LINEAR ILL-POSED PROBLEMS ARE SOLVABLE ON THE AVERAGE FOR ALL GAUSSIAN MEASURES

Technical Report CUCS-035-92

J. F. TRAUB AND A. G. WERSCHULZ

July, 1992

1. INTRODUCTION

The purpose of this paper is to bring a series of rather surprising results which have appeared in the technical literature to the attention of a wider audience.

We describe the results informally here; precise definitions and results will be given below. Our subject is the solution of an ill-posed problem on a digital computer. Such a problem can at best only be solved approximately. We say a problem is *solvable* if we can compute an ε-approximation for any positive ε and is *unsolvable* if we cannot compute an ε-approximation, even for arbitrarily large ε.

It was shown in [16] that a linear problem is unsolvable iff it is ill-posed. Are there circumstances under which one can avoid this negative conclusion?

It is common to associate an operator S with an ill-posed problem. If S is linear, then a problem is ill-posed iff S is unbounded. Note that unboundedness is a worst case concept; the supremum of S operating on elements in the unit ball of its domain is infinite. This suggests introducing the concept of a problem's being *ill-posed on the average* if the expectation of S with respect to a measure is infinite and being *well-posed on the average* if this expectation is finite.

These concepts were introduced in [16], where it was shown that if the measure is Gaussian and the linear operator S is measurable, then a linear problem is unsolvable on the average iff it is ill-posed on the average.

A natural next step is to find a linear ill-posed problem which is also ill-posed on the average. Several attempts to do this for Gaussian measures were unsuccessful. This suggested the question: Is every linear problem well-posed on the average for any Gaussian measure? This question was answered in the affirmative in [5] and [14].

The conclusion in the title follows immediately from this and the preceding result.

What if the measure is non-Gaussian? Now the situation is more complicated. In particular, there are problems that are even ill-posed on the average. In this paper, we round out the picture by presenting examples of such problems for non-Gaussian measures.

The work of the the first author was supported in part by NSF Grant CCR-91-14042 and AFOSR Grant 91-0347. Part of this work was done while the first author was a Sherman Fairchild Distinguished Scholar at the California Institute of Technology and a visiting researcher at the Santa Fe Institute. The work of the second author was supported in part by NSF Grant CCR-91-01149.

The reader may note a similarity between the result that a linear ill-posed problem is unsolvable and a result from computability theory. In [9], it is shown that a linear operator preserves computability iff it is bounded. There has been considerable interest in this result; see, for example, [1] and [8, pp. 188, 215].

The results in [9] and [16] indicate a reassuring robustness. In the context of computability theory, Pour-El and Richards show that a linear unbounded transformation causes noncomputability. In the context of information-based complexity, Werschulz shows that a linear unbounded transformation causes unsolvability. (See [13] or [18] for the formulation of information-based complexity.)

Note however that there is no counterpart in the work of Pour-El and Richards to the average case setting, which is the primary focus of this paper. It might be of interest to study computability on the average.

We emphasize that this paper deals with solvability, not computational complexity. Computing an ε-approximation for a solvable problem may still require arbitrarily large computing resources. Indeed, there is no *model of computation* in this paper. We only require that certain functionals can be computed. (This is sometimes called an oracle or black-box model.)

We summarize the contents of this paper. After introducing the basic concepts in Section 2, worst case results are reported in Section 3. The average case is discussed in Section 4. We report three new results for non-Gaussian measures. In the concluding section, much stronger conclusions are given for Gaussian measures. The backwards heat equation is used as a running illustrative example. Proofs of the new results are sketched in the Appendix.

2. Basic concepts

Many problems that arise in science and engineering are *inverse problems*. That is, we are given a subset D of a normed linear space F_1, a normed linear space G, and a mapping $A: G \to F_1$. For an element $f \in D$, our goal is to compute $u \in G$ such that $Au = f$.

Recall that Hadamard ([2], [3]) said that such a problem is *well-posed*[1] if the following three conditions are satisfied:

(1) *Existence.* For every $f \in D$, there exists an element $u \in G$ such that $Au = f$.
(2) *Uniqueness.* There is at most one element u such that $Au = f$.
(3) *Continuity.* The mapping $f \mapsto u$ is continuous.

