book [5, §5.7]. Let

$$\phi \in L^2(X) = \sum_{\pi_\nu \in \widehat{G}} V_{\pi_\nu} \otimes (V_{\pi_\nu}^*)^K = \sum_{\pi_\nu \in \widehat{G}} A(\pi_\nu) \tag{6.1}$$

and decompose ϕ as the sum of its components,

$$\phi = \sum_{\pi_\nu \in \widehat{G}} \phi_\nu \quad \text{with} \quad \phi_\nu \in V_{\pi_\nu} \otimes (V_{\pi_\nu}^*)^K = A(\pi_\nu)\,. \tag{6.2}$$

For each real number $s \geqq 0$, the s^{th} *Sobolev norm* is given by

$$\|\phi\|_s^2 = \sum_{\pi_\nu \in \widehat{G}} (1 + \|\nu\|^2)^s \|\phi_\nu\|_{L^2(X)}^2 \tag{6.3}$$

and the s^{th} *Sobolev space* is defined to be

$$H^s(X) = \{\phi \in L^2(X) \mid \|\phi\|_s^2 < \infty\} . \tag{6.4}$$

The Sobolev space $H^s(X)$ is a Hilbert space with inner product

$$\langle \phi , \psi \rangle_s = \sum_{\pi_\nu \in \widehat{G}} (1 + \|\nu\|^2)^s \langle \phi , \psi \rangle_{L^2(X)} . \tag{6.5}$$

In particular $H^0(X) = L^2(X)$. As in the classical euclidean case, if $s < t$ then the inclusion $H^t(X) \to H^s(X)$ is completely continuous (compact).

As usual if k is a non-negative integer we write $C^k(X)$ for the space of functions $f : X \to \mathbb{C}$ that are k times differentiable, with all k^{th} derivatives continuous. We write $C^\infty(X)$ for $\bigcap_{k>0} C^k(X)$. Differentiability and Sobolev norms are related by

Sobolev Lemmas 6.6. *Let $n = \dim_{\mathbb{R}} G$ and let k be an integer $\geqq 0$. If $\phi \in C^k(X)$ then $\phi \in H^{k-n/2-\epsilon}(X)$ for every $\epsilon > 0$. If $\phi \in H^{k+n/2+\epsilon}(X)$ for some $\epsilon > 0$ then $\sum_{\pi_\nu \in \widehat{G}} \phi_\nu$ converges absolutely and uniformly to an element of $C^k(X)$. In particular, $\phi \in H^s(X)$ for all $s \geqq 0$ if, and only if, $\phi \in C^\infty(X)$.*

We now consider the implications for the initial data function $b \in L^2(X)$. Here, for book keeping purposes, we renumber the highest weights to eliminate repetitions. Thus ν_h occurs just once, but π_{ν_h} occurs with multiplicity $mult(1_K, \pi_{\nu_h}|_K)$ on $L^2(X)$. Here it is possible that $mult(1_K, \pi_{\nu_h}|_K) = 0$.

If $b \in H^s(X)$, then

$$\|b\|_s^2 = \sum_{h=1}^{\infty} (1 + \|\nu_h\|^2)^s \|b_{\nu_h}\|^2 < \infty \tag{6.7}$$

where b_{ν_h} is the orthogonal projection of b to $A(\pi_{\nu_h})$ and $\|b_{\nu_h}\|$ stands for $\|b_{\nu_h}\|_{L^2(X)}$. It follows that

$$\|b_{\nu_h}\|^2 = o((1 + \|\nu_h\|^2)^{-s}) \text{ as } h \to \infty. \tag{6.8}$$

The degree $d(\nu_h) = deg(\pi_{\nu_h})$ is a polynomial function of degree bounded by the number m of positive roots, as a function of ν_h, according to Weyl's degree formula. So[4]

$$d(\nu_h) = O((1 + \|\nu_h\|^2)^{m/2}) \text{ as } h \to \infty. \tag{6.9}$$

Combine (6.8) and (6.9):

$$d(\nu_h)^3 \|b_{\nu_h}\|^2 = o((1 + \|\nu_h\|^2)^{-s+(3m/2)}) \text{ as } h \to \infty. \tag{6.10}$$

