Supporting Information

Dokmanić et al. 10.1073/pnas.1221464110

SI Text

1. Proof of the Theorem

Theorem 1. *Consider a room with a loudspeaker and $M \geq 4$ microphones placed uniformly at random inside the feasible region.* Then the set of first-order echoes uniquely specifies the room with probability 1. In other words, almost surely exactly one assignment of first-order echoes to walls describes a room.

Proof: It is sufficient to prove the claim for $M = 4$. Cases when $M > 4$ follow by considering any subset of four microphones. Draw independently and uniformly at random microphone locations $\mathbf{r}_1, \ldots, \mathbf{r}_4$ in the feasible region. To this particular choice of microphone locations we correspond vectors $\mathbf{y}_k$ and $\tilde{\mathbf{y}}_k$ as follows,

$$y_{k,m} \overset{def}{=} \left\|\tilde{\mathbf{s}}_k - \mathbf{r}_m\right\|^2 = \left\|\tilde{\mathbf{s}}_k\right\|^2 - 2\,\tilde{\mathbf{s}}_k^\top \mathbf{r}_m + \left\|\mathbf{r}_m\right\|^2, \qquad [S1]$$

and

$$\tilde{y}_{k,m} \overset{def}{=} -\frac{1}{2}\left(y_{k,m} - \left\|\mathbf{r}_m\right\|^2\right), \qquad [S2]$$

where $\tilde{\mathbf{s}}_k$ is the location of the image source with respect to wall k. We have in vector form

$$\begin{bmatrix} \tilde{y}_{k,1} \\ \tilde{y}_{k,2} \\ \tilde{y}_{k,3} \\ \tilde{y}_{k,4} \end{bmatrix} = \left[\begin{array}{c|c} \mathbf{r}_1^\top & -\frac{1}{2} \\ \mathbf{r}_2^\top & -\frac{1}{2} \\ \mathbf{r}_3^\top & -\frac{1}{2} \\ \mathbf{r}_4^\top & -\frac{1}{2} \end{array}\right] \begin{bmatrix} \tilde{\mathbf{s}}_k \\ \left\|\tilde{\mathbf{s}}_k\right\|^2 \end{bmatrix}, \quad \text{or} \quad \tilde{\mathbf{y}}_k = \mathbf{R}\tilde{\mathbf{u}}_k. \qquad [S3]$$

Thanks to the condition $\sum_{m=1}^4 \mathbf{r}_m = 0$, we have that

$$1^\top \tilde{\mathbf{y}}_k = -\frac{M}{2}\left\|\tilde{\mathbf{s}}_k\right\|^2 \quad \text{or} \quad \left\|\tilde{\mathbf{s}}_k\right\|^2 = -\frac{2}{M}\sum_{m=1}^{M} \tilde{y}_{k,m}. \qquad [S4]$$

The image source is found as

$$\tilde{\mathbf{s}}_k = \mathbf{S}\tilde{\mathbf{y}}_k, \qquad [S5]$$

where $\mathbf{S}$ is a matrix satisfying

$$\mathbf{SR} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \qquad [S6]$$

It follows that

$$\frac{2}{4}1^\top \tilde{\mathbf{y}}_k + \left\|\mathbf{S}\tilde{\mathbf{y}}_k\right\|^2 = 0. \qquad [S7]$$

Vector $\tilde{\mathbf{y}}_k$ corresponds to the kth wall, or kth image source (it is the correct permutation). We now show that wrong permutations cannot satisfy Eq. **S7**. We do it by replacing one, two, or three entries in $\tilde{\mathbf{y}}_k$ by wrong values and arguing that these are not good combinations. We choose

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{R}^\dagger, \qquad [S8]$$

where $\mathbf{R}^\dagger$ is the Moore–Penrose pseudoinverse of $\mathbf{R}$. With this choice, any column submatrix of $\mathbf{S}$ with $n \leq 3$ columns is rank n with probability 1.

