Received: 27 September 2022 | Revised: 22 February 2023 | Accepted: 19 April 2023

DOI: 10.1111/sapm.12581

ORIGINAL ARTICLE

STUDIES IN APPLIED MATHEMATICS WILEY

Can a ground-based vehicle hear the shape of a room?

Mireille Boutin[1] | **Gregor Kemper**[2]

[1]Department of Mathematics, Purdue University, West Lafayette, Indiana, USA

[2]Technische Universität München, Department of Mathematics, Garching, Germany

Correspondence
Mireille Boutin, Department of Mathematics, Purdue University, 150 N. University St., West Lafayette, IN 47907, USA.
Email: mboutin@purdue.edu

Abstract

Assume that a ground-based vehicle moves in a room with walls or other planar surfaces. Can the vehicle reconstruct the positions of the walls from the echoes of a single sound event? We assume that the vehicle carries some microphones and that a loudspeaker is either also mounted on the vehicle or placed at a fixed location in the room. We prove that the reconstruction is almost always possible if (1) no echoes are received from floors, ceilings, or sloping walls and the vehicle carries at least three noncollinear microphones, or if (2) walls of any inclination may occur, the loudspeaker is fixed in the room and there are four noncoplanar microphones. The difficulty lies in the echo-matching problem: How to determine which echoes come from the same wall. We solve this by using a Cayley–Menger determinant. Our proofs use methods from computational commutative algebra.

KEYWORDS

echo sorting, geometry from echoes, shape reconstruction

JEL CLASSIFICATION

51K99, 13P10, 13P25 (Primary)

1 INTRODUCTION

This paper is concerned with the problem of reconstructing the position of walls and other planar surfaces using echoes. More specifically, an omnidirectional loudspeaker produces a short, high-frequency impulse. The echoes of the impulses are captured by some microphones. The loudspeaker and the microphones are assumed to be synchronized. The walls are modeled using *mirror points*, which are the reflections of the loudspeaker with respect to the planar surfaces. The times of arrival of an echo give us the total distance traveled by the sound, which is equal to the distance from the microphone to a mirror point. An illustration can be found in Figures 2, 3. We are interested in determining to what extent it is theoretically possible to reconstruct the wall positions from the distance measurements obtained in this fashion. All measurements are assumed to be exact, and the computations are assumed to be performed with infinite precision. Yet, even in this theoretical scenario, it may be impossible, at least in some circumstances, to correctly reconstruct the wall positions. One of the goals of our work is to narrow down the problematic cases and show that one can expect to avoid them in some common application scenarios. Here, we are thinking of a vehicle carrying the microphones and having the ability to move more or less freely on the ground, for example, a robot rolling on the floor inside a warehouse.

In previous work,[1] we considered the case of a drone or, more generally, a vehicle with all six degrees of freedom (the three coordinate axes and yaw, pitch, and roll). If there are four microphones mounted on the vehicle and there is a loudspeaker either on the vehicle or at a fixed location in the room, then our algorithm[1] reconstructs the position of every wall from which an echo is received by all microphones. But it cannot be avoided that for certain “bad” vehicle positions, a “ghost wall,” one which is not really there, is detected. Our main result states that the bad positions are rare. However, to get out of a possibly bad position, the vehicle may need to use all six degrees of freedom, including pitch and roll. For a drone, this means that being in a good (i.e., not bad) position may result in receiving a thrust in some direction, so the drone cannot hover while carrying out the wall reconstruction. So we are interested in investigating whether our results carry over to a situation where the degrees of freedom are restricted to the coordinate axes and yaw. Perhaps more interesting is the case of a ground-based vehicle. These are the scenarios considered in this paper.

The restriction of degrees of freedom not only makes the mathematical arguments harder but also yields somewhat weaker results. In fact, in Section 3, we exhibit “unlucky” wall arrangements for which it is not true that almost all vehicle positions are good, if the degrees of freedom are restricted as stated above. So the cases of a “hovering drone” and a ground-based vehicle really are different from the case of a freely moving drone, and more difficult. The first main result (Theorem 3) deals with the case of a ground-based vehicle and a loudspeaker placed at a fixed position in the room, and gives a precise characterization of the wall arrangements for which it is not true that almost all vehicle positions are good. These are the arrangements where an “unlucky stack of mirror points,” as defined in Section 3, occurs, and they are themselves very rare. The second main result (Theorem 4) treats the case of a hovering drone (four degrees of freedom) and, perhaps surprisingly, reaches the same conclusion: The additional degree of freedom does not result in fewer exceptional wall arrangements. Both results are first stated and proved in purely mathematical terms (Theorem 1). The proofs of these results are similar in spirit to the proofs in Ref. 1, and in particular involve large computations in ideal theory done by computer. But apart from the new aspect of exceptional wall arrangements, there arise further theoretical difficulties that are explained after the statement of Theorem 1. Analyzing our proof method shows that for each degree of freedom taken away, the final computation requires roughly two additional variables. This is why computational bottlenecks prevented us from achieving general results in the case that the loudspeaker is mounted on the vehicle.

In the first result of this paper, the vehicle moves in two dimensions, but still lives in three-dimensional space, meaning that the walls it detects may be floors, ceilings, and sloping walls. The last result (Theorem 5) of this paper concerns the truly two-dimensional case. In real-world terms, this means that no echoes should be received from floors, ceilings, and sloping walls, since they cannot be dealt with. This case is much easier, has no exceptional wall arrangements, only requires three microphones, and also allows for the loudspeaker to be mounted on the vehicle.

In this paper, a room is just a spacial arrangement of walls, and a wall is just a plane surface. We only consider first-order echoes, meaning that the sound has not bounced twice or more often. A further assumption is that the sound frequency is high enough to justify the use of ray acoustics. When we say that almost all vehicle positions are good, we mean that within the configuration space of all vehicle positions, the bad ones form a subset of volume zero. In fact, we use this language only if the bad positions are contained in a subvariety of lower dimension. Intuitively "almost all" can be thought of as "with probability one."

2 | RELATED WORK

The relationship between the geometry of a room and the source-to-receiver acoustic impulse response has long been a subject of interest, in particular for the design of concert halls with desirable acoustic properties. For example, the two-point impulse response within rectangular rooms was analyzed in Ref. 2. The results were later extended to arbitrarily polyhedral rooms in Ref. 3. Subsequent work used the impulse response of a room to reconstruct its geometry, including that of objects within the room (e.g., Ref. 4).

Some reconstruction methods are focused on 2D room geometry (e.g., Refs. 5–7) while others consider a more general but still constrained 3D geometry. For example, Park and Choi[8] reconstruct the wall surface for a convex, polyhedral, and bounded room using several loudspeakers and microphones in known positions. Other methods are applicable to reconstructing of arbitrarily planar surfaces in 3D (e.g., Ref. 9).

Some methods reconstruct wall points directly (e.g., Refs. 7, 8, 10–12) while others, like us, reconstruct mirror points representing the reflection of a loudspeaker with respect to a wall plane (e.g., Refs. 9, 13–16). Some, like us, assume that the microphones are synchronized (e.g., Refs. 9, 13, 17, 18), while others do not (e.g., Refs. 6, 19, 20).

In certain setups, the microphones are placed on a vehicle such as a robot (e.g., Ref. 21) or a drone.[1] Recent work uses cell phones as a cheaper alternative. For example, Shih and Rowe[22] place an omnidirectional speaker at a known position in a polyhedral room and use a cell phone to record the echoes while Zhou et al.[23] assume a smart phone is held horizontally by a person walking along rooms and corridors.

The vast majority of the work in the literature is focused on the numerical reconstruction of the room, including the difficult problem of labeling echoes coming from the same surface. However, from a theoretical standpoint, there is still a lot to be understood regarding the well-posedness of different reconstruction setups, especially for nongeneric room geometries or receiver and source placements.

3 | UNLUCKY STACKS OF MIRROR POINTS

As mentioned above, the primary challenge in wall detection from echoes is the matching of echoes heard by different microphones, that is, determining which echoes come from the same wall and which ones do not.

Let us recall an example from Ref. 1 where this matching inevitably goes wrong, leading to a ghost wall. As it turns out, this type of example will play a crucial role in this paper.

