

On the Gravitational Inverse Problem

Ulvi Yurtsever

MathSense Analytics, 1273 Sunny Oaks Circle
Altadena, CA 91001 USA

and

Hearne Institute for Theoretical Physics
Louisiana State University
Baton Rouge, LA 70803, USA
ulvi@phys.lsu.edu

Caren Marzban

Applied Physics Laboratory
University of Washington
Seattle, WA 98195, USA

and

Department of Statistics
University of Washington
Seattle, WA 98195, USA

Marina Meilă

Department of Statistics
University of Washington
Seattle, WA 98195, USA

Abstract

We discuss some mathematical aspects of the problem of inverting gravitational field data to extract the underlying mass distribution. While the forward problem of computing the gravity field from a given mass distribution is mathematically straightforward, the inverse of this forward map has some interesting features that make inversion a difficult problem. In particular, the forward map has an infinite-dimensional kernel which makes the inversion fundamentally non-unique. We characterize completely the kernels of two gravitational forward maps, one map-

ping mass density to the Newtonian scalar potential, and the other mapping mass density to the gravity gradient tensor, which is the quantity most commonly measured in field observations. In addition, we present some results on unique inversion under constrained conditions, and comment on the roles the kernel of the forward map and non-uniqueness play in discretized approaches to the continuum inverse problem.

Keywords: Gravitational Inverse Problem, Ill-Posed Problems, Discretization

Weighing the shape of a gravitating body

More than thirty years ago, Mark Kac asked “Can you hear the shape of a drum?” meaning: do two distinct planar domains always have distinct spectra of eigenvalues for their respective Laplace operators (acting on functions) with the usual Dirichlet (or Neumann) boundary conditions? If the answer is yes, the shape of a “drum” can be inferred by hearing its spectrum (characteristic sound), if the answer is no, then two distinctly shaped drums may have identical spectra (in which case they are called “isospectral domains”) [7, 3].

Kac’s article [7] stimulated a long line of research which eventually settled his question in the negative: There do exist isospectral domains (and, more generally, isospectral Riemann surfaces and isospectral Riemannian manifolds in higher dimensions) which are *not* isometric. In other words, the spectral inverse problem is ill-defined, subject to a fundamental ambiguity which can be precisely characterized [2].

A similar ambiguity plagues the gravitational inverse problem, that is, the problem of inferring the precise shape of a mass distribution by observing its distant gravitational field.

The gravitational inverse problem is the problem of inverting the gravitational forward map, which we take to be a map sending a compact supported mass distribution to a gravity observable: in practice, the observable could be either the Newtonian gravitational potential or gravity gradients.

More precisely, and focusing on the gravity potential Φ for the moment, what we will mean by the gravitational inverse problem is the following: Given a spherical region $B_R = \{\vec{r} : |\vec{r}| < R\}$ of radius R in \mathbb{R}^3 , and a solution $\Phi(\vec{r})$ (the gravitational potential in free space) of the Laplace equation $\nabla^2\Phi = 0$ outside the region B_R (i.e. for $|\vec{r}| > R$) which vanishes at infinity, find a mass

density distribution $\rho(\vec{r}')$ supported inside B_R which gives rise to $\Phi(\vec{r})$ in the exterior region outside B_R . In plainer language, find a $\rho(\vec{r}')$ with support inside B_R such that

$$\Phi(\vec{r}) = -G \int_{B_R} \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r' \quad \text{for } r > R. \quad (1)$$

Kernel of the forward map onto the gravitational potential

Equation (1) of course represents the unique solution to the “forward problem” where one searches for a solution Φ to $\nabla^2\Phi = 4\pi G\rho$ with vanishing boundary conditions at infinity. Of key interest is the “kernel” of this (linear) forward map, i.e. the set of mass distributions ρ supported inside B_R that are mapped to a potential Φ via Eq. (1) which identically vanishes outside B_R .

Theorem 1: The kernel of the forward map Eq. (1) mapping mass distributions ρ supported in B_R to solutions of Laplace equation outside the region B_R (i.e. for $|\vec{r}| > R$) is precisely functions ρ satisfying

$$\rho = \nabla^2\chi, \quad (2)$$

where $\chi(\vec{r})$ is any (sufficiently smooth) function on \mathbb{R}^3 with support inside B_R (i.e. $\chi(\vec{r}) = 0$ for $r > R$). In other words, if ρ is a solution of the inverse problem for a given exterior potential Φ , then $\rho + \nabla^2\chi$ is also a solution for any $\chi \in C_0^\alpha(B_R)$, where α is a sufficiently large integer. Normally, $\alpha \geq 2$ should be sufficient, but smoothness is not a key issue; in particular, χ can even be a distribution if point-mass (delta-function) singularities need to be allowed in the problem.