1. *(1 replacement).* Without loss of generality, let us replace the fourth entry of $\tilde{\mathbf{y}}_k(\tilde{y}_{k,4})$, by $\tilde{y}_{k',4}$, $k' \neq k$, and plug it into Eq. **S7**. We can rewrite the equation as

$$\alpha + \beta\tilde{y}_{k',4} + \gamma\tilde{y}_{k',4}^2 = 0, \qquad [S9]$$

where α, β, and γ do not functionally depend on $\tilde{y}_{k',4}$, and $\gamma \neq 0$ with probability 1. For any realization of $\tilde{y}_{k,1}, \ldots, \tilde{y}_{k,3}$, the distribution of $\tilde{y}_{k',4}$ is continuous, thus the probability that it assumes any given value is zero (note that this is not true for $\tilde{y}_{k,4}$—for echoes coming from the same wall, knowing three of them constraints the fourth to two possible values). Therefore, the probability that $\tilde{y}_{k',4}$ is one of at most two real roots of Eq. **S9** is zero.

2. *(2 replacements).* Now we replace $\tilde{y}_{k,3}$ and $\tilde{y}_{k,4}$ by $\tilde{y}_{k',3}$ and $\tilde{y}_{k'',4}$. We can have either (i) $k' = k''$ or (ii) $k' \neq k''$. We rewrite Eq. **S7** as

$$\begin{bmatrix} \tilde{y}_{k',3} & \tilde{y}_{k'',4} \end{bmatrix} \mathbf{A} \begin{bmatrix} \tilde{y}_{k',3} \\ \tilde{y}_{k'',4} \end{bmatrix} + \mathbf{a}^\top \begin{bmatrix} \tilde{y}_{k',3} \\ \tilde{y}_{k'',4} \end{bmatrix} + a = 0, \qquad [S10]$$

where $\mathbf{A} = \mathbf{S}[:, 3:4]^\top \mathbf{S}[:, 3:4]$ (with Matlab notation) is full rank with probability 1 and is positive semidefinite. Also, $\mathbf{A}$, $\mathbf{a}$ and a do not functionally depend on $\begin{bmatrix} \tilde{y}_{k',3} & \tilde{y}_{k'',4} \end{bmatrix}^\top$. Locus of the roots of Eq. **S10** is an ellipse. However, for any realization of $\tilde{y}_{k,1}$ and $\tilde{y}_{k,2}$ the distribution of $\begin{bmatrix} \tilde{y}_{k',3} & \tilde{y}_{k'',4} \end{bmatrix}^\top$ is continuous over some 2D subset of $\mathbb{R}^2$ both in cases i and ii. Therefore, the probability that it takes a value on the root ellipse of Eq. **S10** is zero.

3. *(3 replacements).* Here we replace $\tilde{y}_{k,2}, \tilde{y}_{k,3}, \tilde{y}_{k,4}$ with $\tilde{y}_{k',2}, \tilde{y}_{k'',3}, \tilde{y}_{k''',4}$. If $k' = k'' = k'''$, then the argument is the same as in the case of one replacement. If $k' = k''$ or $k' = k'''$ or $k'' = k'''$, but not all three are equal, then we can just repeat the argument for the case of 2 replacements (ii). Finally if they are all different, we write

$$\begin{bmatrix} \tilde{y}_{k',2} & \tilde{y}_{k'',3} & \tilde{y}_{k''',4} \end{bmatrix} \mathbf{B} \begin{bmatrix} \tilde{y}_{k',2} \\ \tilde{y}_{k'',3} \\ \tilde{y}_{k''',4} \end{bmatrix} + \mathbf{b}^\top \begin{bmatrix} \tilde{y}_{k',2} \\ \tilde{y}_{k'',3} \\ \tilde{y}_{k''',4} \end{bmatrix} + b = 0. \qquad [S11]$$

Again $\mathbf{B} = \mathbf{S}[:, 2:4]^\top \mathbf{S}[:, 2:4]$ is full rank with probability 1, so the locus of the roots of Eq. **S11** is an ellipsoid. The set of values that $\begin{bmatrix} \tilde{y}_{k',2} & \tilde{y}_{k'',3} & \tilde{y}_{k''',4} \end{bmatrix}$ takes is again some 3D region in $\mathbb{R}^3$ and the probability that the triplet takes value on an ellipsoid is zero.

In conclusion, almost surely only one (correct) combination of echoes satisfies Eq. **S7**, so almost surely only one room corresponds to collected first-order echoes.

2. Experimental Setup

2.1. Equipment. We used a Lange D12A dodecahedron omnidirectional loudspeaker (Fig. S1*A*) and a two-way directional active monitoring loudspeaker Genelec 8030A (Fig. S1*B*). The

horizontal beam patter of Genelec 8030A is depicted in Fig. S2. The horizontal directivity sonogram of Lange D12A is shown in Fig. S3.