We have enumerated the ν_h by increasing length, ignoring the multiplicities that occur in the decomposition $L^2(X) = \sum_{\pi \in \widehat{G}} mult(1_K, \pi|_K) V_\pi$. So the growth of the $\|\nu_h\|$ is given by the growth of the euclidean norms of the lattice points in a Weyl chamber. This growth has the same order $h^{1/\ell}$ as that of the non–negative integral ℓ-tuples where $\ell = rank\ G$. Note that G has dimension $n = \ell + 2m$. So (6.10) says

Lemma 6.11. *Let $b \in H^s(X)$ with $s \geqq 0$, let $\ell = rank\ G$, and let m be the number of positive roots. Then*

$$d(\nu_h)^3 \|b_{\nu_h}\|^2 = o(h^{(-2s+3m)/\ell}) \text{ as } h \to \infty. \tag{6.12}$$

Now we try to proceed as for the heat equation on the sphere. Combine Proposition 4.10 and Lemma 6.11:

Lemma 6.13. *The typical entry of the column vector $N_r \cdot V_r'(x_0 : t)$ of (4.6) has growth*

$$\sum_{\|\nu_h\| \geqq r} \ \sum_{1 \leqq j \leqq d(\nu_h)} |m_{h,j}(s_k) a_{h,j} e^{-t\lambda_{h,j}} \phi_{h,j}(x_0)|^2 = o\left(\sum_{\|\nu_h\| \geqq r} h^{(-2s+3m)/\ell} \cdot e^{-t\lambda(\nu_h)} \right) \text{ as } r \to \infty. \tag{6.14}$$

The column vector $N_r \cdot V_r'(x_0 : t)$ of (4.6) has height n_r given by

$$n_r = \sum_{\|\nu_h\| < r} \dim A(\pi_{\nu_h}) \leqq \sum_{\|\nu_h\| < r} d(\nu_h)^2 .$$

The number of irreducible representations of G with highest weight of a given length v grows as a constant multiple of $v^{\ell-1}$. So $\sum_{\|\nu_h\| = v} \dim A(\pi_{\nu_h})$ grows at most as a constant multiple of $v^{2m+\ell-1}$. Thus

n_r is bounded by a polynomial of degree $n = 2m + \ell$ in r. (6.15)

Proposition 3.15 and Lemma 6.13 combine with (6.15) to give

[4] Here too we use the standard definition: $p = O(q)$ as $s \to s_0$ if there exist a neighborhood U of s_0 and a number $B > 0$ such that $|p(s)| \leq B\, q(s)$ for $s \in U \setminus \{s_0\}$.

Theorem 6.16. *Let $b \in H^s(X)$ with $s \geqq 0$. Then head error*

$$|V_r(x_0:t) - \widetilde{V}_r(x_0:t)| = o\left(\|M_r^{-1}\|\, r^{2m+\ell} \sum_{\|\nu_h\| \geqq r} h^{(-2s+3m)/\ell} \cdot e^{-t\lambda(\nu_h)}\right) \quad (6.17)$$

as $r \to \infty$.

In the next two Sections we will specialize Theorem 6.16 to reflect various sorts of spectral behavior of D, i.e. various growth rates of the $\lambda(\nu_h)$. Then in Section 9 we will sharpen all these results for the case of symmetric spaces.

7 Spectrum of Logarithmic or Faster Growth

In this Section we show that the head error goes to zero, for sufficiently smooth initial data and also for sufficiently large time, provided that the spectrum of D has at least logarithmic growth. This growth condition is that the $\lambda(\nu_h) = min_{1 \leqq j \leqq d(\nu_h)} \operatorname{Re} \lambda_{h,j}$ satisfy

$$\lambda(\nu_h) \geqq c \log h \qquad \text{for} \qquad h \gg 0 \quad (7.1)$$

where $c = c_{X,D} > 0$ is a constant that depends only on X and D.

