THE RANGE OF THE HEAT OPERATOR

BRIAN C. HALL

ABSTRACT. This paper describes results characterizing the range of the time-t heat operator on various manifolds, including Euclidean spaces, spheres, and hyperbolic spaces. The guiding principle behind these results is this: The functions in the range of the heat operator should be, roughly, those functions having an analytic continuation to the appropriate complexified manifold with growth in the imaginary directions at most like that of the time-t heat heat kernel.

1. INTRODUCTION

The heat equation is very smoothing. Suppose, for example, that we consider some initial function f in $L^2(\mathbb{R}^d)$ and run the heat equation for some fixed time $t > 0$. Then the resulting function $F(x) := u(x, t)$ (suppressing the dependence on t, which is fixed) will certainly be C^∞, even real-analytic. Is there some regularity condition that characterizes which functions F arise in this way? If we define the time-t heat operator to be the operator taking f to F, we may formulate the question in this way: Which functions are in the range of the time-t heat operator? To put it a different way, for which initial conditions F can we solve the *backward* heat equation for time t? Clearly, F must be real-analytic, but this is not nearly sufficient. The question also makes sense for the heat equation on other manifolds, with a similar situation: Functions in the range of the heat operator are real-analytic, but this is not sufficient to characterize the image.

One approach to characterizing the range of the heat operator is to use the Fourier transform. If we consider initial conditions in $L^2(\mathbb{R}^d)$ and normalize the heat equation as $\partial u/\partial t = \frac{1}{2}\Delta u$, then the range of the time-t heat operator consists of precisely those functions F such that

$$\int_{\mathbb{R}^d} \left| \hat{F}(\xi) \right|^2 e^{t|\xi|^2} d\xi < \infty,$$

where $\hat{F}$ denotes the Fourier transform of F. Although this is a very simple condition, it is desirable to have a condition expressible in terms of F itself, rather than in terms of the Fourier transform of F. On other manifolds, similarly, one could characterize the image of the heat operator in terms of the expansion of a function in eigenfunctions of the Laplacian. Nevertheless, it would be desirable to have a characterization in terms of the function itself.

The premise of this paper is that the right way to characterize the range of the heat operator (in terms of the functions themselves) is by means of analyticity properties. That is, I wish to formulate the question as: How analytic are the functions in the range of the heat operator? The answer to this question should

be provided by the (ubiquitous) heat kernel. The heat operator is computed by integration against the heat kernel, so functions in the range of the time-t heat operator should pick up the analyticity properties of the time-t heat kernel. This paper, then, will investigate the validity of the following (deliberately vague) idea.

> The functions in the range of the time-t heat operator are the functions having the same analyticity properties as the time-t heat kernel.

I do not know how to make this idea work for general manifolds. I will consider mainly certain special manifolds, the Euclidean spaces, spheres, hyperbolic spaces and other compact and noncompact symmetric spaces. The cases of Euclidean space and compact symmetric spaces (including spheres) are by now well understood. The case of a noncompact symmetric space is just beginning to be understood, and I concentrate on the most tractable example, hyperbolic 3-space.

Let us now be a bit more precise about how we will try to use analyticity properties to characterize the range of the heat operator. For each manifold $\mathcal{M}$ that we consider, we will introduce an appropriate "complexification" $\mathcal{M}_{\mathbb{C}}$. The time-t heat kernel for $\mathcal{M}$ will then have an analytic continuation to $\mathcal{M}_{\mathbb{C}}$ with a certain t-dependent rate of growth in the "imaginary directions." The rate of growth of the analytically continued heat kernel will lead to a "metatheorem" characterizing, roughly, the range of the time-t heat operator as those functions having analytic continuations to $\mathcal{M}_{\mathbb{C}}$ with at most the specified growth rate in the imaginary direction. For each manifold, the metatheorem will be translated into various precisely stated theorems characterizing the image under the time-t heat operator of various spaces of initial conditions.

The strategy in the previous paragraph can be implemented in a straightforward way for Euclidean spaces, spheres, and compact symmetric spaces. The case of hyperbolic space or a noncompact symmetric space is substantially more subtle, because of singularities that arise in the analytically continued heat kernel. Nevertheless, recent results allow for progress in these cases, as described in Section 5. Section 6 gives a brief account of the issues that are involved in the case of general manifolds.

2. The $\mathbb{R}^d$ case

We consider the Laplacian Δ on $\mathbb{R}^d$, given by

$$\Delta f = \sum_{k=1}^{d} \frac{\partial^2 f}{\partial x_k^2}.$$

Note that our Laplacian is a *negative* operator, which means that $e^{t\Delta/2}$ is the *forward* heat operator. We may view Δ as an unbounded operator in $L^2(\mathbb{R}^d, dx)$, with domain initially taken to be, say, C^∞ functions of compact support. The Laplacian is essentially self-adjoint on this domain. We let Δ also denote the closure of this operator, which is self-adjoint and negative-semidefinite. The spectral theorem then allows us to define the heat operator $e^{t\Delta/2}$ as a contraction operator on $L^2(\mathbb{R}^d)$. We will let "time-t heat operator" denote $e^{t\Delta/2}$ (with a factor of 2 in the exponent).

Given f in $L^2(\mathbb{R}^d, dx)$, if we define

$$u(x, s) = (e^{s\Delta/2} f)(x),$$

then u satisfies the heat equation

$$\frac{\partial u}{\partial s} = \frac{1}{2}\Delta u, \quad s > 0,$$

subject to the initial condition

$$\lim_{s \searrow 0} u(x, s) = f(x).$$

Here the limit is in the norm topology of $L^2(\mathbb{R}^d, dx)$.

It is well known that the heat operator can be computed by convolution against the *heat kernel*, namely, the function

$$(1) \qquad \rho_t(x) = (2\pi t)^{-d/2} e^{-x^2/2t},$$

where $x^2 = x_1^2 + \cdots + x_d^2$. This means that

$$(2) \qquad (e^{t\Delta/2} f)(x) = \int_{\mathbb{R}^d} \rho_t(x - x') f(x') \, dx'.$$

In this expression, we initially think of x as being in $\mathbb{R}^d$. However, it is easy to see that the heat kernel (1) has an entire analytic continuation to $\mathbb{C}^d$. Furthermore, on the right-hand side of (2), the variable x appears only in ρ_t and not in f. Thus, even if f itself is not analytic, we can analytically continue $e^{t\Delta/2} f$ from $\mathbb{R}^d$ to $\mathbb{C}^d$ by setting

$$(e^{t\Delta/2} f)(z) = \int_{\mathbb{R}^d} \rho_t(z - x') f(x') \, dx'$$

$$(3) \qquad\qquad = (2\pi t)^{-d/2} \int_{\mathbb{R}^d} e^{-(z-x')^2/2t} f(x') \, dx', \quad z \in \mathbb{C}^d.$$

Here $(z - x')^2$ is to be interpreted as $(z_1 - x_1')^2 + \cdots + (z_d - x_d')^2$ (no absolute values). It is not hard to prove, using Morera's Theorem, that this integral really depends holomorphically on z, for any f in $L^2(\mathbb{R}^d)$.

Note that the analytically continued heat kernel $\rho_t(z)$ grows like $e^{|y|^2/2t}$ in the imaginary directions. Specifically, if $z = x + iy$, with x and y in $\mathbb{R}^d$, then

$$(4) \qquad |\rho_t(z - x')| = (2\pi t)^{-d/2} e^{-|x-x'|^2/2t} e^{|y|^2/2t}.$$

Suppose we take a function f, convolve f with ρ_t, and then analytically continue $e^{t\Delta/2} f$ to $\mathbb{C}^d$, as in (3). Then the analytic continuation of $e^{t\Delta/2} f$ will have at most the same rate of growth in the imaginary directions as ρ_t itself. This suggests the following "metatheorem," which will serve as the unifying idea behind the theorems stated in this section.

Metatheorem 1. *The functions in the range of the heat operator $e^{t\Delta/2}$ for $\mathbb{R}^d$ are those functions F having an analytic continuation to $\mathbb{C}^d$ with growth at most like $e^{|y|^2/2t}$ in the imaginary directions.*

This metatheorem is worded in a deliberately vague way and is not intended to be taken too literally. In the first place, the metatheorem does not say what sort of functions we take in the *domain* of the heat operator. We will consider various spaces of initial conditions: $L^2(\mathbb{R}^d)$ with respect to Lebesgue measure, $L^2(\mathbb{R}^d)$ with respect to various Gaussian measures, Sobolev spaces, the Schwarz space, and the space of tempered distributions. In the second place, the metatheorem does not say precisely how "growth at most like $e^{|y|^2/t}$ in the imaginary directions" is to

be interpreted. We will be flexible in how we interpret this condition, sometimes in an L^2 sense, sometimes in a pointwise sense, and in either case with different possible dependence of the estimates on the real variable x. We will also polynomial variations around the "basic" growth rate of $e^{|y|^2/t}$ to account for different levels of smoothness in the initial data. I hope that, despite all this equivocation, each of the precisely formulated theorems in this section can be seen to be compatible with our metatheorem.

2.1. **Isometry formulas.** In this subsection, we consider mainly initial data in $L^2(\mathbb{R}^d)$, either with respect to Lebesgue measure or with respect to a Gaussian measure. This case is the easiest to deal with because it allows the use of Hilbert space methods. This case is also natural from the point of view of quantum mechanics, where the heat equation arises in the so-called Segal–Bargmann transform [Se1, Se2, Se3, B1].

We now come to our first theorem, characterizing the image of $L^2(\mathbb{R}^d, dx)$ under the time-t heat operator, where dx denotes Lebesgue measure on $\mathbb{R}^d$. This result is a slight modification of results of Segal [Se3] and Bargmann [B1]. See [H6, Sect. 6.3] or [H5] for how to adapt the results of Segal and Bargmann to the present setting.

Theorem 2 (Isometry 1). *A function F on $\mathbb{R}^d$ is of the form $e^{t\Delta/2}f$, with f in $L^2(\mathbb{R}^d, dx)$, if and only if F has an analytic continuation to $\mathbb{C}^d$ with the property that*

$$(5) \qquad \int_{\mathbb{C}^d} |F(x+iy)|^2 \, \frac{e^{-|y|^2/t}}{(\pi t)^{d/2}} \, dy \, dx < \infty.$$

If F is such a function then

$$(6) \qquad \int_{\mathbb{C}^d} |F(x+iy)|^2 \, \frac{e^{-|y|^2/t}}{(\pi t)^{d/2}} \, dy \, dx = \int_{\mathbb{R}^d} |f(x)|^2 \, dx.$$

The isometry theorem may be restated as saying that F is in the range of the time-t heat operator if and only if $F(x+iy)e^{-|y|^2/2t}$ is square-integrable with respect to Lebesgue measure on $\mathbb{C}^d$. This condition is just an L^2 version of the growth condition in our metatheorem. That is to say, in Theorem 2, "F has growth at most like $e^{|y|^2/2t}$" is interpreted to mean that $F(x+iy)$ divided by $e^{|y|^2/2t}$ is square-integrable with respect to Lebesgue measure.

It is instructive to see how this L^2 growth condition relates to pointwise growth conditions. By adapting a result of Bargmann [B1] to our normalization of the isometry theorem (compare [H1, Thm. 6]) we see that if a holomorphic function F on $\mathbb{C}^d$ satisfies (5), then F automatically satisfies the pointwise bound

$$(7) \qquad |F(x+iy)| \le C e^{|y|^2/2t}$$

where

$$C = \|f\|_{L^2(\mathbb{R}^d, dx)} \, (4\pi t)^{d/4}.$$

Meanwhile, direct calculation shows that if F satisfies a polynomially stronger bound, say,

$$(8) \qquad |F(x+iy)| \le \frac{C e^{|y|^2/2t}}{(1+|x|^2+|y|^2)^{d/2+\varepsilon}},$$

then F satisfies (5). Thus, the L^2 version of "growth at most like $e^{|y|^2/2t}$" in (5) is closely related to (but not equivalent to) the pointwise version of this condition in (7).

We could also consider initial conditions that are square integrable with respect to some measure besides Lebesgue measure. The most important such measures from the point of view of physics (and also the infinite-dimensional case) are the Gaussian measures. We consider, then, $L^2(\mathbb{R}^d, \rho_s(x)dx)$, where ρ_s is the Gaussian density defined in (1). A simple calculation shows that the convolution (2) is convergent for all f in $L^2(\mathbb{R}^d, \rho_s(x)dx)$, provided that $t < 2s$. We now state a result from [DH, H5].