To record the responses, we used five nonmatched Behringer ECM 8000 omnidirectional measurement microphones (Fig. S1*C*). The microphones and the loudspeaker were interfaced with a PC through a Motu 896HD unit (Fig. S1*D*) operating at a sampling frequency of 96 kHz.

2.2. Microphone Arrays. Table S1 contains the distances between the microphones in the experiments. Distances were measured between the tips of the omnidirectional microphones using a tape measure.

2.3. Measurement Technique. We measured the room impulse responses by the swept-sine technique (1). An excitation signal is played back over the chosen loudspeaker while simultaneously recording the signals arriving at the microphones. The played signal is a sine sweep with an instantaneous frequency varying exponentially with time,

$$x(t) = \sin\left[\frac{\omega_1 T}{\ln(\omega_2/\omega_1)}\left(e^{\frac{t}{T}\ln(\omega_2/\omega_1)} - 1\right)\right], \qquad \textbf{[S12]}$$

where ω_1 is the start frequency, ω_2 is the end frequency, and T is the total duration of the sweep in seconds. The recorded signals $(y(t))$ can be written in Fourier domain as $Y(\omega) = H(\omega)S(\omega)$. Hence, the room transfer function $(H(\omega))$ can be computed by spectral division,

$$H(\omega) = \frac{Y(\omega)}{S(\omega)}. \qquad \textbf{[S13]}$$

Inverse Fourier transforming $H(\omega)$ yields the impulse response.

2.4. Remark About Loudspeakers. Peak-picking and RIR measurement techniques are out of the scope of this paper. Nevertheless, the loudspeaker size, build, and impulse response affect the quality of the estimation. This effect is indirect through the peak shape. Size is relevant, as we assume a point source. This assumption is satisfied for the directional speaker that has a well-defined acoustic center. However, the omnidirectional loudspeaker has widely placed drivers, so it is a poor approximation of a point source. We can partially compensate for the speaker size by assuming it is spherical and using the Huygens principle, but the structure of the impulse response still reflects the distributed drivers.

2.5. Delay of the Processing Chain. We measured the total delay of the processing chain to be 365 samples at the sampling frequency of 96 kHz. This offset must be accounted for when processing the impulse responses. In the case of the Lange omnidirectional loudspeaker, we used a smaller offset of 338 samples. The reason for this is to compensate for the loudspeaker size: At time 0, the sound wave is already at a distance R from the center of the loudspeaker, where R is the speaker radius. Therefore, the delay until the sound reaches some point in space is smaller than for the point source. Of course, this would be correct if the loudspeaker was a perfect sphere. In practice, we can only compensate the radius "in the mean." That is also why in general, the results obtained with the omnidirectional loudspeaker are slightly less accurate than with the directional one.

2.6. Experiment in the Lausanne Cathedral. The measurements in the side portal of the Lausanne cathedral were challenging as a large part of the boundary surfaces are not flat (as assumed by the algorithm, and as has been the case in the classroom measurements).

The glass front (numeral 1 in Fig. 6*F*) and the floor beneath the microphone array can be considered flat surfaces. For all of the other boundaries of the room, this assumption does not hold. The arched roof cannot be represented by a single height estimate. The side windows (numerals 2 and 3 in Fig. 6*F*) with pillars placed in front and erratic structural elements at the height of the microphones, the rear wall, and the angled corners with large pillars and large statues, all present irregular surfaces creating diffuse reflections. Fig. 4 shows the details of the sidewall structure and the microphone arrangement. The waveform of a reflection from such diffusive architectural surfaces exhibits distinct differences compared with one from a large, flat surface. In general, such a reflection response is temporally spread and has a lower peak amplitude than an impulse containing the same energy (2). These characteristics are unfavorable for our algorithm because it is harder to detect peaks that actually belong to walls. Many of the detected peaks stem from reflections off small structural elements. The purpose of this experiment was to get an idea about the robustness of the echo-sorting algorithm to inputs from measurements made in environments that violate the assumptions made by the proposed algorithm.

The measurement procedure and equipment was the same as in the EPFL classroom measurements. As has been the case before, the microphones were not calibrated; the single channel preamp potentiometers had approximately equal settings.