Figure 1 shows three microphones in a plane at positions $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$ that hear echoes from three walls W_i, but the time elapsed between sound emission and echo detection is the same as if they were hearing echoes from one single wall W_{ghost}, which does not exist. This arises because (1) the walls W_i are all horizontal, (2) the distances between $\mathbf{m}_i$ and W_i are the same for all i, and (3) each $\mathbf{m}_i$ can hear the echo from W_i. It is easy to add a fourth microphone outside of the drawing plane, together with a wall possibly also outside of the plane, such that (1)–(3) extend to the fourth microphone and wall. (We find it harder to include that in our two-dimensional sketch in Figure 1.) Since the echoes heard by the microphones behave precisely like echoes from the single wall W_{ghost}, any matching procedure will falsely assume that they have, in fact, come from W_{ghost}, giving rise to a ghost wall. The situation is particularly severe since the ghost wall has some persistence properties. In fact, the conditions (1)–(3), and therefore the emergence of a ghost wall, will be preserved if

(a) the microphones are moved, independently and within a certain range, in directions parallel to the walls;
(b) the microphones are moved, *together* and within a certain range, along the z-axis, or, less importantly;
(c) the loudspeaker is moved, within a certain range, in any direction.

Our sketch does not contain any vehicle that carries the microphones. Let us now imagine there is such a vehicle, and that it is *ground based*, in the sense that its movement (rotation and translation) is restricted to the x-y-plane, where the y-axis is perpendicular to the drawing plane. The vehicle may also be allowed to move up or down along the z-axis, like a hovering drone. Then, the persistence properties imply that within a certain range, vehicle movements will not

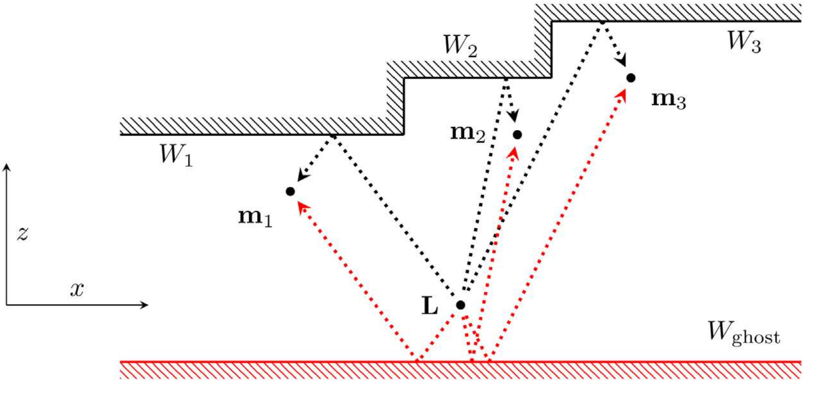

FIGURE 1 The microphones $\mathbf{m}_i$ think they are hearing echoes from the wall W_{ghost}. The dotted lines stand for sound rays, $\mathbf{L}$ for the loudspeaker.

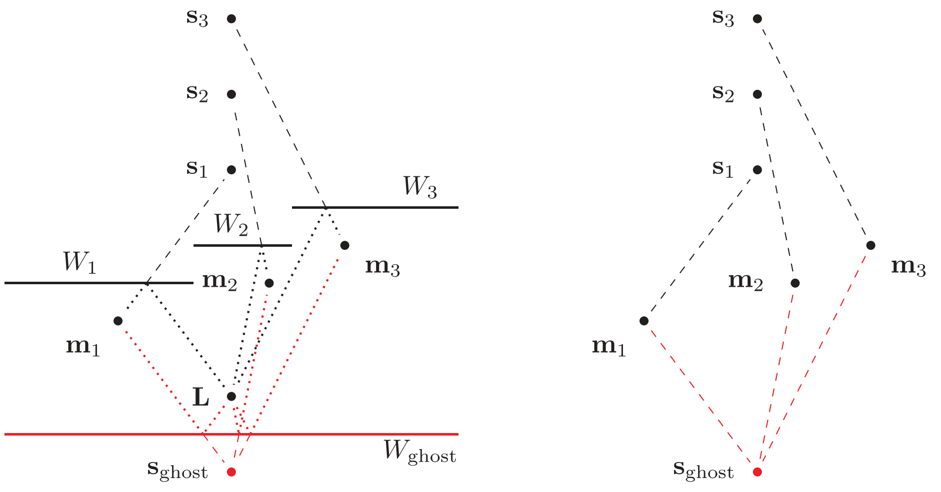

FIGURE 2 The sound comes from the loudspeaker **L** and is reflected to the microphones $\mathbf{m}_i$. But virtually, it comes from the mirror points $\mathbf{s}_i$. So the loudspeaker and walls might as well be left out of the sketch.

make the ghost wall disappear. For this, it does not matter whether the loudspeaker is mounted on the vehicle or at a fixed position. In all cases, there is a region of positive volume within the space of all vehicle movements, in which the vehicle remains in a bad position. So it is not true that almost all vehicle positions are good. Thus, here we hit a limitation to what can possibly be shown for microphones mounted on ground-based vehicles, or vehicles with an additional degree of freedom of moving up or down. (In comparison, a drone has two additional degrees of freedom: pitch and roll.)

Since we are using ray acoustics, an echo reflected at a wall arrives with the same time delay as if it were emitted at what we call the **mirror point**, which is the reflection of the loudspeaker position at the wall or, more precisely, at the plane containing the wall surface. Figure 2 shows the mirror points $\mathbf{s}_i$ in the situation of Figure 1.

If we have mirror points as in the right sketch of Figure 2, then we have the persistence properties (a) and (b) of the ghost wall discussed above. For this, it is not necessary that the loudspeaker be situated on the same vertical line as the mirror points, so the walls may be nonhorizontal. Figure 3 illustrates that. In fact, it is only if the loudspeaker is moving as well that we need horizontal walls to achieve the persistence properties of the ghost wall. It is useful to introduce a catchphrase for the situation shown in Figures 2, 3, which is done in the following definition. From now on, we will only consider the case in which the loudspeaker is at a fixed position and not moving with the vehicle.

Definition 1. Assume that four microphones are mounted on a vehicle that can move within the x-y-plane, and possibly up and down along the z-axis. Assume that the vehicle moves in a scene that contains several flat walls and a loudspeaker at a fixed position. We say that the microphones are **deceived by an unlucky stack of mirror points** if among the walls there are four whose mirror points $\mathbf{s}_1, \dots, \mathbf{s}_4$ are contained in a common vertical line, such that for $i, j \in \{1, \dots, 4\}$, the z-coordinate of the vector $\mathbf{s}_j - \mathbf{s}_i$ is twice the z-coordinate of $\mathbf{m}_j - \mathbf{m}_i$. In particular, this means that if for some $i \neq j$ the positions $\mathbf{m}_i$ and $\mathbf{m}_j$ share the same z-coordinate, then $\mathbf{s}_i = \mathbf{s}_j$; so for deception by an unlucky stack of mirror points, the mirror points (and the walls) need not be pairwise distinct.

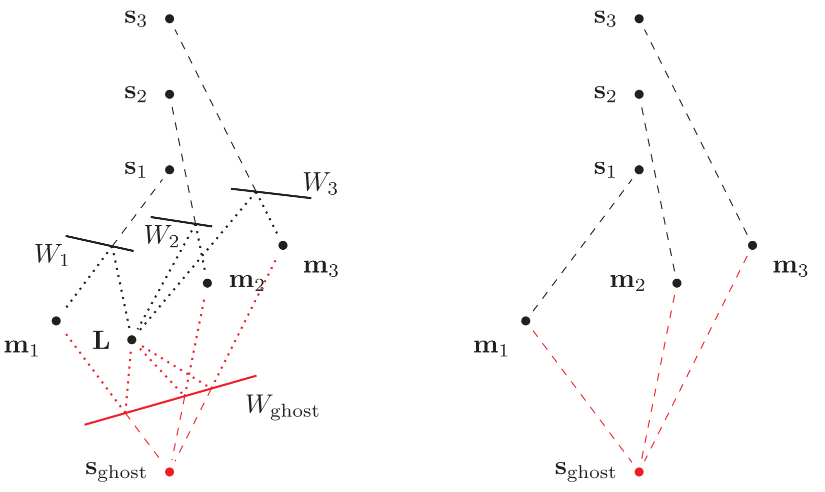

FIGURE 3 A different loudspeaker position and different walls produce the exact same unlucky stack of mirror points as in Figure 2. But here the walls are not horizontal.

A formal statement and proof that an unlucky stack of mirror points leads to deception (i.e., to ghost wall) is given in Theorem 1(a) in the following section.