Theorem 3 (Isometry 2). *For $t < 2s$, a function F on $\mathbb{R}^d$ is of the form $F = e^{t\Delta/2}f$, with $f \in L^2(\mathbb{R}^d, \rho_s(x)dx)$, if and only if F has an analytic continuation to $\mathbb{C}^d$ satisfying*

$$(9) \qquad \int_{\mathbb{C}^d} |F(x+iy)|^2 \frac{e^{-|y|^2/t}}{(\pi t)^{d/2}} \frac{e^{-|x|^2/r}}{(\pi r)^{d/2}} \, dy \, dx < \infty,$$

where

$$r = 2s - t.$$

If F is such a function then

$$(10) \qquad \int_{\mathbb{C}^d} |F(x+iy)|^2 \frac{e^{-|y|^2/t}}{(\pi t)^{d/2}} \frac{e^{-|x|^2/r}}{(\pi r)^{d/2}} \, dy \, dx = \int_{\mathbb{R}^d} |f(x)|^2 \frac{e^{-|x|^2/2s}}{(2\pi s)^{d/2}} \, dx.$$

The condition (9) is another L^2 version of the condition "growth like $e^{|y|^2/2t}$ in the imaginary directions," except that we now allow the constant in that growth rate to depend on x. The growth rate in the x-direction, namely, $e^{|x|^2/2r}$, is what one would expect based on estimating the rate of growth of $(e^{t\Delta/2}f)(x)$ for $x \in \mathbb{R}^d$ and $f \in L^2(\mathbb{R}^d, \rho_s(x)dx)$.

Using the methods of [H1, Sect. 9], one can compute the reproducing kernel for the holomorphic L^2 space in (9) (i.e., the space of holomorphic functions satisfying (9)). The result is that if F is holomorphic and satisfies (9) then F satisfies a pointwise bound of the form

$$(11) \qquad |F(x+iy)| \le Ce^{|y|^2/2t}e^{|x|^2/2r}.$$

Conversely, a holomorphic function satisfying a polynomially stronger bound than this will, by direct calculation, satisfy (9). Thus, as in the case of the image of $L^2(\mathbb{R}^d, dx)$, the L^2 condition in (9) is closely related to, but not identical to, the analogous pointwise condition.

The case $s = t$ is just a standard form of the Segal–Bargmann transform (essentially just the finite-dimensional version of the transform in [Se3] or [BSZ]). The general case is developed in [DH] and [H5]. The case of $L^2(\mathbb{R}^d, dx)$ can be obtained, at least heuristically, by rescaling the norms on both sides of (10) by a factor of $(2\pi s)^{d/2}$ and then letting s tend to infinity. In the other direction, if we allow s to approach $t/2$, then the quantity $r = 2s - t$ tends to zero. This means that as s approaches $t/2$, the measure on $\mathbb{C}^d$ collapses onto the imaginary axis. In this limit we obtain a finite-dimensional version of the Fourier–Wiener transform.

Finally, let us consider briefly initial data in $L^p(\mathbb{R}^d, dx)$, with $1 \le p \le \infty$. (See also [H9].) For $p \ne 2$, we take (2) as the definition of $e^{t\Delta/2}$. It is unlikely that there is any simple exact description of the image of the heat operator for $p \ne 2$.

Nevertheless, we can give a necessary condition and a sufficient condition for F to be in the image of $L^p(\mathbb{R}^d, dx)$, with not too much of a gap between the two conditions. From (4) we see that

$$\|\rho_t(z - x')\|_{L^{p'}(\mathbb{R}^d, dx)} = Ce^{|\operatorname{Re} z|^2/2t},$$

where p' is the conjugate exponent to p and where C is a constant depending on p, d, and t, but not on z. From this it follows that if $F = e^{t\Delta/2}f$ with $f \in L^p(\mathbb{R}^d, dx)$ we have

$$(12) \qquad\qquad |F(x + iy)| \leq C \|f\|_{L^p(\mathbb{R}^d, dx)} e^{|y|^2/2t}.$$

Conversely, using the inversion formula (14) from Section 2.2 and moving the L^p norm inside the integral, we can show that if

$$(13) \qquad\qquad \int_{\mathbb{C}^d} |F(x + iy)| e^{-|y|^2/2t} \, dx \, dy < \infty,$$

then there exists a function f on $\mathbb{R}^d$ such that $F = e^{t\Delta/2}f$ and such that $f \in L^p(\mathbb{R}^d, dx)$ for all $1 \leq p \leq \infty$. We may say, then, that the image of $L^p(\mathbb{R}^d, dx)$ is somewhere between "growth like $e^{|y|^2/2t}$" interpreted in the pointwise sense (condition (12)) and "growth like $e^{|y|^2/2t}$" interpreted in the L^1 sense (condition (13)).

By a simple interpolation argument (as in [H9]) we can improve the results in the previous two paragraphs as follows. For $1 \leq p \leq 2$, if f is in $L^p(\mathbb{R}^d, dx)$ then $Fe^{-|y|^2/2t}$ is in $L^{p'}(\mathbb{C}^d, dz)$. For $2 \leq p \leq \infty$, if $Fe^{-|y|^2/2t}$ is in $L^{p'}(\mathbb{C}^d, dz)$, then f is in $L^p(\mathbb{R}^d, dx)$.

2.2. **Inversion formula.** In this section, we consider the process of recovering f from (the analytic continuation of) $F := e^{t\Delta/2}f$. The inversion formula in Theorem 4 will also give another piece of evidence for the validity of our metatheorem. Specifically, the inversion formula will help explain why a growth rate of $e^{|y|^2/2t}$ should be sufficient to imply that F is in the range of the heat operator. (Analysis of the convolution (2) shows that this condition is necessary; sufficiency is not as obvious.) Of course, the isometry theorems of the previous subsection already tell us that the growth rate of $e^{|y|^2/2t}$ (interpreted in an L^2 sense) is sufficient. Nevertheless, the inversion formula will give us a more intuitive reason for sufficiency: A growth rate of $e^{|y|^2/2t}$ allows a direct construction of the function f for which $F = e^{t\Delta/2}f$.

It is important to note that there are many different possible inversion formulas for the heat operator. After all, since F is a holomorphic function, there can be many different integrals involving F that all give the same answer. For example, in the Cauchy integral formula, integrals over many different contours all yield the value $F(z)$. See [H9] for a application of multiple inversion formulas for the heat operator.

The most obvious inversion formula is the one obtained by thinking of the heat operator (followed by analytic continuation) as a unitary map from $L^2(\mathbb{R}^d, dx)$ onto the Hilbert space of L^2 holomorphic functions on $\mathbb{C}^d$ with respect to the Gaussian measure in (5). Since the adjoint of a unitary map is its inverse and since the transform may be thought of as an integral operator with integral kernel $\rho_t(z - x)$,

we easily obtain the following inversion formula, valid for any $f \in L^2(\mathbb{R}^d, dx)$:

$$(14) \qquad f(x) = \lim_{R \to \infty} \int_{|z| \le R} \overline{\rho_t(z - x)} F(z) \frac{e^{-|\operatorname{Re} z|^2/t}}{(\pi t)^{d/2}} \, dz,$$

where the limit is in the norm topology of $L^2(\mathbb{R}^d, dx)$. (Compare Section 9 of [H1].)

It is tempting to think of (14) as *the* inversion formula for the heat operator. However, there is another inversion formula that is, in some ways, more fundamental.

Theorem 4 (Inversion). *Let f be in $L^2(\mathbb{R}^d, dx)$ and let $F = e^{t\Delta/2} f$. Then F has an analytic continuation to $\mathbb{C}^d$ and f may be recovered from this analytic continuation by the formula*

$$(15) \qquad f(x) = \lim_{R \to \infty} \int_{|y| \le R} F(x + iy) \frac{e^{-|y|^2/2t}}{(2\pi t)^{d/2}} \, dy,$$

where the limit is in the norm topology of $L^2(\mathbb{R}^d, dx)$.

This inversion formula is special because the value of f at a point x is computed from the values of F at points of the form $x + iy$ (at least in those cases where the limit exists pointwise). This inversion formula is also special because it is easy (as we shall see) to prove directly that the integral in (15) undoes the heat equation. The analog of (15) in the case of compact symmetric spaces (Sections 3 and 4 below) is an important part of the theory and helps elucidate the role of duality between compact and noncompact symmetric spaces.

In general, one cannot expect pointwise convergence of the limit in (15). After all, f is an arbitrary L^2 function, which can have singularities. Any inversion formula must have a mechanism for reproducing those singularities, and in (15) that mechanism is the possible failure of convergence of the integral. An alternative means of regularizing the integral is

$$(16) \qquad f(x) = \lim_{s \nearrow t} \int_{\mathbb{R}^d} F(x + iy) \frac{e^{-|y|^2/2s}}{(2\pi s)^{d/2}} \, dy,$$

where the pointwise bounds (7) show that the integral (16) is absolutely convergent for all $s < t$. Here, again, the limit is to be taken in the norm topology of $L^2(\mathbb{R}^d, dx)$. As we will see in the next subsection, if we assume sufficient smoothness for f, then we may take $s = t$ in (16) (or $R = \infty$ in (15)), with absolute convergence of the resulting integral.

Let us think about why the inversion formula works. We give now an intuitive argument; a rigorous proof follows. The key observation is that the function

$$(17) \qquad \frac{e^{-|y|^2/2s}}{(2\pi s)^{d/2}}$$

occurring in (16) is just the heat kernel for $\mathbb{R}^d$ (compare (1)), except now thought of a living on the imaginary axis in $\mathbb{C}^d$. The integral (15) is then effectively computing the (forward) heat operator in the y-directions. However, because F is holomorphic, the Laplacian in the x-direction is the negative of the Laplacian in the y-direction. Thus, doing the *forward* heat operator in the y-variables amounts to doing the *backward* heat operator in the x-variables, thus undoing the heat equation.

More precisely, let us fix a holomorphic function F and define

$$u(x, s) = \int_{\mathbb{R}^d} F(x + iy) \frac{e^{-|y|^2/2s}}{(2\pi s)^{d/2}} \, dy$$

whenever the integral converges. Assume the integral has reasonable convergence properties at least for $s < t$. Then the following argument shows that u satisfies the *backward* heat equation on $\mathbb{R}^d \times (0, t)$: (1) differentiate under the integral with respect to s; (2) change the s-derivative on the Gaussian factor (17) into $\frac{1}{2}\partial^2/\partial y^2$; (3) integrate by parts to move the y-derivatives onto F; (4) use the Cauchy–Riemann equations for F to change $\partial^2/\partial y^2$ into $-\partial^2/\partial x^2$, which then comes outside the integral. Meanwhile, as s approaches zero, $u(x, s)$ will converge to $F(x)$. If we assume also that $u(x, s)$ approaches some function $f(x)$ as s approaches t, then we will have that $f = e^{-t\Delta/2}F$ and $F = e^{t\Delta/2}f$. Thus (16) indeed serves to recover f from F.

Now, the above argument also helps us see how the inversion formula fits into our metatheorem. What conditions should F satisfy so that we can run the above argument and produce a function f with $F = e^{t\Delta/2}f$? We need convergence of the integral (16) for $s < t$ and reasonable behavior of the integral as s approaches t. That is, we need (roughly) F to have growth at most like $e^{|y|^2/2t}$ in the imaginary directions. Thus, even if this reasoning does not easily translate into a precise "if and only if" result, the inversion formula still gives another confirmation of the validity of our metatheorem.

Note that in the isometry formula (15) we have $e^{-|y|^2/t}$, whereas in the inversion formula we have $e^{-|y|^2/2t}$. This factor of two change in the variance is necessary because the isometry formula involves $|F|^2$ (which grows like $e^{|y|^2/t}$) whereas the inversion formula involves F itself (which grows like $e^{|y|^2/2t}$).