To measure an impulse response of a room, a high-level broadband excitation signal is needed. Using an impulse as excitation signal, the recorded signal is immediately the impulse response. Unfortunately, impulsive sources (e.g., popping a balloon or firing a starter pistol) have poor repeatability, produce unpredictable spectra, and do not guarantee omnidirectionality (3). Assuming that the side portal is a linear and time invariant system, the required energy can be spread over time. We excite the room with a deterministic signal, and the room impulse response can be calculated from the signal recorded in the room. We applied the swept-sine technique as described in *SI Text*, section 2.3.

The loudspeaker used in this experiment is not omnidirectional. Therefore, in addition to the position of the loudspeaker and the microphones, the measured impulse responses depend on the orientation of the loudspeaker. This has been considered by placing the directional loudspeaker close to one wall. The microphone array was positioned in the lowered center part of the portal, which was surrounded on all four sides by stairs. We recorded the impulse responses with a randomly setup microphone arrangement.

The described effects of the architectural structures in the cathedral portal become apparent in the recorded impulse responses. Fig. S5 shows a comparison of impulse responses recorded in the lecture room and in the cathedral. We can see that the number of distinct peaks in the cathedral impulse responses is smaller than in the classroom measurement, and that the peaks in the cathedral RIR have lower amplitudes compared with the direct sound than the peaks in the classroom (the floorplan dimensions are comparable between the two cases, and the timescale was chosen accordingly).

3. Distances in Fig. 5

For aesthetic reasons the distances in Fig. 5 of the manuscript were specified to a single decimal place. Assuming the left lower corner of the room as origin, the exact microphone and loudspeaker positions are as follows

The upper wall is at a distance 200/15 from the origin. Higher precision entries for the distance matrices are as follows,

$$\mathbf{D}_{aug,1} = \begin{bmatrix} 000.000 & 011.000 & 002.000 & 010.000 & 002.000 & 161.111 \\ 011.000 & 000.0000 & 005.000 & 005.000 & 009.000 & 178.778 \\ 002.000 & 005.000 & 000.000 & 008.000 & 004.000 & 149.778 \\ 010.000 & 005.000 & 008.000 & 000.000 & 004.000 & 224.444 \\ 002.000 & 009.000 & 004.000 & 004.000 & 000.000 & 196.444 \\ 161.111 & 178.778 & 149.778 & 224.444 & 196.444 & 000.000 \end{bmatrix}$$

and

$$\mathbf{D}_{aug,2} = \begin{bmatrix} 000.000 & 011.000 & 002.000 & 010.000 & 002.000 & 349.556 \\ 011.000 & 000.0000 & 005.000 & 005.000 & 009.000 & 178.778 \\ 002.000 & 005.000 & 000.000 & 008.000 & 004.000 & 149.778 \\ 010.000 & 005.000 & 008.000 & 000.000 & 004.000 & 224.444 \\ 002.000 & 009.000 & 004.000 & 004.000 & 000.000 & 196.444 \\ 349.556 & 178.778 & 149.778 & 224.444 & 196.444 & 000.000 \end{bmatrix}.$$

4. Multidimensional Scaling

As pointed out, in the presence of noise it is not favorable to use the rank test on $\mathbf{D}_{\text{aug}}$. A very good way (as verified through simulations) to deal with this nuisance is to measure how close $\mathbf{D}_{\text{aug}}$ is to a true $\mathbb{EDM}$. To measure the distance, we use multidimensional scaling (MDS) to construct a point set in a given dimension (either 2D or 3D), which produces the $\mathbb{EDM}$ "closest" to $\mathbf{D}_{\text{aug}}$.

MDS was originally proposed in psychometrics as a method for data visualization (4). Many variations have been proposed to adapt the method for sensor localization. We use the s-stress criterion as proposed by Takane et al. (5). Given an observed noisy matrix $\tilde{\mathbf{D}}$, the s-stress criterion is

$$s(\tilde{\mathbf{D}}) = \text{minimize} \sum_{i,j} \left(d_{i,j}^2 - \tilde{d}_{i,j}^2\right)^2$$

$$\text{subject to } \mathbf{D} \in \mathbb{EDM}^2.$$

We call $s(\tilde{\mathbf{D}})$ the score of matrix $\tilde{\mathbf{D}}$. By $\mathbb{EDM}^2$ we denote the set of $\mathbb{EDM}$s with embedding dimension 2 (produced by point sets in 2D). In the 3D case, $\mathbb{EDM}^2$ is replaced by $\mathbb{EDM}^3$.