In order to combine (7.1) with Theorem 6.16 we will need

Lemma 7.2. *If the eigenvalue patterns of D and the M_r give us numbers s and $\epsilon > 0$ such that*

$$\frac{-2s+3m}{\ell} - \frac{t\lambda(\nu_h)}{\log h} \leqq -(1+\epsilon) \text{ for } h \gg 0 \text{ and} \quad (7.3a)$$

$$\|M_r^{-1}\|\, r^{2m+\ell} \text{ is bounded for } r \gg 0 \quad (7.3b)$$

then

$$\lim_{r\to\infty} |V_r(x_0:t) - \widetilde{V}_r(x_0:t)| = 0. \quad (7.3c)$$

Proof. Note that (7.3a) ensures convergence of the sum in (6.17), in fact implies

$$\sum_{\|\nu_h\| \geqq r} h^{(-2s+3m)/\ell} \cdot e^{-t\lambda(\nu_h)} \leqq \frac{1}{\epsilon}\,(h(r)-1)^{-\epsilon} \text{ for } r \gg 0 \quad (7.4)$$

where $h(r)$ is chosen so that summation over $h \geqq h(r)$ is the same as summation over $\|\nu_h\| \geqq r$. In effect, (7.3a) says

$$h^{(-2s+3m)/\ell} \cdot e^{-t\lambda(\nu_h)} \leqq h^{-1-\epsilon} \text{ for } h \gg 0,$$

and of course

$$\sum_{h \geqq h(r)} h^{-1-\epsilon} \leqq \int_{h(r)-1}^{\infty} h^{-1-\epsilon} dh = \frac{1}{\epsilon}\,(h(r)-1)^{-\epsilon}.$$

That proves (7.4).

Combine (6.17) and (7.4) to obtain

$$|V_r(x_0:t) - \widetilde{V}_r(x_0:t)| = o\left(\|M_r^{-1}\|\, r^{2m+\ell}\, \frac{1}{\epsilon}\,(h(r)-1)^{-\epsilon}\right) \quad (7.5a)$$

as $r \to \infty$. Since $(h(r)-1)$ is of the order of r^ℓ, also $r^{2m+\ell}(h(r)-1)^{-\epsilon}$ is of the order of $r^{2m+\ell} r^{-\epsilon\ell}$, and we can write (7.5a) in the form

$$|V_r(x_0:t) - \widetilde{V}_r(x_0:t)| = o\left(\|M_r^{-1}\|\, r^{2m+(1-\epsilon)\ell}\right) \text{ as } r \to \infty \quad (7.5b)$$

for some $\epsilon > 0$. That, of course, implies

$$|V_r(x_0:t) - \widetilde{V}_r(x_0:t)| = o\left(\|M_r^{-1}\|\, r^{2m+\ell}\right) \text{ as } r \to \infty. \quad (7.6)$$

But (7.3b) says, exactly, that the term $\|M_r^{-1}\|\, r^{2m+\ell}$ in the right hand side of (7.6) is bounded as $r \to \infty$. This completes the proof of (7.3), thus of Lemma 7.2. □

Lemma 7.7. *Fix $t, \epsilon > 0$ and assume (7.1). If b is sufficiently[5] differentiable, or if t is sufficiently[6] large, then (7.3a) holds.*

Proof. For $h \gg 0$ we have $\lambda(\nu_h) \geqq c \log h$. Write that as

$$\frac{-2s'+3m}{\ell} - \frac{t\lambda(\nu_h)}{\log h} \leqq -(1+\epsilon) \text{ for } h \gg 0$$

[5]Differentiability of any class $k > \max\{\frac{1}{2}(5m + (2+\epsilon-tc)\ell), \frac{1}{2}(2m+\ell)\}$ is sufficient. This will come out of the proof. Here we use $n = 2m+\ell$. In particular C^∞ will always ensure (7.3a).

[6]The proof will show that time $t \geq (1+\epsilon+\frac{3m}{\ell})/c$ is sufficient to ensure (7.3a).

where $s' = \frac{1}{2}(3m + (1+\epsilon - tc)\ell)$. Let $n = dim_{\mathbb{R}}\, G$. Suppose $b \in C^k(X)$ for some integer $k > \max\{s' + n/2, n/2\}$. Then the Sobolev Lemmas say that $b \in H^s(X)$ for some real $s \geqq \max\{s', 0\}$. Now

$$\frac{-2s+3m}{\ell} - \frac{t\lambda(\nu_h)}{\log h} \leqq -(1+\epsilon) \text{ for } h \gg 0,$$

which is just (7.3a).

