A Theoretical Analysis of Compactness of the Light Transport Operator

CYRIL SOLER, INRIA - Grenoble University, France
RONAK MOLAZEM, INRIA - Grenoble University, France
KARTIC SUBR, University of Edinburgh, UK

Rendering photorealistic visuals of virtual scenes requires tractable models for the simulation of light. The rendering equation describes one such model using an integral equation, the crux of which is a continuous integral operator. A majority of rendering algorithms aim to approximate the effect of this light transport operator via discretization (using rays, particles, patches, etc.). Research spanning four decades has uncovered interesting properties and intuition surrounding this operator. In this paper we analyze compactness, a key property that is independent of its discretization and which characterizes the ability to approximate the operator uniformly by a sequence of finite rank operators. We conclusively prove lingering suspicions that this operator is not compact and therefore that any discretization that relies on a finite-rank or nonadaptive finite-bases is susceptible to unbounded error over arbitrary light distributions. Our result justifies the expectation for rendering algorithms to be evaluated using a variety of scenes and illumination conditions. We also discover that its lower dimensional counterpart (over purely diffuse scenes) is not compact except in special cases, and uncover connections with it being noninvertible and acting as a low-pass filter. We explain the relevance of our results in the context of previous work. We believe that our theoretical results will inform future rendering algorithms regarding practical choices.

CCS Concepts: • Computing methodologies → Rendering; • Mathematics of computing → Functional analysis; Integral equations.

Additional Key Words and Phrases: Light Transport Operator, Compactness, Fredholm Equations

ACM Reference Format:

1 INTRODUCTION

The simulation of light transport has a long history in the computer graphics literature. Several numerical approximations have been devised such as the use of rays [Ward et al. 1988], stochastocity [Cook 1986], finite patches [Goral et al. 1984; Hanrahan et al. 1991], particles [Jensen 1996] and paths of light [Veach 1997]. Elegant theories have confirmed that these methods all serve as approximations to a central equation that governs radiative transfer [Heckbert and Winget 1991; Kajiya 1986; Lessig 2012]. This rendering equation is an integral equation which is sometimes misrepresented as a Fredholm’s integral equation of the second kind [Kajiya 1986]. In this paper, we confirm suspicions [Arvo 1995] that the light transport operator in the rendering equation is not a Fredholm operator in the general (common) case. Arvo analyses the specific case of specular transport and explains that specular reflections result in a non-compact operator. We show that the light transport operator is generally non-compact, barring special cases involving specific geometric arrangement and materials. This finding explains why several classes of light-agnostic approximation techniques result in unbounded approximation error.


Alongside theoretical results, a vast amount of practical knowledge and intuition has been amassed regarding the numerical simulation of light transport. Approaches that employed light-agnostic finite dimensional approximations—any basis such as Fourier, wavelets, radial basis functions—posed significant challenges for high-quality, artifact-free rendering [Gortler et al. 1993; Hanrahan et al. 1991; Sloan et al. 2005]. Galerkin methods that were useful for Lambertian scenes proved difficult to extend to scenes that are not Lambertian [Christensen et al. 1996; Immel et al. 1986]. Adaptive computation has evolved to be a key contributor to limiting approximation error in many scenarios such as refinement of radiosity meshes, irradiance or radiance caching [Silvennoinen and Lehtinen 2017], progressive photon mapping [Kaplanyan and Dachsbacher 2013] and baked light maps [Seyb et al. 2020; Silvennoinen and Sloan 2019]. The simplicity and efficacy of Monte Carlo methods has led to their dominance of rendering solutions for offline-rendering and their emerging relevance to real-time rendering. We hypothesize that these seemingly disparate observations stem from a key theoretical property of the transport operator—that it is not compact.

Despite tremendous developments it remains a challenge to bound approximation errors, even in simple settings such as purely Lambertian scenes, across arbitrary lighting conditions. In practice, any newly proposed approximation therefore must be validated using a variety of lighting conditions and scenes. We explain these trends from a theoretical perspective, by analyzing properties of the global light transport operator independent of its discretization or any
Contributions. In this paper, we:

- prove that the light transport operator $T$ is not compact;
- identify criteria when a simpler operator $T_k$ that transports radiant exitance in Lambertian scenes is compact;
- demonstrate that $T_k$ is noninvertible and acts as a low-pass filter.