From now on, we assume that the target space is $\mathbb{R}^2$. The 3D adaptation is immediate. If we associate to each point in $\mathbb{R}^2$ a coordinate vector $\mathbf{x}_i = (x_i, y_i)^\top$, we have that $d_{ij}^2 = \|\mathbf{x}_i - \mathbf{x}_j\|_2^2 = (x_i - x_j)^2 + (y_i - y_j)^2$. Thus, the s-stress criterion can be rephrased as

$$s(\tilde{\mathbf{D}}) = \underset{x_i, y_i \in \mathbb{R}}{\text{minimize}} \sum_{i,j} \left[(x_i - x_j)^2 + (y_i - y_j)^2 - \tilde{d}_{i,j}^2\right]^2. \qquad \textbf{[S14]}$$

The objective function in Eq. **S14** is not convex. However, it has been shown to have less local minima compared with other MDS criteria (5). Furthermore, it yields a meaningful definition of the distance of a matrix from an optimal $\mathbb{EDM}$.

To further skip the local minima of Eq. **S14**, we use coordinate alternation for finding the optimal $\mathbb{EDM}$: we compute Eq. **S14**, by first minimizing over x_i and then over y_i. Although this approach is suboptimal compared with simultaneous minimization with respect to $\mathbf{x}_i$, it leads to simpler computations.

Assuming that x_i has to be updated by Δx_i to give the minimum of $s(\tilde{\mathbf{D}})$, we will have

$$s(\tilde{\mathbf{D}})_i^{(k+1)} = \sum_{j=1}^{n} \left[\left(x_i^{(k)} + \Delta x_i^{(k+1)} - x_j^{(k)}\right)^2 + \left(y_i^{(k)} - y_j^{(k)}\right)^2 - \tilde{d}_{i,j}^2\right]^2, \qquad \textbf{[S15]}$$

where $(\cdot)^{(k)}$ returns the value at iteration k. Taking the derivative of $s(\tilde{\mathbf{D}})_i^{(k+1)}$ with respect to $\Delta x_i^{(k+1)}$, we will have

$$\begin{aligned} \frac{\partial\, s(\tilde{\mathbf{D}})_i^{(k+1)}}{\partial \Delta x_i^{(k+1)}} &= 4n\left(\Delta x_i^{(k+1)}\right)^3 + 3\sum_{j=1}^{n}\left(x_i^{(k)} - x_j^{(k)}\right)\left(\Delta x_i^{(k+1)}\right)^2 \\ &+ \sum_{j=1}^{n}\left[3\left(x_i^{(k)} - x_j^{(k)}\right)^2 + \left(y_i^{(k)} - y_j^{(k)}\right)^2 - \tilde{d}_{i,j}^2\right]\Delta x_i^{(k+1)} \\ &+ \sum_{j=1}^{n}\left[\left(x_i^{(k)} - x_j^{(k)}\right)^3 + \left(x_i^{(k)} - x_j^{(k)}\right)\left(y_i^{(k)} - y_j^{(k)}\right)^2 - \left(x_i^{(k)} - x_j^{(k)}\right)\tilde{d}_{i,j}^2\right]. \end{aligned} \qquad \textbf{[S16]}$$

Setting Eq. **S16** to zero yields at most three real solutions, and comparing the value of $s(\tilde{\mathbf{D}})^{(k+1)}$ for the results gives the optimal value for $\Delta x_i^{(k+1)}$.

The complete optimization procedure is summarized in Algorithm S1.

Algorithm S1. Coordinate alternation for s-stress optimization

Input: Symmetric and zero-diagonal matrix $\tilde{\mathbf{D}}$
Output: Estimate positions: $\mathbf{x}$ and $s(\tilde{\mathbf{D}})$

```
Assume an initial configuration for the points x^0
repeat
  for i = 1 to n do
    Assume the configuration of the points different from i fixed,
    Update x_i using the i^th row of D̃,
    Update y_i using the i^th row of D̃,
  end for
until convergence or maximum number of iterations is reached.
```