Proof in one direction is easy: Every function in the kernel is given by the forward image of a function of the kind Eq. (2). To prove this, let Φ be a function belonging to the kernel, i.e. let Φ vanish outside B_R . Put

$$\chi \equiv \frac{1}{4\pi G}\Phi.$$

Then $\chi \in C_0^\alpha(B_R)$ and $\rho \equiv \nabla^2\chi$ satisfies the Laplace equation $\nabla^2\Phi = 4\pi G\rho$ everywhere (with vanishing boundary conditions at infinity). Therefore, Φ satisfies Eq. (1) with this ρ , which is what we needed to prove.

Conversely, let ρ be a density distribution supported inside B_R such that $\rho = \nabla^2\chi$ for some $\chi \in C_0^\alpha(B_R)$. Then, according to Eq. (1), the gravitational potential Φ which the forward map sends ρ onto satisfies

$$\Phi(\vec{r}) = -G \int_{B_R} \frac{\nabla^2\chi(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r' \quad \text{for } r > R. \quad (3)$$

To show that the right hand side of Eq. (3) is in the kernel of the forward map, i.e., that it vanishes for $r > R$, use Green's identity:

$$\int_B (U \nabla^2 V - V \nabla^2 U) d^3 r = \int_{\partial B} \left(U \frac{\partial V}{\partial n} - V \frac{\partial U}{\partial n} \right) d\sigma, \quad (4)$$

where B is any region bounded by the surface ∂B , and U, V are arbitrary functions on \mathbb{R}^3 . Applying Eq. (4) with B taken as the region B_R , $U(\vec{r}') \equiv 1/|\vec{r} - \vec{r}'|$, and $V(\vec{r}') \equiv \chi(\vec{r}')$, and noting that $\nabla^2(1/|\vec{r} - \vec{r}'|) = 0$ when $r > R$ and $r' < R$, it immediately follows that the right hand side of Eq. (3) vanishes outside B_R (i.e. for $r > R$). This completes the proof of Theorem 1.

A geometric interpretation of the kernel:

One way to conceptualize the kernel of the gravitational (potential) forward map is to note that the geometric freedom of choice in the inverse “datum” $\Phi(\vec{r})$ is that of choosing an arbitrary function on the two-sphere S^2 (compare also [5], Chapter 1):

Theorem 2: Let $\Phi(\vec{r})$ be any solution of the free-space Laplace equation $\nabla^2 \Phi = 0$ outside the region B_R which vanishes at infinity, as in the formulation of the gravitational inverse problem. Then Φ is completely determined outside B_R by its values on any two-sphere S_{R_1} of radius $R_1 > R$ (or, more generally, by its values on any closed surface which encloses B_R).

Proof: This is really a restatement of a standard result in potential theory (uniqueness of solutions to the Dirichlet problem): There exists a unique Green's function $G(\vec{r}, \vec{r}_0)$, defined for \vec{r}, \vec{r}_0 outside B_{R_1} , such that G satisfies $\nabla^2 G(\vec{r}, \vec{r}_0) = \delta(\vec{r} - \vec{r}_0)$ and vanishes for $\vec{r} \in S_{R_1}$ and for $\vec{r} \in S_\infty$ [as discussed, e.g., in [6], for the two-sphere S_{R_1} $G(\vec{r}, \vec{r}_0)$ can be constructed explicitly using the classic “method of images”]. Plugging such a $G(\vec{r}, \vec{r}_0)$ into Eq. (4) as V and taking $\Phi(\vec{r})$ as U and the region B as the region *outside* B_{R_1} we obtain, by virtue of the vanishing boundary conditions at infinity,

$$\Phi(\vec{r}_0) = \int_{S_{R_1}} \Phi(\vec{r}) \frac{\partial G(\vec{r}, \vec{r}_0)}{\partial n} d\sigma. \quad (5)$$

Therefore Φ everywhere outside B_{R_1} is determined uniquely by its values on the two-sphere S_{R_1} .

We can now understand the kernel Eq. (2) in the following way: Since the data for Φ consist of the values of a function defined on a two-surface S_{R_1} , we can infer from these data uniquely at the most another function of two

variables, and not the full three-dimensional density field $\rho(\vec{r})$. In fact, the forward kernel (or the ambiguity in the corresponding inversion) as described by Eq. (2) corresponds precisely to this geometric statement.