We explain the consequences of these results and use them to reconcile challenges faced in previous works. We believe that the analysis of the continuous operator is useful to inform choices in emerging trends, such as discretizations used in approximations for real-time global illumination [Seyb et al. 2020; Silvennoinen and Lehtinen 2017] and the choice of adaptive neural bases [Müller et al. 2021].

2 REVIEW AND NOTATION

We use upper-case boldface notation for operators and scripted upper-case letters for function spaces. We denote dot products using $\langle \cdot, \cdot \rangle$ and norms using $\| \cdot \|$.

2.1 Linear operator algebra

We briefly recall various mathematical properties used in the paper.

Space and operator norms. Given a Hilbert space $\mathcal{H}$ (with dot product and norm) of square-integrable functions from a domain $\mathcal{L}$ to $\mathbb{C}$, the norm$^1$ of the linear operator $A : \mathcal{H} \to \mathcal{H}$ is (see Figure 1a)

$$
\|A\| = \sup_{\|f\|=1} \sqrt{\langle Af, Af \rangle} = \sup_{\|f\|=1} \|Af\|. \tag{1}
$$

Pointwise vs. operator-norm convergence. The convergence of a sequence of linear operators $A_n$ to some desired operator $A$ can be measured in several ways. Pointwise convergence (a.k.a. strong convergence) requires that the outputs of the sequence of operators converge to the output of $A$ on all elements of $\mathcal{H}$ (see Chatelin [2011] p122):

$$
\forall f \in \mathcal{H} \lim_{n \to \infty} \|(A_n - A)f\| = 0. \tag{2}
$$

Despite convergence everywhere, the convergence rate may be different at each $f \in \mathcal{H}$. An even stronger requirement is imposed by operator norm convergence (a.k.a. uniform convergence) which requires that

$$
\lim_{n \to \infty} \|A_n - A\| = 0. \tag{3}
$$

The key distinction between these two definitions is evident when errors $\varepsilon$ are considered for some finite $n$. Strong convergence leads to an error bound $\varepsilon_f$ which depends on $f$ (see Figure 1c). Uniform convergence on the other hand guarantees that the same $\varepsilon$ holds for all $f \in \mathcal{H}$. Uniform convergence implies pointwise convergence and we use the former by default when referring to convergence.

In the context of light transport, this implies uniform convergence irrespective of the specific radiance field being transported.

Integral operators and kernels. An integral operator $A$ given by

$$
\forall f \in \mathcal{H} \forall x \in \mathcal{L} \quad (Af)(x) = \int f(y) \kappa(x, y) \, d\mu(y) \tag{4}
$$

is a linear operator with a kernel $\kappa$ (Note that we do not use the term kernel to describe the null space $N$ of $A$ where $Af = 0$ for all $f \in N$). $A$ is known as a partial integral operator when the integral is over a linear subspace of the integration domain $\mathcal{L}$.

Compact sets and operators. A subset $W \subseteq \mathcal{L}$ is called compact if every sequence in $W$ has a converging subsequence in $W$ (assuming $\mathcal{L}$ is a metric space). Any finite set is trivially compact. A linear operator is called compact$^2$ if it maps the unit ball into a subset whose closure is compact. Equivalently, a compact operator $A : \mathcal{H} \to \mathcal{H}$ transforms any bounded sequence $f_n$ into a sequence $Af_n$ for which at least one subsequence converges in $\mathcal{H}$ (see Figure 1b). Finite-rank operators are always compact. The identity operator over infinite-dimensional spaces is not compact. The set of compact operators is closed w.r.t. the operator norm (See Golberg [1978] p67).

Spectral properties$^3$. A linear operator is normal when it commutes with its adjoint. Being normal is the necessary and sufficient condition for an operator with countable eigenvalues to have real eigenvalues and orthogonal eigenvectors. A normal operator is always self-adjoint. Compact operators in infinite dimensions are never invertible (see Conway [1990] p214). Yet their set of eigenvalues is countable with 0 as a unique limit point (see Barry [2000] section 1.2). On Hilbert spaces, compact operators can be uniformly approximated (in the operator norm) using finite rank linear operators. That is, compact operators can be reliably approximated by their action over a finite dimensional space which in practice allows approximations such as finite element and collocation methods [Chatelin 2011] and kernel-Nyström methods [Wang et al. 2009].

