Applied Analysis and Optimization
Volume 5, Number 3, 2021, 357–389

Yokohama Publishers
ISSN 2189-1664 Online Journal

ON THE STEP 2 GRUSHIN OPERATOR

LEVON SHMAVONYAN

Abstract. The Hamilton-Jacobi theory provides the necessary tools to explore the geometric and analytic properties of the Grushin operator (Δ_G). After brief discussions of sub-Riemannian geometry, which the Grushin operator integrates, and Hamiltonian formalism, this survey paper introduces techniques that enable (i) finding all the geodesics of Δ_G, (ii) formulating a modified complex action, and (iii) constructing the heat kernel of the operator.

1. Introduction

In this section, we define *subelliptic operators* and discuss the *sub-Riemannian* geometry they induce. Our discussion of these operators will be restricted to $\mathbb{R}^n$.

For an open subset Ω of $\mathbb{R}^n$, let L be a second order partial differential operator on Ω.

Definition 1.1 (Subelliptic estimate)**.** L is said to satisfy the *subelliptic estimate* property if there exists $C > 0$ and $0 < \epsilon < 2$ such that $\forall u \in C_0^\infty(\Omega)$ the following inequality holds:

$$\|u\|_{2-\epsilon} \leq C(\|Lu\|_0 + \|u\|_0). \tag{1.1}$$

$\|u\|_s$ stands for the Sobolev norm of order s:

$$\|u\|_s = \left(\int_\Omega |\hat{u}(\xi)|^2(1+|\xi|^2)^s\, d\xi\right)^{\frac{1}{2}},$$

with $\hat{u}$ being the Fourier transform of u (see [10]).

The above property is key in differentiating between elliptic and subelliptic operators. Specifically, if (1.1) holds for $\epsilon = 0$, then L is an elliptic operator. Whereas if an operator L satisfies (1.1) for $\epsilon > 0$, then the operator is said to be *subelliptic*.

Significant progress in understanding subelliptic operators was made by Lars Hörmander who, in one of his major works (see [6]), examined a class of operators called "sum of squares of vector fields," generally written as:

$$L = \sum_{j=1}^m X_j^2 + X_0, \tag{1.2}$$

where X_i are vector fields on some n-dimensional manifold $\mathcal{M}_n$ for $0 \leq i \leq m$ and $m \leq n$.

One of Hörmander's results that we consult in our discussion of the Grushin operator is that the operator L is subelliptic iff the vector fields X_i $(0 \leq i \leq m)$ satisfy the *bracket generating condition*, also referred to as *Hörmander's condition*.

2020 *Mathematics Subject Classification.* Primary: 53C17; Secondary: 35H20.
Key words and phrases. Grushin operator, geodesic, heat kernel.

Theorem 1.1. Vector fields $\{X_1, X_2, \dots X_m\}$ on a $n-$dimensional manifold $\mathcal{M}_n$ satisfy the *bracket generating property* if a finite number of their Lie brackets span the tangent space $\mathrm{T}\mathcal{M}_n$.

The Lie bracket of two vector fields X and Y is defined as

$$[X, Y] = XY - YX. \tag{1.3}$$

Definition 1.2 (Step k)**.** $X = \{X_1, X_2, \dots X_m\}$ on $\mathcal{M}_n$ is said to be *step k*, if the vector fields satisfy the bracket generating property and at least $k-1$ Lie brackets are needed to cover the tangent space.

Corollary 1.1. Elliptic operators belong to step 1, while subelliptic operators belong to higher (> 1) steps.

Other common operators that can be realized as *sum of square of vector fields* are:

$$L_1 = \frac{1}{2}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_k}\right)^2 + \frac{1}{2}\sum_{k=1}^{n}\left(x_k^{m_k}\frac{\partial}{\partial y_k}\right)^2 = \sum_{k=1}^{n}\left(X_k^2 + Y_k^2\right),$$

$$L_2 = \frac{1}{2}\left(\frac{\partial}{\partial x}\right)^2 + \frac{1}{2}\sum_{k=1}^{n}\left(x^m\frac{\partial}{\partial y_k}\right)^2 = \frac{1}{2}X_1^2 + \frac{1}{2}\sum_{k=1}^{n}Y_k^2,$$

for $m, m_k \in \mathbb{N}$ and $k = 1, \dots n$.

These operators are elliptic except when $\{x_k = 0, k = 1, \dots n\}$. We refer to them as the *missing directions*. Meanwhile, the distribution $\mathrm{span}\{X_1, X_2, \dots X_m\}$ is called the *horizontal subspace*. In the example above, L_1 is defined on the product space $\underbrace{\mathbf{R}^2 \times \dots \times \mathbf{R}^2}_{n}$ and has n missing directions.

The role of geometry is significant in understanding subelliptic operators and studying their fundamental solutions and heat kernels. Wei-Liang Chow and Petr Konstanovich Rashevskii (1938, 1939) have shown that if given vector fields do not cover the tangent space yet satisfy the bracket generating condition—meaning that the vector fields and their Lie brackets cover the tangent space of the manifold on which they are defined—then two points on the manifold can be connected by a horizontal curve.

Theorem 1.2 (Chow-Rashevskii 1938, 1939)**.** If a manifold $\mathcal{M}_n$ is topologically connected and the vector fields $X_1, X_2, \dots X_m$ on $\mathcal{M}_n$ satisfy the bracket generating condition, then any two points on $\mathcal{M}_n$ can be connected by a horizontal curve.

Definition 1.3 (Horizontal Curve)**.** A curve is said to be *horizontal* if the tangents of the curve are linear combinations of $X_1, X_2, \dots X_m$ in an n-dimensional manifold $\mathcal{M}_n$ with $m \leq n$.

Following definition 1.3, we may conclude that for any given two points A and B on manifold $\mathcal{M}_n$ there exists a piecewise differentiable horizontal curve $\gamma : [0, \tau] \to \mathcal{M}_n$ such that:

$$\gamma(0) = A, \quad \gamma(\tau) = B$$

and

$$\dot{\gamma}(s) = \sum_{k=1}^{m} a_k(s) X_k.$$

Now, we are ready to define the sub-Riemannian geometry over a manifold.

Definition 1.4. A *sub-Riemannian* structure over a manifold $\mathcal{M}_n$ is a pair $(\mathcal{H}, \langle\cdot,\cdot\rangle)$, where $\mathcal{H}$ is a bracket generating distribution and $\langle\cdot,\cdot\rangle$ is a fibre inner product defined on $\mathcal{H}$.

Following the definitions (1.3) and (1.4), the length of a horizontal curve in the sub-Riemannian geometry can be obtained as follows:

$$\ell(\gamma) = \int_0^\tau \sqrt{\langle \dot{\gamma}(s), \dot{\gamma}(s)\rangle}\, ds. \tag{1.4}$$

In minimizing the lengths of the horizontal curves, we obtain the *Carnot-Carathéodory* distance in sub-Riemannian sense, which between two points $A, B \in \mathcal{M}_n$ is given and denoted by

$$d_{cc}(A, B) = \inf \ell(\gamma), \tag{1.5}$$

with the infimum being taken over all absolutely continuous horizontal curves connecting A and B.

Remarkably, while every point O of a Riemannian manifold is connected to every other point in a sufficiently small neighborhood by a *unique* geodesic, on a sub-Riemannian manifold there will be points arbitrarily near the point O that are connected to O by *infinitely many* geodesics.

For a more thorough treatment of sub-Riemannian geometry, see [1, 5, 7].

2. Hamiltonian formalism

In this section we discuss some geometric mechanics that will be essential in finding both the geodesics of the Grushin operator and the modified complex action, which will be essential in formulating the heat kernel of the operator. This section will also elucidate the development of concepts, such as geodesics.

2.1. Newtonian Mechanics.

Let us consider a particle of mass m moving in three-dimensional space under an external force $F\big(x(t)\big)$, where $x(t)$ denotes the position of the particle m at time t. Then, $x(t)$ satisfies *Newton's equation*:

$$m\frac{d^2x(t)}{dt^2} = F\big(x(t)\big).$$

Moreover, if the force $F\big(x(t)\big)$ can be expressed as $-\nabla V(x)$, then the force is called a *conserved force* and $V(x)$ is referred to as the *potential energy.* In fact, by Newton's law, one can show that the *total energy* E is conserved; i.e. $dE/dt = 0$, which we prove for the one-dimensional case.

Under the absence of any damping or friction, by Newton's second law, we have:

$$F\big(x(t)\big) = m\frac{d^2x}{dt^2}.$$

Assuming that the force is conserved, we have:

$$m\frac{d^2x(t)}{dt^2} + \frac{dV}{dx} = 0.$$

Multiplying both sides of the equation by dx/dt, we get the time derivative of the total energy as shown below:

$$m\frac{d^2x(t)}{dt^2}\frac{dx}{dt} + \frac{dV}{dx}\frac{dx}{dt} = \frac{d}{dt}\left(\frac{m}{2}\Big(\frac{dx}{dt}\Big)^2 + V(x)\right) = \frac{dE}{dt} = 0.$$

Thus, we may conclude that E is a constant function over time. Indeed, it can be shown that in higher dimensions,

$$\frac{dE}{dt} = \sum_k \left(m\frac{d^2x_k}{dt^2} + \frac{\partial V}{\partial x_k}\right)\frac{dx_k}{dt} = 0.$$

For properties and applications of conservative systems, see [9].

2.2. **Lagrangian Formalism.** While fundamental, Newtonian mechanics presents difficulties, such as in accounting for constraints or analyzing global properties of the system, that are better dealt with by using Lagrangian formalism.

Let N parameters $\{q_i, i = 1, \dots, N\}$ describe the state of a system. A parameter is an element of some manifold $\mathcal{M}$. Then, the manifold $\mathcal{M}$ is referred to as the *configuration space*, while the parameters q_i are called the *generalized coordinates*, and their respective time derivatives, $q_i' = dq_i/dt$, are referred to as the *generalized velocities*. The function $L(q, q')$ to be defined later in the Hamilton's principle is the *Lagrangian* of a system.

Let us consider a trajectory $q(t)$ for $t \in [t_0, t_f]$ and the functional

$$S\big(q(t), q'(t)\big) = \int_{t_0}^{t_f} L(q, q')\, dt,$$

which is called the *action*. *Hamilton's principle*, or the *principle of the least action*, states that the physically realized trajectory corresponds to an extremum of the action. Thus, the Lagrangian must be determined in a way to satisfy the Hamilton's principle.

Let $q(t)$ be a path realizing the extremum of the action and consider a variation $\delta q(t)$ of the trajectory such that $\delta q(t_0) = \delta q(t_f) = 0$. Under this variation, we have

$$\begin{aligned}\delta S(t) &= \int_{t_0}^{t_f} L(q + \delta q, q' + \delta q')\, dt - \int_{t_0}^{t_f} L(q, q')\, dt\\ &= \int_{t_0}^{t_f} \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial q'}\right)\delta q \delta t,\end{aligned}$$

which must equal to 0. This condition allows us to derive the *Euler-Lagrange equation*:

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial q'} = 0.$$

Letting $\partial L/\partial q'$ be the *generalized momentum* p, one has

$$\frac{dp}{dt} = \frac{\partial L}{\partial q}.$$

Moreover, accounting for Newton's equation, one finds that

$$L(q,q') = \frac{m}{2}\left(\frac{dq}{dt}\right)^2 - V(q)$$

and its substitution into the Euler-Lagrange equation results in

$$m\frac{d^2q_k}{dt^2} + \frac{\partial V}{\partial q_k} = 0.$$

For further discussion of analytic geometry, we need to introduce the notion of *functional derivative*, which will also allow us to derive a new form of the Euler-Lagrange equation. The functional derivative of S with respect to q is defined as follows:

$$\frac{\delta S(q,q')}{\delta q(s)} = \lim_{\epsilon\to 0} \frac{S\Big(q(t)+\epsilon\delta(t-s), q'(t)+\epsilon\frac{d}{dt}\delta(t-s)\Big) - S\Big(q(t),q'(t)\Big)}{\epsilon},$$

where

$$S\Big(q(t)+\epsilon\delta(t-s), q'(t)+\epsilon\frac{d}{dt}\delta(t-s)\Big) = S(q,q') + \epsilon\left(\frac{\partial L}{\partial q}(s) - \frac{d}{dt}\frac{\partial L}{\partial q'}(s)\right) + \mathcal{O}(\epsilon^2).$$

Hence, the Euler-Lagrange equation can be written as

$$\frac{\delta S}{\delta q(s)} = \frac{\partial L}{\partial q}(s) - \frac{d}{dt}\left(\frac{\partial L}{\partial q'}\right)(s) = 0.$$

2.3. **Hamiltonian Formalism.** As shown above, the Lagrangian formalism yields a second-order differential equation. The Hamiltonian formalism, on the other hand, will provide us with equations of motion that are of first order in derivative with respect to time.