A problem that is not well-posed is said to be *ill-posed*.

Suppose without essential loss of generality that A is injective, and let D be the range of A. Define the operator $S: D \subseteq F_1 \to G$ by $S = A^{-1}$. Then the solution u of the original problem $Au = f$ is given by $u = Sf$. In this paper, we shall consider this latter form of the problem.

Note that many other mathematical problems (besides inverse problems) can be stated as obtaining Sf for some $f \in D$, where S is a linear or nonlinear operator, see [11], [13], or [18]. We say that S is the *solution operator* for our problem.

In this paper we further confine ourselves to *linear problems*. That is, we assume that the operator S is linear, in which case D may be taken as a linear subspace of F_1 without loss of generality.

[1] Hadamard's original term was "correctement posé;" his translator used "correctly set."

Recall that a linear operator S on a normed linear space is continuous iff it is *bounded*, i.e., if there exists a positive constant M such that

$$\sup_{\substack{f\in D\\ \|f\|\le 1}} \|Sf\| \le M.$$

If S is not bounded, it is said to be *unbounded*. Hence for linear S, a problem is well-posed if S is bounded, and is ill-posed otherwise.

Throughout this paper we will illustrate our results with the following running example, which is a standard instance of an ill-posed problem:

Main example: The backwards heat equation. Let $I = (0, 1)$. Consider the model one-dimensional heat equation

$$u_t = u_{xx} \qquad \text{for } (x, t) \in I \times \mathbb{R},$$

with zero boundary conditions and initial condition $u(\cdot, 0) = f$. Let $F_1 = G = L_2(I)$. Write $\langle \cdot, \cdot \rangle$ for the usual $L_2(I)$ inner product. For each positive integer j, we let

$$u_j(x) = \sqrt{2}\,\sin j\pi x \qquad (x \in I)$$

be the jth eigenfunction of the heat equation. For $t \in \mathbb{R}$, we let

$$D = D_t = \left\{ f \in F_1 : \sum_{j=1}^{\infty} \langle f, u_j\rangle^2 e^{-2j^2\pi^2 t} < \infty \right\}.$$

Then for $f \in D$,

$$Sf = S_t f = \sum_{j=1}^{\infty} \langle f, u_j\rangle e^{-j^2\pi^2 t} u_j$$

is the solution to this problem at time t. Suppose that $t < 0$, i.e., we are trying to solve this problem backwards in time. Since $S_t u_j = e^{-j^2\pi^2 t} u_j$ for all indices j, we see that the operator S_t is unbounded for any $t < 0$. □

Assume now that we want to approximate Sf by using a digital computer. In applications, f is typically a multivariate function defined on an uncountable set of points, whereas a digital computer can only store a finite set of numbers. Hence the mathematical input f must be replaced by a computer input. That is, the computer input Nf has the form

$$Nf = [\lambda_1(f), \ldots, \lambda_n(f)] \qquad \forall f \in D,$$

where the λ_i are functionals. Here, $n = n(f)$ is the number of functionals. We assume that $\lambda_1, \ldots, \lambda_n \in \Lambda$, with Λ a given set of *permissible information functionals* on F_1. Note that the number $n(f)$ of functionals, as well as the choice of which functionals λ_i to evaluate, may be determined adaptively. That is, they may depend on the previously-computed information; see [13] for further details.

The many-to-one operator N is called an *information operator*. We shall refer to the computer input Nf as the *information* about f. See Figure 1.

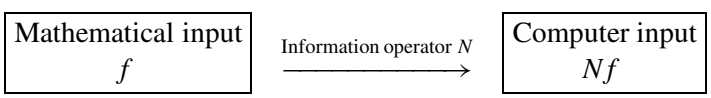

FIGURE 1

What are typical examples of Λ?

(1) Let F_1 be a function space. For some applications, only function evaluations (*standard* information) can be computed. The class Λ^{std} consists of function values, and
$$Nf = [f(x_1), \ldots, f(x_n)].$$

(2) Let F_1 be a normed linear space. The class Λ^* consists of all continuous linear functionals on F_1. In particular, if F_1 is a Hilbert space, then the continuous linear functionals are inner products of f with elements of F_1.