Proof. (Of Theorem 4) The problem with the direct argument above for the inversion formula is that it does not easily yield convergence in L^2 as $R \to \infty$. We use, then, another argument involving the Fourier transform. Let us normalize the Fourier transform in such a way that the Fourier inversion formula takes the form $f(x) = \int e^{i\xi x}\hat{f}(\xi)d\xi$. A simple calculation then shows that
(18)

$$\int_{|y| \leq R} F(x+iy)\frac{e^{-|y|^2/2t}}{(2\pi t)^{d/2}} \, dy = \int_{\mathbb{R}^d} e^{i\xi x}\left[\hat{f}(\xi)\left(\int_{|y| \leq R} e^{-\xi y}\frac{e^{-|y|^2/2t}}{(2\pi t)^{d/2}} \, dy\right)e^{-|\xi|^2 t/2}\right] d\xi.$$

Now, the expression in round parentheses converges monotonically to $e^{|\xi|^2 t/2}$ as R tends to infinity. It follows that the expression in square brackets is in L^2 for all R and converges to $\hat{f}(\xi)$ in L^2. Thus, the left-hand side of (18) converges in L^2 to f as R tends to infinity. $\square$

2.3. **The effect of smoothness of the initial data.** Up to now, we have considered the behavior of $F := e^{t\Delta/2}f$ when f is in some L^2 or L^p space. It is worth considering how the behavior of F will change if we assume some differentiability for f. Our main result will concern the image under $e^{t\Delta/2}$ of the Sobolev space $H^n(\mathbb{R}^d)$, which consists of functions whose partial derivatives up to order n (computed in the distributional sense) lie in $L^2(\mathbb{R}^d, dx)$. We will also record results of Bargmann concerning the image under $e^{t\Delta/2}$ of the Schwarz space and of the space of tempered distributions.

Metatheorem 5. *Smoothness properties of f are reflected in polynomial fluctuations of F around the "basic" growth rate of $e^{|y|^2/2t}$ in the imaginary directions. Specifically, each derivative that f has in L^2 gives an improvement in the behavior of $F(x + iy)$ by one power of $|y|$.*

This gets interpreted, as in the isometry theorems, in an L^2 sense.

Theorem 6. *Suppose n is a positive integer. Then a function F on $\mathbb{R}^d$ is of the form $F = e^{t\Delta/2}f$, with f in the Sobolev space $H^n(\mathbb{R}^d)$, if and only if F has an analytic continuation to $\mathbb{C}^d$ satisfying*

$$(19) \qquad \int_{\mathbb{C}^d} |F(x + iy)|^2 \, e^{-|y|^2/t}(1 + |y|^{2n}) \, dy \, dx < \infty.$$

If F is such a function then F satisfies the pointwise bounds

$$(20) \qquad |F(x + iy)| \leq C \frac{e^{|y|^2/2t}}{1 + |y|^n}$$

for some constant C.

We are *not* asserting that the norm in H^n is equal to the integral in (19). Rather, the norm in (19) is different from but equivalent to the H^n norm on f. (Compare (22) in the proof of Theorem 6.)

Note that if $n > d$, then (20) implies that $\int_{\mathbb{R}^d} |F(x + iy)| \, e^{-|y|^2/2t} \, dy < \infty$. This means that we may take $R = \infty$ in (15) and obtain

$$(21) \qquad f(x) = \int_{\mathbb{R}^d} F(x + iy) \frac{e^{-|y|^2/2t}}{(2\pi t)^{d/2}} \, dy,$$

with absolute convergence of the integral for all x. However, the condition $n > d$ is not sharp. By a better argument (that does not use pointwise bounds as an intermediate step) we can get the same result whenever $n > d/2$. This is presumably the optimal condition, since this is also the condition in the Sobolev embedding theorem needed to guarantee that $f \in H^n(\mathbb{R}^d)$ is continuous. (If $n \leq d/2$ then $f \in H^n(\mathbb{R}^d)$ can have singularities, in which case we cannot expect convergence of the integral (21) for all x.)

Theorem 7. *If $f \in H^n(\mathbb{R}^d)$ and $n > d/2$, then for all $x \in \mathbb{R}^d$ we have*

$$f(x) = \int_{\mathbb{R}^d} F(x + iy) \frac{e^{-|y|^2/2t}}{(2\pi t)^{d/2}} \, dy$$

with absolute convergence of the integral for all x.

Proof. (Of Theorem 6) I adapt here the arguments in [HL] from the compact group case to the (easier) $\mathbb{R}^d$ case. We consider the Sobolev space $H^n(\mathbb{R}^d)$. This space consists of those functions in $L^2(\mathbb{R}^d)$ all of whose partial derivatives of order at most n (computed in the distributional sense) are in $L^2(\mathbb{R}^d)$. Let c_n be a positive constant, whose value will be fixed later. We consider the inner product on $H^n(\mathbb{R}^d)$ given by

$$\langle f_1, f_2 \rangle_n = c_n \langle f_1, f_2 \rangle_{L^2} + (-1)^n \langle f_1, \Delta^n f_2 \rangle_{L^2}.$$

(Technically, this formula assumes that f_2 is in $H^{2n}(\mathbb{R}^d)$, but a simple integration-by-parts argument shows that the inner product extends continuously to f_2 in $H^n(\mathbb{R}^d)$.) For any choice of c_n, $H^n(\mathbb{R}^d)$ is a Hilbert space with respect to this inner product.

Suppose now that f is "nice," say, having Fourier transform that is a C^∞ function of compact support. In that case, all of the integrations by parts that follow are easily justified. Using the isometry result (Theorem 2) and the fact that Δ^n commutes with the heat operator, we obtain

$$\langle f, \Delta^n f \rangle_{L^2} = \int_{\mathbb{C}^d} \overline{F(z)} \left(\Delta_{\mathbb{C}}^n F(z) \right) \frac{e^{-|y|^2/t}}{(\pi t)^{d/2}} \, dy \, dx,$$

where

$$\Delta_{\mathbb{C}} = \sum_{k=1}^d \frac{\partial^2}{\partial z_k^2}.$$

We now integrate by parts. Since $\overline{F(z)}$ is antiholomorphic, the z-derivatives do not "see" this factor and all of the derivatives go onto the Gaussian factor, giving

$$\langle f, \Delta^n f \rangle_{L^2} = \int_{\mathbb{C}^d} |F(z)|^2 \left(\Delta_{\mathbb{C}}^n \frac{e^{-|y|^2/t}}{(\pi t)^{d/2}} \right) \, dy \, dx.$$

Since, now, the Gaussian factor is independent of x, applying $\Delta_{\mathbb{C}}$ to this factor gives the same result as applying $\Sigma[(-i/2)\partial/\partial y_k]^2 = -\frac{1}{4}\Delta_y$. Multiplying and dividing by the Gaussian factor then gives

(22) $$\langle f, f \rangle_n = \int_{\mathbb{C}^d} |F(z)|^2 \left[c_n + h_{n,t}(y) \right] \frac{e^{-|y|^2/t}}{(\pi t)^{d/2}} \, dy \, dx,$$

where

$$h_{n,t}(y) = \frac{1}{4^n} e^{|y|^2/t} \Delta^n (e^{-|y|^2/t}).$$

The function $h_{n,t}$ is a polynomial of degree $2n$ whose leading term is a positive multiple of $|y|^{2n}$. (Essentially, $h_{n,t}$ is the Wick ordering of the function $|y|^{2n}$.) If we choose the constant c_n large enough, then $c_n + h_{n,t}(y)$ will be strictly positive for all y.

Up to now we have been assuming that f_1 and f_2 are "nice." However, once (22) is established for nice functions, it is not hard to show that it holds for all functions in $H^n(\mathbb{R}^d)$. Furthermore, it can be shown that the map sending f to the analytic continuation of $F := e^{t\Delta/2} f$ is an isometry of $H^n(\mathbb{R}^d)$ *onto* the Hilbert space of holomorphic functions for which (22) is finite. (Compare the reasoning in Section 10 of [H1].) Since $h_{n,t}$ behaves like $|y|^{2n}$ at infinity, the functions satisfying (22) are precisely the functions satisfying (19). This establishes the first assertion in Theorem 6.

To establish the second assertion in Theorem 6, we need to estimate the reproducing kernel for the space of holomorphic functions satisfying (19). This can be done by the argument in Section 2 of [HL]. The details are omitted. $\square$

Proof. (Of Theorem 7) If we write out the formula (3) for F we obtain

(23) $$F(x + iy) = (2\pi t)^{-d/2} e^{|y|^2/2t} e^{-iy \cdot x/t} \int_{\mathbb{R}^d} e^{iy \cdot x'/t} e^{-|x-x'|^2/2t} f(x') \, dx'.$$

Since f is in $H^n(\mathbb{R}^d)$, it is easily seen that the function $g_x(x') := e^{-|x-x'|^2/2t} f(x')$ is in $H^n(\mathbb{R}^d)$ for each x. Now, the integral on the right-hand side of (23) is (up to a constant) the Fourier transform of g_x, evaluated at the point $\xi = -y/t$. Since the

Fourier transform $\hat{g}$ of a function g in $H^n(\mathbb{R}^d)$ satisfies $\int_{\mathbb{R}^d} |\hat{g}(\xi)(1 + |\xi|^n)|^2 \, d\xi < \infty$, we conclude that

$$\int_{\mathbb{R}^d} \left| F(x+iy)e^{-|y|^2/2t}(1+|y|^n) \right|^2 \, dy < \infty$$

for all x. The Schwarz inequality then tells us that

$$\int_{\mathbb{R}^d} |F(x+iy)| \, e^{-|y|^2/2t} \, dy$$

$$(24) \qquad \leq \left[\int_{\mathbb{R}^d} \left(\frac{1}{1+|y|^n} \right)^2 dy \int_{\mathbb{R}^d} \left| F(x+iy)e^{-|y|^2/2t}(1+|y|^n) \right|^2 \, dy \right]^{1/2},$$

where the right-hand side of (24) is finite provided that $2n > d$. This means that we may take $R = \infty$ in (15) with absolute convergence of the resulting integral. $\quad\square$

We conclude this section by recording two results of Bargmann, about the Schwarz space and the space of tempered distributions. Note that tempered distributions are *less* smooth than L^2 (they have, roughly, a negative degree of differentiability) so we get polynomially *worse* behavior. These results are obtained by adapting Theorem 1.7 of [B2] to our normalization conventions.

Theorem 8 (Bargmann). *A function F on $\mathbb{R}^d$ is of the form $F = e^{t\Delta/2}f$, with f in the Schwarz space, if and only if for every n there exists a constant c_n with*

$$|F(x+iy)| \leq c_n \frac{e^{|y|^2/2t}}{(1+|x|^2+|y|^2)^n}.$$

A function F on $\mathbb{R}^d$ is of the form $F = e^{t\Delta/2}f$, with f in the space of tempered distributions, if and only if for some n and some constant c we have

$$|F(x+iy)| \leq ce^{|y|^2/2t}(1+|x|^2+|y|^2)^n.$$

2.4. **Multiplication properties.** Our metatheorem says that the functions in the range of the time-t heat operator are those having an analytic continuation to $\mathbb{C}^d$ with growth in the imaginary directions like $e^{|y|^2/2t}$. Thus if F_1 is in the range of the time-t heat operator and F_2 is in the range of the time-s heat operator, we expect that the product function F_1F_2 will have growth in the imaginary directions like $e^{|y|^2/2r}$, where r satisfies

$$\frac{1}{r} = \frac{1}{s} + \frac{1}{t}.$$

Thus we expect F_1F_2 to be in the range of the time-r heat operator. Indeed, a straightforward application of the preceding results yields a precise result of that sort.

Theorem 9. *Suppose that F_1 is of the form $F_1 = e^{t\Delta/2}f_1$ with $f_1 \in L^2(\mathbb{R}^d, dx)$ and that F_2 is of the form $e^{s\Delta/2}f_2$ with $f_2 \in L^2(\mathbb{R}^d, dx)$. Let r be such that $1/r = 1/s + 1/t$. Then there exists a unique f in $L^2(\mathbb{R}^d, dx)$ such that*

$$F_1F_2 = e^{r\Delta/2}f.$$

Furthermore, if either f_1 or f_2 is in the Schwarz space then f is also in the Schwarz space.

Proof. In light of Theorem 2 (with t replaced by r), we need to verify that the holomorphic function $F_1 F_2$ is square-integrable against the measure $e^{-|y|^2/r} \, dy \, dx$. To do this, we use the pointwise bound (7) on the function F_1 together with the square-integrability property that F_2 has from Theorem 2. This establishes the first assertion in the theorem. The assertion about Schwarz spaces follows from Theorem 8 and the pointwise bound (7). □

3. The sphere case

The main results of this section are special cases of results of Stenzel [St], who considered arbitrary compact symmetric spaces. Nevertheless, I think it is instructive to consider this case separately, since in this case one does not need the machinery of Lie group theory to state the results. See [H8] for a survey of related results.