For a given Lagrangian L, the corresponding *Hamiltonian* is introduced by the following transformation:

$$H(q,p) = \sum_k p_k q'_k - L(q,q'),$$

where p_k stands for the momentum and is equal to $\partial L/\partial q'_k$.

Note that for the transformation above to be well-defined, the following condition must be imposed:

$$det\left(\frac{\partial p_i}{\partial q'_j}\right) = det\left(\frac{\partial^2 L}{\partial q'_i q'_j}\right) \neq 0.$$

Next, we consider an infinitesimal change in the Hamiltonian:

$$\begin{aligned}\delta H &= \sum_k \left(\delta p_k q'_k + p_k\delta q'_k - \frac{\partial L}{\partial q_k}\delta q_k - \frac{\partial L}{\partial q'_k}\delta q'_k\right)\\ &= \sum_k \left(\delta p_k q'_k - \frac{\partial L}{\partial q_k}\delta q_k\right).\end{aligned}$$

The relation above implies that

$$\frac{\partial H}{\partial p_k} = q_k' \quad \text{and} \quad \frac{\partial H}{\partial q_k} = -\frac{\partial L}{\partial q_k}.$$

Applying the Euler-Lagrange equation to the above equations, one derives the *Hamilton's equations of motion*,

$$q_k' = \frac{\partial H}{\partial p_k} \quad \text{and} \quad p_k' = -\frac{\partial H}{\partial q_k}.$$

We will use Hamilton's equations of motion to construct *Hamiltonian systems* that will allow us to find bicharacteristic curves solving the system. For a more detailed treatment of the subject, see [2, 8].

3. The Grushin operator and its geodesics

In this section, we begin our study of the step 2 Grushin operator. The Grushin operator is the partial differential operator broadly defined as follows:

Definition 3.1 (Grushin operator)**.**

$$\Delta_k = \frac{1}{2}(X_1^2 + X_2^2)$$

with vector fields $X_1 = \frac{\partial}{\partial x}$ and $X_2 = x^m \frac{\partial}{\partial y}$ in $\mathbb{R}^2$ and $(x, y) \in \mathbb{R}^2$.

When $m = 1$, we obtain the step 2 Grushin operator.

Definition 3.2 (Step 2 Grushin operator)**.**

$$\Delta_G = \frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}x^2\frac{\partial^2}{\partial y^2}.$$

From definition (3.2), it is easy to observe that X_1 and X_2 are linearly independent vector fields everywhere except when $x = 0$, which renders $X_2 = 0$. As a result, Δ_G is an elliptic operator on the y-axis, where X_2 vanishes and Δ_G does not possess Riemannian geometry. Nevertheless, since

$$\begin{aligned}[X_1, X_2] = X_1(X_2) - X_2(X_1) &= \frac{\partial}{\partial x}\left(x\frac{\partial}{\partial y}\right) - x\frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}\right) \\ &= \frac{\partial}{\partial y} + x\frac{\partial^2}{\partial x \partial y} - x\frac{\partial^2}{\partial y \partial x} = \frac{\partial}{\partial y},\end{aligned}$$

which indicates that $\{X_1, X_2\}$ satisfy the *bracket generating property*, *Chow-Rashevskii's Theorem* (1.2 on p. 358) may be invoked to conclude that any two points in $\mathbb{R}^2$ can be connected by a piecewise differentiable *horizontal curve*. Specifically, we have the following:

Theorem 3.1. For any two points $P(x_0, y_0)$ and $Q(x_1, y_0)$ on the same horizontal line $y = y_0$, there exists a single geodesic connecting P and Q. If $y_0 \neq y_1$, there are infinitely many geodesics connecting points $P(0, y_0)$ and $Q(x_1, y_1)$.

To examine the geometric and analytic properties of the Grushin operator, we will rely on the *Hamilton-Jacobi theory of bicharacteristics*. The Hamiltonian function associated with the step 2 Grushin operator can be expressed as follows:

$$H(x, y, \xi, \eta) = \frac{1}{2}\left(\xi^2 + x^2\eta^2\right), \tag{3.1}$$

where (x, y, ξ, η) are the coordinates of $T^*\mathbb{R}^2$—the cotangent bundle of $\mathbb{R}^2$. We are primarily interested in obtaining the *geodesics* of Δ_G, which are the projections of bicharacteristic curves of H into the (x, y) plane. To do so, we seek to find the solutions of the Hamiltonian system below associated with (3.1), which represent the geodesics between two points (x_0, y_0) and (x, y) in $\mathbb{R}^2$:

$$\begin{aligned} &\frac{dx}{ds} = H_\xi = \xi, && \frac{d\xi}{ds} = -H_x = -x\eta^2 \quad \left(\text{so } \frac{d^2x}{ds^2} = -x\eta^2\right), \\ &\frac{dy}{ds} = H_\eta = x^2\eta, && \frac{d\eta}{ds} = -H_y = 0, \end{aligned} \tag{3.2}$$

with, for our purpose, $s \in [0, 1]$ and boundary conditions being

$$x(0) = x_0, \quad y(0) = y_0, \quad x(1) = x, \quad y(1) = y. \tag{3.3}$$

The vector fields and the Hamiltonian associated with the Grushin operator present us with some properties that allow our local investigation of the operator to have a global effect. For example, since $\frac{\partial}{\partial x}$ and $x\frac{\partial}{\partial y}$ are translation invariant in the y-direction, we will have $y_0 = 0$ for the rest of this work. Moreover, the Hamiltonian system (3.2) is invariant under the transformation

$$(x, y, \xi, \eta) \mapsto (x, -y, \xi, -\eta) \tag{3.4}$$

allowing us to concentrate only on the case of $y \geq 0$ and $\eta \geq 0$. Additionally, $\frac{d\eta}{ds}$ is 0, which implies that η is a constant. Intuitively, considering possible values of η is a convenient starting point for developing a rough idea about the solution to the system (3.2).

3.1. **Case I:** $\eta = 0$.

When $\eta = 0$, we have $\frac{d\xi}{ds} = -x * 0$, which implies that $\xi = \xi(0) := \xi_0$ and, after integrating and adjusting for the initial conditions, we have

$$x(s) = \xi_0 s + x_0, \quad y(s) = y_0 = C. \tag{3.5}$$

Since the vector fields are translation invariant in the y-direction, we consider $C = 0$. Thus, the geodesic in this case is a straight line segment connecting $(x_0, 0)$ and $(x, 0)$. To validate the claim of theorem (3.1), we assume conversely that $y = 0$:

$$0 = y = y(1) = \eta \int_0^1 x^2(u)\, du. \tag{3.6}$$

For (3.6) to hold, it must be the case that either $\eta = 0$ or $x(s) = 0$ for all s. For the latter case, one would also have that $\xi = 0$ and $x_0 = x = 0$. In this trivial case the two points coincide and η is arbitrary. We state these results in the following lemma.

Lemma 3.1. In the Hamiltonian system (3.2), the constant η is zero if and only if the initial conditions are $(x_0, 0)$ and $(x, 0)$ with $x_0 \neq x$ with the straight line connecting $(x_0, 0)$ to $(x, 0)$ being the only geodesic.

3.2. **Case II: $\eta > 0$.** From our discussion of the case $\eta = 0$ and Lemma 3.1, it follows that in the case of $\eta > 0$, so is $y > 0$. From the first row of the Hamiltonian system, we have

$$\frac{d^2x}{ds^2} = \frac{d\xi}{ds} = -x\eta^2,$$

which results in the following second order differential equation:

$$x'' + x\eta^2 = 0.$$

Solving the characteristic equation associated with the differential equation, we have the complex conjugate roots with $\Re(r) = 0$ and $\Im(r) = \eta$ and a general solution

$$x(s) = C_1 \cos \eta s + C_2 \sin \eta s.$$

Applying the initial conditions, we get

$$x(0) = x_0 = C_1 \cos(0) + C_2 \sin(0) \Rightarrow C_1 = x_0,$$

$$x'(0) = \xi(0) = \xi_0 = -\eta x_0 \sin(0) + C_2 \eta \cos(0) \Rightarrow C_2 = \frac{\xi_0}{\eta},$$

and conclude that

$$x(s) = x_0 \cos \eta s + \frac{\xi_0}{\eta} \sin \eta s. \tag{3.7}$$

The calculation of $y(s)$ is more involved and requires some use of trigonometric identities. For ease of readability, we let $B = \xi_0/\eta$.

$$\begin{aligned} y(s) &= \eta \int_0^s x^2(u)\, du \qquad (3.8)\\ &= \eta\Bigg(x_0^2 \int_0^s \frac{1+\cos(2\eta u)}{2}\, du \\ &\quad + x_0 B \int_0^s \sin(2\eta u)\, du + B^2 \int_0^s \frac{1-\cos(2\eta u)}{2}\, du \Bigg). \qquad (3.9) \end{aligned}$$

The three integrals of (3.8) are calculated below:

$$\begin{aligned} x_0^2 \int_0^s \frac{1+cos(2\eta u)}{2}\, du &= x_0^2 \left(\frac{s}{2} + \int_0^s \frac{\cos(2\eta u)}{2}\, du \right) \\ &= x_0^2 \left(\frac{s}{2} + \frac{\sin(2\eta s)}{4\eta} \right) \qquad (3.10) \end{aligned}$$

$$x_0 B \int_0^s \sin(2\eta u)\, du = x_0 B \left(\frac{1}{2\eta} - \frac{\cos(2\eta s)}{2\eta} \right) \tag{3.11}$$

$$B^2 \int_0^s \frac{1-\cos(2\eta u)}{2}\, du = B^2 \left(\frac{s}{2} - \frac{\sin(2\eta s)}{4\eta} \right) \tag{3.12}$$

Adding (3.10), (3.11), (3.12) together and multiplying by η, gives us

$$y(s) = \eta \left(\frac{1}{2}(x_0^2 + B^2)s + \frac{1}{4\eta}(x_0^2 - B^2)\sin(2\eta s) + \frac{x_0 B}{\eta} \sin^2(\eta s) \right). \tag{3.13}$$

Remark 3.1. Our solution for x (3.7) hints at further categorization of η. Namely, when $\eta = k\pi$ for some $k \in \mathbb{N}$, then $x(s)$ reduces to $(-1)^k x_0$.

3.3. **Case II.A:** $\eta \neq k\pi$**.** For this case, after adjusting for the initial conditions we get

$$x(1) = x = x_0 \cos\eta + B\sin\eta,$$

from which it follows that

$$B = \frac{x - x_0\cos\eta}{\sin\eta}. \tag{3.14}$$

Then, we may find y by applying (3.14) to (3.13) to obtain

$$\begin{aligned} y = y(1) &= \frac{1}{2}\left((x^2 + x_0^2)\Big(\frac{\eta - \sin\eta\cos\eta}{\sin^2\eta}\Big) + 2xx_0\Big(\frac{\sin\eta - \eta\cos\eta}{\sin^2\eta}\Big)\right) \\ &= \frac{1}{4}\Bigg((x + x_0)^2\left(\frac{(\eta + \sin\eta)(1-\cos\eta)}{\sin^2\eta}\right) \\ &\qquad\qquad + (x - x_0)^2\left(\frac{(\eta - \sin\eta)(1+\cos\eta)}{\sin^2\eta}\right)\Bigg). \end{aligned}$$

We write this result more succinctly, as

$$y = \frac{1}{4}\left((x + x_0)^2\tilde{\mu} + (x - x_0)^2\mu\right), \tag{3.15}$$

$$\tilde{\mu} = \frac{\eta + \sin(\eta)}{1 + \cos(\eta)}, \tag{3.16}$$

$$\mu = \frac{\eta - \sin(\eta)}{1 - \cos(\eta)}. \tag{3.17}$$

The functions $\tilde{\mu}$ and μ and their properties will be of special interest in our discussion. While η-s solving (3.15) give geodesics connecting $(x_0, 0)$ and (x, y), not all of the geodesics stem from solving (3.15). For instance, when $x = x_0 = 0$ and $y_0 \neq y > 0$, the solutions of (3.15) do not provide us with information about the geodesics connecting the two points. As a matter of course, we will refer to the geodesics arising from (3.15) as *generic*, and *exceptional* otherwise.