Because N is many-to-one, we cannot compute Sf exactly. We have to settle for an approximation. We use the information Nf to obtain an approximation $\phi(Nf)$ to Sf. We say that ϕ is an *algorithm* using the information N. See [11], [13], or [18] for more regarding this formulation.

The quality of an algorithm ϕ using information N is given by its *error* $e(\phi, N)$. How this error is defined depends on which *setting* is used. We consider the worst case setting in Section 3, and the average case setting in Sections 4 and 5.

Finally, we define the degree of *solvability* of a problem. Let us say that we can *compute* an ε-*approximation* if there exist information N and an algorithm ϕ using N such that $e(\phi, N) \le \varepsilon$. There are then (at most) three mutually exclusive possibilities:

(1) The problem is *solvable* if we can compute an ε-approximation for any $\varepsilon > 0$, no matter how small.
(2) The problem is *weakly solvable*[2] if there exists an $\varepsilon_0 > 0$ such that we can compute an ε-approximation iff $\varepsilon \ge \varepsilon_0$.
(3) The problem is *unsolvable* if we cannot compute an ε-approximation, no matter how large ε is chosen to be.

Note that the extent to which a problem is solvable depends, in particular, on the setting (i.e., on how $e(\phi, N)$ is defined) and the class Λ of permissible information functionals.

3. Worst case results

We first consider the *worst case setting*, in which the error of an algorithm ϕ using information N is its largest value over all possible f, i.e.,
$$e^{\text{worst}}(\phi, N) = \sup_{f \in D \cap F} \|Sf - \phi(Nf)\|.$$

Here $F = \{ f \in F_1 : \|f\| \le 1 \}$ is the class of *problem elements* for which we seek to find approximate solutions, our restriction to the unit ball being merely a normalization. We assume that $\sup_{f \in D \cap F} n(f) < \infty$, so that we cannot use an arbitrarily-large number of information evaluations at any problem element f. Moreover, we assume that $\Lambda = \Lambda^*$, i.e., the permissible information functionals are precisely the continuous linear functionals. Finally, in this section, the terms "solvable," "weakly solvable," and "unsolvable" will always mean "in the worst case setting."

The following theorem for the worst case setting was proven in [16]:

[2]This is not to be confused with the weak solution of an operator equation.

Theorem 1. *A linear problem is unsolvable iff it is ill-posed.* □

Main example (continued). Suppose we wish to compute the numerical solution of the backwards heat equation. Let $t < 0$. Recall that our solution operator S is $S_t \colon D_t \subseteq F_1 \to G$, where $F_1 = G = L_2(I)$ and $D = S_{-t}\big(L_2(I)\big)$. Since S is unbounded, the problem is ill-posed. Thus, this backwards heat equation problem is unsolvable in the worst case setting. □

Since we are using a worst case setting, Theorem 1 means that for any information N and any algorithm ϕ using N, there exist bad problem elements $f \in D \cap F$ for which $\|Sf - \phi(Nf)\|$ is arbitrarily large. We might hope that there are only a "few" such bad elements. This hope is demolished by the next theorem, which says that no algorithm, no matter how sophisticated, can do substantially better than the *zero algorithm* ϕ_0, which is defined by

$$\phi_0(Nf) \equiv 0 \qquad \forall f \in D.$$

(The zero algorithm is arguably the most naive algorithm imaginable.) Then we have the following result from [16]:

Theorem 2. *Let $q \in [0, 1)$. For any information N and for any algorithm ϕ using N, let*

$$A_q = \left\{ f \in D \cap F : \frac{\|Sf - \phi(Nf)\|}{\|Sf - \phi_0(Nf)\|} \le q \right\}.$$

If the problem is ill-posed, then the relative interior of A_q in the unit ball of F_1 is empty.□

Note that q can be chosen arbitrarily close to one. The set A_q consists of all problem elements at which our clever algorithm ϕ beats the naive zero algorithm by a factor of at least q. Theorem 2 tells us that A_q can contain no ball, no matter how small its radius. In short, there are many bad elements for any algorithm that attempts to solve an ill-posed linear problem in the worst case setting.