4 | MATHEMATICAL FORMULATION

Assume we are given four noncoplanar points $\mathbf{m}_1, \dots, \mathbf{m}_4 \in \mathbb{R}^3$ (which in our application are the positions of the microphones on a vehicle). With $D_{i,j} := \|\mathbf{m}_i - \mathbf{m}_j\|^2$ the squared Euclidean distances, define the polynomial f_D in four variables as the Cayley–Menger determinant

$$f_D(u_1, \dots, u_4) := \det \begin{pmatrix} 0 & u_1 & \cdots & u_4 & 1 \\ u_1 & D_{1,1} & \cdots & D_{1,4} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ u_4 & D_{4,1} & \cdots & D_{4,4} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix}. \tag{1}$$

So by Cayley,[24] if $\mathbf{s} \in \mathbb{R}^3$, then

$$f_D\left(\|\mathbf{s} - \mathbf{m}_1\|^2, \dots, \|\mathbf{s} - \mathbf{m}_4\|^2\right) = 0. \tag{2}$$

We need two groups (which in our application model the positions of a hovering drone or a ground-based vehicle). The first is the group A_3SO_2 of all maps

$$\phi = \phi_{a_1,\dots,a_5}\colon \mathbb{R}^3 \to \mathbb{R}^3,\ v \mapsto \begin{pmatrix} a_1 & a_2 & 0 \\ -a_2 & a_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot v + \begin{pmatrix} a_3 \\ a_4 \\ a_5 \end{pmatrix}, \tag{3}$$

with $a_i \in \mathbb{R}$ such that $a_1^2 + a_2^2 = 1$ (so the matrix is a rotation acting on the first two coordinates). The second one is the subgroup ASO_2 given by the additional condition $a_5 = 0$. Simultaneously applying a $\phi \in A_3SO_2$ to the $\mathbf{m}_i$ preserves the squared distances $D_{i,j}$.

Now in addition to the points $\mathbf{m}_i$, let $\mathcal{S} \subset \mathbb{R}^3$ be a finite set of points (which in our application are the mirror points). Let $\phi \in A_3SO_2$. We say that ϕ is **good** if for all choices of four points $\mathbf{s}_1, \dots, \mathbf{s}_4 \in \mathcal{S}$, not necessarily distinct, the condition

$$f_D\left(\|\mathbf{s}_1 - \boldsymbol{\phi}(\mathbf{m}_1)\|^2, \dots, \|\mathbf{s}_4 - \boldsymbol{\phi}(\mathbf{m}_4)\|^2\right) = 0 \tag{4}$$

implies that there is an $\mathbf{s} \in \mathcal{S}$ such that

$$\|\mathbf{s} - \phi(\mathbf{m}_i)\| = \|\mathbf{s}_i - \boldsymbol{\phi}(\mathbf{m}_i)\| \quad \text{for} \quad i = 1, \dots, 4. \tag{5}$$

As we will see in Proposition 1(a), such an $\mathbf{s} \in \mathbb{R}^3$, even if it lies outside $\mathcal{S}$, is unique. (In our application, $\mathbf{s}$ will be detected as a mirror point, so no ghost walls are detected if the vehicle is in a position given by a good ϕ.) In contrast, we call ϕ **bad** if it is not good. (In our application, this means that a ghost wall will be detected, provided the echoes from the walls making the position bad are heard.) Notice that if for some $\mathbf{s}_1, \dots, \mathbf{s}_4 \in \mathcal{S}$, there is an $\mathbf{s} \in \mathbb{R}^3 \setminus \mathcal{S}$ satisfying (5), then ϕ is bad. In fact the converse of this is also true, but we do not need this. Moreover, we call $\phi \in A_3SO_2$ **really good** if it is good and for all choices of four points $\mathbf{s}_1, \dots, \mathbf{s}_4 \in \mathcal{S}$, the $\mathbf{s} \in \mathcal{S}$ satisfying (5) is actually one of the $\mathbf{s}_i$. Finally, we call ϕ **very good** if (4) implies $\mathbf{s}_1 = \mathbf{s}_2 = \mathbf{s}_3 = \mathbf{s}_4$. So we have implications

$$\text{“very good”} \implies \text{“really good”} \implies \text{“good.”}$$

Example 1. Here are examples for the various flavors of “good.”

(1) From the right sketch of Figure 2 or 3 (which is drawn in two dimensions and therefore with only three $\mathbf{m}_i$'s and three $\mathbf{s}_i$'s), various examples can be gleaned. In the “likely” case that $\mathbf{s}_{\text{ghost}} \notin \mathcal{S}$, the group element $\phi = \text{id}$ is bad, and so is any horizontal shift. However, in the “unlikely” case that $\mathbf{s}_{\text{ghost}} \in \mathcal{S}$, these group elements are all good, but not really good. Now simultaneously shifting the $\mathbf{m}_i$ vertically until one of them is directly to the left or right of the corresponding $\mathbf{s}_i$ will shift $\mathbf{s}_{\text{ghost}}$ into this $\mathbf{s}_i$ and thus give a group element that is really good, but not very good. The group element $\phi = \text{id}$, or, in fact, any other, can be turned into a very good one simply by changing the $\mathbf{s}_i$ in such a way that the relation $f_D\left(\|\mathbf{s}_1 - \boldsymbol{\phi}(\mathbf{m}_1)\|^2, \dots, \|\mathbf{s}_3 - \boldsymbol{\phi}(\mathbf{m}_3)\|^2\right) = 0$ is no longer satisfied. Notice that in this example we only considered one choice of points $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3 \in \mathcal{S}$, while these should really be picked arbitrarily for checking goodness, and so forth.

(2) Assume that $\mathcal{S}$ consists of only the two points $\begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ -z \end{pmatrix}$, and that one of the $\mathbf{m}_i$, say $\mathbf{m}_1$, lies on the x-y-plane, that is, $\mathbf{m}_1 = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$. With $\mathbf{s}_1 = \begin{pmatrix} 0 \\ 0 \\ -z \end{pmatrix}$ and $\mathbf{s}_2 = \mathbf{s}_3 = \mathbf{s}_4 = \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$, $\mathbf{s}_1$ and $\mathbf{s}_4$ have the same distance to $\mathbf{m}_1$, and this remains true after applying a $\phi \in ASO_2$ to the $\mathbf{m}_i$. So

$$\|\mathbf{s}_4 - \phi(\mathbf{m}_i)\| = \|\mathbf{s}_i - \boldsymbol{\phi}(\mathbf{m}_i)\| \text{ for } i = 1, \dots, 4 \text{ and for all } \phi \in ASO_2. \tag{6}$$

By (2), this implies (4), so we see that no element of ASO_2 is very good. But for this choice of $\mathbf{s}_1, \dots, \mathbf{s}_4$, all elements are really good. To be indeed really good, they must also pass the test for other choices of the $\mathbf{s}_i \in \mathcal{S}$, and that may depend on the specific locations of the $\mathbf{m}_i$. If, for instance, also $\mathbf{m}_2$ and $\mathbf{m}_3$ lie on the x-y-plane, then (6) holds for all choices of $\mathbf{s}_i \in \mathcal{S}$, so then all $\phi \in \mathrm{ASO}_2$ are really good, but none is very good. It is due to examples like this that we need to consider the idea of really good elements.

When we say that **almost all** elements of $\mathrm{A_3SO_2}$ are really good, we mean that there is a subvariety of $\mathrm{A_3SO_2}$ of lower dimension, which contains all $\phi \in \mathrm{A_3SO_2}$ that are *not* really good. We use the same language for other properties and for the group $\mathrm{ASO_2}$. Of course, for $\mathrm{ASO_2}$, this means that the exceptional set has lower dimension than $\mathrm{ASO_2}$.

Finally, writing x, y, and z for the coordinates of a vector in $\mathbb{R}^3$, we say that $\mathcal{S}$ contains an **unlucky stack of points** if there are $\mathbf{s}_1, \dots, \mathbf{s}_4 \in \mathcal{S}$, not necessarily distinct, that share the same x- and y-coordinates (i.e., they lie on a vertical line), such that for $i, j \in \{1, \dots, 4\}$, the z-coordinate of the vector $\mathbf{s}_j - \mathbf{s}_i$ is twice the z-coordinate of $\mathbf{m}_j - \mathbf{m}_i$.

Our main result, in purely mathematical language, is the following dichotomy.