Our general principle is that we expect the functions in the range of the time-t heat operator to have analytic continuations with the same rate of growth in the imaginary directions as the time-t heat kernel. Meanwhile, the Minakshisundaram–Pleijel asymptotic series for the heat kernel on a manifold suggests that the heat kernel should behave roughly like a Gaussian times $j^{-1/2}$, where j is the Jacobian of the exponential mapping. Specifically, the zeroth term in the series is precisely of this form. If we are on a nice manifold such as a sphere, we may hope that this leading-order approximation will be accurate even after analytic continuation. Thus, in the sphere case, the growth rate in our metatheorem will come directly from the analytic continuation of a Gaussian times $j^{-1/2}$. The Jacobian factor $j^{-1/2}$ contributes a vitally important exponentially decaying factor to our metatheorem in the sphere case, a factor that we do not have in the Euclidean case.

Before stating our metatheorem precisely, we need a bit of set-up. First, we need the appropriate set to which functions in the range of the heat operator for S^d should be analytically continued. This will be the "complexified sphere" $S_{\mathbb{C}}^d$. If the ordinary sphere S^d is

$$S^d = \left\{ x \in \mathbb{R}^{d+1} \,\middle|\, x_1^2 + \cdots + x_{d+1}^2 = 1 \right\},$$

then we define the complexified sphere $S_{\mathbb{C}}^d$ to be

$$S_{\mathbb{C}}^d = \left\{ z \in \mathbb{C}^{d+1} \,\middle|\, z_1^2 + \cdots + z_{d+1}^2 = 1 \right\}.$$

Note that there are no complex-conjugates in the definition of $S_{\mathbb{C}}^d$; it is *not* the unit sphere in $\mathbb{C}^{d+1} \cong \mathbb{R}^{2(d+1)}$. Rather, $S_{\mathbb{C}}^d$ is a noncompact complex manifold with complex dimension d, defined by a single holomorphic condition in $\mathbb{C}^{d+1}$. The real sphere S^d sits inside $S_{\mathbb{C}}^d$ as a totally real submanifold of maximal dimension (i.e., it looks locally like $\mathbb{R}^d$ inside $\mathbb{C}^d$). Thus it makes sense to speak of analytically continuing (sufficiently regular) functions from S^d to $S_{\mathbb{C}}^d$.

Second, we need the right way to measure the rate of growth of functions on $S_{\mathbb{C}}^d$. This comes by identifying $S_{\mathbb{C}}^d$ with the tangent bundle $T(S^d)$. We think of the tangent bundle to S^d as the set

$$T(S^d) = \left\{ (x, Y) \in \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \,\middle|\, |x| = 1, \ Y \cdot x = 0 \right\}.$$

Then we have the "exponential map" for S^d, in the differential geometric sense. Given a point x in S^d and a tangent vector Y at x, let $\gamma : \mathbb{R} \to S^d$ be the unique geodesic with $\gamma(0) = x$ and $d\gamma(\tau)/d\tau|_{\tau=0} = Y$. Then $\exp_x(Y)$ is defined to be $\gamma(1)$.

In the case of the sphere, the geodesics are simply great circles and the exponential map may be computed explicitly as

$$\exp_x(Y) = \cos|Y|\ x + \frac{\sin|Y|}{|Y|}Y.$$

For each $x \in S^d$, $\exp_x$ is a smooth map of $T_x(S^d)$ into S^d. For each fixed x, we may analytically continue $\exp_x$ into a holomorphic map of the complexified tangent space $T_x(S^d)_{\mathbb{C}}$ into the complexified sphere $S^d_{\mathbb{C}}$, where $T_x(S^d)_{\mathbb{C}} = \left\{ Y \in \mathbb{C}^{d+1} \big| Y \cdot x = 0 \right\}$. To see this, note that both $\cos\theta$ and $\sin\theta/\theta$ are even functions of θ. Thus, $\cos|Y|$ and $\sin|Y|/|Y|$ can be expressed as power series (with infinite radius of convergence) in powers of $|Y|^2 = Y_1^2 + \cdots + Y_{d+1}^2$. For Y in the complexified tangent space at x, we compute the complex-valued quantity $Y_1^2 + \cdots + Y_{d+1}^2$ and plug this into the same power series, in order to compute $\exp_x(Y)$ as an element of $\mathbb{C}^{d+1}$. Since $\exp_x$ maps the real tangent space into $S^d \subset \mathbb{R}^{d+1}$, its analytic continuation will map the complexified tangent space into $S^d_{\mathbb{C}} \subset \mathbb{C}^{d+1}$.

We then define a map $\Phi : T(S^d) \to S^d_{\mathbb{C}}$ by the formula

$$\Phi(x, Y) = \exp_x(iY).$$

This may be computed explicitly as

$$\Phi(x, Y) = \cosh|Y|\ x + i\frac{\sinh|Y|}{|Y|}Y$$

It is not hard to show that Φ is a diffeomorphism of $T(S^d)$ with $S^d_{\mathbb{C}}$. (Compare [St].) This identification of $T(S^d)$ with $S^d_{\mathbb{C}}$ can be obtained naturally from the theory of "adapted complex structures" as developed in [GS1, GS2, LS, Sz]. (See also Section 3 of [H3].)

We are now ready to state our metatheorem for the S^d case.

Metatheorem 10. *The functions in the range of the heat operator $e^{t\Delta/2}$ for S^d are those functions F having an analytic continuation to $S^d_{\mathbb{C}}$ such that $F(\exp_x(iY))$ grows at most like*

$$(25) \qquad\qquad e^{|Y|^2/2t}\left(\frac{|Y|}{\sinh|Y|}\right)^{(d-1)/2}.$$

As in the $\mathbb{R}^d$ case, we will allow ourselves some flexibility in how we interpret the metatheorem, sometimes in an L^2 sense and sometimes in a pointwise sense and always allowing polynomial variations depending on the *domain* of the heat operator (L^2, C^∞, etc.).

Note that (except when $d = 1$) the function $(|Y|/\sinh|Y|)^{(d-1)/2}$ has exponential decay as $|Y|$ goes to infinity. Although the exponential decay of this factor in (25) might seem insignificant compared to the Gaussian growth of the factor $e^{|Y|^2/2t}$, the exponential factor is in fact very important. After all, we expect, based on our experience in the $\mathbb{R}^d$ case, that smoothness properties of the initial data f will reflect themselves in polynomial fluctuations around the "basic" rate of growth for F. If we miss an exponential factor in the basic growth rate, we are not going to be able to see the polynomial fluctuations.

The factor $(|Y|/\sinh|Y|)^{(d-1)/2}$ is nothing but the analytic continuation of the Jacobian factor $j^{-1/2}$ referred to at the beginning of this section. Specifically, the

Jacobian of the exponential mapping at a point x in S^d is given by

$$j(\exp_x Y) = \left(\frac{\sin |Y|}{|Y|}\right)^{d-1}.$$

Thus the extra factor in our metatheorem is nothing but $j^{-1/2}(\exp_x(iY))$. We expect (based on the Minakshisundaram–Pleijel expansion) the heat kernel $\rho_{t,x}$ at a point x in S^d to satisfy

$$\rho_{t,x}(\exp_x Y) \approx (2\pi t)^{-d/2} e^{-|Y|^2/2t} j^{-1/2}(\exp_x Y).$$

Thus we expect the analytic continuation of ρ_t to grow in the imaginary directions like (a constant times) $e^{|Y|^2/2t} j^{-1/2}(\exp_x(iY))$. This line of reasoning accounts for the form of our metatheorem.

3.1. **Isometry and inversion formulas.** Consider the spherical Laplacian Δ on S^d, as described, for example, in [Ta]. This is just the Laplace–Beltrami operator for S^d as a Riemannian manifold, except that we choose the sign so that Δ is a negative operator. This operator is essentially self-adjoint as an unbounded operator on $L^2(S^d)$ with domain $C^\infty(S^d)$. We let Δ also denote the unique self-adjoint extension of this operator. The spectral theorem allows us define $e^{t\Delta/2}$ as a contraction operator on $L^2(S^d)$. For each point $x \in S^d$, there is the "heat kernel based at x", denoted $\rho_{t,x}$ with the property that

$$(e^{t\Delta/2} f)(x) = \int_{S^d} \rho_{t,x}(y) f(y) \, dy.$$

Note that under the diffeomorphism Φ of $T(S^d)$ with $S^d_{\mathbb{C}}$, the fiber directions in $T(S^d)$ correspond to what we may call the "imaginary directions" in $S^d_{\mathbb{C}}$, namely, the directions corresponding to the exponential map with an imaginary argument. The isometry and inversion formula involve an identification of fiber directions in $T(S^d)$, and thus of the imaginary directions in $S^d_{\mathbb{C}}$, with d-dimensional hyperbolic space. The *heat kernel measure* for hyperbolic space will then play the role of the Gaussian measure in the imaginary direction that we have in the $\mathbb{R}^d$ case.

Concretely, this means that we consider the function $\nu_t(R)$ satisfying the differential equation

$$(26) \qquad \frac{\partial}{\partial t}\nu_t(R) = \frac{1}{2}\left[\frac{\partial^2}{\partial R^2} - (d-1)\frac{\cosh R}{\sinh R}\frac{\partial}{\partial R}\right]\nu_t(R)$$

subject to the initial condition

$$(27) \qquad \lim_{t \searrow 0} c_d \int_0^\infty f(R)\nu_t(R)(\sinh R)^{d-1} dR = f(0)$$

for all continuous functions f on $[0,\infty)$ with at most exponential growth. Here c_d is the volume of the unit sphere in $\mathbb{R}^d$. The function $\nu_t(R)$ is just the heat kernel function on d-dimensional hyperbolic space, as a function of the distance R from the basepoint. The operator in square brackets in (26) is just the hyperbolic Laplacian acting on radial functions. The integral of $f(R)$ against $c_d(\sinh R)^{d-1} dR$ in (27) is the formula for the integral of a radial function on hyperbolic space. See [Da] or [HM] for more information and for formulas for ν_t in various dimensions.

We are now ready to state the isometry and inversion formulas for the sphere case. These results are special cases of results of Stenzel [St], who considers general compact symmetric spaces. The case of S^3 (which is isometric to the group $SU(2)$

with a bi-invariant Riemannian metric) falls under the earlier work of [H1, H2] on the compact group case. A self-contained treatment of the isometry formula in the sphere case was given (subsequently to [St]) in [HM]. See also [KR, HM2] for more on the sphere case and [TW, Sect. IV.D] for the special case of $S^3 \cong SU(2)$.

Theorem 11 (Isometry). *A function F on S^d is of the form $F = e^{t\Delta/2}f$, with $f \in L^2(S^d)$, if and only if F has an analytic continuation to $S^d_{\mathbb{C}}$ satisfying*

$$(28) \quad \int_{x \in S^d} \int_{Y \in T_x(S^d)} |F(\exp_x(iY))|^2 \, \nu_{2t}(|2Y|) \, \left(\frac{\sinh|2Y|}{|2Y|} \right)^{d-1} 2^d dY \, dx < \infty.$$

Here dx denotes the natural volume measure on S^d and dY denotes Lebesgue measure on $T_x(S^d)$. If F is such a function then the integral in (28) is equal to $\int_{S^d} |f(x)|^2 \, dx$.

Theorem 12 (Inversion). *For any f in $L^2(S^d)$, let $F = e^{t\Delta/2}f$. Then*

$$f(x) = \lim_{R \to \infty} \int_{\substack{Y \in T_x(S^d) \\ |Y| \leq R}} F(\exp_x(iY)) \, \nu_t(|Y|) \left(\frac{\sinh|Y|}{|Y|} \right)^{d-1} dY,$$

where the limit is in the norm topology of $L^2(S^d)$ and where dY denotes Lebesgue measure on $T_x(S^d)$. Furthermore, if f is sufficiently smooth then

$$f(x) = \int_{Y \in T_x(S^d)} F(\exp_x(iY)) \, \nu_t(|Y|) \left(\frac{\sinh|Y|}{|Y|} \right)^{d-1} dY$$

with the integral on the right-hand side being absolutely convergent for all $x \in S^d$.

These theorems take on an especially explicit form in the case $d = 3$; see Section 3.4. The expression $(\sinh|Y| / |Y|)^{d-1} \, dY$ is the formula for the Riemannian volume measure on hyperbolic space, written in global exponential coordinates. Thus the quantity $\nu_t(Y)(\sinh|Y| / |Y|)^{d-1} \, dY$ is the *heat kernel measure* for hyperbolic space (i.e., the heat kernel function times the Riemannian volume measure).