We consider the consequences of Theorem 1. If the solution operator is bounded, then the problem is either weakly solvable or solvable. Since weak solvability will generally not be enough for our needs (we might not be able to find an algorithm whose error is small enough), we want to distinguish between these classes of problems. Using Theorem 1, along with results found in Chapter 2 of [11], we have the following "trichotomy theorem" characterizing the solvability of linear problems in a Hilbert setting:

Theorem 3. *Suppose that the domain and codomain of the linear solution operator S are Hilbert spaces and that $\Lambda = \Lambda^*$.*

(1) *If S is unbounded, then the problem is unsolvable.*
(2) *If S is bounded but non-compact, then the problem is only weakly solvable.*
(3) *If S is compact, then the problem is solvable.* □

Note that bounded non-compact solution operators often appear in practice. One example is the identity operator on any infinite-dimensional space. Another is the solution of the Poisson equation $-\Delta u = f$ under zero Dirichlet boundary conditions, with error being measured in the H_0^1-norm (H_0^1 is the space of functions having one L_2 derivative and zero boundary conditions), with f belonging to the dual space of H_0^1. These two problems illustrate the general principle that a solution operator will be bounded but non-compact

when there is just enough smoothness in the elements of F to guarantee existence and uniqueness of solutions.

We close this section by mentioning that the unsolvability of ill-posed problems in the worst case setting was proved under the assumption that the *absolute error criterion* $\|Sf - \phi(Nf)\|$ was used to measure the error at a particular f. This negative result may not hold if we use a different error criterion. In particular, suppose that our original problem was a linear inverse problem $Au = f$, and that we use the *residual error criterion* $\|A(\phi(Nf)) - f\|$. Then the problem is either weakly solvable or solvable. For further details, see [17]. We also note that regularization methods (see, e.g., [7]), which are often used to solve ill-posed problems, do not claim to solve these problems in the worst case setting. Instead, they provide a solution in an *asymptotic setting*, this setting being described in Chapter 10 of [13].

4. Average case results

We have seen that ill-posed problems are unsolvable. Since ill-posedness involves a supremum, it is a worst case concept. This suggests that a positive result might be achieved if one introduced the concept of being well-posed on the average. One might be able to show that if an ill-posed problem is well-posed on the average, then it is "solvable on the average." This motivated Werschulz to initiate the average case study of ill-posed problems, see [16].

In the rest of this paper, we consider an average case setting, where we replace the requirement that the error be always at most ε by the weaker requirement that the expected error be at most ε. The terms "solvable," "weakly solvable," and "unsolvable" will always mean "in the average case setting" in what follows.

We make the following assumptions:

(1) D is a linear subspace of a Polish space[3] F_1, the space F_1 being equipped with a Borel probability measure μ and the subspace D being of full measure. (See [6], [10], [13], [15] for background material about probability measures over infinite-dimensional spaces.)
(2) G is a separable Hilbert space.
(3) $S\colon D \to G$ is a measurable linear operator.
(4) The class Λ is a subset of the class Λ_μ of *measurable linear functionals* on D.

In the *average case setting*, the error of an algorithm ϕ using information N is its expected L_2 error, i.e.,

$$e^{\text{avg}}(\phi, N) = \left[\int_D \|Sf - \phi(Nf)\|^2 \,\mu(df)\right]^{1/2}.$$

Remark. Note that there are several changes when we replace the worst case setting by the average case setting. The most obvious, of course, is that the error is given by an average, rather than a supremum. Moreover, we take our average over all elements in D, and not merely in the unit ball. Often, this does not affect the results (see, e.g., [13, pp. 257–268]). There is one other important difference between the two settings. In the worst case setting, we assumed that for any information N, the number $n(f)$ of linear functionals in Nf was uniformly bounded over all problem elements f. In the average case setting, we can relax

[3] A *Polish space* is a complete separable metric space.

this requirement; we only require that the integral $\int_D n(f)\,\mu(df)$ be finite. See [13] for further discussion. □

Recall that in the worst case setting, a problem is unsolvable iff the solution operator is unbounded, i.e, iff the problem is ill-posed. This suggests that solvability in the average case setting may depend on whether the problem is "well-posed on the average." We say that S is *bounded on the average* if $\int_D \|Sf\|^2\,\mu(df) < \infty$. We say a problem is *well-posed on the average* if S is bounded on the average, and is *ill-posed on the average*, otherwise.