Hilbert-Schmidt and trace-class operators. Some compact operators satisfy stronger properties: Hilbert-Schmidt operators are the integral operators with a square-integrable kernel. For such operators, the sum $\sum_{i \geq 0} \| \lambda_i \|^2$ is finite for any orthogonal basis $\{ \epsilon_i \}_{i \geq 0} \in \mathcal{H}$ (Golberg [1978] p106). Hilbert-Schmidt operators whose trace $\sum_{i \geq 0} \langle \lambda_i, \epsilon_i \rangle$ converges are called trace-class or nuclear operators (see Golberg [1978] p95). We show that the radiant exitance transport operator in Lambertian scenes is sometimes Hilbert-Schmidt or even trace-class depending on the geometry. Trace class implies Hilbert-Schmidt, and Hilbert-Schmidt implies compactness.

2.2 The light transport operator

We denote the set of surfaces in the scene using $S$ and the hemisphere of outgoing directions expressed in the local coordinate system

$^1$We use the same notation for the operator norm (LHS of eq. 1) and the norm over $\mathcal{H}$ (RHS of eq. 1) since there is no ambiguity.

$^2$A broader definition exists for non linear operators

$^3$Throughout this paper, we use ‘spectral’ to refer to harmonic analysis rather than to the color spectrum of visible light. Radiance fields are assumed to be monochromatic.
at each point on \( S \) using \( \Omega \). The space of square-integrable half-dome distributions \( O = L_2(\Omega) \) and the space of square integrable functions \( B = L_2(S) \), with respective measures \( \cos \theta \, d\omega \) and \( dx \), are Hilbert spaces\(^4\). Their tensor product space \( \mathcal{H} \) is the Hilbert space we use to represent radiance distributions, with Lebesgue measure \( d\mu = \cos \theta \, d\omega \, dx \) (See notations on Fig. 1.d), and inner product
\[
\langle L_1, L_2 \rangle_{\mathcal{H}} = \int_{\Omega} L_1(x, \omega) L_2(x, \omega) \, d\mu.
\]
(5)

A Hilbert space allows us to consider compactness properties for adjoint operators (which need inner products, thus preferable over Banach spaces) while remaining less restrictive than Sobolev spaces.

Radiance transport operators. The reflectance operator \( \mathbf{R} : \mathcal{L} \to \mathcal{L} \) transforms a field of incident radiance into a field of reflected radiance. To simplify notations (without altering the properties of the operators), we parameterize both the incident and outgoing radiance in the upper hemisphere of directions \( \Omega \). With these settings (See Fig. 1.d), we have
\[
(KL)(x, \omega) = \int_{\Omega} L(x, \omega) \rho(x, \omega, \omega') \cos \theta \, d\omega',
\]
(6)
where \( \rho : S \times \Omega \times \Omega \to \mathbb{R}^+ \) is the bidirectional reflectance distribution (BRDF).

The light transport operator \( T \) transports a field of radiance in the scene to the next visible surface where it reflects it to obtain the transported and scattered radiance field. Arvo [1996] defined a linear operator \( G \) to transform a radiance field \( L \) into the field after transport to the nearest surface. Defining \( p \) to turn a point and direction \( (x, \omega') \) into the corresponding point and direction \( (y, \omega''') \) as traced from \( x \) along direction \( \omega'' \) (see Figure 1.d), the transported radiance before reflection will be
\[
L_t(x, \omega') = L(p(x, \omega')) = (GL)(x, \omega').
\]

Using this, the light transport operator \( T \) is expressed
\[
(TL)(x, \omega) = \int_{\Omega} L_t(x, \omega) \rho(x, \omega, \omega') \cos \theta' \, d\omega' = (KGL)(x, \omega).
\]
(7)
At equilibrium, the radiance field \( L \) is the sum of the emitted radiance field (from light sources) \( E \) and transported radiance (bounced once) in the scene, leading to the light transport equation
\[
L = TL + E.
\]
(8)

Comments on transport operators. \( \mathbf{K} \) and \( T \) are partial integral operators, since the integral is defined over a subspace of \( \mathcal{L} \). Since \( \mathbf{K} \) only acts over directions, Equation 6 may also be written using the local reflectance operator \( \mathbf{K}_x : O \to O \), considering \( x \) as a constant. Defined as above, \( \mathbf{K} \) and \( \mathbf{G} \) act and take values in the same space \( \mathcal{H} \). With this convention, both \( \mathbf{K} \) and \( \mathbf{G} \) are self-adjoint [Veach 1997]

\footnotetext[4]{Both spaces are complete inner product spaces [Royden and Fitzpatrick 2010].}

---

**Fig. 2.** Notations used in this paper.
w.r.t. the inner product in Equation 5, and wherever p is defined, we have $(G^2L)(x, \omega) = L(x, \omega)$.