3.4. **Case II.B:** $\eta = k\pi$**.** In this case, by remark 3.1, we have $x = (-1)^k x_0$. Moreover, setting $s = 1$ in (3.13), we get

$$y = \frac{k\pi}{2}(x_0^2 + B^2) \Rightarrow B = \pm\left(\frac{2y}{k\pi} - x_0^2\right)^{\frac{1}{2}}. \tag{3.18}$$

From (3.18), we must have

$$\frac{2y}{k\pi} - x_0^2 \geq 0 \Rightarrow 2y \geq k\pi x_0^2. \tag{3.19}$$

The cases when $x_0 = 0$ and $x_0 \neq 0$ wil be treated separately later. We will now focus on the properties of functions $\tilde{\mu}$ and μ.

3.5. **Functions $\tilde{\mu}(\eta)$ and $\mu(\eta)$.** In this section we concentrate on $\tilde{\mu}(\eta)$, $\mu(\eta)$, and a function $F(\eta)$ defined as

$$F(\eta) = a^2\tilde{\mu}(\eta) + b^2\mu(\eta) \tag{3.20}$$

for $a, b \in \mathbb{R}$. Because of (3.4) we consider $\eta \geq 0$ only. Below, we summarize some of the properties of $\tilde{\mu}$ and μ that will be of use in our discussion.

Lemma 3.2. For $\eta \geq 0$ and $k \in \mathbb{N}$:

(1) Functions $\tilde{\mu}$ and μ are positive functions, vanishing at only $\eta = 0$.
(2) Function μ has poles when $\eta = 2k\pi$, for $k \in \mathbb{N}$. Function $\tilde{\mu}$, on the other hand, has poles when $\eta = (2k-1)\pi$.
(3) Function μ is *convex* in each interval $(2(k-1)\pi, 2k\pi)$ and function $\tilde{\mu}$ is *convex* in each interval $(2(k-1)\pi, 2(k+1)\pi)$.
(4) Function μ assumes the minimum value of α'_k in each interval $(2k\pi, 2(k+1)\pi)$, which satisfies $\tan(\alpha'_k/2) = \alpha'_k/2$. Function $\tilde{\mu}$ assumes the minimum value of $\alpha''_k \in ((2k-1)\pi, 2k\pi)$ in each interval $(2(k-1)\pi, 2(k+1)\pi)$, which satisfies $-\cot(\alpha''_k/2) = \alpha''_k/2$.
(5) At the respective minimums, we have:

$$\mu(\alpha'_k) = \frac{\alpha'_k}{2}, \quad \text{and} \quad \tilde{\mu}(\alpha''_k) = \frac{\alpha''_k}{2}.$$

Moreover,

$$\begin{aligned}\ldots < \mu((2k-1)\pi) = \left(k - \frac{1}{2}\right)\pi < \mu(\alpha'_k) < \mu((2k+1)\pi) = \\ = \left(k + \frac{1}{2}\right)\pi < \mu(\alpha_{k+})' < \ldots\end{aligned}$$

and, similarly,

$$\begin{aligned}\ldots < \tilde{\mu}(2(k-1)\pi) = (k-1)\pi < \tilde{\mu}(\alpha''_k) < \tilde{\mu}(2k\pi) = \\ = k\pi < \tilde{\mu}(\alpha_{k+})'' < \ldots\end{aligned}$$

Proof. We prove parts 3 through 5 of the lemma for $\mu(\eta)$. The cases for $\tilde{\mu}(\eta)$ are similar.

(3) For $\eta \geq 0$ and $\eta \neq k\pi, k \in \mathbb{N}$, the second derivative of function μ is

$$\mu'' = \frac{2\eta + \eta\cos\eta - 3\sin\eta}{(1-\cos\eta)^2} = \frac{\eta(1 + \cos\eta + \eta - 3\sin\eta)}{(1-\cos\eta)^2}.$$

Let $\psi(\eta)$ denote the numerator of μ''. To prove the convexity of $\tilde{\mu}$, it suffices to show that $\psi(\eta) > 0$ for $\eta \in (0, 3)$. Calculating the derivatives of $\psi(\eta)$ up to third order, we have

$$\begin{aligned}\psi'(\eta) &= 2 - \eta\sin\eta - 2\cos\eta\,,\\ \psi''(\eta) &= \sin\eta - \eta\cos\eta\,,\\ \psi'''(\eta) &= \eta\sin\eta\,.\end{aligned}$$

First, we note that $\psi'''(\eta)$ is always positive for $\eta \in (0, 3)$. This fact, coupled with $\psi''(0) = 0$, implies that $\psi''(\eta) > 0$ for $\eta \in (0, 3)$. Similarly, $\psi''(\eta) > 0$

for $\eta \in (0,3)$, coupled with $\psi'(0) = 0$, implies that $\psi'(\eta) > 0$ for $\eta \in (0,3)$. Finally, $\psi'(\eta) > 0$ for $\eta \in (0,3)$, coupled with $\psi(0) = 0$, implies that $\psi(\eta) = \mu''(\eta) > 0$ for $\eta \in (0,3)$. The convexity of the functions μ and $\tilde{\mu}$ can be observed in *Figure 1* below.

(4) We first calculate the first derivative of $\mu(\eta)$ and use the half-angle trigonometric identity to reach the desired result.

$$\mu'(\eta) = \frac{2 - 2\cos\eta - \eta\sin\eta}{(1-\cos\eta)^2}.$$

At its minimum, the derivative of μ equals 0, implying that

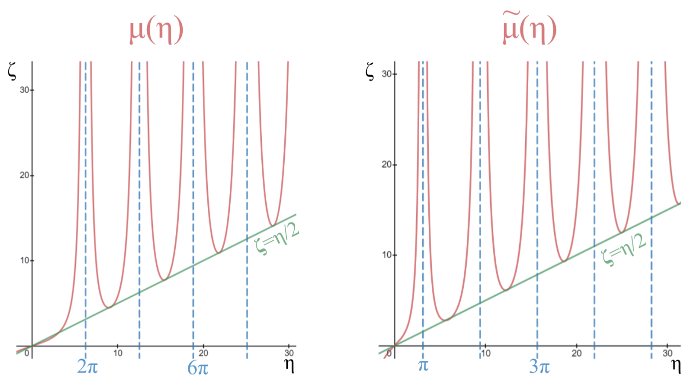
$$2 - 2\cos\eta - \eta\sin\eta = 0,$$
$$\tan\frac{\eta}{2} = \frac{1-\cos\eta}{\sin\eta} = \frac{\eta}{2}$$

(5) Using our previous result that at its minimum α_k', $\tan\alpha_k'/2 = \alpha_k'/2$, we have

$$\mu(\alpha_k') = \frac{\frac{\alpha_k'}{2} - \sin\frac{\alpha_k'}{2}\cos\frac{\alpha_k'}{2}}{\sin^2\frac{\alpha_k'}{2}} = \frac{1-\cos^2\frac{\alpha_k'}{2}}{\sin\frac{\alpha_k'}{2}\cos\frac{\alpha_k'}{2}}$$
$$= \frac{\sin^2\frac{\alpha_k'}{2}}{\sin^2\frac{\alpha_k'}{2}\cos\frac{\alpha_k'}{2}} = \tan\frac{\alpha_k'}{2} = \frac{\alpha_k'}{2}.$$

□

FIGURE 1. Graphical Representations of functions μ and $\tilde{\mu}$



The graphs above highlight some of the features of functions μ and $\tilde{\mu}$, such as their convexity, positivity, and unique minimums, for $\eta \geq 0$.
Lemma 3.2 serves as the underpinning for our next theorem concerning (3.20).

Theorem 3.2. For $ab \neq 0$ and $k \in \mathbb{N}$, the function $F(\eta) = a^2\tilde{\mu}(\eta) + b^2\mu(\eta)$ has the following properties:

(1) $F(\eta) \geq 0$ for $\eta \geq 0$ and vanishes only at $\eta = 0$.
(2) $F(\eta)$ has poles at $\eta = k\pi$.
(3) $F(\eta)$ is strictly convex in each interval $((k-1)\pi, k\pi)$.
(4) $F(\eta)$ takes a unique minimum at α_k in each interval $(k\pi, (k+1)\pi)$ such that $\alpha_{2k-1} < \alpha_k''$, $\alpha_{2k} < \alpha_k'$, and

$$F(\alpha_k) > \begin{cases} a^2\alpha_k', & k \text{ odd} \\ b^2\alpha_k'', & k \text{ even.} \end{cases}$$

(5) $F(\alpha_k) < F(\alpha_{k+1})$. Moreover, when $a = 0$ and b is nonzero, F reduces to $b^2\mu$ and when $b = 0$ and a is nonzero, F reduces to $a^2\tilde{\mu}$.

Proof. Properties (1) through (4) are immediate consequences of Lemma 3.2. Thus, we concentrate on proving property (5). Let

$$g(\eta) = \sin \frac{a^2}{1+\cos\eta} - \frac{b^2}{1-\cos\eta}$$

and, accordingly,

$$g'(\eta) = \left(\frac{a^2}{1+\cos\eta} + \frac{b^2}{1-\cos\eta}\right),$$

which is positive for all $\eta \in \mathbb{R} - \{k\pi | k \in \mathbb{Z}\}$. Note that F can be expressed in terms of g using the product rule as follows:

$$F(\eta) = \big(\eta g(\eta)\big)' = g(\eta) + \eta g'(\eta) = a^2\frac{\eta + \sin\eta}{1+\cos\eta} + b^2\frac{\eta - \sin\eta}{1-\cos\eta}. \tag{3.21}$$

Function g's periodicity of 2π, which can be observed in *Figure 2* below, will aid us in reaching our goal. Specifically,

$$g(\eta) = -g(2k\pi - \eta), \quad \text{for } \eta \in (k\pi, (k+1)\pi). \tag{3.22}$$

By part (4) of the theorem, $F(\eta)$ achieves a unique minimum at $\alpha_k' \in (k\pi, (k+1)\pi)$. By (3.22), for $\eta \in (k\pi, (k+1)\pi)$ we have:

$$g(\eta) = -g(2k\pi - \eta) \quad \text{and} \quad g'(\eta) = g'(2k\pi - \eta).$$

Combining this result with (3.21) we get

$$\begin{aligned} F(\eta) &= -g(2k\pi - \eta) + \eta g'(2k\pi - \eta) \\ &= g(2k\pi - \eta) + (2k\pi - \eta)g'(2k\pi - \eta) \\ &\quad + 2\big(-g(2k\pi - \eta) + (\eta - k\pi)g'(2k\pi - \eta)\big) \\ &= F(2k\pi - \eta) + 2\big(g(\eta) + (\eta - k\pi)g'(\eta)\big). \end{aligned} \tag{3.23}$$

If k is even, then, for $\eta \in (k\pi, (k+1)\pi)$, (3.23) reduces to

$$F(\eta) = F(2k\pi - \eta) + 2F(\eta - k\pi),$$

which implies that

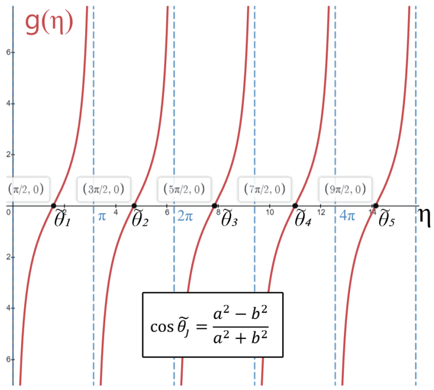
$$F(\alpha_k) = F(2k\pi - \alpha_k) + 2F(\alpha_k - k\pi) > F(2k\pi - \alpha_k) \geq F(\alpha_{k-1}),$$

since $2F(\alpha_k - k\pi)$ is positive for all $k \in \mathbb{N}$. A similar argument shows that when k is odd, (3.23) implies that $F(\eta) > F(2k\pi - \eta)$, for $\eta \in (k\pi, (k+1)\pi))$ and, as a result,

$$F(\alpha_k) > F(2k\pi - \alpha_k) \geq F(\alpha_{k-1}).$$

□

FIGURE 2. Graphical Representation of $g(\eta)$ for $a = b = 1$.



This theorem has now equipped us with the necessary tools for studying the geodesics of the Grushin operator.