We first consider problems that are well-posed on the average. Clearly any such problem is at least weakly solvable, since

$$e^{\mathrm{avg}}(\phi_0, N) = \left(\int_D \|Sf\|^2\,\mu(df) \right)^{1/2},$$

i.e., the average case error of the zero algorithm ϕ_0 is finite. Furthermore, for some restricted classes Λ, the problem is only weakly solvable. One example of a problem that is only weakly solvable is the inversion of a compact injection on a Hilbert space, the permissible information consisting of inner products of f with elements that are orthogonal to a non-trivial eigenspace.

On the other hand, we have the following positive result:

Theorem 4. *If a linear problem is well-posed on the average and $\Lambda^* \subseteq \Lambda \subseteq \Lambda_\mu$, then the problem is solvable.*

(The proofs of the theorems in this section may be found in the Appendix.)

Hence a problem that is well-posed on the average is solvable on the average, assuming that any continuous linear functional is permissible. If the problem is not well-posed on the average, it may or may not be solvable, depending on the measure μ and the class Λ of permissible information functionals, as seen in the following theorems. Our first result states that there exist problems that are ill-posed on the average, but are solvable on the average, assuming only that any continuous linear functional is permissible:

Theorem 5. *Let S be a diagonal operator on a separable Hilbert space G, let μ be an atomic measure on the eigenvectors of S, and suppose that $\Lambda = \Lambda^*$. Then the problem is solvable, regardless of whether the problem is well-posed on the average.*

Our second result is that one can find problems that are ill-posed on the average, but for which the average case solvability depends strongly on the class of permissible information functionals:

Theorem 6. *There exists a problem that is ill-posed on the average, such that the following holds:*

(1) *If $\Lambda = \Lambda^*$, then the problem is unsolvable.*
(2) *If $\Lambda = \Lambda_\mu$, then the problem is solvable.*

It is natural to ask whether there exists a problem that is so badly ill-posed on the average that it is unsolvable for $\Lambda = \Lambda_\mu$. We have not yet succeeded in finding such a problem. It is tempting to conjecture that such a problem exists.

5. Average case results for Gaussian measures

We have seen that problems in the average case setting can be either well-posed or ill-posed on the average. If a problem is well-posed on the average, it is solvable (assuming that any continuous linear functional is permissible). However, if the problem is ill-posed on the average, the problem can be either solvable, weakly solvable, or unsolvable, depending on the measure μ and on the class Λ of permissible information. In short, there is no simple criterion for deciding solvability in the average case setting. We will now discuss a recent result that settles the question of average case solvability for any Gaussian measure μ.

We assume that the measure μ is a *Gaussian measure* on a separable Banach space F_1. This means that there exists a linear *correlation operator* $C_\mu : F_1^* \to F_1$ and a *mean element* $m_\mu \in F_1$ such that

$$\int_{F_1} \exp\{i\lambda(f)\}\,\mu(df) = \exp\left\{i\lambda(m_\mu) - \tfrac{1}{2}\lambda(C_\mu\lambda)\right\} \qquad \forall\,\lambda \in F_1^*,$$

where $i = \sqrt{-1}$. (Equivalently, a measure on F_1 is Gaussian iff all continuous linear functionals on F_1 are normally distributed.) Assume that the solution operator S is a measurable linear operator defined on a subspace D of F_1 of full measure. Finally, we assume that the class Λ of permissible information functionals satisfies $\Lambda^* \subseteq \Lambda \subseteq \Lambda_\mu$.

Consider what happens when the measure for our backwards heat equation problem is Gaussian:

Main example (continued). Recall that we are trying to compute approximations to the problem whose solution operator is $S_t : D_t \subset F_1 \to G$, where $S_t f$ is the solution at time t to the heat equation with initial condition f, with $F_1 = G = L_2(I)$ and $D_t = S_t(F_1)$. We suppose that μ is a Gaussian measure on F_1; there is no essential loss of generality in assuming that the mean element of μ is zero. To simplify the analysis, we assume that the eigenvectors of C_μ are precisely the eigenvectors $u_1, u_2, \ldots$ of the heat equation, and that the corresponding eigenvalues are $\alpha_1 \geq \alpha_2 \geq \ldots > 0$. We will also require that S_t be measurable with respect to μ.