Weighing the shape of a body of known radial density

One might hope that practical (physical) prior constraints on the three-dimensional density distribution $\rho(\vec{r})$ might make it uniquely recoverable from its far-zone gravity field despite the fundamental non-uniqueness of the inverse problem. For example, we want the density to be positive everywhere, which is a requirement that constrains the ambiguity Eq. (2) to some extent. However, simple spherically-symmetric counterexamples show that positivity is not a sufficiently strong constraint to help provide us with a unique inversion. As the next step in a series of physically-reasonable constraints on ρ , we might assume a known positive radial density distribution with profile $\rho(\vec{r}) \equiv \rho_0(r) > 0$ distributed on some arbitrary compact three-dimensional region D in \mathbb{R}^3 . Put another way, such a density profile represents a body of arbitrary shape carved out of a spherically symmetric (hence spherical) mass distribution. Again, counterexamples based on hollow spherical shells show that this is not quite enough for unique inversion. Nevertheless, it turns out that if we further constrain the region D such that it is connected and has no “holes” (i.e., if D is topologically a ball), and, furthermore, if D is “radially convex” in a sense made precise below, then unique inversion is possible:

Theorem 3: Let D be compact region in \mathbb{R}^3 such that its boundary ∂D is a connected and simply-connected surface (in other words, ∂D is a topological two-sphere) which is *radially convex* in the sense that any straight line in \mathbb{R}^3 passing through the center-of-mass of the volume D intersects ∂D at precisely two points. Assume that D is filled with material of a *known* positive mass density $\rho(\vec{r})$ which is distributed spherically-symmetrically with respect to the coordinate origin given by the center of mass. That is, if \vec{r} lies inside D , then $\rho(\vec{r}) = \rho_0(r) > 0$, and if \vec{r} is outside D , then $\rho(\vec{r}) = 0$. Under these conditions, D itself (or, equivalently, its boundary ∂D) is uniquely recoverable from the far-zone gravity field of this radial density distribution.[†]

This result is not too surprising in view of Theorem 2, since the specification of ∂D entails just a single real function on the two-sphere S^2 (measuring just

[†]The assumption that ∂D is a topological two-sphere is redundant since it follows from the assumption of radial convexity as formulated in the theorem. However, it is perhaps useful to emphasize this assumption in a redundant statement since the theorem is certainly false without it.

how much we need to deform S^2 in order to stretch it onto ∂D). The forward map Eq. (1) can then be interpreted as a nonlinear map from real functions on S^2 (representing the deformations of S^2 needed to obtain ∂D) to real functions on S^2 (representing the values of the potential Φ on S_R), and we will now show that this map is locally one-to-one.

Proof of Theorem 3: The main idea of the proof is simple: explicitly write down, in terms of spherical-harmonic coefficients, the forward transform mapping the “shape function” of ∂D to the exterior potential Φ , and show that the derivative of this nonlinear forward map is nonsingular. The result then follows from the inverse function theorem as generalized to infinite-dimensional spaces [1, 8] given appropriate topologies on the domain and target spaces of the map to make it continuous. In this paper we will give a detailed proof that the forward map has nonsingular (and bounded) derivative at the point (shape) which corresponds to a perfect sphere, so the result holds for shapes which are nearby distortions of a perfect sphere (in other words, we will explicitly prove that the forward map is invertible in some open neighborhood of the perfect sphere in the space of all shapes D which satisfy the conditions of the theorem). This case covers most planetary bodies at the levels of resolution we are interested in. Nevertheless, the statement that the forward map is nonsingular everywhere remains valid, although we are not going to give an explicit proof of it here. The proof of this more general case is substantially similar apart from requiring more careful estimates.

To proceed with the proof, introduce a spherical coordinate system (r, θ, ϕ) centered at the center of mass of the volume D . In this coordinate system, let $\mathbf{n}(\theta, \phi)$ denote the half-line which starts at the origin and expands outward in the direction (θ, ϕ) . Let $\psi(\theta, \phi)$ be the (positive) function which gives the length of the radial vector which starts at the origin and ends at the intersection point of the line $\mathbf{n}(\theta, \phi)$ with the boundary ∂D as (θ, ϕ) ranges over the unit two-sphere S^2 of all possible directions [by the radial convexity assumption, there exists a unique such intersection point for each direction (θ, ϕ)]. The function $\psi(\theta, \phi)$ can then be taken to be the “shape function” which specifies D , and, explicitly, we can write

$$D = \{(r, \Omega) \mid r \leq \psi(\Omega)\} , \quad (6)$$

and

$$\partial D = \{(r, \Omega) \mid r = \psi(\Omega)\} , \quad (7)$$

where we introduced the short-hand notation $\Omega \equiv (\theta, \phi)$ for the angular coordinates. The exterior gravitational potential Φ can be expanded in spherical