3.6. **The geodesics.**

3.6.1. *Generic geodesics.* Recall that the generic geodesics connecting $(x_0, 0)$ and (x, y) arise from η-s that solve (3.15) (page 365), which is equivalent to $F(\eta)$ with $a = \frac{x+x_0}{2}$ and $b = \frac{x-x_0}{2}$. Within the generic case, we consider the geodesics connecting $(x_0, 0)$ to points (x, y) with varying x and y.

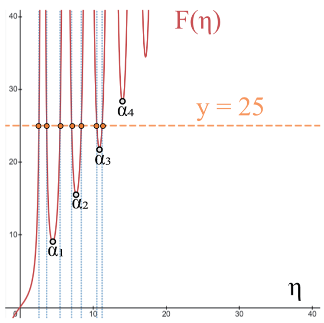
Case I: y is positive, and $x^2 \neq x_0^2$.

By Remark 3.1, $\eta = k\pi$ would imply that $x = (-1)^k x_0$, meaning that $x^2 = x_0^2$. Hence, $\eta = k\pi$ is not a solution. In this case, however, Theorem 3.2.5 suggests that for $y > 0$ there exists $N \in \mathbb{N}$ such that $F(\alpha_{N-1}) \leq y \leq F(\alpha_N)$. Furthermore, by Theorem 3.2.1-3, we may conclude that there are $2N - 1$ solutions to (3.15). If $y = F(\alpha_{N-1})$, then α_{N-1} is counted as a solution of multiplicity two. Thus, in either case there are always $2N - 1$ geodesics connecting $(x_0, 0)$ to (x, y).

Example 1. In this example, we find the number of geodesics of the step 2 Grushin operator connecting point $(0, 0)$ to $(2, 25)$. Since $\alpha_3 < y = 25 < \alpha_4$, the number of geodesic connecting the two points are 2(4)-1=7, which can also be observed in *Figure 3* in the form of intersections of $F(\eta)$ and $y = 25$. Plugging in the values of $x_0 = 0$ and $x = 2$ into (3.15) results in

$$F(\eta) = \frac{(2+0)^2}{4}\frac{\eta + \sin\eta}{1 + \cos\eta} + \frac{(2-0)^2}{4}\frac{\eta - \sin\eta}{1 - \cos\eta} = \frac{\eta + \sin\eta}{1 + \cos\eta} + \frac{\eta - \sin\eta}{1 - \cos\eta},$$

Figure 3. Graph of $F(\eta)$ from Example 1



the graph of which illustrates the logic behind counting the number of geodesics.

Case II: y is positive and $x = x_0 \neq 0$.
In this case,

$$F(\eta) = \frac{(x_0 + x_0)^2}{4}\tilde{\mu}(\eta) + \frac{(x_0 - x_0)^2}{4}\mu(\eta) = \frac{4x_0^2}{4}\tilde{\mu}(\eta) = x_0^2\tilde{\mu}(\eta).$$

As a result, we are interested in finding η-s that satisfy $\tilde{\mu}(\eta) = y/x_0^2$. From our discussion of function $\tilde{\mu}$ and Lemma 3.2, it follows that there exists $N \in \mathbb{N}$ such that $\tilde{\mu}(\alpha''_{N-1}) \leq y/x_0^2 = \tilde{\mu}(\eta) < \tilde{\mu}(\alpha''_N)$. Similar to Case I, we conclude that there are $2N - 1$ geodesics connecting $(x_0, 0)$ and (x_0, y).

Case III: y is positive and $x = -x_0 \neq 0$.
In this case,

$$F(\eta) = \frac{(-x_0 + x_0)^2}{4}\tilde{\mu}(\eta) + \frac{(-x_0 - x_0)^2}{4}\mu(\eta) = \frac{4x_0^2}{4}\mu(\eta) = x_0^2\mu(\eta).$$

Again, by Lemma 3.2, we have that there exists $N \in \mathbb{N}$ such that $\mu(\alpha'_{N-1}) \leq y/x_0^2 = \mu(\eta) < \mu(\alpha'_N)$. Thus, similar to the previous cases, we conclude that there are $2N - 1$ geodesics connecting $(x_0, 0)$ to $(-x_0, y)$.

Case IV: $y = 0$ and $x \neq x_0$.
In this case, $\eta = 0$ is the only solution to (3.15). While we have derived the explicit form of the geodesic in section 2.1 (see (3.5) and (3.6) on page 363), we now show its consistency with (3.7) for $y \to 0^+$ and $\eta \to 0^+$. Since $x \neq x_0$ and y is positive, we have by Case I that there exists $\eta \in (0, \pi)$ solving (3.7). Then, (3.7) can be written as:

$$x(s) = x_0 \cos \eta s + \frac{\eta(x - x_0 \cos \eta)}{\sin \eta} \frac{\sin \eta s}{\eta s} s.$$

Letting $y \to 0^+$ and $\eta \to 0^+$, we have:

$$x(s) = x_0 + (x - x_0)s,$$

which is consistent with our finding summarized in Lemma 3.1.

To develop an all-encompassing theorem on geodesics, we now turn to the exceptional geodesics.

3.6.2. *Mild exceptional case: $y > 0$, $x^2 = x_0^2 \neq 0$ with η being an integer multiple of π but not a solution to* (3.15). As mentioned before, these geodesics are called exceptional, because they are not solutions of (3.15). In the case when $\eta \equiv 0 \bmod \pi$, either $\mu(\eta)$ or $\tilde{\mu}(\eta)$ disappears in $F(\eta)$; hence, the name mild exceptional. In this subsection, we advance our discussion of Case II.B and show that geodesics defined by (3.18) are limits of the generic ones.

Lemma 3.3. For a given $y > 0$, $x_0 \neq 0$, and a positive even integer k such that $y > k\pi x_0^2/2$, we have that for any $\epsilon \in (0, (k\pi - \alpha''_{k/2})/2)$, there exists $\delta > 0$ such that for any x_0, for which $0 < |x - x_0| < \delta$, there exist two geodesics connecting $(x_0, 0)$ to (x, y) with $\eta = \eta^+$ or η^-, $\eta^- < k\pi < \eta^+$, and $|\eta^\pm - k\pi| < 2\epsilon$.

Proof. We will only concentrate on the case when $x = x_0 \neq 0$, as the case of $x = -x_0 \neq 0$ can be treated in a similar way.

Let $y > 0$ and $x_0 \neq 0$ be fixed. Recall from Remark 3.1 that when $x = x_0$ and for $\eta = k\pi$ to define a geodesic connecting $(x_0, 0)$ to (x_0, y), k must be an even integer and (3.19) must hold. As a result, we have:

$$y > \frac{k\pi x_0^2}{2} = x_0^2 \tilde{\mu}(\eta).$$

Then, by Lemma 3.2, we have that $\alpha''_{k/2} < k\pi$ and $\tilde{\mu}(\eta)$ is increasing for $\eta \in (\alpha''_{k/2}, (k+1)\pi)$. Given $\epsilon > 0$ and $\epsilon \in (0, (k\pi - \alpha''_{k/2})/2)$, let $x \neq x_0$ but x close to x_0. By definitions of $\mu(\eta)$ and $\tilde{\mu}(\eta)$, there exists $\delta > 0$ such that

$$\begin{cases} \frac{(x+x_0)^2}{4}\tilde{\mu}(\eta) < \frac{k\pi x_0^2}{2} + \frac{1}{2}\left(y - \frac{k\pi x_0^2}{2}\right) = \frac{y}{2} + \frac{k\pi x_0^2}{4} \\ \qquad\qquad \text{if } |x - x_0| < \delta \text{ and } |\eta - k\pi| < 2\epsilon \\ \frac{(x-x_0)^2}{4}\mu(\eta) < \frac{1}{2}\left(y - \frac{k\pi x_0^2}{2}\right) \\ \qquad\qquad \text{if } |x - x_0| < \delta \text{ and } \epsilon < |\eta - k\pi| < 2\epsilon \end{cases}.$$

Remark 3.2. Note that the above results easily follow from the assumptions made. For example, since $x = x_0$, k is even, and $y - k\pi x_0^2/2 > 0$, we have that

$$\frac{(x+x_0)^2}{4}\tilde{\mu}(\eta) = \frac{4x_0^2}{4}\frac{k\pi}{2} = \frac{k\pi x_0^2}{2} < \frac{k\pi x_0^2}{2} + \frac{1}{2}\left(y - \frac{k\pi x_0^2}{2}\right).$$

Thus, it follows that if $0 < |x - x_0| < \delta$,

$$\min_{\eta \in (k\pi+\epsilon, (k+1)\pi)} F(\eta) < y \quad \text{and} \quad \max_{\eta \in (k\pi - 2\epsilon, k\pi - \epsilon)} F(\eta) < y.$$

Invoking Theorem 3.2, we have two solutions, η^+ and η^-, with $\eta^- < k\pi < \eta^+$ that solve $y = F(\eta^\pm)$. Moreover, the inequalities above imply that $|\eta^\pm - k\pi| < 2\epsilon$, thus proving the lemma. □

Now, for simplicity, we may rewrite (3.7) as:

$$x(s) = A\sin(\eta s + \alpha) \tag{3.24}$$

where

$$A = \sqrt{x_0^2 + B^2} \quad \text{and} \quad \sin\alpha = \frac{x_0}{\sqrt{x_0^2 + B^2}}.$$

Consequently, (3.13) becomes:

$$y(s) = \eta A^2 \int_0^s \sin(\eta u + \alpha)\, du = \eta A^2 \left(\frac{s}{2} - \frac{\sin 2(\eta s + \alpha)}{4\eta} + \frac{\sin 2\alpha}{4\eta} \right). \tag{3.25}$$

Note that when $s = 1$, we have that

$$y = y(1) = \frac{\eta}{2}(x_0^2 + B^2) \left(1 - \frac{\sin\eta \cos 2\alpha + \eta}{\eta} \right). \tag{3.26}$$

Solving (3.26) for B, we get that

$$B = \pm\sqrt{\frac{2y}{\eta - \sin\eta\cos\eta + 2\alpha} - x_0^2}. \tag{3.27}$$

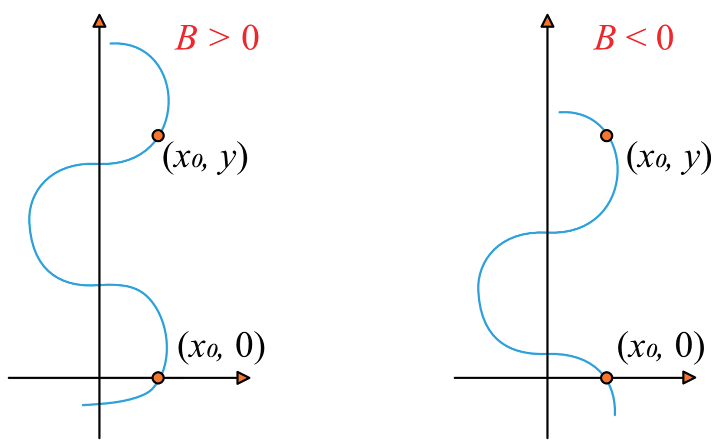
Remark 3.3. B in (3.27) is in C^∞ if $\frac{2y}{\eta - \sin\eta\cos\eta + 2\alpha} > x_0^2$. Moreover, the sign of B is determined by relations (3.7) and (3.14).

For $x_0, y, \eta = \eta^\pm$ as in Lemma 3.3, we have:

$$B = \pm\sqrt{\frac{2y}{k\pi} - x_0^2}, \quad \text{as} \quad \eta \to k\pi,$$

thus recovering (3.18). Hence, if $2y > k\pi x_0^2$ and $x \to x_0$, the two geodesics with $\eta = \eta^+$ or η^- tend to *two distinct geodesics* with $B > 0$ and $B < 0$, respectively (see *Figure 4*). When $2y = k\pi x_0^2$, consider $\tilde{y} > k\pi x_0^2/2$ and let $x = x_0$. As $\tilde{y}$ decreases to $k\pi x_0^2/2$, $B \to 0$ and the two distinct geodesics become one. However, this geodesic does not have multiplicity of two, but rather a multiplicity of three, since another geodesic merges by Case I.

FIGURE 4. Two Distinct Geodesics for $B > 0$ and $B < 0$.