Since μ is Gaussian, the Kolmogorov zero-one principle holds. This means that S_t is measurable iff the series $\sum_{j=1}^{\infty} \alpha_j e^{-2j^2\pi^2 t}$ converges. Moreover, an easy calculation establishes that

$$\int_D \|S_t f\|^2\,\mu(df) = \sum_{j=1}^{\infty} \alpha_j e^{-2j^2\pi^2 t}.$$

Thus the measurability of S_t implies that S_t is bounded on the average, and so the problem is well-posed on the average. As we will discuss below, this is an instance of a general result. For this example, we can draw some additional conclusions.

We now apply the techniques in [18, Chapter 7] to find optimal information N_n using at most n functionals and an optimal error algorithm using N_n. That is, $e^{\mathrm{avg}}(\phi_n, N_n)$ is the infimum of $e^{\mathrm{avg}}(\phi, N)$ over all algorithms ϕ using information N consisting of at most n functionals. We find that

$$N_n f = [\langle f, u_1\rangle, \ldots, \langle f, u_n\rangle] \qquad \forall f \in D$$

is nth optimal information, and that the algorithm

$$\phi_n(N_n f) = \sum_{j=1}^{n} e^{-j^2\pi^2 t}\langle f, u_j\rangle u_j \qquad \forall f \in D$$

is an optimal error algorithm using N_n. It is straightforward to check that

$$e^{\mathrm{avg}}(\phi_n, N_n) = \left[\sum_{j=n+1}^{\infty} \alpha_j e^{-2j^2\pi^2 t} \right]^{1/2}.$$

Since S_t is bounded on the average, we see that

$$\lim_{n\to\infty} e^{\mathrm{avg}}(\phi_n, N_n) = 0.$$

To be even more specific, suppose that $\alpha_j = e^{-2j^2\pi^2 t_0}$ for some $t_0 > 0$, i.e., that $C_\mu^{1/2} = S_{t_0}$. We then find that if $t > -t_0$, then S_t is bounded on the average, and S_t is measurable with domain of full measure. Hence the problem is solvable for any $t > -t_0$. On the other hand, if $t \le -t_0$, then the domain of S_t has measure zero. □

Hence we have an example for which measurability with respect to a Gaussian measure implies that the problem is well-posed on the average, and hence solvable on the average. This is an instance of a general property, as stated in the following theorem, which was proven independently in [5] and [14]:

Theorem 7. *Let $S\colon D \subseteq F_1 \to G$ be a measurable linear transformation of separable Banach spaces. If F_1 is equipped with a Gaussian measure μ, then the problem is well-posed on the average.* □

From Theorems 4 and 7, we immediately obtain the result mentioned in the title of this paper:

Corollary. *Let $S\colon D \subseteq F_1 \to G$ be a measurable linear transformation, with F_1 a separable Banach space and G a separable Hilbert space. Let F_1 be equipped with a Gaussian measure μ. Then the problem having S as its solution operator is solvable in the average case setting.* □

Appendix

We sketch the proofs of Theorems 4, 5, and 6 in this Appendix.

Proo fo fTheorem 4. Suppose first that $\Lambda = \Lambda_\mu$, i.e., all measurable linear functionals are permissible. If $g_1, g_2, \ldots$ is an orthonormal basis for G, define information N_n as

$$N_n f = [\langle Sf, g_1\rangle, \ldots, \langle Sf, g_n\rangle] \qquad \forall f \in D,$$

and an algorithm ϕ_n using N_n as

$$\phi_n(N_n f) = \sum_{j=1}^{n} \langle Sf, g_j\rangle g_j \qquad \forall f \in D.$$

(Note that the functionals making up N_n are measurable linear functionals, but are not necessarily bounded.) Using the Lebesgue dominated convergence theorem, we find that

$$e^{\mathrm{avg}}(\phi_n, N_n)^2 = \sum_{j=n+1}^{\infty} \gamma_j,$$

where

$$\gamma_j = \int_D \langle Sf, g_j \rangle^2 \, \mu(df).$$

Since S is bounded on the average, we have

$$\sum_{j=1}^{\infty} \gamma_j = \int_D \|Sf\|^2 \, \mu(df) < \infty.$$

Since the remainder in approximating a convergent series by its partial sums converges to zero, we have

$$\lim_{n \to \infty} e^{\text{avg}}(\phi_n, N_n) = 0,$$

as required.