Algorithm S2. Room reconstruction procedure

Input: Candidate images $\tilde{\mathbf{s}}_1, \ldots, \tilde{\mathbf{s}}_P$, loudspeaker location $\mathbf{s}_0$, distance threshold ϵ
Output: Room vertices

```
[s̃_1,⋯,s̃_P] ← SortByDistanceFromLoudspeaker([s̃_1,⋯,s̃_P])
deleted[1 : P] ← false
for i = 1 to P do
  if ∃j,k < i, j ≠ k s.t. ‖Combine(s̃_j, s̃_k) − s̃_i‖ < ϵ then
    deleted[i] ← true
  else if Plane(s̃_i) intersects the current room then
    Add Plane(s̃_i) to the current set of planes
  else
    deleted[i] ← true
  end if
end for
```

5. Room Reconstruction Procedure

The echo-sorting algorithm outputs a list of image sources. Some of these image sources are first-order images that we use to reconstruct the room. Some of the output image sources are higher-order sources, and we need to detect them and remove them from the list. As explained in the text, higher-order image sources are obtained as certain "combinations" of lower-order ones—a fact that we use to discriminate between them, as explained below.

We process the candidate image sources in the order of increasing distance from the loudspeaker. If the current image

source cannot be obtained as a combination of closer sources, we add the corresponding plane (halfspace) to the list of halfspaces whose intersection determines the final room.

Beyond the "combining criterion," if the halfspace (which is really an inequality) that we are adding does not change the room, we discard the corresponding image source. We also do it if the new inequality perturbs the room only slightly.

This procedure is summarized in Algorithm S2. The following definition is used in the algorithm (**s** is the loudspeaker):

$$\mathrm{Combine}\left(\tilde{\mathbf{s}}_1, \tilde{\mathbf{s}}_2\right) \stackrel{\mathrm{def}}{=} \tilde{\mathbf{s}}_1 + 2\left\langle \mathbf{p}_2 - \tilde{\mathbf{s}}_1, \mathbf{n}_2 \right\rangle \mathbf{n}_2, \qquad \textbf{[S17]}$$

where $\mathbf{p}_2 = (\mathbf{s} + \tilde{\mathbf{s}}_2)/2$ is a point on the (hypothetical) wall defined by $\mathbf{s}_2$, that is, a point on the median plane between the loudspeaker and $\mathbf{s}_2$. The outward pointing unit normal is defined as $\mathbf{n}_2 = (\tilde{\mathbf{s}}_2 - \mathbf{s})/\|\tilde{\mathbf{s}}_2 - \mathbf{s}\|$.

Room is defined as the intersection of halfspaces generated by the first-order image sources. With the above notation, halfspace corresponding to the image source $\mathbf{s}_i$ is defined by

$$\{\mathbf{x} : \langle \mathbf{n}_i^\top, \mathbf{x} \rangle \leq \langle \mathbf{n}_i, \mathbf{p}_i \rangle\}. \qquad \textbf{[S18]}$$

The plane corresponding to the image source $\mathbf{s}_i$ is denoted simply by plane($\mathbf{s}_i$).

6. Negative Answer to Kac's Question

The reader might be interested by the construction of the counterexample to Kac's question. Here, we explain a counterexample presented by Gordon and Webb (6). The beauty of their example is that elementary means suffice to understand why the two geometrically distinct drums (the same as those shown in Fig. 1) have the same resonant frequencies. However, how to systematically arrive at this construction (or other isospectral drums) is far more involved and requires the knowledge of advanced group representation theory (see the references in the paper).

The homogeneous Helmholtz (time-harmonic wave) equation on a domain D with clamped boundary is given as

$$\Delta\varphi + \lambda\varphi = 0 \qquad \textbf{[S19]}$$

$$\varphi(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial D. \qquad \textbf{[S20]}$$

The solution needs to satisfy both Eq. **S19** and Eq. **S20**. On compact domains, this equation admits the solution only for countably many eigenvalues λ, and the set of all admissible λ's is denoted as the spectrum. Two domains are called "isospectral" if their spectra coincide (counting multiplicities). We note that the actual frequency is proportional to $\sqrt{\lambda}$, not to λ.

To understand the counterexample we need to use two properties of the solutions to the above equation,

i) Linearity: Linear combination of solutions is again a solution; and

ii) Reflection principle: If we have a solution on a domain bounded by a straight line segment with the clamped (Dirichlet) boundary condition, we can extend the domain and the solution by mirroring it over the line segment and changing the sign. This procedure ensures that the solution continues smoothly into the mirrored domain.