Going back to (3.14), when $\eta^{\pm} \to 2k\pi^{\pm}$ (as opposed to simply specifying k to be an even integer) and $x - x_0 \to 0$, we have

$$\lim_{\substack{\eta^{\pm}\to 2k\pi^{\pm} \\ x-x_0\to 0}} \frac{x-x_0}{\sin\eta^{\pm}} = \lim_{\substack{\eta^{\pm}\to 2k\pi^{\pm} \\ x-x_0\to 0}} \frac{x - x_0\cos\eta^{\pm} - x_0(1-\cos\eta^{\pm})}{\sin\eta^{\pm}}$$

$$(3.28) \qquad = \lim_{\substack{\eta^{\pm}\to 2k\pi^{\pm} \\ x-x_0\to 0}} \frac{x - x_0\cos\eta^{\pm}}{\sin\eta^{\pm}} = B,$$

where

$$B > 0 \text{ if } \begin{cases} \eta^{+} \to 2k\pi^{+}, x > x_0 \text{ or} \\ \eta^{-} \to 2k\pi^{-}, x < x_0, \end{cases}$$

and

$$B < 0 \text{ if } \begin{cases} \eta^{+} \to 2k\pi^{+}, x < x_0 \text{ or} \\ \eta^{-} \to 2k\pi^{-}, x > x_0 \end{cases}.$$

Next, we consider (3.14) with $\eta^{\pm} \to (2k-1)\pi^{\pm}$ and $x + x_0 \to 0$:

$$\lim_{\substack{\eta^{\pm}\to (2k-1)\pi^{\pm} \\ x+x_0\to 0}} \frac{x+x_0}{\sin\eta^{\pm}} = \lim_{\substack{\eta^{\pm}\to (2k-1)\pi^{\pm} \\ x+x_0\to 0}} \frac{x - x_0\cos\eta^{\pm} + x_0(1+\cos\eta^{\pm})}{\sin\eta^{\pm}}$$

$$(3.29) \qquad = \lim_{\substack{\eta^{\pm}\to (2k-1)\pi^{\pm} \\ x+x_0\to 0}} \frac{x - x_0\cos\eta^{\pm}}{\sin\eta^{\pm}} = B,$$

where

$$B > 0 \text{ if } \begin{cases} \eta^{+} \to (2k-1)\pi^{+}, x + x_0 > 0 \text{ or} \\ \eta^{-} \to (2k-1)\pi^{-}, x - x_0 < 0, \end{cases}$$

and

$$B < 0 \text{ if } \begin{cases} \eta^{+} \to (2k-1)\pi^{+}, x + x_0 < 0 \text{ or} \\ \eta^{-} \to (2k-1)\pi^{-}, x + x_0 > 0 \end{cases}.$$

Thus, the total number of mild exceptional geodesics connecting $(x_0, 0)$ to (x_0, y) is $2(N-1)$, whereas the total number of mild exceptional geodesics connecting $(x_0, 0)$ to $(-x_0, y)$ is $2N$. From Case II of the generic geodesics, we have $\alpha''_N \in ((2N-1)\pi, 2N\pi)$ and $\alpha'_N \in (2N\pi, (2N+1)\pi)$. As a result, the number of geodesics connecting $(x_0, 0)$ to (x_0, y) and $(x_0, 0)$ to $(-x_0, y)$ is $2N-1$. For the former case, the total number of geodesics is $2N - 1 + 2(N-1) = 2(2N-1) - 1$; whereas in the latter case, the total number is $2N - 1 + 2N = 2(2N) - 1$.

For the cases when $x = x_0 \neq 0$ and for the mild exceptional case, *Figure 5* and *Figure 6* illustrate η's. In *Figure 5*, η_1, η_2, and η_5 correspond to generic geodesics, while η_3 and η_4 correspond to the mild exceptional geodesics. Note that as y decreases to πx_0^2, η_5 approaches and merges with η_3 and $\eta_4 = 2\pi$, which corresponds to a geodesic of multiplicity three.

3.6.3. *The exceptional case: $y > 0, x = x_0 = 0$.* Making a similar argument as with the mild exceptional case, for $x_0 = 0$, we have, for $k \in \mathbb{N}$,

$$B = \pm\sqrt{\frac{2y}{k\pi}}.$$

Hence, there exist infinitely many geodesics connecting the origin, $(0,0)$ to $(0,y)$, with $\eta_{2k} = \eta_{2k-1} = k\pi$ for $k \in \mathbb{N}$. The geodesics are defined by the following characteristic curves:

$$\begin{cases} x(s) = \pm\sqrt{\frac{2y}{k\pi}} \sin(k\pi s) \\ y(s) = y\left(s - \frac{\sin(2k\pi s)}{2k\pi}\right) \end{cases}.$$

Thus, every $k\pi$, for $k \in \mathbb{N}$ corresponds to two distinct geodesics, as illustrated in *Figure 7.*

3.6.4. *The main theorem on geodesics.* Now, we are ready to summarize our findings and state them in the following theorem.

Theorem 3.3. Given points $(x_0, 0)$ and (x, y) in the plane and $y > 0$,

(1) If $x \neq x_0$, then (3.15) has finitely many solutions η_j for $j \in \{1, 2 \ldots N\}$, where N is odd and

$$0 < \eta_1 < \pi < \eta_2 \leq \eta_3 < 2\pi < \ldots < \frac{N-1}{2}\pi < \eta_{N-1} \leq \eta_N < \frac{N+1}{2}\pi,$$

where $\eta_{N-1} = \eta_N$ occurs when $\eta_N = \alpha_{(N-1)/2}$.
The geodesics connecting $(x_0, 0)$ to (x, y) are defined by the curves with the following parametric equations, with $s \in [0, 1]$:

$$\mathcal{C}_j : \begin{cases} x(s) = x_0 \cos(\eta_j s) + \frac{x - x_0 \cos\eta_j}{\sin\eta_j} \sin(\eta_j s) \\ y(s) = \eta_j \int_0^s x^2(u)\, du \end{cases}. \tag{3.30}$$

FIGURE 5

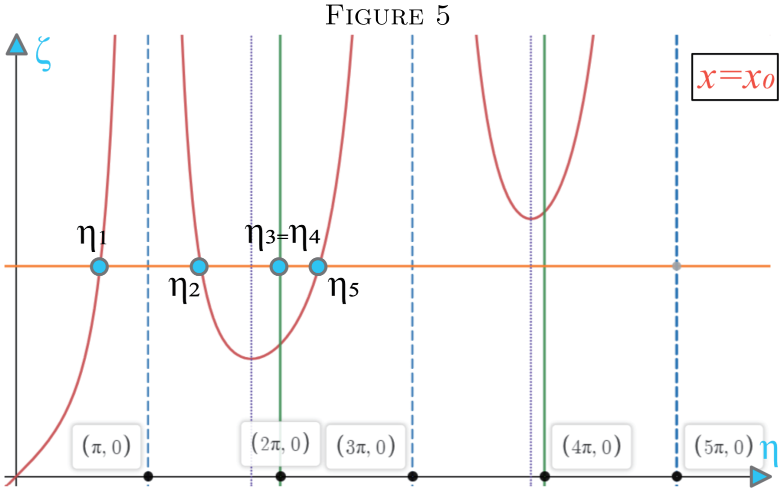


FIGURE 6

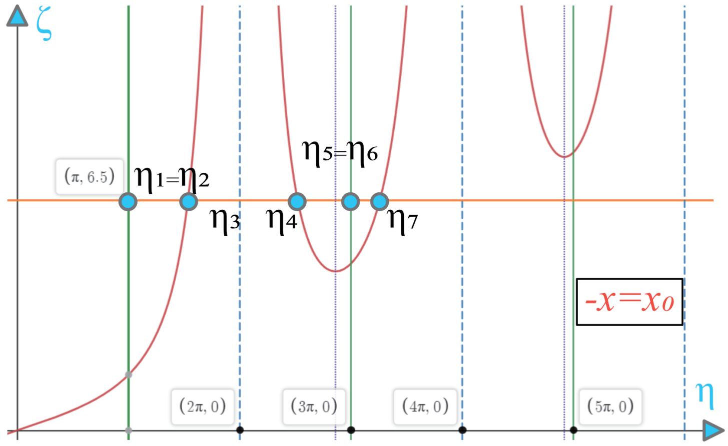
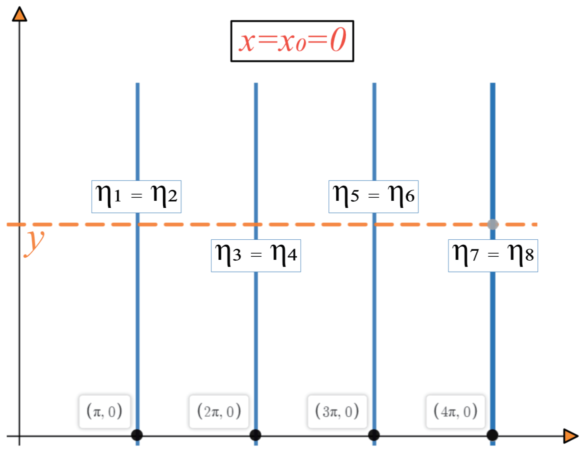


FIGURE 7



(2) If $x = x_0 \neq 0$, then there exist finitely many geodesics $\mathcal{C}_j$ for $j \in \{1, 2 \dots N\}$, N odd, connecting $(x_0, 0)$ to (x_0, y). Moreover, the corresponding η_j-s satisfy

$$0 < \eta_1 < \pi < \eta_2 < \eta_3 = 2\pi = \eta_4 < \eta_5 < \dots < \eta_{4k-2} < \eta_{4k-1} = 2k\pi = \eta_{4k}$$
$$\eta_{4k+1} < \dots < \eta_{N-2} \leq \eta_{N-1} \leq \eta_N \leq \frac{N+1}{2}\pi,$$

where if $N \equiv 3 \mod 4$ and $\eta_N = \alpha''_{(N-1)/2}$, then $\eta_{N-2} < \frac{N-1}{2}\pi < \eta_{N-1} = \eta_N < \frac{N+1}{2}\pi$; if $N \equiv 1 \mod 4$ and $\eta_N = \frac{N-1}{2}\pi$, then $\frac{N-3}{2}\pi < \eta_{N-3} < \eta_{N-2} = \eta_{N-1} = \eta_N < \frac{N+1}{2}\pi$; and if $\eta_{N-2} = \eta_{N-1} = \frac{N-1}{2}\pi$ and $N \equiv 1 \mod 4$, then $\eta_{N-3} < \eta_{N-2} = \eta_{N-1} < \eta_N < \frac{N+1}{2}\pi$.

If $\eta \neq 2k\pi$, the corresponding geodesic has the form (3.30). When $y > k\pi x_0^2$, we have $\eta_{4k-1} = \eta_{4k} = 2k\pi$ with the two related geodesics being:

$$\mathcal{C}_{4k-1}, \mathcal{C}_{4k} : \begin{cases} x(s) = x_0 \cos(2k\pi s) \pm \frac{x_0 \sin(2k\pi s)}{\sqrt{2k\pi}} \sqrt{\frac{2y}{x_0^2} - 2k\pi} \\ y(s) = 2k\pi \int_0^s x^2(u)\, du \end{cases} . \tag{3.31}$$

If $y = k\pi x_0^2$, then we have $N \equiv 1 \mod 4$ and $C_{N-2} = C_{N-1} = C_N$. Thus, in this case, the geodesic has the same form as (3.31) with $k = (N-1)/4$ and $x(s) = x_0 \cos\left(\frac{N-1}{2}\pi s\right)$.

(3) If $x = -x_0 \neq 0$, then there exist finitely many geodesics $\mathcal{C}_j$ for $j \in \{1, 2 \ldots N\}$, N odd, connecting $(x_0, 0)$ to $(-x_0, y)$. Moreover, the corresponding η_j-s satisfy

$$\begin{aligned} 0 < \eta_1 = \pi = \eta_2 < \eta_3 < \eta_4 \ldots \eta_{4k-4} < \eta_{4k-3} = (2k-1)\pi = \eta_{4k-2} \\ < \eta_{4k-1} < \ldots < \eta_{N-2} \leq \eta_{N-1} \leq \eta_{N-1} < \frac{N+1}{2}\pi, \end{aligned}$$

where if $N \equiv 1 \mod 4$ and $\eta_N = \alpha'_{(N-1)/4}$, then $\eta_{N-2} < \eta_{N-1} = \eta_N$; if $N \equiv 3 \mod 4$ and $\eta_N = \frac{N-1}{2}\pi$, then $\eta_{N-2} = \eta_{N-1} = \eta_N$; and if $N \equiv 3 \mod 4$ and $\eta_{N-2} = \eta N - 1 = \frac{N-1}{2}\pi$, then $\eta_{N-2} = \eta N - 1 < \eta_N$.
If $\eta \neq (2k-1)\pi$, then the geodesic has the same form as (3.30).
If $2y > (2k-1)\pi x_0^2$, we have $\eta_{4k-3} = \eta_{4k-2} = (2k-1)\pi$ with the two related geodesics, $\mathcal{C}_{4k-3}$ and $\mathcal{C}_{4k-2}$, being:

$$\begin{cases} x(s) = x_0 \cos((2k-1)\pi s) \pm \frac{x_0 \sin((2k-1)\pi s)}{\sqrt{(2k-1)\pi}} \sqrt{\frac{2y}{x_0^2} - (2k-1)\pi} \\ y(s) = (2k-1)\pi \int_0^s x^2(u) du \end{cases} . \tag{3.32}$$

If $2y = (2k-1)\pi x_0^2$ we have $C_{N-2} = C_{N-1} = C_N$. In this case, the geodesic has the same form as in (3.32) with $k = \frac{N+1}{4}$ and $x(s) = x_0 \cos(\frac{N-1}{2}\pi s)$.