Now look at (7.3a) with $s = 0$. It reduces to $\lambda(\nu_h) \geqq \frac{1}{t}(1+\epsilon+\frac{3m}{\ell})\log h$. This is automatic whenever $c \geqq (1+\epsilon+\frac{3m}{\ell})/t$, i.e. whenever $t \geqq (1+\epsilon+\frac{3m}{\ell})/c$. □

Combine Lemmas 7.2 and 7.7 to obtain

Proposition 7.8. *Suppose that the spectrum of D has logarithmic or faster growth* (7.1). *Fix $t > 0$. Suppose that either the initial data function b is sufficiently differentiable or the time t is sufficiently large. If*

$$||M_r^{-1}||\, r^{2m+\ell} \text{ is bounded for } r \gg 0 \tag{7.9a}$$

then the head error at time t tends to zero,

$$\lim_{r\to\infty} |V_r(x_0 : t) - \widetilde{V}_r(x_0 : t)| = 0. \tag{7.9b}$$

8 Spectrum of Polynomial Growth

In this Section we suppose that the spectrum of D has polynomial growth of degree $q \geqq 1$. In other words, the $\lambda(\nu_h) = min_{1\leqq j\leqq d(\nu_h)}$ Re $\lambda_{h,j}$ satisfy

$$\lambda(\nu_h) \geqq ch^q \quad \text{for} \quad h \gg 0 \tag{8.1}$$

where $q = q_{X,D} \geqq 1$ and $c = c_{X,D} > 0$ are constants that depend only on X and D. Given polynomial growth (8.1) we show that the head error goes to zero exponentially fast.

The polynomial growth condition (8.1) says $e^{-t\lambda(\nu_h)} \leqq e^{-tch^q}$, so (6.14)

becomes

$$\begin{aligned}\sum_{h\geqq h(r)} \sum_{1\leqq j\leqq d(\nu_h)} &|m_{h,j}(s_k)a_{h,j}e^{-t\lambda_{h,j}}\phi_{h,j}(x_0)|^2 \\ &= o\left(\sum_{h\geqq h(r)} h^{(-2s+3m)/\ell}\cdot e^{-tch^q}\right) \text{ as } r\to\infty \\ &= o\left(\sum_{h\geqq h(r)} e^{-tch^q+((-2s+3m)/\ell)\log h}\right) \text{ as } r\to\infty.\end{aligned} \tag{8.2}$$

Fix $t > 0$. If $1 > \eta > 0$ then for $h \gg 0$ we have

$$((-2s+3m)/\ell)\ \log h \ \leqq\ \eta tch^q,$$

so

$$e^{-tch^q+((-2s+3m)/\ell)\log h} \leqq e^{-tc'h^q} \text{ where } c' = (1-\eta)c > 0. \tag{8.3}$$

Now the last sum in (8.2) is bounded by

$$\begin{aligned}\sum_{h\geqq h(r)} e^{-tc'h^q} &\leqq \sum_{h\geqq h(r)} h^{q-1}e^{-tc'h^q} \\ &\leqq \int_{h(r)-1}^{\infty} h^{q-1}e^{-tc'h^q}\,dh = \frac{1}{qtc'}e^{-tc'h(r)^q}.\end{aligned} \tag{8.4}$$

Thus Lemma 6.13 becomes

Lemma 8.5. *Assume the polynomial growth condition* (8.1). *Then typical entry of the column vector $N_r \cdot V_r'(x_0 : t)$ of* (4.6) *has growth*

$$\sum_{h\geqq h(r)} \sum_{1\leqq j\leqq d(\nu_h)} |m_{h,j}(s_k)a_{h,j}e^{-t\lambda_{h,j}}\phi_{h,j}(x_0)|^2 = o(e^{-tc'(h(r)-1)^q}) \tag{8.6}$$

as $r\to\infty$, whenever $0 < c' < c$, where c is given by (8.1).

Note that the functions $e^{-tc'(h(r)-1)^q}$ and $r^{2m+\ell}\, e^{-tc'(h(r)-1)^q}$ have the same growth rate as $r\to\infty$, because $e^{tc'(h(r)-1)^q}$ grows faster than any polynomial function of r. With this in mind, we combine (8.6), (6.15) and Proposition 3.15 to obtain the following specialized form of Theorem 6.16.

Theorem 8.7. *Assume the polynomial growth condition* (8.1) *on the spectrum of D. Then the head error*

$$|V_r(x_0:t) - \widetilde{V}_r(x_0:t)| = o\left(||M_r^{-1}||\ e^{-tc'(h(r)-1)^q}\right) \text{ as } r \to \infty \tag{8.8}$$

whenever $0 < c' < c$ *with* c' *as in* (8.1).