harmonics[6]:

$$\Phi(\vec{r}) = \sum_{l,m} d_{lm} \frac{Y_{lm}(\Omega)}{r^{l+1}} \quad \text{for } r > R > \max_{\Omega} \psi(\Omega) , \quad (8)$$

where we can regard $\{d_{lm}\} \equiv \mathbf{D}$ as an infinite sequence (vector) of “observables” which completely describes the data for the inverse problem in view of Theorem 2. On the other hand, according to Eq. (1), for $r > R > \max_{\Omega} \psi(\Omega)$ we have

$$\begin{aligned} \Phi(\vec{r}) &= G \int_{r' < \psi(\Omega')} \rho_0(r') \frac{d^3 r'}{|\vec{r} - \vec{r}'|} \\ &= G \int_0^{\psi(\Omega')} \rho_0(r') r'^2 dr' \int_{S^2} \sum_{l,m} \frac{r'^l}{r^{l+1}} Y_{lm}(\Omega) Y_{lm}^*(\Omega') d\Omega' \\ &= G \sum_{l,m} \frac{Y_{lm}(\Omega)}{r^{l+1}} \int_{S^2} Y_{lm}^*(\Omega') \mu_{l+2}(\psi(\Omega')) d\Omega' \\ &= G \sum_{l,m} f_{lm}[\psi] \frac{Y_{lm}(\Omega)}{r^{l+1}} , \end{aligned} \quad (9)$$

where

$$\mu_n(w) \equiv \int_0^w \rho_0(r) r^n dr , \quad (10a)$$

and

$$f_{lm}[\psi] \equiv \int_{S^2} Y_{lm}^*(\Omega) \mu_{l+2}(\psi(\Omega)) d\Omega \quad (10)$$

is a vector functional of the shape function ψ which represents the forward map in the same way as $\mathbf{D} = \{d_{lm}\}$ represents the data. In fact, introducing the notation $\mathbf{F}[\psi] \equiv \{f_{lm}[\psi]\}$ and combining Eqs.(8) and (9), the forward equation for the shape function ψ takes the simple form

$$\mathbf{F}[\psi] = \frac{1}{G} \mathbf{D} . \quad (11)$$

(Due to our choice of origin as the center of mass, both f_{lm} and d_{lm} vanish for $l = 1$, but this fact will not be of any consequence in what follows.) It is also convenient to introduce a coordinatization of the space of shape functions via a spherical harmonic expansion

$$\psi(\Omega) \equiv \sum_{l,m} s_{lm} Y_{lm}(\Omega) , \quad (12)$$

and consider the coordinate vector $\mathbf{S} \equiv \{s_{lm}\}$ as the representation of the function $\psi(\Omega)$. In this coordinate system the forward map Eq. (11) takes the form

$$\mathbf{F}[\mathbf{S}] = \frac{1}{G} \mathbf{D} , \quad (13)$$

where

$$f_{lm}[\mathbf{S}] \equiv \int_{S^2} Y_{lm}^*(\Omega) \mu_{l+2} \left(\sum_{p,q} s_{pq} Y_{pq}(\Omega) \right) d\Omega . \quad (14)$$

Assume now, contrary to the conclusion of Theorem 3, that two distinct domains D_1 and D_2 constrained as in the statement of the theorem give rise to identical external gravitational potentials when filled with the given radial density distribution $\rho_0(r)$. First of all, since the monopole and dipole moments of the two mass distributions must agree, they must have the same center of mass, therefore we can set up a common spherical coordinate system for both volumes with their shared center of mass chosen as the origin of coordinates. It then follows that there exist two distinct shape functions ψ_1 and ψ_2 , corresponding to the two distinct volumes D_1 and D_2 , which satisfy Eq. (11) with the same data \mathbf{D} ; in other words

$$\mathbf{F}[\psi_2] = \mathbf{F}[\psi_1] . \quad (15)$$

We will now show that Eq. (15) is impossible as long as ψ_1 and ψ_2 belong to some fixed open neighborhood of a perfect sphere $\{\psi(\Omega) \equiv a_0 = \text{const}\}$ in the infinite-dimensional nonlinear function space of all ψ 's. Using the inverse function theorem as generalized to such infinite-dimensional manifolds [1, 8], it is sufficient to show that the derivative of the map \mathbf{F} at the point $\psi(\Omega) \equiv a_0$ is a nonsingular, bounded (continuous) linear map. In general, at an arbitrary point $\psi = \psi_0(\Omega)$, this derivative is given by