(4) If $x = x_0 = 0$ and $y > 0$, then there exist infinitely many geodesics connecting $(0, 0)$ to $(0, y)$. The η_j-s satisfy the relation and have the form as stated below:

$$\begin{aligned} \eta_1 = \pi = \eta_2 < \eta_3 = 2\pi < \eta_4 < \ldots < \eta_{4k-3} = (2k-1)\pi = \eta_{4k-2} \\ < \eta_{4k-1} = 2k\pi = \eta_{4k} < \ldots \end{aligned}$$

The geodesics are of the following form:

$$\mathcal{C}_{2k-1}, \mathcal{C}_{2k} : \begin{cases} x(s) = \pm\sqrt{\frac{2y}{k\pi}} \sin(k\pi s) \\ y(s) = y\left(s - \frac{\sin(2k\pi s)}{2k\pi}\right) \end{cases} . \tag{3.33}$$

where $\eta_{2k-1} = \eta_{2k} = k\pi$.

(5) If $y = 0$ and $x = x_0$, then there exists a unique geodesic connecting $(x_0, 0)$ to $(x, 0)$:

$$\mathcal{C} : \begin{cases} x(s) = x_0 + (x - x_0)s \\ \eta = 0;\ y(s) = 0 \end{cases} .$$

Remark 3.4. The geodesics in cases (2)-(5) are limits of the geodesics in case (1).

4. The modified complex action

The heat kernels of differential operators that model physical phenomena, such as heat conduction and diffusion, are especially of interest to applied scientists and their derivation can be highly non-trivial. We use the geometric mechanics method to find a modified complex action, which is essential in constructing the heat kernel for the Grushin operator. To give some insight into the problem, we look at the definition of the heat kernel in the context of the Grushin operator:

Definition 4.1. P_t is said to be the heat kernel for the Grushin operator if:

$$\begin{cases} \Delta_G P_t - \frac{\partial}{\partial t} P_t = 0, \;\; t > 0 \\ \lim_{t\to 0} P_t(x, x_0, y) = \delta(x - x_0)\delta(y) \\ P_t > 0 \end{cases} . \tag{4.1}$$

4.1. **Intuitive Approach.** To approach the problem of identifying the complex action for the Grushin operator, we look at the heat kernel of the Laplace-Beltrami operator, which admits Riemannian geometry.

Definition 4.2 (Laplace-Beltrami operator)**.** For $X_1, X_2, \ldots, X_n$ n linearly independent vector fields defined on an n-dimensional manifold $\mathcal{M}_n$, the Laplace-Beltrami operator is defined as:

$$\Delta = \frac{1}{2}\sum_{j=0}^{n} X_j^2 + \ldots,$$

where the "..." stands for lower order terms.

It has been shown (see [2–5]) that the heat kernel, at least on a local scale, is of the following form:

$$P_t(\mathbf{x}, \mathbf{x_0}) = \frac{1}{(2\pi t)^{n/2}} e^{-\frac{d(\mathbf{x},\mathbf{x_0})^2}{2t}} (a_0 + a_1 t + a_2 t^2 + \ldots),$$

where $d(\mathbf{x}, \mathbf{x_0})$ is the Riemannian distance between $\mathbf{x}$ and $\mathbf{x_0}$ on $\mathcal{M}_n$, given that the metric is induced by the orthonormal set $\{X_1, X_2, \ldots, X_n\}$, and a_j stand for functions of $\mathbf{x}$ and $\mathbf{x}_0$. Moreover,

$$\frac{\partial}{\partial t}\left(\frac{d(\mathbf{x}, \mathbf{x_0}^2)}{2t}\right) + \frac{1}{2}\sum_{j=1}^{n}\left(X_j \frac{d(\mathbf{x}, \mathbf{x_0})^2}{2t}\right)^2 = 0, \tag{4.2}$$

meaning that $d\frac{(\mathbf{x},\mathbf{x_0})^2}{2}$ solves the Hamilton-Jacobi equation.

In the case of the Grushin operator, the heat kernel will be of a similar form:

$$\frac{1}{t^\alpha} e^{-h}, \tag{4.3}$$

where h, similar to (4.2), solves

$$\frac{\partial h}{\partial t} + \frac{1}{2}\left(\frac{\partial h}{\partial x}\right)^2 + \frac{x^2}{2}\left(\frac{\partial h}{\partial y}\right) = 0. \tag{4.4}$$

Hence, we are interested in solving

$$\frac{\partial z}{\partial t} + H\left(x, y, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) = 0, \tag{4.5}$$

for z, with H being

$$H(x, y, \xi, \eta) = \frac{1}{2}\xi^2 + \frac{1}{2}x^2\eta^2,$$

for $\xi = \frac{\partial z}{\partial x}$ and $\eta = \frac{\partial z}{\partial y}$. Letting $\gamma = \frac{\partial z}{\partial t}$, we have

$$\tilde{H}(x, y, t, z, \xi, \eta, \gamma) = \gamma + H(x, y, \xi, \eta) = 0. \tag{4.6}$$

Accordingly, we are interested in finding the bicharacteristic curves solving the following system, where ($'$) denotes the derivative with respect to s:

$$\begin{aligned}
x' &= \tilde{H}_\xi = \xi, \\
y' &= \tilde{H}_\eta = \eta x^2, \\
t' &= \tilde{H}_\gamma = 1, \\
\xi' &= -\tilde{H}_x = -\eta^2 x, \\
\eta' &= -\tilde{H}_y = 0, \\
\gamma' &= -\tilde{H}_t = 0, \\
z' &= \xi\tilde{H}_\eta + \eta\tilde{H}_\eta + \gamma\tilde{H}_\gamma.
\end{aligned} \tag{4.7}$$

This system is an extended system of (3.2). Note that for $0 \le s \le t$,

$$\begin{aligned}
\eta(s) &= \eta(0) = \eta, \\
\gamma(s) &= \gamma = -H, \\
t(s) &= s,
\end{aligned}$$

meaning that η and γ remain constant along the characteristic curves. From the system, we also have:

$$x'' = \xi' \quad \Rightarrow \quad x'' + \eta^2 x = 0. \tag{4.8}$$

Hence, by solving (4.8) and accounting for the initial conditions, we get:

$$x(s) = x(0)\cos\eta s + \frac{\xi(0)}{\eta}\sin\eta s. \tag{4.9}$$

Solving for $\xi(0)/\eta$, we get:

$$\frac{\xi(0)}{\eta} = \frac{x(t) - x(0)\cos(\eta t)}{\sin(\eta t)}, \tag{4.10}$$

for $\sin(\eta t) \neq 0$. For y, we have:

$$y'(s) = \eta x^2(s), \tag{4.11}$$

where

$$x^2(s) = x^2(0)\left(\frac{1+\cos(2\eta s)}{2}\right) + \frac{2x(0)\xi(0)}{\eta}\sin(\eta s)\cos(\eta s) + \frac{\xi^2(0)}{\eta^2}\left(\frac{1-\cos(2\eta s)}{2}\right).$$

Then, using (4.10), (4.11), and some trigonometric identities, we get:

$$\begin{aligned}
y(s) - y(0) &= \eta\Bigg(\frac{x^2(0)}{2}\left(s + \frac{\sin(2\eta s)}{2\eta}\right) + x(0)\frac{\eta(0)}{\eta^2}\sin^2(\eta s) \\
&\quad + \frac{\xi^2(0)}{\eta^2}\left(s + \frac{\sin(2\eta s)}{2\eta}\right)\Bigg) \\
&= \frac{\eta}{2}\left(x^2(0) + \frac{\xi^2(0)}{\eta^2}\right)s + \frac{\sin(2\eta s)}{4}\left(x^2(0) - \frac{\xi^2(0)}{\eta^2}\right) \\
&\quad + \frac{x(0)\xi(0)}{2\eta}(1-\cos(2\eta s)) \\
&= \frac{x^2(0)}{2}\left(\eta s + \frac{\sin(2\eta s)}{2}\right) \\
&\quad + \frac{x(0)\big(x(t) - x(0)\cos(\eta t)\big)}{2\sin(\eta t)}(1-\cos(2\eta s)) \\
&\quad + \frac{x^2(t) - 2x(t)x(0)\cos(\eta t) + x^2(0)\cos^2(\eta t)}{2\sin^2(\eta t)}\left(\eta s - \frac{\sin(2\eta s)}{2}\right) \\
&= \frac{x^2(0)}{4\sin^2(\eta t)}(2\eta s - \sin(2\eta t) + \sin(2\eta(t-s))) \\
&\quad + \frac{x^2(t)}{4\sin^2(\eta t)}(2\eta s - \sin(2\eta s)) \\
&\quad + \frac{x(0)x(t)}{2\sin^2(\eta t)}(\sin(2\eta s - \eta t) + \sin(\eta t) - 2\eta s\cos(\eta t)). \qquad (4.12)
\end{aligned}$$

Moreover, since $\gamma = -H$, we have:

$$z' = \xi x' + \eta y' - H,$$

$$z(t) = z(0) + \int_0^t z'(s)\,ds = z(0) + S(t), \qquad (4.13)$$

where $S(t)$ will be referred to as *the classical action*, which is given by

$$\begin{aligned}
S(t) &= \int_0^t (\xi x' + \eta y' - H)\,ds \\
&= \eta(y(t) - y(0)) + \int_0^t (\xi^2(s) - H(s))\,ds. \qquad (4.14)
\end{aligned}$$

From the system and (4.9), we have:

$$\xi(s) = x'(s) = \xi(0)\cos(\eta s) - \eta x(0)\sin(\eta s).$$

As for H, we may refer to the fact that $\gamma(s)$ is constant, to conclude that:

$$H = H(0) = \frac{1}{2}\left(\xi^2(0) + \eta^2 x^2(0)\right),$$

and that $S(t)$ may be rewritten as follows:

$$\begin{aligned} S(t) &= \eta(y(t) - y(0)) \\ &+ \int_0^t \left(\frac{\cos(2\eta s)}{2}(\xi^2(0) - \eta^2 x^2(0)) - \eta x(0)\xi(0)\sin(2\eta s)\right) ds \\ &= \eta(y(t) - y(0)) \\ &+ \frac{\sin(2\eta t)}{4\eta}\left(\xi^2(0) - \eta^2 x^2(0)\right) + \eta x(0)\xi(0)\frac{\cos(2\eta t) - 1}{2\eta}. \end{aligned}$$

Finally, after replacing $\xi(0)$ with (4.10) and some simplifications, we get:

$$S(t) = \eta(y(t) - y(0)) - \frac{\eta}{4}\left((x(t) + x(0))^2 \tan(\frac{\eta t}{2}) - (x - x(0))^2 \cot\frac{\eta t}{2}\right).$$

Remark 4.1. Notably, the Hamiltonian formalism yields in varying ways of calculating distances depending on the operator under the question. For example, consider (and compare to the $S(t)$ for the Grushin operator) the Laplace operator restricted to $\mathbb{R}^3$, which is defined as

$$\Delta = \frac{\partial^2}{\partial^2 x_1^2} + \frac{\partial^2}{\partial^2 x_2^2} + \frac{\partial^2}{\partial^2 x_3^2}.$$