Corollary 8.9. *Assume* (8.1) *and suppose that* $||M_r^{-1}||$ *has at worst exponential growth in the sense*

$$||M_r^{-1}|| = O(e^{c''h(r)}) \text{ for some } c'' > 0 \text{ as } r \to \infty. \tag{8.10a}$$

If t is sufficiently large[7] *then the head error at time t tends to zero,*

$$\lim_{r\to\infty} |V_r(x_0:t) - \widetilde{V}_r(x_0:t)| = 0. \tag{8.10b}$$

Proof. The o bound on the head error $|V_r(x_0:t) - \widetilde{V}_r(x_0:t)|$ in (8.8) is

$$||M_r^{-1}||\ e^{-tc'(h(r)-1)^q} \leqq Ce^{c''h(r)-tc'(h(r)-1)^q},$$

which tends to zero as $r \to \infty$ whenever we can maintain $t > (c''/c')h(r)^{1-q}$ for all $r \gg 0$. Since $q \geqq 1$ the function $h(r)^{1-q}$ is non-increasing, so we just need $t > (c''/c')h(r_0)^{1-q}$ for some number $r_0 > 0$. With r_0 fixed, we may increase c' to c for this condition on t. □

9 The Case of Riemannian Symmetric Spaces

In this Section we see how the general results of Sections 6, 7 and 8 take considerably sharper form when our homogeneous space is a symmetric space. In effect, when $X = G/K$ is a riemannian symmetric space, we will see that ℓ can be reduced from *rank* G to the symmetric space rank *rank* G/K, and that n_r is bounded by a polynomial of degree $m+\ell$, which is somewhat less the bound for the general case. Furthermore, nonconstant invariant *differential* operators D turn out to have spectra of polynomial growth in the sense of (8.1).

Let $X = G/K$ be a compact riemannian symmetric space. Thus G is a compact Lie group. We assume that G, and thus also X, is connected. We also assume that the riemannian metric is induced by a positive definite symmetric bilinear form on the Lie algebra $\mathfrak{g}$ of G of the form $\beta = \beta_0 + \beta_1$ where β_0 annihilates the derived algebra (the "semisimple part") of $\mathfrak{g}$ and β_1 is the negative of the Killing form of the derived algebra. In other words, the riemannian metric is normalized as in [1].

[7] The proof shows that it suffices to have $t > (c''/c)h(r_0)^{1-q}$ for $r_0 > 0$.

We recall some standard facts about symmetric spaces; see [1] and [3]. The algebra $\mathcal{D}(G/K)$ of G-invariant differential operators on X is commutative. If $\pi \in \widehat{G}$ there are two possibilities for the π-isotypic subspace $A(\pi) \subset L^2(G)$: either the multiplicity $m(1_K, \pi|_K) = 0$ and $A(\pi) = 0$, or the multiplicity $m(1_K, \pi|_K) = 1$ and $A(\pi) \cong V_\pi$ as G-module. In the latter case we say that π is a *class 1* representation of G. Then the algebra $\mathcal{D}(G/K)$ acts on $A(\pi)$ by scalars, and the corresponding associative algebra homomorphism

$$\chi_\pi : \mathcal{D}(G/K) \to \mathbb{C} \tag{9.1}$$

specifies $A(\pi)$. In particular, (4.8a) simplifies to

$$\lambda(\nu) = \chi_{\pi_\nu}(D) \tag{9.2}$$

Note here that $\lambda(\nu)$ is the actual eigenvalue, not just its real part.

We saw [4, Lemma 4.1] that every $D \in \mathcal{D}(G/K)$, viewed as having domain $C^\infty(G/K)$, has unique closure $\widetilde{D}$ as densely defined linear operator on $L^2(G/K)$, and that $\widetilde{D}$ is a normal operator. A glance at the proof shows that we could start with domain the algebraic sum of the $A(\pi_\nu)$, where the invariance just means scalar action on each $A(\pi_\nu)$. So in general we have the analog

$$\chi_{\pi_\nu} : D \mapsto \mathbb{C} \text{ by } Df = \chi_{\pi_\nu}(D)f \text{ for all } f \in A(\pi_\nu) \tag{9.3}$$

of (9.1) and (9.2).