The idea behind the inversion formula is roughly that the integration against the heat kernel is computing the forward hyperbolic heat operator in the imaginary variables. A fairly simple analytic continuation argument shows that on holomorphic functions, the *forward* hyperbolic heat operator in the imaginary directions is the same as the *backward* spherical heat operator in the real directions. Thus the integration in the inversion formula undoes the heat equation that takes us from f to F. This line of reasoning should be compared to the heuristic argument for the inversion formula (Theorem 4) in the $\mathbb{R}^d$ case. Stenzel makes this argument rigorous and then uses the inversion formula to obtain the isometry formula (as in [H2] in the compact group case).

Observe that in the inversion formula we have $\nu_t(|Y|)$, whereas in the isometry formula we have $\nu_{2t}(|2Y|)$. In [St], the inversion formula is obtained first and the isometry formula is reduced to the inversion formula. In the process of this reduction, the scalings $Y \to 2Y$ and $t \to 2t$ arise naturally. It is hard to predict the correct way to scale the variables by looking at the Euclidean case, since in that case a scaling of the space variables can be absorbed into a scaling of the time variables. That is, in the $\mathbb{R}^d$ case we have $2^d \nu_{2t}(2y) = \nu_{t/2}(y)$, which definitely does not hold in the hyperbolic case.

Let us think about how the isometry formula fits with our metatheorem in the sphere case. It is known [DM, HS] that the heat kernel ν_t satisfies, for each t,

$$(29) \qquad \nu_t(R) \approx e^{-R^2/2t} \left(\frac{R}{\sinh R} \right)^{(d-1)/2}.$$

Thus

$$(30) \qquad \nu_{2t}(|2Y|) \left(\frac{\sinh|2Y|}{|2Y|} \right)^{d-1} \approx e^{-|Y|^2/t} \left(\frac{\sinh|2Y|}{|2Y|} \right)^{(d-1)/2}.$$

Here $\approx$ means that the ratio of the two sides is bounded and bounded away from zero as a function of R for each fixed t. Thus if $|F|^2$ is to be *square*-integrable against the measure in (28) we must have, roughly,

$$(31) \qquad |F(\exp_x(iY))| \lesssim e^{|Y|^2/2t} \left(\frac{|2Y|}{\sinh|2Y|} \right)^{(d-1)/4}.$$

Since $(\sinh|2Y|)^{(d-1)/4}$ has the same behavior at infinity as $(\sinh|Y|)^{(d-1)/2}$, the right-hand side of (31) has the same behavior as the quantity in our metatheorem, to within a polynomial factor in $|Y|$.

Let us think, in turn, about how the inversion formula fits with our metatheorem. To get convergence we need something like

$$|F(\exp_x(iY))| \lesssim \left[\nu_t(|Y|) \left(\frac{\sinh|Y|}{|Y|} \right)^{d-1} \right]^{-1}.$$

Putting in the estimate (29), this comes to

$$|F(\exp_x iY)| \lesssim e^{|Y|^2/2t} \left(\frac{|Y|}{\sinh|Y|} \right)^{(d-1)/2},$$

which is precisely what we have in the metatheorem.

3.2. **Pointwise bounds.** The optimal pointwise bounds for holomorphic functions F on $S_{\mathbb{C}}^d$ satisfying (28) are obtained from the reproducing kernel for the corresponding L^2 space of holomorphic functions. It follows easily from the results of Stenzel (using the method in Section 9 of [H1]) that the reproducing kernel for the space of holomorphic functions satisfying (28) is given in terms of the analytic continuation of the heat kernel ρ_t on S^d. This leads to the following.

Theorem 13. *For $f \in L^2(S^d)$, let $F = e^{t\Delta/2}f$. Then the analytic continuation of F satisfies*

$$|F(\exp_x(iY)| \leq \|f\|_{L^2(S^d)} \sqrt{\rho_{2t}(\exp_x(2iY))}.$$

I conjecture that for each d and t, there exists a constant $a_{t,d}$ such that

$$(32) \qquad \rho_t(\exp_x(iY)) \leq a_{t,d} e^{|Y|^2/2t} \left(\frac{|Y|}{\sinh|Y|} \right)^{(d-1)/2}.$$

Such bounds are known for the cases $d = 1$ and $d = 3$. (Apply Theorem 2 of [H3] in the case $K = S^1$ or $K = SU(2) \cong S^3$.) If this conjecture were true, we would obtain the following precise pointwise version of the estimate (31):

$$(33) \qquad |F(\exp_x(iY))| \leq \|f\|_{L^2(S^d)} \sqrt{a_{2t,d}} e^{|Y|^2/2t} \left(\frac{|2Y|}{\sinh|2Y|} \right)^{(d-1)/4}.$$

Note that this bound is better by a polynomial factor than the one in our metatheorem, where the bound in our metatheorem is nothing but what one would expect from the behavior of ρ_t in (32). This improvement presumably reflects a slight smoothing that takes place when integrating a function f in $L^2(S^d)$ against the heat kernel: The behavior of $e^{t\Delta/2}f$, for $f \in L^2(S^d)$, is slightly better than the behavior of the heat kernel ρ_t.

We also expect, based on our experience in the $\mathbb{R}^d$ case, that assuming smoothness for f gives polynomial improvement in the pointwise behavior of F (as compared to (33)). This, again, is known in the cases $d = 1$ and $d = 3$, by applying the results of [HL] to the case $K = S^1$ or $K = SU(2) \cong S^3$. See the next Section 3.4 for more on the S^3 case.

3.3. **Multiplication properties.** The metatheorem in the sphere case suggests a multiplication result similar to what we have in the Euclidean case (Theorem 9). However, the "non-Euclidean" factor involving $\sinh|Y|$ works to our advantage in this case. If we multiply to functions in the range of the heat operator, the Gaussian factors combine in the same way as in the $\mathbb{R}^d$ case, but we now get the exponentially decaying sinh factors twice. This suggests the following result, which can be proved by lifting to the group $SO(d+1)$ and using results of [HL]. (See the proof of Theorem 18 in Section 4.)

Theorem 14. *Assume $d \geq 2$. Given f_1, f_2 in $L^2(S^d)$, let $F_1 = e^{t\Delta/2}f_1$ and $F_2 = e^{s\Delta/2}f_2$, for two positive numbers s and t. Let r be such that $1/r = 1/s + 1/t$. Then there exists a unique f in $C^\infty(S^d)$ such that*

$$e^{r\Delta/2}f = F_1 F_2.$$

Note that although f_1 and f_2 are assumed only to be in L^2, we obtain that f is in C^∞.

Although this theorem has the look of the sort of soft result that might hold quite generally, it is actually quite special. For example, nothing remotely like this holds in the hyperbolic case. Indeed, the class of manifolds for which a result of this type holds is probably very small.

3.4. **The S^3 case.** In the case of the 3-sphere, the isometry and inversion formulas take on a particularly explicit form. Furthermore, in this case, we can obtain results involving smoothness of the initial data similar to what we have in the $\mathbb{R}^d$ case. Such results are expected to hold for other spheres, but the proofs are more difficult when $d \neq 3$, due to the less explicit formulas for the relevant heat kernels. What is special about the S^3 case is that S^3 is a group. That is, S^3 with the standard metric is isometric to the group $SU(2)$ with a bi-invariant metric. The results of this subsection come from [H1, H2, H3, HL].

Theorem 15. *In the S^3 case, the isometry and inversion formulas take the form*
(34)
$$\int_{S^d} |f(x)|^2\, dx = \int_{x \in S^3} \int_{Y \in T_x(S^3)} |F(\exp_x(iY))|^2\, e^{-t} \frac{\sinh|2Y|}{|2Y|} \frac{e^{-|Y|^2/t}}{(\pi t)^{3/2}}\, dY\, dx$$

and

(35)
$$f(x) = \int_{Y \in T_x(S^3)} F(\exp_x(iY)) e^{-t/2} \frac{\sinh|Y|}{|Y|} \frac{e^{-|Y|^2/2t}}{(2\pi t)^{3/2}}\, dY.$$

Here, as usual, F is the analytic continuation of $e^{t\Delta/2}f$. The isometry formula (34) holds for all f in $L^2(S^d)$, and F is in the image of $L^2(S^d)$ under $e^{t\Delta/2}$ if and only if the integral on the right-hand side of (34) is finite. The inversion formula (35) holds for sufficiently smooth f, with absolute convergence of the integral for each x.

In addition, the expected results (32) and (33) are known to hold in the $d = 3$ case.

By specializing the results of [HL] to the case $K = SU(2)$, we obtain the following counterpart to Theorem 6.

Theorem 16. *The function f is in the Sobolev space $H^{2n}(S^3)$ if and only if F satisfies*

$$(36) \quad \int_{x \in S^3} \int_{Y \in T_x(S^3)} |F(\exp_x(iY))|^2 (1+|Y|^{4n}) e^{-t} \frac{\sinh|2Y|}{|2Y|} \frac{e^{-|Y|^2/t}}{(\pi t)^{3/2}} \, dY \, dx < \infty.$$

For such an f we have the pointwise bounds

$$(37) \quad |F(\exp_x(iY))| \le C e^{|Y|^2/2t} \left(\frac{|2Y|}{\sinh|2Y|} \right)^{1/2} \frac{1}{1+|Y|^{2n}}.$$

A function f is in $C^\infty(S^3)$ if and only if F satisfies a bound of the form (37) for each $n = 1, 2, \ldots$ (with the constant C depending on n).

4. Compact symmetric spaces

The concerning spheres in the previous section have straightforward extensions to arbitrary compact symmetric spaces. I briefly summarize those extensions here. For more details, see [St] or Section 5 of [HM3]. I will for simplicity restrict to the case of simply connected compact symmetric spaces. These spaces may be represented in the form $X = U/K$, where U is a simply connected compact Lie group (automatically semisimple), K is the fixed-point subgroup of an involution of U, and the metric on X is invariant under the action of U. Conversely, any Riemannian manifold constructed this way is a simply connected compact symmetric space. (See [He] for details.)

Let $\mathfrak{u}$ denote the Lie algebra of U and let $\mathfrak{u}_\mathbb{C} := \mathfrak{u} + i\mathfrak{u}$ be the complexification of $\mathfrak{u}$. Let $U_\mathbb{C}$ be the unique simply connected Lie group with Lie algebra $\mathfrak{u}_\mathbb{C}$. Let $K_\mathbb{C}$ be the connected Lie subgroup of $U_\mathbb{C}$ whose Lie algebra is $\mathfrak{k}_\mathbb{C} := \mathfrak{k} + i\mathfrak{k}$, where $\mathfrak{k}$ is the Lie algebra of K. Then we define the "complexification" $X_\mathbb{C}$ of X to be the complex manifold $U_\mathbb{C}/K_\mathbb{C}$. As in the sphere case, we have a diffeomorphism Φ of $T(U/K)$ with $U_\mathbb{C}/K_\mathbb{C}$ given by

$$\Phi(x, Y) = \exp_x(iY),$$

where $\exp_x(iY)$ refers to the analytically continued exponential map for U/K.

The involution of U induces a involution of $\mathfrak{u}$, and $\mathfrak{u}$ then decomposes as $\mathfrak{u} = \mathfrak{k} + \mathfrak{p}$, where $\mathfrak{p}$ is the -1 eigenspace for the involution. If we set $\mathfrak{g} = \mathfrak{k} + i\mathfrak{p}$, then $\mathfrak{g}$ is a subalgebra of $\mathfrak{u}$. Let G be the connected Lie subgroup of $U_\mathbb{C}$ whose Lie algebra is $\mathfrak{g}$. Then G/K (with an appropriately normalized G-invariant Riemannian metric) is the a noncompact symmetric space, called the "dual" of U/K. For example, if U/K is a d-sphere then G/K will be hyperbolic d-space.

Let x_0 denote the identity coset in $U/K \subset U_\mathbb{C}/K_\mathbb{C}$ and let us identify the tangent space $T_{x_0}(U/K)$ with $\mathfrak{p}$. Since the geometric and group-theoretic exponential

mappings coincide for symmetric spaces, we have

$$\exp_{x_0}(iY) = e^{iY} \cdot x_0$$

for all $Y \in \mathfrak{p}$, where e^{iY} is the group-theoretic exponential of iY inside $U_{\mathbb{C}}$. Note that $iY \in i\mathfrak{p} \subset \mathfrak{g}$ and therefore $e^{iY} \in G$. From this it is not hard to see that the image of $T_{x_0}(U/K)$ under Φ is precisely the G-orbit of x_0. But the stabilizer of x_0 inside G is just K, and so the image of $T_{x_0}(U/K)$ is naturally identified with the noncompact symmetric space G/K. This identification is actually just the geometric exponential mapping for G/K, viewing $\mathfrak{p} \cong i\mathfrak{p}$ as the tangent space at the identity coset to G/K.