$$(\mathbf{F}'[\psi_0(\Omega)] \cdot \delta\psi)_{lm} = \int_{S^2} Y_{lm}^*(\Omega) \rho_0(\psi_0(\Omega)) \psi_0(\Omega)^{l+2} \delta\psi(\Omega) d\Omega , \quad (16)$$

where $\mathbf{F}'[\psi_0(\Omega)]$ denotes the derivative evaluated at the point $\psi = \psi_0$, acting (as a linear map) on the tangent vector (linear perturbation) $\delta\psi$, and we have used Eq. (10) to derive this explicit form. Specializing to the perfect sphere $\psi_0(\Omega) = a_0 (= \text{const})$ and using the coordinate representation [cf. Eq. (12)]

$$\delta\psi(\Omega) \equiv \sum_{p,q} \delta s_{pq} Y_{pq}(\Omega) , \quad \delta\mathbf{S} \equiv \{\delta s_{pq}\} , \quad (17)$$

Eq. (16) takes the form

$$\begin{aligned}
 (\mathbf{F}'[a_0] \cdot \delta \mathbf{S})_{lm} &= \rho_0(a_0) a_0^{l+2} \int_{S^2} Y_{lm}^*(\Omega) \left[\sum_{p,q} \delta s_{pq} Y_{pq}(\Omega) \right] d\Omega \\
 &= \rho_0(a_0) a_0^{l+2} \delta s_{lm} ,
 \end{aligned}
 \tag{18}$$

where we made use of the fact that the Y_{lm} 's form an orthonormal basis for $L^2(S^2)$. According to Eq. (18), the derivative $\mathbf{F}'[a_0]$ is a *diagonal* linear map with only nonzero entries (eigenvalues) on the diagonal; therefore, $\mathbf{F}'[a_0]$ is clearly nonsingular.

To complete the proof of Theorem 3, all we have left to do is to choose topologies on the tangent space to the space of shapes spanned by $\delta \mathbf{S} \equiv \{\delta s_{pq}\}$ and the tangent space to the space of data $\{d_{lm}\} \equiv \mathbf{D}$ between which the linear map $\mathbf{F}'[a_0]$ acts, such that $\mathbf{F}'[a_0]$ is a continuous, or, equivalently, bounded map. But since $\mathbf{F}'[a_0]$ is a diagonal linear map in the coordinate basis Eq. (12) with only nonzero eigenvalues, it is trivial to choose norms in its domain and range so that $\mathbf{F}'[a_0]$ is bounded with, say, unit norm (just choose inner products under which the image and target bases are orthonormal). Upon completing these two spaces under these norms to make them into Banach (indeed, Hilbert) spaces, the inverse function theorem as formulated, e.g., in Theorem 5.2 of [8] applies, and we are done.[‡]

The kernel of the forward map onto gravity gradient observables

The gravitational gradient tensor is (apart from a minus sign) simply the (symmetric) tensor of second derivatives of the potential Φ in a Cartesian coordinate system:

$$T_{ij} \equiv -\frac{\partial^2 \Phi}{\partial x^i \partial x^j} .
 \tag{19}$$

So, for example, we have

$$T_{xx} = -\frac{\partial^2 \Phi}{\partial x^2} , \quad T_{yz} = -\frac{\partial^2 \Phi}{\partial y \partial z}
 \tag{20}$$

etc. Independently of coordinates, the gradient tensor can be defined as the double covariant derivative $\nabla \nabla \Phi$ (in general relativity, T_{ij} corresponds to the Riemann curvature tensor R_{0i0j} describing tidal gravitational forces). One can

[‡]Notice, however, that the inverse map is only guaranteed to be continuous with respect to the Banach-manifold topologies determined by the (somewhat arbitrarily) chosen norms.

also define the gradient tensor explicitly in terms of the source mass distribution as:

$$T_{ij}(\vec{x}) = G \int \frac{\rho(\vec{y}) [3(x^i - y^i)(x^j - y^j) - \delta_{ij} |\vec{x} - \vec{y}|^2]}{|\vec{x} - \vec{y}|^5} d^3y, \quad (21)$$

where all coordinates are Cartesian. The gradient tensor is a particularly useful observable in precision gravimetry since it is better isolated from local non-gravitational acceleration noise compared to other observables, and a large roster of instruments (gradiometers) are available for measuring it.