The symmetry of $X = G/K$ at the base point $1 \cdot K = x_0$ defines an involutive automorphism θ of the Lie algebra $\mathfrak{g}$. Decompose $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$, sum of the ± 1-eigenspaces of θ. Here $\mathfrak{k}$ is the Lie algebra of K. Choose

$$\mathfrak{a}: \text{ maximal abelian subspace of } \mathfrak{p}. \tag{9.4}$$

It is unique up to conjugation by an element of K, and $\mathfrak{a}$ extends uniquely to

$$\mathfrak{t} = \mathfrak{t}_\mathfrak{k} + \mathfrak{a} : \text{ Cartan subalgebra of } \mathfrak{g} \tag{9.5}$$

where $\mathfrak{t}_\mathfrak{k} = \mathfrak{t} \cap \mathfrak{k}$ is a Cartan subalgebra of the centralizer of $\mathfrak{a}$ in $\mathfrak{k}$.

The rank $\ell = rank\ G$ is, of course, just $dim\ \mathfrak{t}$. The *symmetric space rank* of $X = G/K$ is denoted $\ell_X = rank\ X = rank\ G/K$ and is defined by

$$\ell_X = \dim \mathfrak{a}. \tag{9.6}$$

For example, in the case of complex projective space

$$P^n(\mathbb{C}) = SU(n+1)/U(n)$$

we have $\ell = n$ and $\ell_X = 1$.

The root system $\Phi = \Phi(\mathfrak{t}_{\mathbb{C}}, \mathfrak{g}_{\mathbb{C}})$ defines the *restricted root system*

$$\Phi_{\mathfrak{a}} = \Phi_{\mathfrak{a}}(\mathfrak{a}_{\mathbb{C}}, \mathfrak{g}_{\mathbb{C}}) = \{\alpha|_{\mathfrak{a}} \mid \alpha \in \Phi \text{ and } \alpha|_{\mathfrak{a}} \neq 0\}. \tag{9.7}$$

Every choice of *positive restricted root system* $\Phi_{\mathfrak{a}}^+$ is of the form

$$\Phi_{\mathfrak{a}}^+ = \{\alpha|_{\mathfrak{a}} \mid \alpha \in \Phi^+ \text{ and } \alpha|_{\mathfrak{a}} \neq 0\}. \tag{9.8}$$

for an appropriate choice of positive root system $\Phi^+ = \Phi^+(\mathfrak{t}_{\mathbb{C}}, \mathfrak{g}_{\mathbb{C}})$. We fix one such choice of positive restricted root system (9.8).

Consider the lattice

$$\Lambda_{\mathfrak{a}} = \left\{\nu \in \sqrt{-1}\mathfrak{a}^* \mid \frac{\langle \nu, \psi\rangle}{\langle \psi, \psi\rangle} \in \mathbb{Z} \text{ for all } \psi \in \Psi\right\} \tag{9.9}$$

and the subset of dominant linear functionals

$$\Lambda_{\mathfrak{a}}^+ = \left\{\nu \in \Lambda_{\mathfrak{a}} \mid \frac{\langle \nu, \psi\rangle}{\langle \psi, \psi\rangle} \geqq 0 \text{ for all } \psi \in \Psi\right\} \tag{9.10}$$

A famous theorem of Cartan [2], made precise by Helgason [3], says that $\Lambda_{\mathfrak{a}}^+$ parameterizes the class 1 representations of G in case G is simply connected and K is connected. See [6, Ch. III] for a concise proof. We formulate the result so that G need not be semisimple or simply connected.

Theorem 9.11 (Cartan, Helgason). *Suppose that G and K are connected. Then the irreducible representation π_ν of G, with highest weight ν relative to Φ^+, is of class 1 if and only if* (i) $\nu|_{\mathfrak{t}_{\mathfrak{k}}} = 0$ *and* (ii) $\nu_{\mathfrak{a}} \in \Lambda_{\mathfrak{a}}^+$.

Here the polynomial growth condition corresponding to (8.1) is

$$|\chi_{\pi_{\nu_h}}(D)| \geqq ch^q \qquad \text{for} \qquad h \gg 0 \tag{9.12}$$

where $c > 0$ and $q \geqq 1$.