Theorem 1. *Let $\mathbf{m}_1, \dots, \mathbf{m}_4 \in \mathbb{R}^3$ be noncoplanar points and let $\mathcal{S} \subset \mathbb{R}^3$ a finite set of points. With the notation and language just introduced, the following is true.*

(a) If $\mathcal{S}$ contains an unlucky stack of points, then almost all elements of A_3SO_2 are bad.
(b) If $\mathcal{S}$ does not contain an unlucky stack of points, then almost all elements of A_3SO_2 are really good. Moreover, almost all elements of ASO_2 are really good.

Remarks on the proof. Before presenting the proof, let us briefly explain how part (b) translates to a problem in computational commutative algebra. For both groups $G = \mathrm{A_3SO_2}$ and $G = \mathrm{ASO_2}$, we need to show that for almost all $\phi \in G$ and for all $\mathbf{s}_1, \dots, \mathbf{s}_4 \in \mathcal{S}$, the implication

$$f_D(\|\mathbf{s}_1 - \phi(\mathbf{m}_1)\|^2, \dots, \|\mathbf{s}_4 - \phi(\mathbf{m}_4)\|^2) = 0 \implies \exists\, i\colon \|\mathbf{s}_i - \phi(\mathbf{m}_j)\| = \|\mathbf{s}_j - \phi(\mathbf{m}_j)\| \text{ for all } j$$

holds. Since we cannot control which points are in the given set $\mathcal{S}$, we need to show this for all $\mathbf{s}_1, \dots, \mathbf{s}_4 \in \mathbb{R}^3$ that do not form an unlucky stack. The implication is true if $f_D(\|\mathbf{s}_1 - \phi(\mathbf{m}_1)\|^2, \dots, \|\mathbf{s}_4 - \phi(\mathbf{m}_4)\|^2) \neq 0$, which is an *open* condition (in the Zariski topology). Since G is irreducible, it is therefore enough to show that this inequality holds for at least one ϕ. In the situation considered in our earlier paper[1], this turned out to be the case for all choices of $\mathbf{s}_i$ that are not all equal. But in the present situation, as shown in Example 1(2), it may happen that equality holds for all ϕ. In such a case, the right-hand side of the above implication must be satisfied for almost all ϕ. But this is a *closed* condition, so there needs to be an $i \in \{1, \dots, 4\}$ such that $\|\mathbf{s}_i - \phi(\mathbf{m}_j)\| = \|\mathbf{s}_j - \phi(\mathbf{m}_j)\|$ holds for all j and *all* $\phi \in G$. In summary, we need to show what is given as Claim 1 in the proof below, and, as the proof shows, this also suffices. Now it is not hard to work out equations (in 24 variables) that determine all 8-tuples $(\mathbf{m}_1, \dots, \mathbf{m}_4, \mathbf{s}_1, \dots, \mathbf{s}_4)$ that satisfy (7) or (8) in Claim 1, respectively, and equations that say that the $\mathbf{s}_i$ form an unlucky stack, and that the $\mathbf{m}_i$ are coplanar. This reduces the proof to a radical containment test for ideals in a polynomial ring with 24 variables.

However, running this test is far out of reach of the currently available hardware and software, and probably remains so for the foreseeable future. Therefore, a large part of the proof goes into "preprocessing" the given $\mathbf{m}_i$ and $\mathbf{s}_i$ in order to reduce the number of variables. This preprocessing can be viewed as a sequence of coordinate transformations, but we frame them group-theoretically. This way we can guarantee that the steps are permissible in view of the problem at hand, which is somewhat subtle. The preprocessing reduces the radical containment test to 16 variables. Even running this test was not possible without using tricks, such as truncated Gröbner bases and homogenization. □

Proof of Theorem 1. We start with part (a), so there is an unlucky stack $\mathbf{s}_1, \dots, \mathbf{s}_4 \in \mathcal{S}$. Choose the coordinates such that $\mathbf{s}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and write $z_1, \dots, z_4$ for the z-coordinates of the $\mathbf{m}_i$, so $\mathbf{s}_i = \begin{pmatrix} 0 \\ 0 \\ 2(z_i - z_1) \end{pmatrix}$. Let $\phi \in G$ be given by (3) with $a_i \in \mathbb{R}$. For a vector $\mathbf{v} \in \mathbb{R}^3$, we write $d_z(\mathbf{v})$ for the squared distance between $\mathbf{v}$ and the z-axis. Then,

$$\|\mathbf{s}_i - \phi(\mathbf{m}_i)\|^2 = d_z\big(\phi(\mathbf{m}_i)\big) + \big(2(z_i - z_1) - (z_i + a_5)\big)^2,$$

and with $\mathbf{s}_\phi := \begin{pmatrix} 0 \\ 0 \\ 2(z_1 + a_5) \end{pmatrix}$

$$\|\mathbf{s}_\phi - \phi(\mathbf{m}_i)\|^2 = d_z\big(\phi(\mathbf{m}_i)\big) + \big(2(z_1 + a_5) - (z_i + a_5)\big)^2,$$

so $\mathbf{s}_\phi$ satisfies (5). This means that ϕ is bad whenever $\mathbf{s}_\phi \notin \mathcal{S}$, which happens for all but finitely many values of a_5. This proves (a). Let us also remark that all elements of ASO_2 are bad unless the point $\begin{pmatrix} 0 \\ 0 \\ 2z_1 \end{pmatrix}$ happens to lie in $\mathcal{S}$.

Part (b) requires much more work. To deal with both statements simultaneously, let G be one of the groups A_3SO_2 or ASO_2. We have already explained before the proof, why it is necessary to prove the following claim.

Claim 1. Let $\mathbf{s}_1, \dots, \mathbf{s}_4 \in \mathbb{R}^3$ be points that do not form an unlucky stack. If

$$f_D\big(\|\mathbf{s}_1 - \phi(\mathbf{m}_1)\|^2, \dots, \|\mathbf{s}_4 - \phi(\mathbf{m}_4)\|^2\big) = 0 \quad \text{for all} \quad \phi \in G, \tag{7}$$

then there exists an $i \in \{1, \dots, 4\}$ such that

$$\|\mathbf{s}_j - \phi(\mathbf{m}_j)\| = \|\mathbf{s}_i - \phi(\mathbf{m}_j)\| \quad \text{for all } \phi \in G \text{ and all } j \in \{1, \dots, 4\}. \tag{8}$$

Before proving the claim, we show that it is sufficient, that is, that it implies part (b) of the theorem. For $\mathbf{s}_1, \dots, \mathbf{s}_4 \in \mathcal{S}$, consider the set

$$\mathcal{U}_{\mathbf{s}_1,\dots,\mathbf{s}_4} := \big\{\phi \in G \mid f_D\big(\|\mathbf{s}_1 - \phi(\mathbf{m}_1)\|^2, \dots, \|\mathbf{s}_4 - \phi(\mathbf{m}_4)\|^2\big) \neq 0\big\}.$$

Since $f_D\big(\|\mathbf{s}_1 - \phi(\mathbf{m}_1)\|^2, \dots, \|\mathbf{s}_4 - \phi(\mathbf{m}_4)\|^2\big)$ depends polynomially on the coefficients of the a_i defining ϕ according to (3), and since G is an irreducible variety, it follows that if $\mathcal{U}_{\mathbf{s}_1,\dots,\mathbf{s}_4} \neq \emptyset$, then its complement in G has a smaller dimension than G. Therefore, the intersection

$$\mathcal{U} := \bigcap_{\substack{\mathbf{s}_1,\ldots,\mathbf{s}_4\in S \text{ such that} \\ \mathcal{U}_{\mathbf{s}_1,\ldots,\mathbf{s}_4}\neq\emptyset}} \mathcal{U}_{\mathbf{s}_1,\ldots,\mathbf{s}_4} \subseteq G$$

also has a complement of smaller dimension than G. To prove part (b), it thus suffices to show that all $\phi \in \mathcal{U}$ are really good under the hypothesis that there is no unlucky stack of points. So let $\phi \in \mathcal{U}$, and let $\mathbf{s}_1,\ldots,\mathbf{s}_4 \in S$ such that $f_D\left(\|\mathbf{s}_1-\phi(\mathbf{m}_1)\|^2,\ldots,\|\mathbf{s}_4-\phi(\mathbf{m}_4)\|^2\right)=0$. Since $\phi\in\mathcal{U}$, this implies $\mathcal{U}_{\mathbf{s}_1,\ldots,\mathbf{s}_4}=\emptyset$. Now we apply Claim 1 (which we are assuming to be true). This tells us that there is an i such that $\|\mathbf{s}_j-\phi(\mathbf{m}_j)\| = \|\mathbf{s}_i-\phi(\mathbf{m}_j)\|$ holds for all j. (The claim makes this assertion for all elements of G, but we only need it for ϕ.) But this is just what it means for ϕ to be really good.