Now suppose that $\Lambda^* \subseteq \Lambda \subseteq \Lambda_\mu$. By Ulam's theorem (see, e.g., [15, p. 29]), μ is a Radon measure, i.e., $\mu(B) = \sup\{ \mu(K) : \text{compact } K \subseteq B \}$ for any measurable set B. It then follows that any measurable linear functional is the μ-a.e. limit of a sequence of continuous linear functionals, the proof being the same as that for the Hilbert case found in [10, Theorem 2.8.1]. We now see that solvability for the case $\Lambda^* \subseteq \Lambda \subseteq \Lambda_\mu$ follows from that for the case $\Lambda = \Lambda_\mu$. □

Proo fo fTheorem 5. Let $F_1 = G$ be a Hilbert space with orthonormal basis $\{e_j\}_{j=1}^{\infty}$, and let $\{\sigma_j\}_{j=1}^{\infty}$ be a non-decreasing sequence of positive numbers. Define

$$D = \left\{ f \in G : \sum_{j=1}^{\infty} \sigma_j^2 \langle f, e_j \rangle^2 < \infty \right\},$$

and define our solution operator $S\colon D \to G$ by

$$Sf = \sum_{j=1}^{\infty} \sigma_j \langle f, e_j \rangle e_j \qquad \forall f \in D.$$

Let μ be an atomic measure on $\{e_j\}_{j=1}^{\infty}$. Note that

$$\int_D \|S!f\|^2 \mu(df) = \sum_{j=1}^{\infty} \sigma_j^2 \mu(e_j),$$

and this may or may not be finite depending how we choose σ_j and $\mu(e_j)$. Hence, S may or may not be well-posed on the average.

To show that this problem is solvable, let

$$\lambda(f) = \sum_{j=1}^{\infty} j^{-1} \langle f, e_j \rangle \qquad \forall f \in D,$$

noting that $\lambda \in \Lambda^*$. Define information N by $Nf = [\lambda(f)]$, and define an algorithm ϕ using N by

$$\phi(Nf) = \begin{cases} \sigma_j e_j & \text{if } \lambda(f) = j^{-1}, \\ 0 & \text{otherwise.} \end{cases}$$

Then $e^{\mathrm{avg}}(\phi, N) = 0$, showing that the problem is solvable. □

Proo fofTheorem 6. Let $F_1 = \mathbb{R}^\infty$, under the metric

$$d(x, y) = \sum_{j=1}^{\infty} 2^{-j} \frac{|x_j - y_j|}{1 + |x_j - y_j|}$$

where x_j denotes the jth component of $x \in F_1$. We let u_j denote the standard jth basis vector for F_1. Next, we choose a non-decreasing sequence $\{\sigma_j\}_{j=1}^{\infty}$ of positive numbers, and let

$$D = \left\{ f \in F_1 : \sum_{j=1}^{\infty} \sigma_j^2 f_j^2 < \infty \right\}.$$

Letting G be a Hilbert space with orthonormal basis $\{e_j\}_{j=1}^{\infty}$, we define our solution operator $S\colon D \to G$ to be

$$Sf = \sum_{j=1}^{\infty} \sigma_j f_j e_j \qquad \forall f \in D.$$

(That is, $Su_j = \sigma_j e_j$ for all indices j.) Finally, we let μ be an atomic measure on $\{u_j\}_{j=1}^{\infty}$, chosen so that

$$\sum_{j=1}^{\infty} \sigma_j^2 \, \mu(\{u_j\}) = \infty.$$

Then S is not bounded on the average.

Suppose first that $\Lambda = \Lambda^*$. It is easy to see that any continuous linear functional λ on D has the form

$$\lambda\left(\sum_{j=1}^{\infty} \xi_j u_j \right) = \sum_{j=1}^{\infty} \xi_j \eta_j,$$

where only finitely many $\eta_1, \eta_2, \ldots$ are nonzero. From this, it then follows that for any information N, there is an index k_0 such that $N(u_k) = 0$ for $k \geq k_0$. Thus if ϕ is any algorithm using N, we have

$$e^{\mathrm{avg}}(\phi, N)^2 \geq \sum_{k=k_0}^{\infty} \|\sigma_k e_k - \phi(0)\|^2 \, \mu(\{u_k\}) = \infty,$$

i.e., the problem is unsolvable if $\Lambda = \Lambda^*$.