Now consider the two drumheads in Fig. S6 (these are the same as in Fig. 1). The drums are segmented and annotated as in Gordon and Webb (6). Let the vibrations of D1 be described by a function φ supported on the drum. The function φ satisfies both Eq. **S19** and Eq. **S20**, for a given λ. Also, as indicated in Fig. S6, let A, B, . . ., G denote the restrictions of φ to corresponding triangular segments.

There happens to be a way to "transplant" the waveform from D1 to D2, so that the resulting waveform on D2 still satisfies both the Helmholtz Eq. **S19** and the boundary condition Eq. **S20**. This transplantation is effected by placing linear combinations of A, B, . . ., G on D2, as indicated in Fig. S6, while observing the edge colors to ensure proper orientations.

We can check that the transplanted waveform indeed satisfies Eq. **S19** and Eq. **S20**. Consider for example triangles A+C+E and −A+D+F on D2. We require that the corresponding waveforms combine smoothly over the blue edge. Triangles C and D share the blue edge on D1. The same holds for triangles E and F. This means that they combine smoothly on D1 so C+E and D+F will combine smoothly on D2 as well. Now observe that the blue edge of A on D1 is the boundary edge, so A vanishes along the blue edge. By reflection principle we can continue A smoothly over the blue edge by mirroring it and multiplying by −1. Finally, this implies that A+C+E and −A+D+F will stitch smoothly. To check that the boundary conditions are satisfied, consider for example the triangle −A+B+G and its red boundary edge. Triangles A and B share the red edge in D1, so they necessarily have the same value on the red edge. Thus, −A+B is zero over the edge. In triangle G, red edge is the boundary edge, so G is zero on that edge, and −A+B+G must be zero on the boundary edge. It is easy to check that all of the triangles in D2 satisfy Eq. **S19** and Eq. **S20**.

We showed that the Eq. **S19** holds with the same λ for both drums. Therefore, every resonant mode of D1 is also a resonant mode of D2. As we can also do a reverse transplantation procedure, every resonant mode of D2 is a resonant mode of D1, thus the two sets coincide, and the drums are isospectral.

1. Farina A (2000) Simultaneous measurement of impulse response and distortion with a swept-sine technique. *108th Convention of the Audio Engineering Society*.
2. Cox TJ, D'Antonio P (2009) *Acoustic Absorbers and Diffusers: Theory, Design and Application* (Taylor & Francis, New York).
3. Müller S, Massarani P (2001) Transfer-function measurement with sweeps. *J Audio Eng Soc* 49:443–471.
4. Torgerson WS (1952) Multidimensional scaling: I. Theory and method. *Psychometrika* 17:401–419.
5. Takane Y, Young FW, Leeuw J (1977) Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. *Psychometrika* 42:7–67.
6. Gordon C, Webb D (1996) You can't hear the shape of a drum. *Am Sci* 84:46–55.

Fig. S1. Equipment used in the experiments: (*A*) Lange D12A omnidirectional loudspeaker. (*B*) Genelec 8030 directional loudspeaker. (*C*) Behringer ECM 8000 omnidirectional microphone. (*D*) Motu 896HD unit.

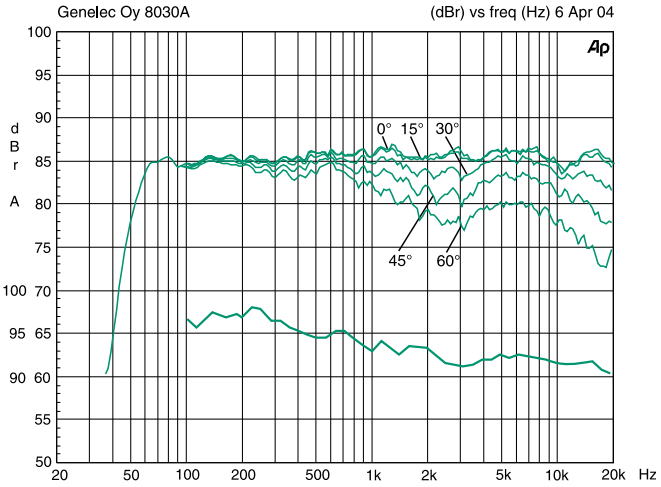

Fig. S2. The upper curve group shows the horizontal directivity characteristics of the 8030A measured at 1 m. The lower curve shows the system's power response.