**Radiant exitance operator.** We will also show that a restricted transport operator which only applies to diffuse surfaces exhibits different properties. Since $L$ in scenes with only Lambertian surfaces is independent of angle, we define a radiant exitance (or radiosity) field instead $B \in \mathcal{B}$ with $B(x) = \int_\Omega L(x, \omega) \cos \theta \, d\omega$. As a result, Equation 7 turns into an equation over exitant radiant exitance $$(T_b B)(x) = \int_S \kappa(x, y) B(y) \, dy,$$ with a kernel that is (notation as in Figures 1.d and 2) $$\kappa(x, y) = \frac{\rho(x) \cos \theta \cos \theta'}{\pi \|x - y\|^2}.$$ While $T$ is a partial integral operator, $T_b : \mathcal{B} \to \mathcal{B}$ is a proper integral operator, since the integration is over $S \subset \Omega$ the entire domain of elements of $\mathcal{B}$. However unlike $T$, operator $T_b$ cannot be written as a nontrivial product of self-adjoint operators both operating over $\mathcal{B}$.

3 THEORY

In this section, we prove that the light transport operator is not compact in general. Compact operators are never invertible, act as low-pass filters, and admit a Schmidt expansion [Gohberg and Krein 1978]. However, since we show that $T_b$ is not compact in general we also address the questions of whether such properties still hold.

3.1 Compactness theorems

**Theorem 3.1:** The light transport operator $T$ is not compact.

**Proof:** In order to express $T$ as an integral operator over the full domain $S \times \Omega$ one needs to rewrite $L_i$ as a function of $L(y, \omega')$ in Equation 7, and restrict the integral to pairs for which $(x, \omega) = p(y, \omega')$. Indeed $\omega'$ is already implicitly determined by the positions of $x$ and $y$. The full-domain integral therefore needs an additional Dirac function, which makes the operator a partial integral operator. Consequently $T_b$ is not compact [Kalitvin and Zabrejko 1991].

Intuitively, partial integral operators are not compact because they leave infinite dimensional subspaces untouched. This is the case for $K$ that ignores the spatial component of light distributions, while $K_S$ (see supplemental material) is compact.

**Theorem 3.2:** If the kernel $\kappa$ of the integral operator $T_b$ in Equation 9 is bounded over $S^2$, then $T_b$ is compact.

**Proof:** If $\kappa$ is bounded over $S^2$, then it is also square-integrable over that domain, which makes $T_b$ a Hilbert-Schmidt operator. This implies that $T_b$ is compact for this case.

**Theorem 3.3:** If $\kappa$ is not bounded, then $T_b$ is not compact.

**Proof:** We prove this (see Section 3 of the supplementary material) by constructing a bounded (in the $L_2$ norm) sequence of light distributions $B_n$ such that $T_b B_n$ will contain arbitrarily high spatial frequencies. Because of this, $u_b = T_b B_n$ is proved to not have any converging subsequence in $\mathcal{B}$. We conclude that $T_b$ is not a compact linear operator on the Hilbert space $\mathcal{H}$. Consequently, the deviation of $T_b B_n$ from finite rank approximations of light distributions will not converge to 0 for arbitrarily large finite function bases.

We provide intuition of Theorem 3.3 via an experiment in Figure 3. Starting with a unit area disk emitter $S_b$ tangent to the edge of two abutting Lambertian half-planes, we render images by progressively halving the emitting zone while doubling its emittance to obtain a sequence of scenes with $L_2$-bounded existant radiance. For each $n$, we transport radiance to the receiver plane via $T_b B_n$, which tends towards a spike as $n \to \infty$. A tabulation of differences in the sequence of $T_b B_n$, calculated via $\|T_b B_n - T_b B_{n+p}\|_2$, shows that it does indeed not converge to zero. Although placing the emitter at the edge is an exaggeration that allows us to illustrate this effect with one-bounce transport, this situation occurs commonly in the multibounce setting. Fortunately, the operator exhibits better behavior away from such ‘edge’ cases as shown by the following theorem.