Then, the related Hamiltonian function is:

$$H = \xi_1 + \xi_2 + \xi_3,$$

and Hamilton's equations for the bicharacteristic curves are:

$$\begin{cases} x_j'(s) = \frac{\partial H}{\partial \xi_j} \\ \xi_j'(s) = -\frac{\partial H}{\partial x_j} \end{cases},$$

with $s \in [0, \tau]$, $x(0) = 0$ and $x(\tau) = x$. Hence,

$$x_j'(s) = 2\xi_j \quad \text{and} \quad \xi_j'(s) = 0,$$

which imply that $\xi_j(s) = C_j$, meaning that ξ_j are constant along the characteristic curves. Solving the system, we get the following:

$$\begin{aligned} x_j'(s) &= 2C_j & \Rightarrow \quad & x_j(s) = 2C_j s + d_j, \\ x_j(0) &= 0 & \Rightarrow \quad & d_j = 0, \\ x_j(\tau) &= x_j & \Rightarrow \quad & C_j = \frac{x_j}{2\tau}, \end{aligned}$$

with

$$x_j(s) = \frac{x_j}{\tau}s, \qquad \text{and} \qquad \xi_j(s) = \frac{x_j}{2\tau}.$$

The action integral, in this case, yields a function of the distance:

$$\begin{aligned} S &= \int_0^\tau \left(\sum_{j=1}^{3} \xi_j(s) x_j'(s) - H\big(x(s), \xi(s)\big) \right) ds \\ &= \int_0^\tau \left(\frac{|x|^2}{2\tau^2} - \frac{|x|^2}{4\tau^2} \right) ds \\ &= \frac{|x|^2}{4\tau}. \end{aligned}$$

4.2. **Solving the Hamilton-Jacobi equation.** Recall that from (4.13) $z(t) = z(0) + S(t)$. While we have found $S(t)$, we lack $z(0)$ for solving the Hamilton-Jacobi equation. Therefore, instead of $z(t)$, we substitute $S(t)$ into (4.5) and find the difference. To do so, we first need to calculate the partial derivatives of $S(t)$ with respect to x, y, and t. Thus, for

$$S = S(x, x_0, y, \eta, t), \quad x(s) = x(x, x_0, y, \eta, t, s), \quad x_0 = x(0),$$

we have:

$$\begin{aligned} \frac{\partial S}{\partial x}(x, x_0, y, \eta, t) &= \int_0^t \bigg(\frac{\partial \xi}{\partial x} + \xi \frac{d}{ds}\frac{\partial x(s)}{\partial x} + \frac{\partial \eta}{\partial x}\frac{dy}{ds} + \eta \frac{d}{ds}\frac{\partial y(s)}{\partial x} \\ &\quad - \frac{\partial H}{\partial \xi}\frac{\partial \xi}{\partial x} - \frac{\partial H}{\partial \eta}\frac{\partial \eta}{\partial x} - \frac{\partial H}{\partial x}\frac{\partial x(s)}{\partial x} - \frac{\partial H}{\partial y}\frac{\partial y(s)}{\partial x} \bigg) ds \\ &= \int_0^t \left(\xi \frac{d}{ds}\frac{\partial x(s)}{\partial x} + \eta \frac{d}{ds}\frac{\partial y(s)}{\partial x} + \xi' \frac{\partial x(s)}{\partial x} + \eta' \frac{\partial y(s)}{\partial x} \right) ds \\ &= \int_0^t \frac{d}{ds}\left(\xi \frac{\partial x(s)}{\partial x} + \eta \frac{\partial y(s)}{\partial x} \right) ds \\ &= \xi(s) \frac{\partial x(s)}{\partial x}\bigg|_{s=0}^{s=t} + \eta(s) \frac{\partial y(s)}{\partial x}\bigg|_0^t \\ &= \xi(t) - \eta(0) \frac{\partial y(0; x, x_0, y, \eta, t)}{\partial x}. \end{aligned} \tag{4.15}$$

Similarly, we may solve for $\partial S / \partial y$ to get:

$$\frac{\partial S}{\partial y}(x, x_0, y, \eta, t) = \eta(t) - \eta(0) \frac{\partial y(0, x, y, x_0, \eta, t)}{\partial y}. \tag{4.16}$$

As for $\partial S/\partial t$, we get:

$$\begin{aligned}\frac{\partial S}{\partial t} &= \xi(t)x'(t) + \eta(t)y'(t) - H(t)\\ &+ \int_0^t \left(\xi \frac{\partial}{\partial t}\frac{dx}{ds} + \eta \frac{\partial}{\partial t}\frac{dy}{ds} - \frac{\partial H}{\partial x}\frac{\partial x}{\partial t} - \frac{\partial H}{\partial y}\frac{\partial y}{\partial t}\right) ds\\ &= \xi(t)x'(t) + \eta(t)y'(t) - H(t)\\ &+ \int_0^t \left(\xi \frac{d}{ds}\frac{\partial x}{\partial t} + \eta \frac{d}{ds}\frac{\partial y}{\partial t} + \xi' \frac{\partial x}{\partial t} + \eta' \frac{\partial x}{\partial t}\right) ds\\ &= \xi(t)x'(t) + \eta(t)y'(t) - H(t) + \xi(s)\frac{\partial x(s)}{\partial t}\Big|_0^t + \eta(s)\frac{\partial y(s)}{\partial t}\Big|_0^t .\end{aligned}$$

Since x_0 and $x(s = t)$ are fixed points, we get:

$$\frac{dx(s=t)}{dt} = 0 = x'(t) + \frac{\partial x(s)}{\partial t}\Big|_{s=t}$$

and that:

$$\frac{\partial x(s)}{\partial t}\Big|_0^t = -x'(t). \tag{4.17}$$

On the other hand,

$$\frac{\partial y(s)}{\partial t}\Big|_{s=0}^{s=t} = -y'(t) - \frac{\partial y}{\partial t}\Big|_{s=0}. \tag{4.18}$$

As a result, $\partial S/\partial t$ reduces to:

$$\frac{\partial S}{\partial t} = -H(t) - \eta(0)\frac{\partial y}{\partial t}(0, \ldots). \tag{4.19}$$

By letting

$$h = \eta y(0, \ldots) + S, \tag{4.20}$$

from (4.15) and (4.16), we get the following relations:

$$\begin{aligned}\frac{\partial h}{\partial x} &= \xi(t),\\ \frac{\partial h}{\partial y} &= \eta(t),\\ \frac{\partial h}{\partial t} + H\big(x(t), y(t), \xi(t), \eta(t)\big) &= 0.\end{aligned} \tag{4.21}$$

Note that with the new assignment (4.20), we can summarize our findings in (4.2) as follows:

$$\frac{\partial h}{\partial t} + H\left(x(t), y(t), \frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}\right) = 0$$

and conclude the following theorem.

Theorem 4.1. When $\eta t \neq \pm k\pi$ for $k \in \mathbb{N}$ (so that $\sin \eta t \neq 0$),

$$h = \eta y - \frac{\eta}{4}\left((x + x_0)^2 \tan\left(\frac{\eta t}{2}\right) - (x - x_0)^2 \cot\left(\frac{\eta t}{2}\right)\right) \tag{4.22}$$

solves the Hamilton-Jacobi equation (4.4).

4.3. **The Modified Complex Action.** Since, by definition, the heat kernel cannot depend on η, we are interested in modifying the action to eliminate η. One way to accomplish this is by summing the kernel over η. Moreover, given that the heat kernel has the form $t^{-\alpha}e^{-h}$, where α has not been determined yet, we may sum over $\eta t = \lambda$, since an extra t can be absorbed by the $t^{-\alpha}$ term. In other words, we are interested in finding a heat kernel of the following form:

$$P = \frac{1}{(2\pi t)^{\alpha}} \int_{\mathbb{R}} e^{-\frac{g(x,x_0,y,\lambda)}{t}} V(\lambda)\, d\lambda, \tag{4.23}$$

where

$$g(x,x_0,y,\lambda) = g(x,x_0,y,\eta t) = h(x,x_0,y,\eta t,1) = th(x,x_0,y,\eta,t) \tag{4.24}$$

and $V(\lambda)$ stands for good measure (see the Appendix).

To derive the modified complex action, we first need to go over two lemmas.

Lemma 4.1. $g(x,x_0,y,\lambda)$ solves the following eikonal equation:

$$\lambda\frac{\partial g}{\partial \lambda} + H\Big(x,y,\frac{\partial g}{\partial x},\frac{\partial g}{\partial y}\Big) = g. \tag{4.25}$$

Proof. From (4.4), we have:

$$\frac{\partial h}{\partial t} = -\frac{1}{2}\Big(\frac{\partial h}{\partial x}\Big)^2 - \frac{x^2}{2}\Big(\frac{\partial h}{\partial y}\Big)^2,$$

which by (4.24) becomes:

$$\frac{\partial h}{\partial t} = -\frac{g}{t^2} + \frac{\eta}{t}\frac{\partial g}{\partial \lambda}.$$

Since

$$-\frac{1}{2}\Big(\frac{\partial h}{\partial x}\Big)^2 - \frac{x^2}{2}\Big(\frac{\partial h}{\partial y}\Big)^2 = -\frac{g}{t^2} + \frac{\eta}{t}\frac{\partial g}{\partial \lambda}$$

for all t, setting $t = 1$ and replacing η by λ proves the claim. □

We are yet to find the path of integration for (4.23). For succinctness, we may write the non-classical action as follows:

$$g(\lambda) = \lambda y - \frac{\lambda}{4}\left(a^2 \tan\Big(\frac{\lambda}{2}\Big) - b^2 \cot\Big(\frac{\lambda}{2}\Big)\right), \tag{4.26}$$

where

$$a^2 = (x+x_0)^2 \quad \text{and} \quad b^2 = (x-x_0)^2.$$

Now we concentrate on the second lemma, which will aid us in determining the integration path.

Lemma 4.2. Let $\lambda = \theta\rho = \lambda_1 + i\lambda_2 \in \mathbb{C}$, where $\rho = |\lambda|$ and $\lambda_1, \lambda_2 \in \mathbb{R}$. For a fixed θ,

$$\lim_{\rho\to\infty} \Re\big(g(\theta\rho)\big) = \infty$$

for all (x, y) off the curve $x^2 + x_0^2 = 0$ iff $\theta \in i\mathbb{R}$.

Proof. First, we observe that

$$\begin{aligned}
a^2\tan\left(\frac{\lambda}{2}\right)-b^2\cot\left(\frac{\lambda}{2}\right) &= \frac{a^2\sin\frac{\lambda}{2}}{\cos\frac{\lambda}{2}}-\frac{b^2\cos\frac{\lambda}{2}}{\sin\frac{\lambda}{2}}\\
&= \frac{a^2(1-\cos\lambda-b^2(1+\cos\lambda))}{\sin\lambda}\\
&= \frac{a^2(1-\cos\lambda-b^2(1+\cos\lambda))}{\sin\lambda}\frac{\sin\bar\lambda}{\sin\bar\lambda}\\
&= \frac{a^2(1-\cos\lambda-b^2(1+\cos\lambda))}{\big(\cos(\lambda-\bar\lambda)-\cos(\lambda+\bar\lambda)\big)/2}\sin\bar\lambda\\
&= \frac{a^2(1-\cos\lambda-b^2(1+\cos\lambda))}{\cosh^2\lambda_2-\cos^2\lambda_1}\sin\bar\lambda.
\end{aligned}$$

Thus,

$$\Re\big(g(\lambda)\big)=\lambda_1 y+\frac{1}{4}\left(a^2\frac{\lambda_2\sinh\lambda_2-\lambda_1\sin\lambda_1}{\cosh\lambda_2+\cos\lambda_1}+b^2\frac{\lambda_2\sinh\lambda_2+\lambda_1\sin\lambda_1}{\cosh\lambda_2-\cos\lambda_1}\right). \tag{4.27}$$

As with λ, we let $\theta=\theta_1+i\theta_2$ for $\theta_1,\theta_2\in\mathbb{R}$. Then,

(1) if $\theta_1=0$ and $\theta_2=\pm1$, it follows that $\theta\in i\mathbb{R}$. For $\rho\approx\infty$, we get:

$$\begin{aligned}
\Re\big(g(\lambda)\big) &\approx \frac{a^2+b^2}{4}\theta_2\rho\tanh(\theta_2\rho)\\
&\approx \frac{x^2+x_0^2}{2}\rho\to\infty.
\end{aligned}$$

(2) if $\theta_1=\pm1$ and $\theta_2=0$, it follows that $\theta\in\mathbb{R}$. Then, for $x^2+x_0^2$ and $\rho\to\infty$, we get:

$$\Re\big(g(\lambda)\big)=\pm\rho y\to\big(\pm\, sgn(y)\big)\infty.$$

(3) if $\theta_1\neq0$ and $\theta_2\neq0$, it follows that $\theta\in\mathbb{C}$. For $\rho\approx\infty$, we get:

$$\begin{aligned}
\Re\big(g(\lambda)\big) &\approx \theta_1\rho y+\frac{1}{4}\big(a^2\lambda_2\tanh\lambda_2+b^2\lambda_2\coth\lambda_2\big)\\
&= \theta_1\rho y+\frac{a^2+b^2}{4}|\theta_2|\rho\\
&= \left(\theta_1 y+\frac{x^2+x_0^2}{2}|\theta_2|\right)\rho.
\end{aligned}$$

Choosing y satisfying

$$|\theta_1 y|>\frac{x^2+x_0^2}{2}|\theta_2|,$$

results in

$$\Re\big(g(\lambda)\big)\to\big(sgn(\theta_1 y)\big)\infty,$$

thus proving the claim.