In practical applications, one often works with a coordinate system where z is the vertical coordinate pointing up from the Earth's center, and the observable of interest is the x – y projection of the gradient tensor in an infinitesimally small neighborhood (tangent plane to the Earth's spherical surface) around $x = y = 0$:

$$T \equiv \begin{pmatrix} T_{xx} & T_{xy} \\ T_{xy} & T_{yy} \end{pmatrix} \quad (22)$$

Typically, a gradiometer takes two kinds of measurements: the component $M_{\times} \equiv 2T_{xy}$ (“crossline”), and the combination $M_{+} \equiv T_{xx} - T_{yy}$ (inline). The choice of Cartesian x, y coordinates is arbitrary upto a rotation R , and T transforms under rotations as

$$T \longrightarrow RTR^t. \quad (23)$$

Neither T nor its crossline or inline components are invariant under rotations, but $\text{Tr}(T)$ and $\text{Det}(T)$ are invariants. In particular, the Euclidean norm of (M_{+}, M_{\times})

$$\sqrt{M_{+}^2 + M_{\times}^2} = \sqrt{\text{Tr}(T)^2 - 4 \text{Det}(T)} \quad (24)$$

is an invariant. More specifically, it is easy to compute that (M_{+}, M_{\times}) transforms under a rotation

$$R_{\theta} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (25)$$

according to the rule

$$\begin{pmatrix} M_{+} \\ M_{\times} \end{pmatrix} \longrightarrow R_{2\theta} \begin{pmatrix} M_{+} \\ M_{\times} \end{pmatrix}. \quad (26)$$

Now, define a new observable

$$V \equiv M_+^2 + M_\times^2. \quad (27)$$

In other words, V is defined at every point of \mathbb{R}^3 by setting up local coordinates x, y, z such that the z -axis goes through the origin, computing the gradient tensor T as in Eqs. (19) and (22) at that point, and then calculating the norm square of the observable (M_+, M_\times) . The key features of V are that (i) it is invariant under rotations, therefore uniquely defined independently of the choice of coordinates, and (ii) it obeys the following lemma:

Lemma: V vanishes at a point if and only if (M_+, M_\times) vanishes there for any allowed choice of coordinates x, y, z .

Proof: By Eq. (11), V is invariant and equal to the norm-square of (M_+, M_\times) for any choice of coordinates.

Therefore the kernel of the coordinate-dependent observable (M_+, M_\times) is precisely the kernel of the nonlinear but coordinate-independent observable V .

Theorem 4: Let $F[\Phi]$ be any analytic functional (linear or nonlinear) of the gravity potential Φ (such as V), and let S be any analytic 2-surface lying outside the spherical region $B_R = \{\vec{r} : |\vec{r}| < R\}$ in \mathbb{R}^3 (such as a sphere of radius $> R$). Then, if $F[\Phi]$ vanishes in any open (two-dimensional) neighborhood on S , it vanishes identically on all of S .

Proof: This just follows from analyticity: Φ is an analytic function outside the spherical region $B_R = \{\vec{r} : |\vec{r}| < R\}$ in \mathbb{R}^3 since it satisfies the homogeneous Laplace equation there (analyticity follows from standard elliptic regularity theorems [4]). Therefore, $F[\Phi]$ is analytic there and so is its restriction to S since S is analytic. Thus vanishing on any open subset is equivalent to vanishing identically on S .

Corollary: If V vanishes in any two-dimensional patch, no matter how small, on any analytic observation surface S lying outside the spherical region $B_R = \{\vec{r} : |\vec{r}| < R\}$, then it vanishes identically on all of S .

This corollary further illustrates the fact that the kernel of the gravity-gradient observable (M_+, M_\times) is precisely the kernel of the observable V , not only locally but also globally.

Theorem 5: Let S be a sphere of radius $> R$. Then V vanishes on S if and only if Φ is spherically symmetric (a function of the radius r only), and hence V vanishes identically everywhere outside the spherical region $B_R = \{\vec{r} : |\vec{r}| < R\}$ in \mathbb{R}^3 .

Proof: The if part is a simple calculation: It is straightforward to compute that both M_+ and M_\times vanish for a radial (monopole) potential function $\Phi(r)$. For the converse, it is easy to see that if V vanishes on a sphere S then this implies that Φ is constant on S . But this implies, according to Eq. (5) (or just by the uniqueness of solutions to the Dirichlet problem), that Φ is a radial function of monopole type:

$$\Phi(\vec{r}) = \frac{C}{r}, \quad (28)$$

where C is a constant.

We can now completely characterize the kernel of the forward map from the mass density to the gravity gradient observables (M_+, M_\times):

Theorem 6: The kernel of the forward map mapping mass distributions ρ supported in B_R to gravity gradient observables (M_+, M_\times) outside the region B_R (i.e. for $|\vec{r}| > R$) is precisely functions ρ supported in B_R satisfying

$$\rho = \rho_0 + \nabla^2 \chi, \quad (29)$$

where $\chi(\vec{r})$ is any (sufficiently smooth) function on \mathbb{R}^3 , and ρ_0 is any *spherically symmetric* function, both supported inside B_R (i.e., both $\chi(\vec{r})$ and $\rho_0(\vec{r})$ vanish for $r > R$).