**Theorem 3.4:** In any region $A$ of the scene for which $\kappa$ is bounded over $A \times S$, the output of $T_b$ coincides with the output of a compact operator $T'_b : S \to A$.

**Proof:** We prove this (Proposition 5.1 in Section 5 of the supplementary material) by constructing a Hilbert-Schmidt operator $T'_b$.

$$\forall B \in \mathcal{B} \quad \forall x \in A \quad (T'_b B)(x) = (T_b B)(x).$$

In summary the above theorems reveal that operators $T$ and $T_b$ exhibit different mathematical properties, because they act between different pairs of spaces. $T$ is never compact, including when all materials are Lambertian. The potential for compactness of $T_b$ arises from the rewriting of $T$ into another space (radiant exitance rather than radiance) within which their values actually match when the scene is Lambertian. While $T_b$ is not compact in the majority of cases, theorem 3.4 states that $T_b$ behaves like a compact operator almost everywhere, except near points where $\kappa$ is unbounded. In particular, $\kappa$ is unbounded next to concave abutting edges, corners, and contact points between surfaces. However $T_b$ is compact in purely curved scenes for which we prove that $\kappa$ is bounded even though it is discontinuous (Sec.4 in the supplementary material). Further, we prove its continuity at points where the two principal curvatures are equal (an uncommon but curious case).

3.2 Consequences

A Fredholm equation defines a compact linear operator. Since neither $T$ nor $T_b$ (generally) is compact, the rendering equation is not a Fredholm equation per se. The lack of compactness of $T$ implies that its finite rank approximations do not converge to $T$ in the operator norm (Equation 2). In other words, for any given $\epsilon > 0$ no finite rank approximation $T_n$ can ensure that $\| (T - T_n) L \| \leq \epsilon$ across all light distributions $L$. However, since pointwise convergence is assured $^3$, it is still possible to provide a well-crafted approximation $T_n L$ to $T L$ for a specific lighting distribution $L$. This result supports the success of adaptive methods for specific light transport problems.

Although a series of compact operators $T_n$ that strongly converge to $T$ or $T_b$ may be constructed – say using a bounded kernel

\footnote{This can be proved by projecting $L$ over an orthogonal basis (such as wavelets) and using the fact that $\|T_b\| < 1$.}
A general factorization technique extends the SVD for all bounded orthogonal, we also find that (see Section 2 of the supplementary material) we show that the radiant exitance transport operator \( T_b \) is not compact. Left: rendered scene with a disk \( S_0 \) emitting diffuse light. The disk is tangential to the abutting edge. Center: measured radiant exitance on the receiver surface for \( n = 7 \). As \( n \) is increased, the spatial frequencies of the emitted energy increases while preserving its \( L_2 \) norm, as does the received energy. In the limit, it escapes any finite rank approximation over the space of functions on the receiver. Right: table showing that the pairwise measured \( L_2 \) distances between \( T_b B_n \) and \( T_b B_{n+p} \) do not depend on \( n \) (for large \( n \), because the receiver is bounded) and converges to non-zero values with \( p \) for a given \( n \). To ensure numerical accuracy in this figure we reduce the disk radius by a factor of 1.2 instead of 2, and use the analytical formula of Naraghi [1988].

\[ \kappa(x, y) = \max(n^2, \kappa(x, y)) \] - pointwise convergence does not yield a compact operator in the limit (unlike with uniform convergence since compact operators form a closed set).

### 3.3 Singular values expansion of \( T \)

A general factorization technique extends the SVD for all bounded linear operators on infinite dimensional spaces. If the operator is compact, a converging infinite sum known as the Schmidt expansion stands for the SVD [Gohberg and Krein 1978]. Otherwise, the factorization involves the use of two unitary operators (standing for the right and left singular vectors) and a function-multiply operator in place of the SVD diagonal matrix [Crane and Gockenbach 2020]. Since \( T \) is not compact, we have no guaranty about the existence of a Schmidt expansion. The following result proves that such an expansion actually exists in a limited set of configurations.