So we are left with proving Claim 1. Thus, we are given four noncoplanar vectors $\mathbf{m}_i$ and four vectors $\mathbf{s}_1,\mathbf{s}_2,\mathbf{s}_3,\mathbf{s}_4\in\mathbb{R}^3$ that do not form an unlucky stack. To make the computations in the final part of the proof feasible, we "preprocess" the given data. This preprocessing is best explained and justified by using yet another group. This is the group $\tilde{G}$ of all maps

$$\psi=\psi_{\sigma,\alpha,v_0}\colon\mathbb{R}^3\to\mathbb{R}^3,\ v\mapsto\alpha\cdot\sigma(v)+v_0,$$

where σ is an orthogonal map acting on the first two coordinates, $0\neq\alpha\in\mathbb{R}$ and $v_0\in\mathbb{R}^3$. It is crucial but straightforward to check that $\tilde{G}$ normalizes G (for both $G=\mathrm{A}_3\mathrm{SO}_2$ and $G=\mathrm{ASO}_2$), that is, $\psi^{-1}\phi\psi\in G$ for $\psi\in\tilde{G}$ and $\phi\in G$.

Now assume that we have chosen a suitable $\psi=\psi_{\sigma,\alpha,v_0}\in G$ and a $\phi_0\in\mathrm{ASO}_2$ such that we can prove Claim 1 for $\tilde{\mathbf{s}}_i := \psi(\mathbf{s}_i)$ and $\tilde{\mathbf{m}}_i := (\phi_0\circ\psi)(\mathbf{m}_i)$. To show that Claim 1 then also follows for the original $\mathbf{s}_i$ and $\mathbf{m}_i$, we first verify that the $\tilde{\mathbf{s}}_i$ do not form an unlucky stack. First, if two of the $\mathbf{s}_i$ have different x- or y-coordinates, then the same is true for the $\tilde{\mathbf{s}}_i$. Moreover, the z-coordinate of $\tilde{\mathbf{s}}_i-\tilde{\mathbf{s}}_j$ is α times the z-coordinate of $\mathbf{s}_i-\mathbf{s}_j$; and, likewise, the z-coordinate of $\tilde{\mathbf{m}}_i-\tilde{\mathbf{m}}_j$ is α times the z-coordinate of $\mathbf{m}_i-\mathbf{m}_j$. It follows that indeed the $\tilde{\mathbf{s}}_i$ do not form an unlucky stack.

Now assume that $f_D\left(\|\mathbf{s}_1-\phi(\mathbf{m}_1)\|^2,\ldots,\|\mathbf{s}_4-\phi(\mathbf{m}_4)\|^2\right)=0$ for all $\phi\in G$. We need to deduce the assertion (8) of Claim 1 from this. Let $\phi'\in G$. Then $\phi := \psi^{-1}\phi'\phi_0\psi\in G$ and $\phi'(\tilde{\mathbf{m}}_i) = (\psi\circ\phi)(\mathbf{m}_i)$. For $i,j\in\{1,\ldots,4\}$, we obtain

$$\|\tilde{\mathbf{s}}_i-\phi'(\tilde{\mathbf{m}}_j)\|^2 = \left\|\psi(\mathbf{s}_i)-\psi\left(\phi(\mathbf{m}_j)\right)\right\|^2 = \alpha^2\left\|\sigma\left(\mathbf{s}_i-\phi(\mathbf{m}_j)\right)\right\|^2 = \alpha^2\|\mathbf{s}_i-\phi(\mathbf{m}_j)\|^2. \tag{9}$$

For the $D_{i,j}=\|\phi(\mathbf{m}_i)-\phi(\mathbf{m}_j)\|^2$ that go into the polynomial f_D, the same calculation shows $\tilde{D}_{i,j} := \|\phi'(\tilde{\mathbf{m}}_i)-\phi'(\tilde{\mathbf{m}}_j)\|^2=\alpha^2 D_{i,j}$. With (9), we obtain

$$f_{\tilde{D}}\left(\|\tilde{\mathbf{s}}_1-\phi'(\tilde{\mathbf{m}}_1)\|^2,\ldots,\|\tilde{\mathbf{s}}_4-\phi'(\tilde{\mathbf{m}}_4)\|^2\right) = \det\begin{pmatrix} 0 & \tilde{d}_1 & \cdots & \tilde{d}_4 & 1 \\ \tilde{d}_1 & \tilde{D}_{1,1} & \cdots & \tilde{D}_{1,4} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \tilde{d}_4 & \tilde{D}_{4,1} & \cdots & \tilde{D}_{4,4} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix}$$

$$= \alpha^8\det\begin{pmatrix} 0 & d_1 & \cdots & d_4 & 1 \\ d_1 & D_{1,1} & \cdots & D_{1,4} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ d_4 & D_{4,1} & \cdots & D_{4,4} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} = \alpha^8 f_D\left(\|\mathbf{s}_1-\phi(\mathbf{m}_1)\|^2,\ldots,\|\mathbf{s}_4-\phi(\mathbf{m}_4)\|^2\right)=0.$$

Since we are assuming Claim 1 for the $\tilde{\mathbf{s}}_i$ and $\tilde{\mathbf{m}}_i$, there is an $i \in \{1, \dots, 4\}$ such that $\|\tilde{\mathbf{s}}_j - \phi'(\tilde{\mathbf{m}}_j)\| = \|\tilde{\mathbf{s}}_i - \phi'(\tilde{\mathbf{m}}_j)\|$ for all j and all $\phi' \in G$.

Now let $\phi \in G$ and set $\phi' := \psi\phi\psi^{-1}\phi_0^{-1} \in G$. Then, (9) yields

$$\|\mathbf{s}_j - \phi(\mathbf{m}_j)\| = |\alpha|^{-1}\|\tilde{\mathbf{s}}_j - \phi'(\tilde{\mathbf{m}}_j)\| = |\alpha|^{-1}\|\tilde{\mathbf{s}}_i - \phi'(\tilde{\mathbf{m}}_j)\| = \|\mathbf{s}_i - \phi(\mathbf{m}_j)\|$$

for all j, so indeed Claim 1 follows for the $\mathbf{s}_i$ and $\mathbf{m}_i$ if it is true for the $\tilde{\mathbf{s}}_i = \psi(\mathbf{s}_i)$ and $\tilde{\mathbf{m}}_i = (\phi_0 \circ \psi)(\mathbf{m}_i)$.

We use this to simplify the $\mathbf{s}_i$ and $\mathbf{m}_i$ in five steps. First, writing $\mathbf{s}_i = \begin{pmatrix} s_{i,1} \\ s_{i,2} \\ s_{i,3} \end{pmatrix}$ and $\mathbf{m}_i = \begin{pmatrix} m_{i,1} \\ m_{i,2} \\ m_{i,3} \end{pmatrix}$, we use $\psi = \psi_{\mathrm{id},1,v_0} \in \tilde{G}$ with $v_0 = \begin{pmatrix} -s_{1,1} \\ -s_{1,2} \\ -m_{1,3} \end{pmatrix}$. Then, $\tilde{\mathbf{s}}_1$ has x- and y-coordinates 0, and $\tilde{\mathbf{m}}_1$ has z-coordinate 0. Hence, without loss, we may assume these things of $\mathbf{s}_1$ and $\mathbf{m}_1$. The second step comes from a QR-decomposition of the top two rows of $S := (s_{i,j}) \in \mathbb{R}^{3\times 4}$. We have

$$\begin{pmatrix} s_{2,1} & s_{3,1} & s_{4,1} \\ s_{2,2} & s_{3,2} & s_{4,2} \end{pmatrix} = QR$$

with $Q \in \mathbb{R}^{2\times 2}$ orthogonal and $R \in \mathbb{R}^{2\times 3}$ upper triangular. Now forming $\sigma \in \mathbb{R}^{3\times 3}$ with Q^{-1} as upper left part, and using $\psi_{\sigma,1,0}$, we may assume $s_{2,2} = 0$. A bit more can be done: If the upper 2×4-part of S is nonzero and $k \in \{2, 3, 4\}$ is the number of the first nonzero column, then we may assume $s_{k,1} \neq 0$ and $s_{k,2} = 0$. As a third step, we use $\psi_{\mathrm{id},\alpha,0}$ with $\alpha = s_{k,1}^{-1}$. This means we can additionally assume $s_{k,1} = 1$. Summing up the first three steps, we may assume