Proof: By Theorem 5, the kernel of the map from mass distributions to outside gradients (M_+, M_\times) consists of those mass distributions that give rise to spherically symmetric potentials Φ outside B_R . If Φ is spherically symmetric outside B_R , then consider Φ_S , the spherical average of Φ (i.e. $\Phi_S(\vec{x}) =$ the average of Φ on the sphere centered at 0 and passing through \vec{x}). Then Φ_S is spherically symmetric everywhere and coincides with Φ outside B_R . Therefore $\chi \equiv \Phi - \Phi_S$ vanishes outside B_R , and thus

$$\nabla^2 \Phi = \nabla^2 \chi + \nabla^2 \Phi_S. \quad (30)$$

Since Φ_S is everywhere spherically symmetric, so is $\nabla^2 \Phi_S$, and Theorem 6 follows.

Kernel of the gravitational forward map and discretization

In practice, gravity inversion is a discrete problem because (i) the measurements of Φ (or of the gradients) are finitely many and discretely distributed in space, and (ii) more problematically, the model for the mass distribution ρ is

some discretized approximation to a continuous distribution. In all contexts, the characterization of ρ would be a finite list of parameters which uniquely specify ρ in some generally non-linear fashion. For example, these ρ parameters could be the masses, locations, and shape parameters of a finite number of tectonic plates in a geophysical model of the Earth's crust. Or, as discussed above, they could be masses of finite blocks into which we divide the source distribution in discretizing it. Or, more straightforwardly, they could be the masses of N point-mass centers distributed throughout the source region, approximating with a discrete configuration the true mass distribution in the limit $N \rightarrow \infty$.

In general, the practical, the discretized gravitational inverse problem is the problem of inverting some (generally nonlinear) forward map:

$$F : \{p_j\} \mapsto \{\Phi_i\}, \quad \Phi_i = F_i[p_j], \quad (31)$$

where p_j are finitely many parameters specifying the mass distribution, and Φ_i are the measurements. It would be conceptually salutary to have the fundamental non-uniqueness in the gravitational inverse problem (the kernel of the forward map) described by Theorem 1 to fall out of the formulation Eq. (31) in a natural way. For example, when we simulate a slab of soil using some large number of mass centers regularly placed at fixed lattice points inside the slab, the parameters p_j are simply the point masses m_j assigned to each center at lattice location j . In this case, the forward map F is in fact linear:

$$\Phi_i = \sum_j F_{ij} m_j, \quad (32)$$

where the matrix F_{ij} is the Green's function in Eq. (1) in discretized form:

$$F_{ij} = -G \frac{1}{|\vec{r}_i - \vec{r}_j|}, \quad (33)$$

with \vec{r}_i being the locations where the measurements $\Phi_i \equiv \Phi(\vec{r}_i)$ are collected. Consider first, for simplicity, a scenario in which we are sampling Φ at the same number of points N as the number of mass centers in the discretization. In other words, F is now a square $N \times N$ matrix. In view of Theorem 1 characterizing the kernel of the forward map, one might expect F to be singular, with the null space corresponding to a discretized version of the kernel, i.e., a discrete approximation to functions of the form $\nabla^2 \chi$ with χ supported inside the slab. It turns out, however, that the matrix F given by Eq. (33) is in fact generically nonsingular. Moreover, the $M \times N$ matrix F with M measurement locations and N mass centers is also nonsingular, in the sense that generically it has maximal rank (i.e. trivial null space).

It is in fact easy to see why this is so, because of the following result:

Theorem 7: Given a solution $\Phi(\vec{r})$ of the Laplace equation $\nabla^2\Phi = 0$ vanishing at infinity and defined for $r > R$, there exists *at most* one configuration $\{m_j, \vec{r}_j\}$ of finitely many point masses placed inside $B_R = \{r \leq R\}$ (i.e. with $m_j \in \mathbb{R}$ and $|\vec{r}_j| \leq R$) that can give rise to this Φ for $r > R$.