Let us agree to look only at class 1 representations of G in the expression (6.1) of $L^2(X)$ and more generally as we run through the considerations of Sections 6, 7 and 8. For, as we discussed just before (9.1), those are precisely the representations of G that occur on $L^2(X)$.

We have the class 1 highest weights ν_h ordered by increasing length. So the growth of the $||\nu_h||$ is given by the growth of the euclidean norms in the parameter space $\Lambda_{\mathfrak{a}}^+$ for the class 1 representations of G. This growth has the same order h^{1/ℓ_X} as that of the non-negative integral ℓ_X-tuples where $\ell_X = rank\ X$. Thus we can convert between growth rates ch^q as used in §8 and in (9.12), and growth rates $c'||\nu_h||^{q'}$ to be used shortly, by means of

$$\text{growth } ch^q \text{ is equivalent to growth } c'||\nu_h||^{\ell_X q}. \tag{9.13}$$

Consider the case where the operator D is differential, i.e. where $D \in \mathcal{D}(X)$. The Helgason–Harish-Chandra correspondence from the algebra of Weyl group invariant polynomials on $\mathfrak{a}$ to $\mathcal{D}(X)$ (see [3, Chapter X, §6.3] where it is done for noncompact symmetric spaces in a way that is valid for compact symmetric spaces) expresses $\lambda'(\nu) = \chi_{\pi_\nu}(D')$ as a polynomial $p_D(\nu)$ such that the degree of p_D as polynomial is equal to the degree $deg\ D$ of D as a differential operator. We combine this with (9.12) and (9.13):

Lemma 9.14. *If $D \in \mathcal{D}(X)$ then its spectrum satisfies* (9.12) *with $q = \ell_X\ deg\ D$, i.e., $|\chi_{\pi_{\nu_h}}(D)| \geqq ch^{\ell_X\ deg\ D}$ for $h \gg 0$.*

Now we incorporate this information into the considerations of Sections 6, 7 and 8. Since we only look at class 1 representations of G, Lemma 6.11 becomes

Lemma 9.15. *Let $b \in H^s(X)$ with $s \geqq 0$, let $\ell_X = rank\ X$, and let m be the number of positive roots of G. Then*

$$d(\nu_h)^3||b_{\nu_h}||^2 = o(h^{(-2s+3m)/\ell_X}) \text{ as } h \to \infty. \tag{9.16}$$

This forces a slight change in Lemma 6.13, which becomes

Lemma 9.17. *The typical entry of the column vector $N_r \cdot V_r'(x_0 : t)$ of* (4.6) *has growth*

$$\sum_{||\nu_h|| \geq r} \sum_{1 \leqq j \leqq d(\nu_h)} |m_{h,j}(s_k) a_{h,j} e^{-t\chi_{\pi_{\nu_h}}(D)} \phi_{h,j}(x_0)|^2 = o\left(\sum_{||\nu_h|| \geq r} h^{(-2s+3m)/\ell_X} |e^{-t\chi_{\pi_{\nu_h}}(D)}|\right) \tag{9.18}$$

as $r \to \infty$.

Now we come to the most important change: The column vector $N_r \cdot V_r'(x_0 : t)$ of (4.6) has height $n_r = \sum_{\|\nu_h\|<r} dim\, A(\pi_{\nu_h}) = \sum_{\|\nu_h\|<r} d(\nu_h)$. The number of irreducible representations of G with highest weight of length v grows as a constant multiple of $v^{\ell_X - 1}$. So $\sum_{\|\nu_h\|=v} dim\, A(\pi_{\nu_h})$ grows at most as a constant multiple of $v^{m+\ell_X-1}$. Thus

$$n_r \text{ is bounded by a polynomial of degree } m + \ell_X \text{ in } r. \tag{9.19}$$

Now combine (9.19) with Lemmas 9.17 and 8.3. Then Theorem 6.16, Proposition 7.8, Theorem 8.7 and Corollary 8.9 become

Theorem 9.20. *Suppose that $X = G/K$ is a riemannian symmetric space of compact type. Let $b \in H^s(X)$ with $s \geqq 0$. Then head error*

$$|V_r(x_0 : t) - \tilde{V}_r(x_0 : t)| = o\left(\|M_r^{-1}\|\, r^{m+\ell_X} \sum_{\|\nu_h\| \geq r} h^{(-2s+3m)/\ell_X} |e^{-t\chi_{\pi_{\nu_h}}(D)}|\right) \tag{9.21}$$

as $r \to \infty$.