□

This observation leads us to choose the imaginary axis as the integration path to have the integrand at its best behavior. Hence, we let

$$\lambda = -i\tau, \quad \tau \in \mathbb{C},$$

and set

$$f(\tau) = g(-i\tau). \tag{4.28}$$

Thus, the modified complex action takes the following form:

$$f(\tau) = -i\tau + \frac{\tau}{4}\left((x+x_0)^2 \tanh\frac{\tau}{2} + (x-x_0)^2 \coth\frac{\tau}{2}\right), \tag{4.29}$$

and the next step would be to find a heat kernel of the following form:

$$P = \frac{1}{(2\pi t)^\alpha}\int_{\mathbb{R}} e^{-\frac{f(\tau)}{t}} V(\tau)\, d\tau, \tag{4.30}$$

where $f(\tau)$ satisfies the eiconal equation:

$$\tau\frac{\partial f}{\partial \tau} + H(x, y, f_x, f_y) = f. \tag{4.31}$$

5. The volume element

In the previous section we conjectured the heat kernel's form. However, we still lack some details in order to state its final form; specifically, the volume element and α remain undetermined. We make it our goal in this section to fill in those gaps. We demonstrate that with the correct choice of $V(\tau)$ (4.30) represents the heat kernel of the step 2 Grushin operator. Recall that according to Definition 4.1, the heat kernel of an operator ought to satisfy three properties. Starting with the first property of (4.1), we show that

$$\left(\Delta_G - \frac{\partial}{\partial t}\right)\frac{e^{-\frac{f}{t}}}{t^\alpha} = \frac{e^{-\frac{f}{t}}}{t^{\alpha+2}}\left(H(\nabla f) - f\right) - \frac{e^{-\frac{f}{t}}}{t^{\alpha+1}}\left(\Delta_G f - \alpha\right).$$

Proof.

$$\begin{aligned}
\left(\Delta_G - \frac{\partial}{\partial t}\right)\frac{e^{-f/t}}{t^\alpha} &= \frac{1}{2}\frac{\partial}{\partial x}\left(-\frac{e^{-f/t}}{t^{\alpha+1}}\frac{\partial f}{\partial x}\right) + \frac{1}{2}\frac{\partial}{\partial y}\left(-\frac{x^2 e^{-f/t}}{t^{\alpha+1}}\frac{\partial f}{\partial y}\right)\\
&\quad - \frac{-t^\alpha e^{-f/t} f/t^2 - \alpha t^{\alpha-1}e^{-f/t}}{t^{2\alpha}}\\
&= \frac{1}{2}\Bigg(\left(\frac{\partial f}{\partial x}\right)^2\frac{e^{-f/t}}{t^{\alpha+2}} - \frac{e^{-f/t}}{t^{\alpha+1}}\frac{\partial^2 f}{\partial x^2}\Bigg)\\
&\quad + \frac{1}{2}\Bigg(\left(\frac{\partial f}{\partial y}\right)^2\frac{x^2e^{-f/t}}{t^{\alpha+2}} - \frac{x^2e^{-f/t}}{t^{\alpha+1}}\frac{\partial^2 f}{\partial y^2}\Bigg)\\
&\quad + \frac{e^{-f/t}(f/t+\alpha)}{t^{\alpha+1}}\\
&= \frac{e^{-f/t}}{t^{\alpha+2}}\Bigg(\underbrace{\frac{1}{2}\left(\frac{\partial f}{\partial x}\right)^2 + \frac{x^2}{2}\left(\frac{\partial f}{\partial y}\right)^2}_{H(\nabla f)} - f\Bigg)\\
&\quad - \frac{e^{-f/t}}{t^{\alpha+1}}\Bigg(\underbrace{\frac{1}{2}\frac{\partial^2 f}{\partial x^2} + \frac{x^2}{2}\frac{\partial^2 f}{\partial y^2}}_{\Delta_G f} - \alpha\Bigg).
\end{aligned}$$

□

By (4.31), we have that:

$$-\tau\frac{\partial f}{\partial \tau} = H(\Delta f) - f.$$

Thus,

$$\begin{aligned}
\left(\Delta_G - \frac{\partial}{\partial t}\right)\frac{e^{-f/t}V(\tau)}{t^\alpha} &= -\frac{e^{-f/t}}{t^{\alpha+2}}\left(\tau\frac{\partial f}{\tau}\right)V(\tau) - \frac{e^{-f/t}}{t^{\alpha+1}}\left(\Delta_G f - \alpha\right)V(\tau)\\
= -\frac{e^{-f/t}}{t^{\alpha+1}}&\left(\tau\frac{dV}{d\tau} + \left(\Delta_G f - \alpha + 1\right)V\right) + \frac{\partial}{\partial \tau}\left(\frac{\tau e^{-f/t}V(\tau)}{t^{\alpha+1}}\right).
\end{aligned}$$

Under the assumption that $\tau e^{-f/t}V(\tau)$ vanishes at infinity, i.e.

$$\lim_{\tau\to\pm\infty}\tau e^{-f/t}V(\tau) = 0,$$

we have that:

$$\begin{aligned}
&\left(\Delta_G - \frac{\partial}{\partial t}\right)\frac{1}{t^\alpha}\int_{\mathbb{R}} e^{-f/t}V(\tau)\,d\tau =\\
&\qquad -\frac{1}{t^{\alpha+1}}\int_{\mathbb{R}} e^{-f/t}\left(\tau\frac{dV}{d\tau} + \left(\Delta_G f - \alpha + 1\right)V(\tau)\right)d\tau = 0,
\end{aligned}$$

for $t > 0$ and

$$\tau\frac{dV}{d\tau} + \left(\Delta_G f - \alpha + 1\right)V(\tau) = 0. \tag{5.1}$$

From (4.29) and using half angle identities for hyperbolic functions, we may easily find $\Delta_G f$ as follows:

$$\Delta_G f = \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2} + \underbrace{x^2\frac{\partial^2 f}{\partial y^2}}_{0}\right) = \frac{\tau}{4}\left(\tanh\frac{\tau}{2} + \coth\frac{\tau}{2}\right)$$

$$= \frac{\tau}{4}(\coth\tau - \operatorname{csch}\tau + \coth\tau + \operatorname{csch}\tau) = \frac{\tau}{2}\coth\tau. \tag{5.2}$$

Hence, substituting (5.2) into (5.1), enables us to solve for $V(\tau)$ using the separation of variables technique as follows:

$$\tau\frac{dV}{d\tau} + \left(\frac{\tau}{2}\coth\tau - \alpha + 1\right)V(\tau) = 0,$$

$$\frac{dV}{d\tau} = \left(\frac{\alpha-1}{\tau} - \frac{1}{2}\coth\tau\right)d\tau,$$

$$\int\frac{1}{V}\,dV = \int\left(\frac{\alpha-1}{\tau} - \frac{1}{2}\coth\tau\right)d\tau,$$

$$\log V = (\alpha-1)(\log\tau + \log C) - \frac{1}{2}\log\sinh\tau.$$

We summarize this result in the following lemma:

Lemma 5.1. The general solution of (5.1) is

$$V(\tau) = \frac{(C\tau)^{\alpha-1}}{\sqrt{\sinh\tau}}. \tag{5.3}$$

Since we want V to be holomorphic near $\tau = 0$ and avoid singularities, we require α to be equal to $n + \frac{1}{2}$ for $n \in \mathbb{N}$. Thus, substituting this choice of α into (5.3), we have that:

$$V(\tau) = C\tau^{n-1}\sqrt{\frac{\tau}{\sinh\tau}}.$$

Expressing $V(\tau)$ as power series,$\sum_{k=0}^{\infty} a_k\tau^k$, and substituting into (5.1), we have that:

$$\sum_{k=1}^{\infty} k\alpha_k\tau^k + \left(\frac{\tau}{2}\coth\tau - \alpha + 1\right)\left(\sum_{k=0}^{\infty} a_k\tau^k\right) =$$

$$a_1\tau + \frac{3}{2}a_0 + \frac{3}{2}a_1\tau - \alpha a_0 - \alpha a_1\tau + \ldots =$$

$$\left(\frac{3}{2} - \alpha\right)a_0 + \mathcal{O}(\tau) = 0,$$

where $\mathcal{O}(\tau)$ denotes the terms of higher order. To avoid getting the trivial heat kernel, we require $\alpha = 3/2$. Now that we have derived all the necessary components, we are ready to state the heat kernel of the step 2 Grushin operator in its entirety:

$$\mathcal{P}_t(x, x_0, y) = \frac{1}{(2\pi t)^{3/2}} \int_{\mathbb{R}} \exp\left(\frac{\tau}{2t}\left(2iy - \frac{1}{2}\left(a^2 \tanh\frac{\tau}{2} + b^2 \coth\frac{\tau}{2}\right)\right)\right)\sqrt{\frac{\tau}{\sinh\tau}}\, d\tau.$$

6. Discussion

In this paper, we have discussed an operator that possesses relatively new geometry, the study of which relied on developments in geometric mechanics. We have demonstrated the application of the Hamilton-Jacobi theory of bicharacteristic curves to identify the geodesics of the operator admitting a unique geometry. Furthermore, Hamiltonian formalism allowed us to find the action and, later, the modified complex action for the operator, which is essential in constructing the heat kernel for the Grushin operator. Aided by a clever choice of a volume element, we were able to fully derive the heat kernel for the step 2 Grushin operator.

As with many operators with different geometries, advances in the study of operators, such as Grushin, have proven useful to applied sciences and industry, especially for problems involving restricted movement in space. Nevertheless, the study of such operators are highly non-trivial. As one may expect, the study of step 3 Grushin operator will be no easier than that of step 2 and, in fact, will rely on numerical solutions, which can be discussed in future research and papers.

Acknowledgement

This survey paper is based on my senior thesis. I would like to thank my mathematics and statistics professors at Georgetown University for their outstanding teaching and support, especially my advisor Professor Der-Chen Chang for his continued mentorship and for introducing me to the study of subelliptic operators and Professor Michael Raney for his thorough reading of my thesis and valuable suggestions.

References

[1] O. Calin and D.-C. Chang, *Sub-Riemannian Geometry: General Theory and Examples*, Cambridge University Press, New York, NY, 2013.
[2] O. Calin, D.-C. Chang and K. Furutani, *Heat Kernels for Elliptic and Sub-Elliptic Operators: Methods and Techniques*, Birkhäuser, Boston, MA, 2010.
[3] D.-C. Chang, C.-H. Chang, P. Greiner and H.-P. Lee, *The Positivity of the Heat Kernel on the Heisenberg Group, Analysis and Applications*, World Sientific 11, #5, 2013.
[4] D.-C. Chang and Y. Li, *Asymptotics for Heat Kernels of SubLaplacian on Heisenberg Group*, The Royal Societ-Proceedings 11, #5, 2015.
[5] D.-C. Chang and Y. Li, *SubRiemannian geodesics in Grushin plane*, Journal of Geometric Analysis **22** (2012), 800–826.
[6] L. Hörmander, *Hypoelliptic Second Order Differential Equations*, Acta Mathematica 119, 1967.
[7] Y. Lim, *Hamilton-Jacobi Theory and Heat Kernels*, City University of Hong Kong, 2010.
[8] M. Nakahara, *Geometry, Topology, and Physics.* Institute of Physics Publishing, London, UK, 2003.
[9] S. H. Strogatz, *Nonlinear Dynamics and Chaos*, 2nd ed. Westview Press, Boulder, CO. 2015.
[10] R. L. Wheeden and A. Zygmund, *Measure and Integral*, CRC Press, Boca Raton, FL, 2015.

Manuscript received May 4 2021
revised August 12 2021

Levon Shmavonyan
Department of Mathematics & Statistics, Georgetown University, Washington, DC 20057
E-mail address: ls1271@georgetown.edu