Proof: Suppose, on the contrary, that there are two configurations of point masses, $\{m_j, \vec{r}_j\}$ and $\{m'_k, \vec{r}'_k\}$, that produce the same Φ for $r > R$. Subtract the second configuration from the first, and correspondingly subtract the Φ fields that they produce. Since the gravity field depends linearly on the mass distribution, what we obtain is a new configuration $\{m_1, \dots, m_N, -m'_1, \dots, -m'_{N'}, \vec{r}_1, \dots, \vec{r}_N, \vec{r}'_1, \dots, \vec{r}'_{N'}\}$ of point masses inside B_R (unless there are some coincident point masses in the two collections, in which case one would simply subtract the corresponding masses and list the location only once), which produces a field Φ that vanishes identically for $r > R$. Could this actually happen? It turns out the answer is no, unless Φ is identically zero everywhere (and therefore the two original point-mass configurations are in fact identical). To see this, observe that Φ produced by a finite set of point masses is a real-analytic function in \mathbb{R}^3 except at the locations of the point masses where it has singularities. Since Φ vanishes for $r > R$ and is analytic, it must vanish everywhere in \mathbb{R}^3 except possibly at the mass centers. But if any of the mass centers had non-zero mass, we could choose points so close to that center that the contribution to Φ from that center would overwhelm the contributions from any other centers (which are discretely spaced since there are finitely many). This clearly contradicts the fact that Φ is identically zero in any small neighborhood of the chosen mass-center. Therefore, none of the mass centers can have nonzero mass; the two original configurations of point masses must be identical, and Theorem 7 is proved.

Here is one way to understand the apparent conflict between Theorem 1 and Theorem 7: Consider the two spaces between which the forward map F acts: the space of density distributions ρ and the space of potentials Φ . Any discretization is an attempt to approximate these spaces via a sequence of finite-dimensional subspaces. For example, when we use N point masses, we have an N -dimensional subspace of the space of all ρ , and as N gets larger and larger this subspace approximates the full space arbitrarily closely, in the sense that for any ρ_0 , we can find a configuration of N point masses (with large enough N) which comes as close as we want to ρ_0 (in some locally averaged sense). The same goes for the corresponding potentials Φ : given any solution Φ_0 , we can find potentials produced by N point masses that get arbitrarily close to Φ_0 as $N \rightarrow \infty$. But the problem is that these approximating subspaces

completely miss the kernel of the true forward map, which is the subspace of mass distributions (and corresponding potentials) given by $\{\nabla^2\chi \mid \chi \in C_0(B_R)\}$. The intersection of the approximating subspaces with this kernel subspace is the zero vector, for any finite N . This is (mathematically) the explanation for the apparent contradiction between Theorem 1 and Theorem 7.

To resolve this apparent conceptual paradox, one might argue that we must choose the approximating finite dimensional subspaces in such a way that they fully intersect the kernel. But in practice, there is no feasible way to discretize the problem that makes sure this property holds. There is, however, a much simpler practical strategy out of this apparent paradox, and this is the strategy we advocate: Realize that true measurements in the real world always have instrumental noise. What this means is that two potentials are indistinguishable in practice if they differ everywhere by less than, say 1σ worth (in some arbitrary units) of instrumental noise. Therefore, e.g. when we look for the intersection between the kernel and our discretized ρ -subspace with N point masses, what we are really looking for are all N -point-mass configurations that produce a potential Φ that differs from zero by less than 1σ throughout the exterior region $r > R$. And in general there are many such configurations. We can see this, for example, in the matrix F_{ij} of Eq. (33): in general this matrix turns out to be highly ill-conditioned (with very small determinant) with lots of eigenvalues close to zero, even though it has no exactly-zero eigenvalues. And the “approximately null” subspace spanned by the small-eigenvalued eigenspaces is precisely the discrete analogue of the kernel of the forward map; it is what corresponds to the subspace $\{\nabla^2\chi \mid \chi \in C_0(B_R)\}$ in this discretization. We expect a similar description for the discrete analogue of the forward map’s kernel in any other practical discretization scenario.

References

- [1] R. Abraham and J. E. Marsden, *Foundations of Mechanics: A Mathematical Exposition of Classical Mechanics With an Introduction to the Qualitative Theory of Dynamical Systems*, Perseus Books, San Francisco 1994.
- [2] M. Berger, Sect. 9.12 in *A Panoramic View of Riemannian Geometry*, Springer-Verlag, 2002.
- [3] P. Buser, *Isospectral Riemann Surfaces*, Ann. Inst. Fourier (Grenoble), **36**:167, 1986.

- [4] D. Gilbarg and N. S. Trudinger, *Elliptic Differential Equations of Second Order*, Springer-Verlag, Berlin, 2001 (reprint of the third 1998 edition).
- [5] V. Isakov, *Inverse Problems for Partial Differential Equations*, Springer-Verlag, New York 2006.
- [6] J. D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, New York 1998.
- [7] M. Kac, Amer. Math. Monthly **73**, 1966.
- [8] S. Lang, *Differential and Riemannian Manifolds*, Springer-Verlag, New York 1995.

Received: February, 2011