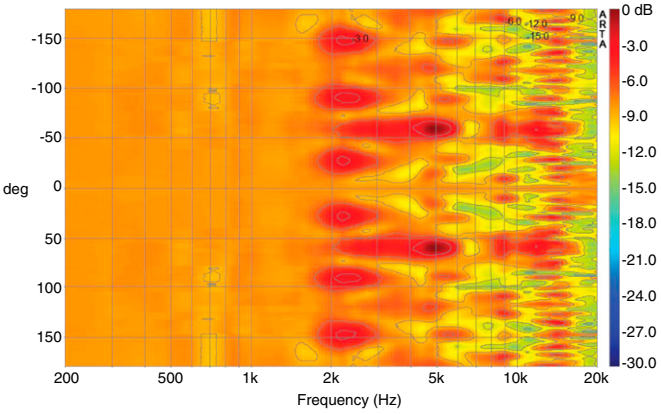

Fig. S3. Directivity sonogram of D12A polar response with 1° resolution. Measurements made at 1 m distance.

Fig. S4. Details of the cathedral side portal (*Left*), microphone arrangement in the cathedral measurements (*Right*).

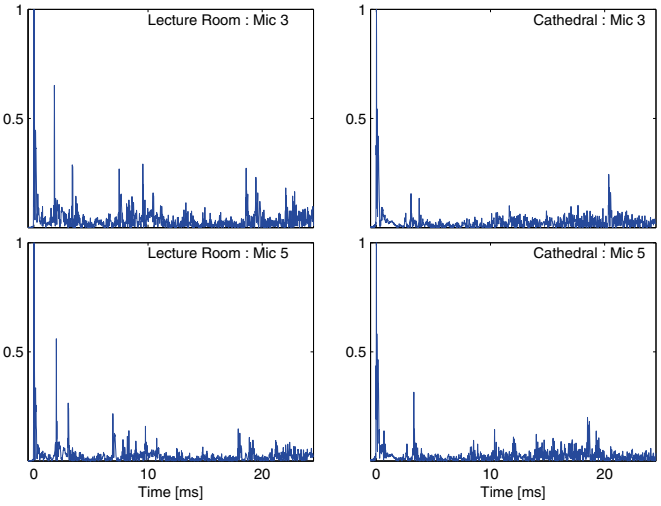

Fig. S5. Comparison of room impulses responses recorded in the lecture room and in portal of the cathedral.

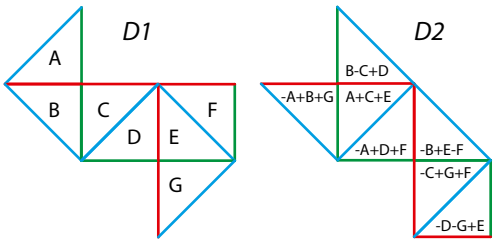

Fig. S6. Isospectral drums with annotations as in (6).

Table S1. Microphone distances in centimeters used in experiments 1, 2, and 3

	M1	M2	M3	M4	M5	M1	M2	M3	M4	M5	M1	M2	M3	M4	M5
M1	00.0	57.0	43.0	51.6	21.0	00.0	19.0	32.5	29.0	24.5	00.0	31.0	42.0	52.5	50.0
M2	57.0	00.0	25.0	43.5	44.0	19.0	00.0	42.5	45.5	40.0	31.0	00.0	54.0	54.5	31.0
M3	43.0	25.0	00.0	29.3	25.0	32.5	42.5	00.0	49.0	25.0	42.0	54.0	00.0	51.5	49.0
M4	51.6	43.5	29.3	00.0	32.3	29.0	45.5	49.0	00.0	27.5	52.5	54.5	51.5	00.0	42.5
M5	21.0	44.0	25.0	32.3	00.0	24.5	40.0	25.0	27.5	00.0	50.0	31.0	49.0	42.5	00.0

Table S2. Exact position of microphones and the loudspeaker in Fig. 5

Object	*x*	*y*	*z*
Mic1	175/15	85/15	150/15
Mic2	220/15	100/15	165/15
Mic3	190/15	100/15	150/15
Mic4	220/15	70/15	150/15
Mic5	190/15	70/15	150/15
LSPK	100/15	140/15	150/15