Corollary 9.22. *Suppose that $X = G/K$ is a riemannian symmetric space of compact type. If the eigenvalue patterns of D and the M_r let us arrange numbers $s > 0$ and $\epsilon > 0$ so that*

$$\frac{-2s+3m}{\ell_X} - \frac{t|\chi_{\pi_{\nu_h}}(D)|}{\log h} \leqq -(1+\epsilon) \text{ for } h \gg 0 \tag{9.23a}$$

and

$$\|M_r^{-1}\|\, r^{m+\ell_X} \text{ is bounded for } r \gg 0 \tag{9.23b}$$

then

$$\lim_{r\to\infty} |V_r(x_0 : t) - \tilde{V}_r(x_0 : t)| = 0. \tag{9.23c}$$

In particular, if the spectrum of D has logarithmic or faster growth

$$|\chi_{\pi_{\nu_h}}(D)| \geqq c \log h \qquad \text{for} \qquad h \gg 0, \tag{9.24}$$

if either the initial data function is sufficiently differentiable or the time t is sufficiently large, and if (9.23b) holds, then the head error at time t tends to zero.

Corollary 9.25. *Suppose that $X = G/K$ is a riemannian symmetric space of compact type. Assume the polynomial growth condition (9.12), $|\chi_{\pi_{\nu_h}}(D)| \geqq ch^q$ for $h \gg 0$, which is automatic if D is a nonconstant invariant differential operator of degree $degD \geqq \ell_X q$. Then the head error*

$$|V_r(x_0 : t) - \tilde{V}_r(x_0 : t)| = o\left(\|M_r^{-1}\|\, e^{-tc'(h(r)-1)^q}\right) \text{ as } r \to \infty \tag{9.26}$$

whenever $0 < c' < c$ with c' as in (9.12).

Corollary 9.27. *Suppose that $X = G/K$ is a riemannian symmetric space of compact type. Assume polynomial growth (9.12) on the spectrum of D, automatic if D is a nonconstant invariant differential operator. Suppose that $\|M_r^{-1}\|$ has at worst exponential growth in the sense*

$$\|M_r^{-1}\| = O(e^{c''h(r)}) \text{ for some } c'' > 0 \text{ as } r \to \infty. \tag{9.29a}$$

If t is sufficiently large then the head error at time t tends to zero,

$$\lim_{r\to\infty} |V_r(x_0 : t) - \tilde{V}_r(x_0 : t)| = 0. \tag{9.29b}$$

We end with an example. Suppose that $X = G/K$ is the sphere $S^n = SO(n+1)/SO(n)$ and D is the Laplace–Beltrami operator Δ as in Section 5. Then $\ell_X = 1$ and $p_D(\nu) = \|\nu + \rho\|^2 - \|\rho\|^2$. So (9.12) holds with $q = 2$, $h(r)$ is a constant multiple of r, and Corollary 9.25 says

$$|V_r(x_0 : t) - \tilde{V}_r(x_0 : t)| = o\left(\|M_r^{-1}\|\, e^{-tc''(r-1)^q}\right) \text{ as } r \to \infty \tag{9.30}$$

for some constant $c'' > 0$. That is equivalent to (5.9).

References

1. R. S. Cahn and J. A. Wolf, *Zeta functions and their asymptotic expansions for compact symmetric spaces of rank one*, Comm. Math. Helv. **51** (1976), 1–21.
2. É. Cartan, *Sur la détermination d'un système orthogonal complet dans un espace de Riemann symétrique clos*, Rendiconti Palermo **53** (1929), 217–252.
3. S. Helgason, *Differential Geometry and Symmetric Spaces*, Academic Press, 1962.
4. D. I. Wallace and J. A. Wolf, *Observability of evolution equations for invariant differential operators*, J. Math. Systems, Estimation, and Control, to appear in 1990..
5. N. R. Wallach, *Harmonic Analysis on Homogeneous Spaces*, Marcel Dekker, 1973.
6. G. Warner, *Harmonic Analysis on Semisimple Lie Groups, I*, Springer Verlag, 1972.
7. J. A. Wolf, *Observability and group representation theory*, in "Computation and Control," (Proceedings, Bozeman, 1988), Birkhäuser, Progress in Systems and Control Theory **1** (1989), 385–391.