Of course, the tangent space at any point x in U/K may be identified (nonuniquely) with $T_{x_0}(U/K)$ by means of the action of U. Thus *every* fiber in the tangent bundle of U/K may be identified with the noncompact symmetric space G/K. This identification is not unique, but is unique up to the action of K on G/K. Once each fiber in $T(U/K)$ is identified with G/K, we may introduce on each fiber the *heat kernel function* ν_t for G/K and the *Jacobian j of the exponential mapping* for G/K.

We are now ready to state the isometry and inversion formulas for U/K.

Theorem 17 (Stenzel). *1. The **isometry formula**. A function F on U/K is of the form $F = e^{t\Delta/2}f$, with $f \in L^2(U/K)$, if and only if F has an analytic continuation to $U_{\mathbb{C}}/K_{\mathbb{C}}$ satisfying*

$$\int_{x \in U/K} \int_{Y \in T_x(U/K)} |F(\exp_x(iY))|^2 \, \nu_{2t}(2Y)j(2y) \, 2^d dY \; dx.$$

If F is such a function, then the above integral is equal to $\int_{U/K} |f(x)|^2 \; dx$. Here $d = \dim(U/K)$, dY is Lebesgue measure on $T_x(U/K)$ and dx is the Riemannian volume measure on U/K.

*2. The **inversion formula**. If f is sufficiently smooth and $F := e^{t\Delta/2}f$, then*

$$f(x) = \int_{T_x(U/K)} F(\exp_x(iY))\nu_t(Y)j(Y) \; dY,$$

with absolute convergence of the integral for all x in U/K.

Note that in the isometry formula we have $\nu_{2t}(2Y)j(2Y)$ whereas in the inversion formula we have $\nu_t(Y)j(Y)$. Stenzel first proves the inversion formula and then (as in [H2] in the compact group case) derives the isometry formula from the inversion formula. In this derivation, the scalings by a factors of 2 in the space and time variables arise naturally.

Among compact symmetric spaces are compact Lie groups with bi-invariant metrics. The compact group case, which was considered prior to [St] in [H1, H2], is special in various ways. In the first place, there are simple explicit formulas for ν_t (and j) in this case. This allows for precise estimates of various quantities of interest [H3, HL] that are not yet known in general. Note that the only spheres that fall into the group case are S^1 and S^3. This accounts for the particularly explicit nature of the formulas for the S^d case when $d = 3$. In the second place, the group case is special because in this case, the isometry and inversion formulas can be obtained by means of geometric quantization [H7]. (See also [FMMN].) The results of [H7] are simply false as soon as one moves outside of the group case.

Of course, one can always lift analysis on U/K up to the compact group U. This strategy works better for some problems than for others. The multiplication

properties for the range of the heat operator is one problem where this lifting strategy works well. We obtain, then, the following result. (In the nonsimply connected case, the same result holds, *provided* that one assumes U/K is of the "compact type" in the sense of Helgason. This last stipulation is the reason for the assumption $d \geq 2$ in Theorem 14. Without the compact-type assumption, one gets only the weaker form of the multiplication theorem one has in the $\mathbb{R}^d$ case.)

Theorem 18. *Let U/K be a simply connected compact symmetric space. Suppose that F_1 is of the form $F_1 = e^{t\Delta/2} f_1$ with $f_1 \in L^2(U/K)$ and that F_2 is of the form $e^{s\Delta/2} f_2$ with $f_2 \in L^2(U/K)$. Let r be such that $1/r = 1/s + 1/t$. Then there exists a unique f in $C^\infty(U/K)$ such that*

$$F_1 F_2 = e^{r\Delta/2} f.$$

Proof. We regard f_1 and f_2 as right-K-invariant functions on U and apply the pointwise bounds of [H3, Thm. x] to F_1 and F_2. Then the characterization in [HL, Thm. 5] of the image under $e^{r\Delta/2}$ of $C^\infty(U)$ shows that $F_1 F_2 = e^{r\Delta/2} f$ for some f in $C^\infty(U)$. Finally, $F_1 F_2$ is right-K-invariant and so, therefore, is f. Thus f may be thought of as a C^∞ function on U/K. $\square$

I conclude this section with a conjecture involving the heat kernel on a compact symmetric space and the heat kernel on the dual noncompact symmetric space. The conjecture is established in [H3] in the compact group case, where it is interpreted as a uniform bound on "phase space probability densities," hence as a form of the uncertainty principle. The conjecture can also be thought of as saying that the sharp pointwise bounds obtained from the reproducing kernel are the same as one would expect from the isometry theorem. In the conjecture, $\rho_{t,x}$ is the heat kernel at a point x in U/K, analytically continued to $U_\mathbb{C}/K_\mathbb{C}$, ν_t is the heat kernel for the dual noncompact symmetric space G/K, and j is the Jacobian of the exponential mapping for G/K.

Conjecture 19 (Heat Kernel Duality)**.** *For any simply connected compact symmetric space, there exist constants a_t and b_t, tending to 1 as t tends to 0, such that*

$$a_t(2\pi t)^{-d} \leq \rho_{t,x}(\exp_x(iY))\nu_t(Y)j(Y) \leq b_t(2\pi t)^{-d}.$$

For applications to the isometry theorem, one would use this conjecture with t replaced by $2t$ and Y by $2Y$. (Compare Theorem 13.) The paper [HS] establishes "half" of this conjecture (the appropriate bounds on ν_t) in the rank-one case and the even multiplicity case.

5. Noncompact symmetric spaces

Symmetric spaces (by which I mean, more specifically, Riemannian globally symmetric spaces) come in three basic types: the compact type, the Euclidean type, and the noncompact type. Every simply connected symmetric space can be decomposed as a product of one space of each of these three types. In the simply connected case, the Euclidean type corresponds simply to manifolds isometric to $\mathbb{R}^d$ with the standard metric. In the simply connected case, "compact type" is equivalent to compact, and this is the case we dealt with in Section 4. This leaves, then, the noncompact type. This case is much trickier is just starting to be understood [HM3, HM4, KS1, KS2, KOS].

In this section, we content ourselves with considering the simplest case, namely, hyperbolic 3-space H^3. This is just the noncompact dual (in the sense described in Section 4) of the S^3 case. The 3-dimensional space is special because this is the only dimension in which hyperbolic d-space is a symmetric space of the "complex type," where the complex-type spaces are precisely the noncompact duals of compact Lie groups. Another way of putting it is that 3 is the dimension for which the unique positive root for hyperbolic d-space has multiplicity 2. The results describe in this section apply more generally to any noncompact symmetric space of the complex type [HM3, HM4]. I begin by describing the results and then consider the issue of what the correct metatheorem might be.

Now, we have already made use in the compact case of the duality between compact and noncompact symmetric spaces. For spheres, duality is reflected in the appearance of the heat kernel for hyperbolic d-space (the dual to S^d) in the isometry and inversion formulas. Roughly, we have spherical geometry in the real directions of $S_{\mathbb{C}}^d$ and hyperbolic geometry in the imaginary directions. The proof of the inversion formula (and thus, indirectly, of the isometry formula) hinges on the interplay between these two geometries.

Now, duality is a symmetric relationship; that is, the dual of the dual is the original symmetric space. (In the notation of Section 4, this is because multiplying $\mathfrak{p}$ by i twice brings us back to $\mathfrak{p}$ again.) Thus (in the simplest case), it is reasonable to attempt to reverse the roles of S^3 and H^3. This means that we now want to let the base manifold be H^3, identify the fibers in the tangent bundle with S^3, and aim for isometry and inversion formulas using the heat kernel for S^3 in the fibers. Unfortunately, further consideration reveals serious problems with this idea. First, the fibers in $T(H^3)$ are diffeomorphic to $\mathbb{R}^3$ and not to S^3. Second, the well-known formula for the heat kernel on H^3 (see (38) below) reveals that the analytic continuation of this heat kernel develops singularities once one travels a sufficient distance in the imaginary direction. The functions F in the range of the heat operator are going to inherit these singularities and thus the expression $F(\exp_x(iY))$ is going to be defined only for small Y.

These problems can be resolved by making use of certain delicate cancellations of singularities. Before stating the results, let us set up some notation and see more precisely what problems arise. We express H^3 as G/K where G is the identity component of $SO(3,1)$ and where K is $SO(3)$. Then we define the complexification $H_{\mathbb{C}}^3$ of H^3 to be the complex manifold $G_{\mathbb{C}}/K_{\mathbb{C}}$, where $G_{\mathbb{C}}$ is $SO(3,1;\mathbb{C})$ and where $K_{\mathbb{C}}$ is $SO(3;\mathbb{C})$. We define a map Φ from $T(H^3)$ into $H_{\mathbb{C}}^3$ by the same formula as in the sphere case,

$$\Phi(x,Y) = \exp_x(iY),$$

where $\exp_x(iY)$ refers to the analytic continuation of the geometric exponential map for H^3.

The map Φ is a well-defined smooth map of $T(H^3)$ into $H_{\mathbb{C}}^3$, but it is not a global diffeomorphism, nor even a local diffeomorphism. The differential of Φ is nonsingular when $|Y| < \pi/2$, but is singular when $|Y| = \pi/2$. The map Φ is a diffeomorphism of the set

$$T^{\pi/2}(H^3) := \left\{ (x,Y) \in T(H^3) \middle|\ |Y| < \pi/2 \right\}$$

with its image in $H_{\mathbb{C}}^3$; this image is the so-called "crown domain" or Akhiezer–Gindikin domain inside $H_{\mathbb{C}}^3$. If one pulls the complex structure on $H_{\mathbb{C}}^3$ back to

$T^{\pi/2}(H^3)$ by means of Φ, the result is the "adapted complex structure" on $T^{\pi/2}(H^3)$, as defined by Guillemin–Stenzel and Lempert–Szőke [GS1, GS2, LS, Sz]. The adapted complex structure on $T^{\pi/2}(H^3)$ degenerates as one approaches the boundary, as one would expect from the degeneration of the map Φ.

Recall that we wish to identify each fiber in $T(H^3)$ with S^3. Although this does not make sense globally, it does make sense locally, using the exponential map for S^3. Under the exponential map, the ball of radius $\pi/2$ in each fiber is identified with one hemisphere of S^3.

Krötz and Stanton show [KS2, Thm. 6.1] that for any f in $L^2(H^3, dx)$ (where dx is the Riemannian volume measure), $e^{t\Delta/2}f$ has an analytic continuation to the crown domain in $H^3_{\mathbb{C}}$. This means that the expression $F(\exp_x(iY))$ is well-defined whenever $|Y| < \pi/2$ (but will generally develop singularities if one tries to go beyond $|Y| = \pi/2$). It is not hard to show, using the explicit form of the spherical functions on H^3, a better result for radial functions. Suppose $f \in L^2(H^3)$ is radial with respect to some basepoint x_0. (Since H^3 is a rank-one symmetric space, this means simply that f is a function of the distance from x_0.) Then the map $Y \to F(\exp_{x_0}(iY))$ has a real analytic extension from the ball of radius $\pi/2$ to the ball of radius π. That is, in the radial case, when working at the basepoint, one does not encounter singularities until $|Y| = \pi$.

Finally, we write down the formula [Ga, Prop. 3.2] for the heat kernel on H^3:

$$(38) \qquad \nu_{x,t}(\exp_x Y) = e^{-t/2} \frac{e^{-|Y|^2/2t}}{(2\pi t)^{3/2}} \frac{|Y|}{\sinh|Y|}.$$

(This heat kernel appeared previously in the isometry and inversion formula for S^3, but is now playing a different role.) The factor of $|Y|/\sinh|Y|$ is $j^{-1/2}(Y)$, where $j(Y)$ is the Jacobian of the exponential mapping $\exp_x$ at the point Y. If we evaluate $\nu_{x,t}$ at $\exp_x(iY)$, the factor of $\sinh|Y|$ becomes $\sin|Y|$ and we encounter the first singularity at $|Y| = \pi$. This reflects the expected behavior of radial functions in the range of the heat operator, as described in the previous paragraph. If, however, one works at some point other than the basepoint, one will encounter singularities sooner. That is, if one looks at $\nu_{x_1,t}(\exp_{x_2}(iY))$ with $x_1 \neq x_2$, one will encounter singularities between $|Y| = \pi/2$ and $|Y| = \pi$. As x_1 and x_2 vary, these singularities will come arbitrarily close to $|Y| = \pi/2$. Thus the analytic continuation of the heat kernel makes sense only on the tube $T^{\pi/2}(H^3)$ and not (as it might superficially appear) on the tube $T^{\pi}(H^3)$.