$$S = \begin{pmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \mathbf{s}_3 & \mathbf{s}_4 \end{pmatrix} = \begin{pmatrix} 0 & 1 & b_1 & b_2 \\ 0 & 0 & b_3 & b_4 \\ b_5 & b_6 & b_7 & b_8 \end{pmatrix} \tag{10}$$

with $b_i \in \mathbb{R}$. (This is for the case $k = 2$; for $k = 3$ or 4, the upper part has more zeroes, and as a last case, which we indicate by setting $k := 5$, it may be all zeroes.) For the $\mathbf{m}_i$ we have only achieved $m_{1,3} = 0$, but so far, we have only used transformations where $\phi_0 = \mathrm{id}$. Now, as the fourth step, we set $\phi_0 \in \mathrm{ASO}_2$ to be the translation by the vector $-\mathbf{m}_1$. This yields $\tilde{\mathbf{m}}_1 = 0$, so we may assume $\mathbf{m}_1 = 0$. The last step uses a QR-decomposition of the upper 2×4-part of $M := (m_{i,j}) \in \mathbb{R}^{3\times 4}$ as we did before for S. (Notice that Q can be assumed to have determinant 1.) So finally we may assume

$$M = \begin{pmatrix} \mathbf{m}_1 & \mathbf{m}_2 & \mathbf{m}_3 & \mathbf{m}_4 \end{pmatrix} = \begin{pmatrix} 0 & c_1 & c_2 & c_3 \\ 0 & 0 & c_4 & c_5 \\ 0 & c_6 & c_7 & c_8 \end{pmatrix} \tag{11}$$

with $c_i \in \mathbb{R}$. The hypothesis that the $\mathbf{m}_i$ are not coplanar translates into $\det \begin{pmatrix} c_1 & c_2 & c_3 \\ 0 & c_4 & c_5 \\ c_6 & c_7 & c_8 \end{pmatrix} \neq 0$. This concludes the "preprocessing" of the given data.

For $a_1, \dots, a_5 \in \mathbb{R}$ with $a_1^2 + a_2^2 = 1$ defining $\phi_{a_1,\dots,a_5} \in \mathrm{A}_3\mathrm{SO}_2$ as in (3), and for S and M as in (10) and (11), there are polynomials $F(x_1, \dots, x_5, y_1, \dots, y_8, z_1, \dots, z_8)$ and $F_{i,j}(x_1, \dots, x_5, y_1, \dots, y_8, z_1, \dots, z_8)$ in 21 indeterminates such that

$$\|\mathbf{s}_i - \phi_{a_1,\dots,a_5}(\mathbf{m}_j)\|^2 = F_{i,j}(a_1, \dots, a_5, b_1, \dots, b_8, c_1, \dots, c_8) \quad (i, j \in \{1, \dots, 4\})$$

and

$$f_D\big(\|\mathbf{s}_1 - \phi_{a_1,\dots,a_5}(\mathbf{m}_1)\|^2, \dots, \|\mathbf{s}_4 - \phi_{a_1,\dots,a_5}(\mathbf{m}_4)\|^2\big) = F(a_1, \dots, a_5, b_1, \dots, b_8, c_1, \dots, c_8).$$

(In the case where $k \geq 3$, there are fewer b_i and hence fewer y_i.) If $G = \mathrm{A}_3\mathrm{SO}_2$, set $\mathcal{G} := \{x_1^2 + x_2^2 - 1\}$, and if $G = \mathrm{ASO}_2$, set $\mathcal{G} := \{x_1^2 + x_2^2 - 1, x_5\}$. In both cases, $\mathcal{G}$ is Gröbner basis, with respect to an arbitrary monomial order, of the vanishing ideal of G as a subvariety of $\mathbb{R}^5$. Consider the normal forms $\tilde{F} := \mathrm{NF}_{\mathcal{G}}(F)$ and $\tilde{F}_{i,j} := \mathrm{NF}_{\mathcal{G}}(F_{i,j})$. Let $J \subseteq \mathbb{R}[y_1, \dots, y_8, z_1, \dots, z_8]$ be the ideal generated by the coefficients of $\tilde{F}$, viewed as a polynomial in $x_1, \dots, x_5$ with coefficients in $\mathbb{R}[y_1, \dots, y_8, z_1, \dots, z_8]$. Moreover, for $i = 1, \dots, 4$, let $D_i \subseteq \mathbb{R}[y_1, \dots, y_8, z_1, \dots, z_8]$ be the ideals generated by the coefficients of all $\tilde{F}_{i,j} - \tilde{F}_{j,j}$ $(j = 1, \dots, 4)$. Also set $d := \det \begin{pmatrix} z_1 & z_2 & z_3 \\ 0 & z_4 & z_5 \\ z_6 & z_7 & z_8 \end{pmatrix}$.

Claim 2. If there is an r such that for each $k \in \{2, \dots, 5\}$

$$\big((d) \cdot D_1 \cdots D_4\big)^r \subseteq J \quad \text{in the case } k \leq 4 \tag{12}$$

or

$$\big((y_6 - y_5 - 2z_6, y_7 - y_5 - 2z_7, y_8 - y_5 - 2z_8) \cdot (d) \cdot D_1 \cdots D_4\big)^r \subseteq J \quad \text{in the case } k = 5, \tag{13}$$

then Claim 1 and therefore part (b) of the theorem follow. Recall that k is the number of the first nonzero column of the upper 2×4-submatrix of S, with $k = 5$ indicating that this submatrix is zero.

In fact, we have already seen by the "preprocessing" that to prove Claim 1, we may assume the $\mathbf{s}_i$ and $\mathbf{m}_i$ to be given by (10) and (11). Assume that the assertion (8) of Claim 1 is not true, so for every $i = 1, \dots, 4$, there is a $j \in \{1, \dots, 4\}$ and a $\phi = \phi_{a_1,\dots,a_5} \in G$ such that $\|\mathbf{s}_j - \phi_{a_1,\dots,a_5}(\mathbf{m}_j)\| \neq \|\mathbf{s}_i - \phi_{a_1,\dots,a_5}(\mathbf{m}_j)\|$. Then,

$$0 \neq F_{i,j}(a_1, \dots, a_5, b_1, \dots, b_8, c_1, \dots, c_8) - F_{j,j}(a_1, \dots, a_5, b_1, \dots, b_8, c_1, \dots, c_8).$$

Since $F_{i,j} - \tilde{F}_{i,j}$ vanishes when setting $x_k = a_k$, we may replace $F_{i,j} - F_{j,j}$ in the above inequality by $\tilde{F}_{i,j} - \tilde{F}_{j,j}$. So there is a j such that some coefficient of $\tilde{F}_{i,j} - \tilde{F}_{j,j}$, viewed as a polynomial in $x_1, \dots, x_5$, does not vanish when evaluated at $y_i = b_i$ and $z_i = c_i$. This coefficient is one of the generators of D_i, so each D_i contains an element that does not vanish at $y_i = b_i$ and $z_i = c_i$.

Moreover, d does not vanish when evaluated at $z_i = c_i$, and we might as well also evaluate it at $y_i = b_i$ since the y_i do not occur in d. So in the case $k \leq 4$, (12) implies that there exists an element of J that does not vanish when evaluated at $y_i = b_i$ and $z_i = c_i$.

On the other hand, if $k = 5$, then all $\mathbf{s}_i$ share the same x- and y-coordinates, so the hypothesis that they do not form an unlucky stack translates into $b_i - b_5 \neq 2c_i$ for some $i \in \{6, 7, 8\}$. So one of the generators of the first ideal in the product in (13) does not vanish when evaluated at $y_i = b_i$. In this case, (13) yields the same conclusion: that J contains an element that does not vanish when evaluated at $y_i = b_i$ and $z_i = c_i$.

Because of the way J was constructed, this means that at least one coefficient of $\tilde{F}$ does not vanish at $y_i = b_i$ and $z_i = c_i$, so

$$\tilde{F}(x_1, \dots, x_5, b_1, \dots, b_8, c_1, \dots, c_8) \neq 0.$$

The linearity of the normal form map implies

$$\mathrm{NF}_{\mathcal{G}}\big(F(x_1, \dots, x_5, b_1, \dots, b_8, c_1, \dots, c_8)\big) = \tilde{F}(x_1, \dots, x_5, b_1, \dots, b_8, c_1, \dots, c_8),$$

so we conclude that $F(x_1, \dots, x_5, b_1, \dots, b_8, c_1, \dots, c_8)$ has nonzero normal form and therefore does not lie in the ideal generated by $\mathcal{G}$. Since this is the vanishing ideal of G, this means that there exists a $\phi = \phi_{a_1,\dots,a_5} \in G$ such that

$$0 \neq F(a_1, \dots, a_5, b_1, \dots, b_8, c_1, \dots, c_8) = f_D\big(\|\mathbf{s}_1 - \phi(\mathbf{m}_1)\|^2, \dots, \|\mathbf{s}_4 - \phi(\mathbf{m}_4)\|^2\big).$$

So the hypothesis (7) of Claim 1 is not true under the assumption that the assertion is not true. So indeed Claim 1 follows from (12) and (13).