Now, suppose that $\Lambda = \Lambda_\mu$. Choosing distinct positive numbers $\eta_1, \eta_2, \ldots$, we define

$$\lambda\left(\sum_{j=1}^{\infty} \xi_j u_j \right) = \sum_{j=1}^{\infty} \xi_j \eta_j.$$

By the Lebesgue dominated convergence theorem, we see that λ is defined almost everywhere and that $\lambda \in \Lambda_\mu$. Define information N by $Nf = [\lambda(f)]$, and define an algorithm ϕ using N by

$$\phi(Nf) = \begin{cases} \sigma_j e_j & \text{if } \lambda(f) = \eta_j, \\ 0 & \text{otherwise.} \end{cases}$$

Then $e^{\mathrm{avg}}(\phi, N) = 0$, and so the problem is solvable for $\Lambda = \Lambda_\mu$. □

ACKNOWLEDGEMENTS

We would like to thank H. Woźniakowski for his insights, comments, and suggestions.

REFERENCES

1. Bridges, D. S., Review of *Computability in Analysis and Physics* (by M. B. Pour-El and J. I. Richards), Bulletin of the American Mathematical Society (New Series) **24** (1991), 216–228.
2. Hadamard, J., *Sur les problèmes aux derivées partielies et leur signification physique*, Princeton Univ. Bulletin **13** (1902), 49-52.
3. ———, *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Dover, New York, 1952.
4. Kelley, J. L. and Srinivasan, T. P., *Measure and Integral, Volume I*, Graduate Texts in Mathematics 116, Springer-Verlag, New York, 1988.
5. Kon, M. A., Ritter, K., and Werschulz, A. G., *On the average case solvability of ill-posed problems*, Journal of Complexity **7** (1991), 220–224.
6. Kuo, H.-H., *Gaussian Measures in Banach Spaces*, Lecture Notes in Mathematics **463**, Springer-Verlag, Berlin, 1975.
7. Morozov, V. A., *Methods for Solving Incorrectly Posed Problems*, Springer-Verlag, New York, 1984.
8. Penrose, R., *The Emperor's New Mind*, Oxford University Press, Oxford, 1989.
9. Pour-El, M. B. and Richards, J. I., *Computability in Analysis and Physics*, Springer-Verlag, Berlin, 1988.
10. Skorohod, A. V., *Integration in Hilbert Space*, Springer-Verlag, Berlin, 1974.
11. Traub, J. F. and Woźniakowski, H., *A General Theory of Optimal Algorithms*, Academic Press, New York, 1980.
13. Traub, J. F., Wasilkowski, G. W., and Woźniakowski, H., *Information-Based Complexity*, Academic Press, New York, 1988.
14. Vakhania, N. N., *Gaussian mean boundedness of densely defined linear operators*, Journal of Complexity **7** (1991), 225–231.
15. Vakhania, N. N., Tarieladze, V. I. and Chobanyan, S. A., *Probability Distributions on Banach Spaces*, Reidel Publishing Company, Dordrecht, 1987.
16. Werschulz, A. G., *What Is The Complexity of Ill-Posed Problems?*, Numerical Functional Analysis and Optimization **9** (1987), 945–967.
17. Werschulz, A. G., *Optimal residual algorithms for linear operator equations*, Journal of Complexity (1991), 150–173.
18. Werschulz, A. G., *The Computational Complexity of Differential and Integral Equations: An Information-Based Approach*, Oxford University Press, Oxford, 1991.

J. F. TRAUB, DEPARTMENT OF COMPUTER SCIENCE, COLUMBIA UNIVERSITY, NEW YORK, NY 10027
E-mail address: traub@cs.columbia.edu

A. G. WERSCHULZ, DIVISION OF SCIENCE AND MATHEMATICS, FORDHAM UNIVERSITY, COLLEGE AT LINCOLN CENTER, NEW YORK, NY 10023, AND DEPARTMENT OF COMPUTER SCIENCE, COLUMBIA UNIVERSITY, NEW YORK, NY 10027
E-mail address: agw@cs.columbia.edu