We are now ready to think about how one might characterize the range of the time-t heat operator for H^3. In light of the Krötz–Stanton result, the image of $L^2(H^3, dx)$ under $e^{t\Delta/2}$ is $some$ space of holomorphic functions on $T^{\pi/2}(H^3)$. Unfortunately, there does not seem to be any condition, involving the behavior of $F(\exp_x(iY))$ on $T^{\pi/2}(H^3)$, that could reasonably characterize this space. After all, if we start with f in $L^2(H^3, dx)$, we already get an extension to $T^{\pi/2}(H^3)$ as soon as we apply the time-ε heat operator for any small, positive ε. Applying the heat operator for some larger time t does not give an extension to a larger domain. Furthermore, the behavior of $F(\exp_x(iY))$ as we approach $|Y| = \pi/2$ does not seem to depend on t. Neither the location nor the type of singularity we get in the analytically continued heat kernel depends on t. On the other hand, the range of the time-t

heat operator depends very strongly on t. It is simply not evident what sort of t-dependent condition we could impose on the values of $F(\exp_x(iY))$ for $|Y| < \pi/2$ that could conceivably characterize the image of the time-t heat operator.

It seems, then, that to make progress we must go beyond the set $T^{\pi/2}(H^3)$ and find a way to work on the whole tangent bundle. Of course, as soon as we leave the safe waters of $T^{\pi/2}(H^3)$, we encounter singularities, both in the complex structure and in the functions of the form $F(\exp_x(iY))$. Nevertheless, I believe that to get a reasonable characterization of the range of the heat operator, we must say, "Damn the singularities! Full speed ahead!"[1] That is, we must move into the dangerous waters of $|Y| \geq \pi/2$ and simply find some way to deal with the singularities.

The key to dealing with the singularities is to recognize that they have a "universal" character. This means that there is a singular quantity that can be factored out such that what remains is nonsingular and such that the singular factor is "universal"—independent of F and t. This is simplest to explain in the radial case; the general case is more subtle, but in the same spirit. Note, to start, that the singularity in the analytically continued heat kernel for H^3 is independent of t; for all t, we have the same factor of $|Y| / \sinh |Y|$ that becomes $|Y| / \sin |Y|$ upon analytic continuation. The same behavior is exhibited by general radial functions in the range of the heat operator. Specifically, if $f \in L^2(H^3, dx)$ is radial with respect to a basepoint x_0 and $F = e^{t\Delta/2}f$, then [HM3] the map

$$Y \to F(\exp_{x_0}(Y))\frac{\sinh |Y|}{|Y|}, \quad Y \in T_{x_0}(H^3),$$

has an entire holomorphic extension to the complexified tangent space at x_0. That is to say, if we factor out from F the universal factor of $|Y| / \sinh |Y|$, what remains (given above) has no singularities when expressed in exponential coordinates at the basepoint. This observation leads [HM3, Sect. 3] to a simple isometry theorem characterizing the image under $e^{t\Delta/2}$ of the radial functions in $L^2(H^3, dx)$.

In the nonradial case, the singularities in F itself are, unfortunately, not of a universal character. Nevertheless, the inversion and isometry theorems can be expressed in terms of certain integrals involving F, and these integral turn out to have universal singularities. Without further ado, let us state the inversion and isometry results for H^3. The inversion formula is proved in [HM3, Sect. 4]. The isometry formula will be addressed in [HM4]; since the details have not all been written down yet, I state the isometry formula as a conjecture. Another, seemingly different isometry formula (for general symmetric spaces of the noncompact type) has been obtained independently by Krötz, Ólafsson, and Stanton [KOS]. It remains to be worked out how this formula relates to the one in Conjecture 21

Theorem 20. *For any $f \in L^2(H^3)$, the function*

$$(39) \qquad L(x, R) := \int_{\substack{Y \in T_x(H^3) \\ |Y| \leq R}} F(\exp_x(iY))e^{t/2}\frac{e^{-|Y|^2/2t}}{(2\pi t)^{3/2}}\frac{\sin |Y|}{|Y|} \; dY,$$

---

[1]I am inspired here by the American naval officer, David Glasgow Farragut, whose strategy of "Damn the torpedoes! Full speed ahead!" led to victory over Confederate forces in the Battle of Mobile Bay in 1864.

initially defined for R in the interval $(0, \pi/2)$, has a real-analytic extension to the interval $(0, \infty)$. If f is sufficiently smooth, then

$$f(x) = \lim_{R \to \infty} L(x, R)$$

and we may write, informally,

$$(40) \qquad f(x) = \text{``} \lim_{R \to \infty} \text{''} \int_{\substack{Y \in T_x(H^3) \\ |Y| \le R}} F(\exp_x(iY)) e^{t/2} \frac{e^{-|Y|^2/2t}}{(2\pi t)^{3/2}} \frac{\sin |Y|}{|Y|} \, dY.$$

Conjecture 21. *For any $f \in L^2(H^3)$, the quantity*

$$(41) \qquad M(R) := \int_{x \in H^3} \int_{\substack{Y \in T_x(H^3) \\ |Y| \le R}} |F(\exp_x(iY))|^2 \, e^t \frac{\sin |2Y|}{|2Y|} \frac{e^{-|Y|^2/t}}{(\pi t)^{3/2}} \, dY \, dx$$

is finite for all $R < \pi/2$, where dY denotes Lebesgue measure on $T_x(H^3)$ and where dx denotes the Riemannian volume measure on H^3. Furthermore, M has a real-analytic continuation from $(0, \pi/2)$ to $(0, \infty)$ that satisfies

$$\lim_{R \to \infty} M(R) = \int_{H^3} |f(x)|^2 \, dx.$$

Thus we may write, informally,

(42)

$$\int_{H^3} |f(x)|^2 \, dx = \text{``} \lim_{R \to \infty} \text{''} \int_{x \in H^3} \int_{\substack{Y \in T_x(H^3) \\ |Y| \le R}} |F(\exp_x(iY))|^2 \, e^t \frac{\sin |2Y|}{|2Y|} \frac{e^{-|Y|^2/t}}{(\pi t)^{3/2}} \, dY \, dx.$$

Conversely, suppose that F is a function on H^3 having an analytic continuation to $T^{\pi/2}(H^3)$, with respect to the adapted complex structure. Suppose further that the integral $M(R)$ defined in (41) is finite for all $R < \pi/2$, that M has an analytic continuation to $(0, \infty)$, and that $\lim_{R \to \infty} M(R)$ exists and is finite. Then there exists a unique f in $L^2(H^3, dx)$ for which $F = e^{t\Delta/2} f$.

Observe how closely these formulas parallel the corresponding results for S^3 in Theorem 15. To get from (40) and (42) to the corresponding S^3 versions (Theorem 15), we simply change sin to sinh; change the exponential factors $e^{t/2}$ and e^t to $e^{-t/2}$ and e^{-t}, respectively; and omit the quotation marks in the limit as R tends to infinity. Furthermore, the quantity

$$e^{t/2} \frac{e^{-|Y|^2/2t}}{(2\pi t)^{3/2}} \frac{\sin |Y|}{|Y|} \, dY$$

in (40) and (scaled appropriately) in (42) is essentially just the heat kernel measure for S^3. More precisely, this is an "unwrapped" version of the heat kernel measure for S^3, in a sense described in [HM3].

Thus, if we are sufficiently imaginative, we can achieve something close to the suggestion made earlier in this section: We let the base manifold be H^3, identify the fibers with S^3, and obtain isometry and inversion formulas involving the heat kernel for S^3 in the fibers. (It is slightly more accurate to say that we identify the fibers with the *tangent space* to S^3 at the basepoint, where this tangent space is then identified *locally* with S^3 itself.)

The idea behind the analytic continuation of the function $L(x, R)$ in Theorem 20 is as follows. The integral in (39) only "sees" the radial part of the function

$Y \to F(\exp_x(iY))$. This radial part, as discussed earlier, has singularities only of the $|Y|/\sin|Y|$ variety, and these are canceled by the factor of $\sin|Y|/|Y|$ in (39).

Let us see if we can formulate a metatheorem that would capture the spirit of these two results. Note that in the sphere case, we actually have two different versions of the metatheorem. There is the one in Metatheorem 10 and the one in (31). In the $d = 3$ case, the first of these conditions will involve $|Y|/\sinh|Y|$ and the second $(|2Y|/\sinh|2Y|)^{1/2}$. The first condition is the one suggested by the inversion formula and by the growth (32) of the analytically continued heat kernel. The second condition is the one suggested by the isometry formula and the sharp pointwise bounds (Theorem 13 or (33)). In the sphere case, these two conditions differ only by a polynomial factor and we do not really need to distinguish them.

In the H^3 case, we could consider the naive hyperbolic analogs of these two conditions, obtained by changing hyperbolic sine to ordinary sine. This means the condition on growth of $F(\exp_x(iY))$ should be either growth at most like

$$(43) \qquad e^{|Y|^2/2t} \frac{|Y|}{\sin|Y|}$$

or at most like

$$(44) \qquad e^{|Y|^2/2t} \left(\frac{|2Y|}{\sin|2Y|} \right)^{1/2}.$$

In the H^3 case, these conditions are seemingly quite different, since they suggest different things about the locations of the singularities. Nevertheless, both conditions suggest reasonable metatheorems; it is just that the two conditions correspond to two different ways of canceling out the singularities. This leads us, then, to two different (hopefully approximately equivalent!) candidates for a metatheorem in the H^3 case.

Metatheorem 22 (Version 1)**.** *The functions in the range of the time-t heat operator $e^{t\Delta/2}$ for H^3 are those functions F such that for each x, the map*

$$Y \to F(\exp_x(iY))^{\mathrm{rad}} \frac{\sin|Y|}{|Y|}$$

has a real-analytic extension from $\{|Y| < \pi/2\}$ to all of $T_x(H^3)$ with growth at most like $e^{|Y|^2/2t}$.

Here $F(\exp_x(iY))^{\mathrm{rad}}$ means the radial part of the function $F(\exp_x(iY))$, with respect to Y with x fixed, that is, the average of $F(\exp_x(iY))$ with respect to the action of the rotation group on the Y variable.

Metatheorem 23 (Version 2)**.** *The functions in the range of the time-t heat operator $e^{t\Delta/2}$ for H^3 are those functions F such that the map*

$$(45) \qquad R \to \frac{\sin(2R)}{2R} \int_{x \in H^3} \int_{\substack{Y \in T_x(H^3) \\ |Y|=R}} |F(\exp_x(iY)|^2 \; dY$$

is finite for all $R < \pi/2$ and has a real-analytic continuation from $R \in (0, \pi/2)$ to $(0, \infty)$ with growth at most like $e^{R^2/t}$.

Metatheorem 23 should be interpreted carefully. After all, if f is a nice function that just happens not to be in $L^2(H^3, dx)$, then F may not be in $L^2(H^3, dx)$ either, and in that case the integral in (45) is likely to be infinite even for small R. Thus

Metatheorem 23 is likely to be most useful for classes of initial conditions where $e^{\varepsilon\Delta/2}f$ is in $L^2(H^3, dx)$ for all positive ε. Such classes include positive and negative Sobolev spaces and L^p spaces for $1 \le p \le 2$. For example, Metatheorem 23 suggests a possible characterization of the image under $e^{t\Delta/2}$ of the intersection over n of the nth Sobolev space on H^3. On the other hand, Metatheorem 22 is more flexible an ought to apply to classes of initial conditions with varying rates of growth or decay at infinity.