It remains to verify Equations (12) and (13), and this can be checked with the help of a computer. For the computation, we used MAGMA[25] and proceeded as follows:

- We ran the following steps for each $k \in \{2, \dots, 5\}$. We will not introduce any notation to indicate the cases for different k below.
- It is straightforward to compute the polynomials F, $F_{i,j}$, $\tilde{F}$, and $\tilde{F}_{i,j}$ according to their definitions, and to pick out the sets of coefficients $C, C_{i,j} \subset \mathbb{R}[y_1, \dots, y_8, z_1, \dots, z_8]$ of $\tilde{F}$ and the $\tilde{F}_{i,j}$.
- Using an additional indeterminate t, we computed the set C^{hom} of homogenizations of the polynomials in C with respect to t.
- We computed truncated Gröbner bases G^{hom} of the ideal generated by C^{hom} of rising degree.
- For each degree, we computed the normal form $\mathrm{NF}_{G^{\mathrm{hom}}}\big(t^i d^j\big)$ for all i and j such that $t^i d^j$ has the degree up to which the Gröbner basis was computed. If at least one of the normal forms is zero, this shows that $t^i d^j$ is an $\mathbb{R}[y_1, \dots, y_8, z_1, \dots, z_8, t]$-linear combination of the polynomials in C^{hom}. Setting $t = 1$ shows that $d^j \in J$, so Equation 12 holds with $r = j$. As it turned out, this was successful for all $k \leq 4$, and we never needed to go beyond degree 11. This means that in these cases, the ideals D_i are not actually needed in Equation 12.
- Finally we dealt with the case $k = 5$. Here, thanks to the reduced number of variables, we were able to directly compute the ideal $J_0 := J : (d)^\infty$, and then $J_i := J_{i-1} : D_i^\infty$ for $i = 1, \dots, 4$. The result is $J_4 = (y_6 - y_5 - 2z_6, y_7 - y_5 - 2z_7, y_8 - y_5 - 2z_8)$, which shows (13).

The total computation time is about 8 min. □

Remark 1. For the group A_3SO_2, our computations actually yielded a slightly stronger result than Theorem 1(b): Almost all elements of A_3SO_2 are very good, not only really good.

5 | A GROUND-BASED VEHICLE AND A HOVERING DRONE

Before stating our main results in a more application-related fashion, it is useful to recall how the wall detection algorithm from Ref. 1 works. So from now on, the points $\mathbf{m}_1, \dots, \mathbf{m}_4 \in \mathbb{R}^3$ will be the positions of the microphones mounted on a vehicle, and points $\mathbf{s}$ or $\mathbf{s}_i$ will stand for mirror points, that is, points given by the reflection of the position $\mathbf{L} \in \mathbb{R}^3$ of a single loudspeaker at various walls W_i in a room. Virtually, the sound reflected by a wall comes from a mirror point. In a nutshell, the detection algorithm uses the relation (2), given by the vanishing of the Cayley–Menger determinant, for matching those echoes received by the microphones that come from the same wall. Then, the following result is used to compute the mirror points and from them the location of the walls. A proof can be found in Ref. 1.

Proposition 1. *Let $\mathbf{m}_1, \dots, \mathbf{m}_4$ and $\mathbf{s}$ be points in $\mathbb{R}^3$, with the $\mathbf{m}_i$ not coplanar. Set*

$$M := \left(\begin{array}{cccc} \mathbf{m}_1 & \mathbf{m}_2 & \mathbf{m}_3 & \mathbf{m}_4 \\ \hline 1 & 1 & 1 & 1 \end{array}\right) \in \mathbb{R}^{4\times 4},$$

which is invertible since the $\mathbf{m}_i$ are not coplanar.

(a) The point $\mathbf{s}$ is uniquely determined by the $\mathbf{m}_i$ and the squared distances $d_i := \|\mathbf{s} - \mathbf{m}_i\|^2$, and can be computed by

$$\mathbf{s} = \frac{1}{2}\tilde{M} \cdot \begin{pmatrix} \|\mathbf{m}_1\|^2 - d_1 \\ \vdots \\ \|\mathbf{m}_4\|^2 - d_4 \end{pmatrix}, \tag{14}$$

where $\tilde{M} \in \mathbb{R}^{3\times 4}$ is the upper 3×4-part of $(M^{-1})^T$, the transpose inverse of M.

(b) Let $\mathbf{L} \in \mathbb{R}^3$ be the position of the loudspeaker. Then, the wall at which the sound was reflected lies on the plane with normal vector $\mathbf{s} - \mathbf{L}$ and passing through the point $\frac{1}{2}(\mathbf{s} + \mathbf{L})$. Four noncollinear points on the wall can be found by intersecting the line between $\mathbf{s}$ and $\mathbf{m}_i$ $(1 \le i \le 4)$ with this plane. These points are given by

$$(1 - \tau_i)\mathbf{s} + \tau_i \mathbf{m}_i \quad \textit{with} \quad \tau_i = \frac{\|\mathbf{s} - \mathbf{L}\|^2}{2\langle \mathbf{s} - \mathbf{L}, \mathbf{s} - \mathbf{m}_i\rangle},$$

where $\langle \cdot, \cdot \rangle$ denotes the standard scalar product.

With this, we can state the wall detection algorithm as Algorithm 1.

Algorithm 1 is guaranteed to detect every wall from which a first-order echo is heard by all four microphones. But as discussed in Section 3, it may happen that the relation $f_D(d_1, \dots, d_4) = 0$ is satisfied by accident even though the squared distances d_i come from sound reflections at different walls. This can deceive the algorithm into outputting walls that do not actually exist. As mentioned above, these are called **ghost walls**. If the algorithm does not output any ghost walls, we say that the vehicle carrying the microphones (and possibly the loudspeaker) is in a **good position**.

Theorem 3 (A ground-based vehicle in a scene with a fixed loudspeaker). *Assume that four microphones are mounted on a ground-based vehicle, which can move on the ground plane within a*

ALGORITHM 1 Detect walls from first-order echoes

Input: The delay times of the first-order echoes recorded by four microphones at positions $\mathbf{m}_1, \dots, \mathbf{m}_4 \in \mathbb{R}^3$, and the distances $D_{i,j}$ between the microphones.

For $i = 1, \dots, 4$, collect the recorded times of the first-order echoes recorded by the ith microphone in the set $\mathcal{T}_i$.

Set $\mathcal{D}_i := \{c^2(t - t_0)^2 \mid t \in \mathcal{T}_i\}$ $(i = 1, \dots, 4)$, where c is the speed of sound and t_0 is the time of sound emission.

for $(d_1, d_2, d_3, d_4) \in \mathcal{D}_1 \times \mathcal{D}_2 \times \mathcal{D}_3 \times \mathcal{D}_4$ **do**

  With f_D defined by Equation 1, evaluate $f_D(d_1, \dots, d_4)$.

  if $f_D(d_1, \dots, d_4) = 0$ **then**

    Use Proposition 1(a) to compute the mirror point $\mathbf{s}$ from $(d_1, \dots, d_4)$. Then use Proposition 1(b) to compute the parameters of the wall.

    Output the parameters of this wall.

  end if

end for

three-dimensional scene containing n walls and a loudspeaker at fixed positions. Assume the microphones do not lie on a common plane. If the microphones are not deceived by an unlucky stack of mirror points, according to Definition 1, then almost all vehicle positions are good.

Moreover, if $l \in \{2, 3, 4\}$ *is the number of distinct z-coordinates of the four microphone positions, then within the* $(3n)$*-dimensional configuration space of all wall arrangements, the ones where an unlucky stack of mirror points occurs are contained in a subvariety of codimension* $3(l - 1)$.

Proof. We are given n walls and the loudspeaker position $\mathbf{L} \in \mathbb{R}^3$. These do not vary as the vehicle moves, so neither do the mirror points, which form a finite set $\mathcal{S} \subset \mathbb{R}^3$. We are also given initial positions $\mathbf{m}_1, \dots, \mathbf{m}_4 \in \mathbb{R}^3$ of the microphones mounted on the vehicle. As the vehicle moves, these are acted on simultaneously by the group ASO_2 (see after (3)). Ray acoustics tells us that the d_i occurring in 1 are the $\|\mathbf{s}_i - \phi(\mathbf{m}_i)\|^2$ with $\mathbf{s}_i \in \mathcal{S}$. So we are exactly in the situation of Theorem 1, and part (b) tells us that almost all $\phi \in \mathrm{ASO}_2$ are good. This means that almost all vehicle positions are good.

Let us prove the last statement about the codimension of the wall arrangements with unlucky stacks of mirror points. To have such a stack, l walls are needed to satisfy the restrictions of Definition 1. These restrictions leave 3 degrees of freedom for the mirror points of these l walls, which gives an affine subspace of codimension $3l - 3 = 3(l - 1)$. The wall arrangements with unlucky stacks are contained in the (finite) union of these subspaces associated to each choice of l walls, so we obtain the claimed codimension. □

In the same way, using Theorem 1(b) for the group $\mathrm{A}_3\mathrm{SO}_2$, we obtain the following result.

Theorem 4 (A hovering drone and a fixed loudspeaker). *Assume the same situation as in Theorem 3, except that the vehicle can take positions in the same way as a hovering drone. Then, the assertions from Theorem 3 hold.*

6 | THE TWO-DIMENSIONAL CASE

In this section, we briefly consider the purely two-dimensional case: A vehicle moves in the plane, and all reflecting walls are also in this plane. A physical example of such a scene is a robot navigating in a warehouse, with no echoes coming back from either the ceiling or the floor, or from any inclined walls. In this case, only three microphones on the vehicle are required. The loudspeaker can be at a fixed position, but will more typically be mounted on the vehicle. In contrast to the case of a ground-based vehicle moving in 3D, we can also handle the case of a mounted loudspeaker here. In fact, this is the two-dimensional variant of the situation considered in Ref. 1, and everything in that paper carries over directly to other dimensions. This includes the relation, given as (2) in this paper, and the wall detection algorithm. Moreover, the MAGMA-programs used for the computational verifications in Ref. 1 were written for general dimension. As it turns out, if the dimension is set to 2, the entire computations require only about 1 s. (They take about 5 min for the three-dimensional case considered in Ref. 1.) Notice that for a two-dimensional modeling to be admissible, the microphones and the loudspeaker need to be contained in a common plane, which is parallel to the plane of motion. The following result emerges.

Theorem 5 (A vehicle in a two-dimensional scene). *Assume a vehicle can move in a two-dimensional scene, which contains a finite number of walls. Assume a loudspeaker is either placed at a fixed position in the scene or mounted on the vehicle, and that three microphones are mounted on the vehicle, but do not lie on a common line. In any case, the microphones and the loudspeaker need to be in a common plane, which is parallel to the plane of motion. Then, almost all vehicle positions are good, again in the sense that all walls whose echoes are heard by every microphone are detected, but no ghost walls are detected.*

In the two-dimensional situation, the issue of unlucky stacks does not occur. Let us also mention that the hypothesis about the microphones and the loudspeaker all lying on a common horizontal plane is not really necessary. In fact, deviations from a common horizontal plane can be taken care of by subtracting a corrective term from the measured squared distances.

The case of a fixed loudspeaker in Theorem 5 is very simple to state in purely mathematical terms: For $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3 \in \mathbb{R}^2$ noncollinear and $\mathcal{S} \subset \mathbb{R}^2$ finite, almost all elements of ASO_2 are very good. This is what we actually proved. Of course the case of a mounted loudspeaker can also be expressed mathematically, but it is messier, and therefore less appealing, since the mirror points are not fixed but depend on the group element, which is applied to the loudspeaker position.

ACKNOWLEDGMENTS

This work has benefited immensely from a research stay of the two authors at the Banff International Research Station for Mathematical Innovation and Discovery (BIRS) under the "Research in Teams" program. We would like to thank BIRS for its hospitality and for providing an optimal working environment.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this paper as no data sets were generated or analyzed during this study.

REFERENCES

1. Boutin M, Kemper G. A drone can hear the shape of a room. *SIAM J Appl Algebra Geometry*. 2020;4:123-140.
2. Alien JB, Berkley DA. Image method for efficiently simulating small-room acoustics. *J Acoust Soc Am*. 1976;60(S1):S9-S9.
3. Borish J. Extension of the image model to arbitrary polyhedra. *J Acoust Soc Am*. 1984;75(6):1827-1836.
4. Moebus M, Zoubir AM. Three-dimensional ultrasound imaging in air using a 2D array on a fixed platform. 2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07. vol. 2, IEEE; 2007:II-961.

5. Dokmanić I, Lu YM, Vetterli M. Can one hear the shape of a room: The 2-D polygonal case. 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2011:321-324.
6. Antonacci F, Filos J, Thomas MRP, et al. Inference of room geometry from acoustic impulse responses. *IEEE Trans Audio Speech Lang Process*. 2012;20(10):2683-2695.
7. El Baba Y, Walther A, Habets EAP. Reflector localization based on multiple reflection points. 2016 24th European Signal Processing Conference (EUSIPCO). IEEE; 2016:1458-1462.
8. Park S, Choi JW. Iterative echo labeling algorithm with convex hull expansion for room geometry estimation. *IEEE/ACM Trans Audio Speech Lang Process*. 2021;29:1463-1478
9. Dokmanić I, Parhizkar R, Walther A, Lu YM, Vetterli M. Acoustic echoes reveal room shape. *Proc Natl Acad Sci*. 2013;110:12186-12191.
10. Canclini A, Antonacci F, Thomas MRP, et al. Exact localization of acoustic reflectors from quadratic constraints. 2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA). IEEE; 2011:17-20.
11. Filos J, Canclini A, Thomas MRP, Antonacci F, Sarti A, Naylor PA. Robust inference of room geometry from acoustic measurements using the Hough transform. 2011 19th European Signal Processing Conference. IEEE; 2011:161-165.
12. Remaggi L, Jackson PJB, Wang W, Chambers JA. A 3D model for room boundary estimation. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2015:514-518.
13. Tervo S, Tossavainen T. 3D room geometry estimation from measured impulse responses. 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2012:513-516.
14. Mabande E, Kowalczyk K, Sun H, Kellermann W. Room geometry inference based on spherical microphone array eigenbeam processing. *J Acoust Soc Am*. 2013;134(4):2773-2789.
15. Remaggi L, Jackson PJB, Coleman P, Wang W. Acoustic reflector localization: novel image source reversion and direct localization methods. *IEEE/ACM Trans Audio Speech Lang Process*. 2016;25(2):296-309.
16. El Baba Y, Walther A, Habets EAP. 3D room geometry inference based on room impulse response stacks. *IEEE/ACM Trans Audio Speech Lang Process*. 2017;26(5):857-872.
17. Jager I, Heusdens R, Gaubitch ND. Room geometry estimation from acoustic echoes using graph-based echo labeling. 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2016:1-5.
18. Rajapaksha T, Qiu X, Cheng E, Burnett I. Geometrical room geometry estimation from room impulse responses. 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2016:331-335.
19. Scheuing J, Yang B. Disambiguation of TDOA estimation for multiple sources in reverberant environments. *IEEE Trans Audio Speech Lang Process*. 2008;16(8):1479-1489.
20. Pollefeys M, Nister D. Direct computation of sound and microphone locations from time-difference-of-arrival data. 2008 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE; 2008:2445-2448.
21. Peng F, Wang T, Chen B. Room shape reconstruction with a single mobile acoustic sensor. 2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP). IEEE; 2015:1116-1120.
22. Shih O, Rowe A. Can a phone hear the shape of a room? 2019 18th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN). IEEE; 2019:277-288.
23. Zhou B, Elbadry M, Gao R, Ye F. BatMapper: Acoustic sensing based indoor floor plan construction using smartphones. Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services. 2017:42-55.
24. Cayley A. On a theorem in the geometry of position. *Camb J Math*. 1841;II:267-271.
25. Bosma W, Cannon JJ, Playoust C. The magma algebra system I: the user language. *J Symb Comput*. 1997;24:235-265.

How to cite this article: Boutin M, Kemper G. Can a ground-based vehicle hear the shape of a room? *Stud Appl Math*. 2023;151:352–368. https://doi.org/10.1111/sapm.12581