I conclude this section with a discussion of multiplication properties in the H^3 case. The sorts of results we had (Theorems 9, 14, and 18) in the case of Euclidean space or a compact symmetric space do *not* extend to the case of hyperbolic 3-space. Of course, if F_1 and F_2 are both holomorphic on $T^{\pi/2}(H^3)$ (with respect to the complex structure obtained by using Φ), then so is $F_1 F_2$. However, the cancellation of singularities we have outside of $T^{\pi/2}(H^3)$ will not be preserved under multiplication. This is seen most easily in the radial case. If F_1 and F_2 are radial functions in the image of the heat operator, then $F_1(\exp_{x_0}(iY))$ and $F_2(\exp_{x_0}(iY))$ will both have singularities like $1/\sin|Y|$ as we approach $|Y| = \pi$. When we multiply them, we will get a singularity like $1/\sin^2|Y|$. A function with a singularity of this sort cannot (in light of Section 3 of [HM3]) be in the range of the heat operator for any positive time.

In the nonradial case, we can see that neither Metatheorem 22 nor Metatheorem 23 describes a class of functions that will be closed under multiplication, even if we permit a change in the time parameter. In Metatheorem 22, the first problem is that the radialization of a product is not the product of the radializations, and the second problem is that even when the functions are already radial, the factor of $\sin|Y|/|Y|$ can only cancel the singularities in one of the two functions. In Metatheorem 23, we expect that the divergence of the integral as R approaches $\pi/2$ to be worse for $F_1 F_2$ than for either function alone. Thus, the divergence of the integral for $F_1 F_2$ is probably not going to be canceled by the factor of $\sin(2R)/2R$.

The results of Krötz and Stanton in [KS1] suggest that $F_1 F_2$ is regular enough to be in the range of the operator $\exp[-(\pi/2)\sqrt{-\Delta}]$, but no better than this.

6. Concluding remarks

I conclude this paper by discussing what little is known (from the point of view of this paper) about the range of the heat operator for manifolds other than symmetric spaces. A recent paper of Krötz, Thangavelu, and Xu [KTX] calculates the image of the heat operator for a left-invariant Riemannian (*not* sub-Riemannian) metric on the Heisenberg group. They characterize this image as a certain space of holomorphic functions on the complexified Heisenberg group. However, the description of this space is a bit complicated; it is the sum of two weighted Bergman spaces with weight functions that assume both positive and negative values.

For a general compact real analytic manifold $\mathcal{M}$, there is a characterization of the range of the Poisson semigroup $\exp[-t\sqrt{-\Delta}]$ that gives an idea of how tricky the range of the heat operator is likely to be in general. According to [GS1, GS2, LS, Sz], there exists a canonically defined "adapted complex structure" on a sufficiently small tube $T^\varepsilon(\mathcal{M})$ inside the tangent bundle. Then a result of [GS2], based in part on earlier work of Boutet de Monville, is as follows.

Theorem 24. *There exists some $\varepsilon_0 \leq \varepsilon$ such that for all $t < \varepsilon_0$, a function F on $\mathcal{M}$ is of the form $F = \exp[-t\sqrt{-\Delta}]f$, with f in $C^\infty(\mathcal{M})$, if and only if F has an analytic continuation to the tube $T^t(\mathcal{M})$ that is smooth up to the boundary.*

That is, roughly, the range of the operator $\exp[-t\sqrt{-\Delta}]$ is the space of functions having an analytic continuation to a tube of radius t, provided that $t < \varepsilon_0$. Note that this theorem says nothing about the range of $\exp[-t\sqrt{-\Delta}]$ when $t > t_0$. Also, it is not known whether it is possible to take t_0 to equal ε (the largest radius for which the adapted complex structure exists), as would be natural to expect.

This result indicates the challenges that await in trying to characterize the range of the heat operator. If F is in the range of $e^{\varepsilon\Delta}$ for any positive ε, then F is automatically in the range of $\exp[-t\sqrt{-\Delta}]$ for all positive t. This means that even for very small t, a function F range of the time-t heat operator will be holomorphic on the tube of radius ε_0. Quite possibly, F is holomorphic on the tube of radius ε, the largest tube on which the adapted complex structure exists. But this condition is unlikely to be sufficient, since holomorphicity on a tube of finite radius characterizes the range of the Poisson semigroup and not the heat semigroup. Conceivably, the way F behaves as on approaches the boundary of the tube of radius ε could characterize the range of the heat operator, but our experience in the H^3 case make this seem unlikely. More likely, if there is a characterization of the range of the heat operator, it will have to involve the nature of the singularities of F once one moves beyond the maximal tube on which the adapted complex structure is defined.

Manifolds where the adapted complex structure exists on all of $T(\mathcal{M})$ might be more tractable to deal with, but these appear to be rare.

References

[BSZ] J. C. Baez, I. E. Segal, and Z.-F. Zhou, Introduction to algebraic and constructive quantum field theory. Princeton Series in Physics. Princeton University Press, Princeton, NJ, 1992.

[B1] V. Bargmann, On a Hilbert space of analytic functions and an associated integral transform, *Comm. Pure Appl. Math.* **14** (1961), 187–214.

[B2] V. Bargmann, On a Hilbert space of analytic functions and an associated integral transform. Part II. A family of related function spaces. Application to distribution theory, *Comm. Pure Appl. Math.* **20** (1967), 1–101.

[Da] E. B. Davies, Heat kernels and spectral theory. Cambridge Tracts in Mathematics, 92. Cambridge University Press, Cambridge, 1990.

[DM] E. B. Davies and N. Mandouvalos, Heat kernel bounds on hyperbolic space and Kleinian groups, *Proc. London Math. Soc.* (3) **57** (1988), 182–208.

[D] B. K. Driver, On the Kakutani-Itô-Segal-Gross and Segal-Bargmann-Hall isomorphisms, *J. Funct. Anal.* **133** (1995), 69–128.

[DG] B. K. Driver and L. Gross, Hilbert spaces of holomorphic functions on complex Lie groups. In: New trends in stochastic analysis (K. Elworthy, S. Kusuoka, and I. Shigekawa, Eds.), 76–106, World Sci. Publishing, River Edge, NJ, 1997.

[DH] B. K. Driver and B. C. Hall, Yang–Mills theory and the Segal–Bargmann transform, *Comm. Math. Phys.* **201** (1999), 249–290.

[FMMN] C. Florentino, P. Matias, J. Mourão, and J. Nunes, Geometric quantization, complex structures, and the coherent state transform, preprint.

[Ga] R. Gangolli, Asymptotic behavior of spectra of compact quotients of certain symmetric spaces, *Acta Math.* **121** (1968), 151–192

[Gr] L. Gross, Uniqueness of ground states for Schrödinger operators over loop groups, *J. Funct. Anal.* **112** (1993), 373–441.

[GS1] V. Guillemin and M. B. Stenzel, Grauert tubes and the homogeneous Monge-Ampère equation, *J. Differential Geom.* **34** (1991), 561–570.

[GS2] V. Guillemin and M. B. Stenzel, Grauert tubes and the homogeneous Monge-Ampère equation. II, *J. Differential Geom.* **35** (1992), 627–641.

[H1] B. C. Hall, The Segal-Bargmann "coherent state" transform for compact Lie groups, *J. Funct. Anal.* **122** (1994), 103–151.

[H2] B. C. Hall, The inverse Segal-Bargmann transform for compact Lie groups, *J. Funct. Anal.* **143** (1997), 98–116.

[H3] B. C. Hall, Phase space bounds for quantum mechanics on a compact Lie group, *Comm. Math. Phys.* **184** (1997), 233–250.

[H5] B. C. Hall, A new form of the Segal-Bargmann transform for Lie groups of compact type, *Canad. J. Math.* **51** (1999), 816–834.

[H6] B. C. Hall, Holomorphic methods in analysis and mathematical physics. In: First Summer School in Analysis and Mathematical Physics (S. P erez-Esteva and C. Villegas-Blas, Eds.), 1–59, Contemp. Math., 260, Amer. Math. Soc., Providence, RI, 2000.

[H7] B. C. Hall, Geometric quantization and the generalized Segal–Bargmann transform for Lie groups of compact type, *Comm. Math. Phys.* **226** (2002), 233-268.

[H8] B. C. Hall, Harmonic analysis with respect to heat kernel measure, *Bull. Amer. Math. Soc. (N.S.)* **38** (2001), 43–78.

[H9] B. C. Hall, Bounds on the Segal-Bargmann transform of L^p functions, *J. Fourier Anal. Appl.* **7** (2001), 553–569.

[H10] B. C. Hall, The Segal-Bargmann transform and the Gross ergodicity theorem. In: Finite and infinite dimensional analysis in honor of Leonard Gross (H.-H. Kuo and A. N. Sengupta, Eds.), 99–116, Contemp. Math., 317, Amer. Math. Soc., Providence, RI, 2003.

[HL] B. C. Hall and W. Lewkeeratiyutkul, Holomorphic Sobolev spaces and the generalized Segal–Bargmann transform, *J. Funct. Anal.* **217** (2004), 192-220.

[HM] B. C. Hall and J. J. Mitchell, Coherent states on spheres, *J. Math. Phys.* **43** (2002), 1211–1236.

[HM2] B. C. Hall and J. J. Mitchell, The large radius limit for coherent states on spheres. In, "Mathematical results in quantum mechanics (Taxco, 2001)," 155–162, Contemp. Math., 307, Amer. Math. Soc., Providence, RI, 2002.

[HM3] B. C. Hall and J. J. Mitchell, The Segal–Bargmann transform on certain noncompact symmetric spaces, preprint.

[HM4] B. C. Hall and J. J. Mitchell, The Segal–Bargmann transform for noncompact symmetric spaces II. In preparation.

[HS] B. C. Hall and M. B. Stenzel, Sharp bounds for the heat kernel on certain symmetric spaces of non-compact type. In: Finite and infinite dimensional analysis in honor of Leonard Gross (H.-H. Kuo and A. N. Sengupta), 117–135, Contemp. Math., 317, Amer. Math. Soc., Providence, RI, 2003.

[He] S. Helgason, "Differential Geometry, Lie Groups, and Symmetric Spaces." Academic Press, 1978.

[KR] K. Kowalski and J. Rembieliński, The Bargmann representation for the quantum mechanics on a sphere, *J. Math. Phys.* **42** (2001), 4138–4147.

[KOS] B. Krötz, G. Ólafsson, and R. Stanton, The image of the heat kernel transform on Riemannian symmetric spaces of the non-compact type, preprint. http://arxiv.org/abs/math.CA/0407391

[KS1] B. Krötz and R. J. Stanton, Holomorpic extensions of representations: (I) automorphic functions, *Ann. Math.* **159** (2004), 641-724.

[KS2] B. Krötz and R. J. Stanton, Holomorphic extension of representations: (II) geometry and harmonic analysis, preprint.

[KTX] B. Krötz, S. Thangavelu, and Y. Xu, The heat kernel transform for the Heisenberg group, preprint http://arxiv.org/abs/math.CA/0401243.

[LGS] E. Leichtnam, F. Golse, and M. B. Stenzel, Intrinsic microlocal analysis and inversion formulae for the heat equation on compact real-analytic Riemannian manifolds, *Ann. Sci. École Norm. Sup.* (4) **29** (1996), 669–736.

[LS] L. Lempert and R. Szőke, Global solutions of the homogeneous complex Monge-Ampère equation and complex structures on the tangent bundle of Riemannian manifolds, *Math. Ann.* **290** (1991), 689–712.

[Se1]   I. E. Segal, Mathematical problems of relativistic physics. In: "Proceedings of the Summer Seminar, Boulder, Colorado, 1960" (M. Kac, Ed.), American Mathematical Society, Providence, RI, 1963.

[Se2]   I. E. Segal, Mathematical characterization of the physical vacuum for a linear Bose-Einstein field. (Foundations of the dynamics of infinite systems. III) *Illinois J. Math.* **6** (1962), 500–523.

[Se3]   I. E. Segal, The complex-wave representation of the free Boson field. In: "Topics in Functional Analysis" (I. Gohberg and M. Kac, Eds.), Advances in Mathematics Supplementary Studies, Vol. 3, Academic Press, New York, 1978.

[St]   M. B. Stenzel, The Segal-Bargmann transform on a symmetric space of compact type, *J. Funct. Anal.* **165** (1999), 44–58.

[Sz]   R. Szőke, Complex structures on tangent bundles of Riemannian manifolds, *Math. Ann.* **291** (1991), 409–428.

[Ta]   M. E. Taylor, Noncommutative harmonic analysis. Mathematical Surveys and Monographs, 22. American Mathematical Society, Providence, RI, 1986.

[TW]   T. Thiemann and O. Winkler, Gauge field theory coherent states (GCS): II. Peakedness properties, *Classical Quantum Gravity* **18** (2001), 2561-2636.