**Theorem 3.5:** In closed scenes with a finite number \( n \) of non-spatially-varying materials \( m_i : \Omega \times \Omega \to \mathbb{R} \), the action of \( T \) over any radiation distribution \( L \) can be written as

\[
TL = \sum_{i=1}^{n} \sum_{j=0}^{\infty} \rho_{ij} \langle \psi_j \rangle_{m_i} L \psi_j,
\]

where \( \rho_{ij} \) is a positive sequence and \( \psi_j \) is the \( j \)-th eigenvalue of \( m_i \), and \( \langle \psi_j \rangle_{j>0} \) are two orthogonal bases of \( H \) defined as

\[
\psi_j(x, \omega) = \phi_{k} \psi_{j}(x) x_{k},
\]

in which \( \{\phi_k\}_{k=0}^{\infty} \) form an orthogonal basis of the subset of \( S \) where the material is \( m_i \), \( \{\psi_j\}_{j=0}^{\infty} \) are the eigenfunctions of \( K_S \) for material \( m_i \), and \( g : j \mapsto (n_1, n_2) \) provides a bijection between \( \mathbb{N} \) and \( \mathbb{N}^2 \).

In the proof (derived in Sections 1 and 2 of the supplementary material) we show that \( \psi_j \) and \( \phi_j \) are the eigenfunctions of \( TT^* \) and \( T^*T \) respectively, and that \( \rho_{ij} \) are the eigenvalues of \( TT^* \). Curiously, in addition to the property that eigenfunctions for each material are orthogonal, we also find that (see Section 2 of the supplementary material) the sets of all eigenfunctions is also orthogonal. Formally, \( \{\psi_j\}_{j>0} \) is orthogonal (as is \( \{\psi_j\}_{j>0} \)) and the set of all \( \{\psi_j\}_{i,j>0} \) is also orthogonal and complete in \( H \) (similarly for \( \{\phi_k\}_{k=0}^{\infty} \)). This result stems from a partitioning of \( S \) into a finite number of regions that have the same material. Theorem 3.5 can be rewritten as

\[
TL = \sum_{n=0}^{\infty} \rho_n \langle \psi_n \rangle_{m_i} \psi_n,
\]

where \( \rho_n \) is the absolute value of an eigenvalue of one of the materials \( m_i \), repeated an infinite number of times within the sum. Equation 11 assumes the familiar form of SVD and is an extension of the concept for \( T \) as an operator over an infinite dimensional space. The above does not apply in the space \( B \) of radiant exitance distributions since, for Lambertian materials, functions \( \psi_n \) in Equation 12 have a non constant angular component.

### 3.4 Non-invertibility of \( T_b \)

When \( T_b \) is compact it is trivially not invertible. The question is solved for the general case by the following theorem:

**Theorem 3.6:** \( T_b \) is not invertible.

**Proof:** We prove this using a counterexample in Section 5 of the supplementary material. We exploit the ‘compact-like’ behavior of \( T_b \) almost everywhere as proved by Theorem 3.4. Using this property, we construct a counterexample - a light distribution that is not in the span of \( T_b \).

Since the output of \( T_b \) matches that of a compact operator in most regions of the scene, the proof leverages the existence of light distributions for which the operator is not-invertible. e.g. inferring diffuse emittance for a discontinuous distribution of radiosity on a receiver in the absence of occluders. In practice evaluating the solution to \( TX = D \) causes \( X \) to have arbitrarily high frequencies when the inverse does not exist. Since \( T \) washes out these high frequencies (See Theorem 3.7), it results in an arbitrarily large norm. This is analogous to deconvolution of a step-function.

### 3.5 \( T_b \) acts as a smoothing operator

We show that \( T_b \) naturally acts as a low-pass filter when it is compact, which also remains true almost everywhere when \( T_b \) is not compact.
Theorem 3.7: In regions $A$ where $x$ is bounded over $A \times S$, $T_b$ acts as a low-pass filter in frequency space.

Proof: The proof (Section 5 of the supplementary material) first shows that compact operators, in general, act as low-pass filters, and combines this result with Theorem 3.4 stating that $T_b$ locally behaves like a compact operator in regions of bounded kernel. That is, given a basis $\{b_n\}_{n \geq 0}$ of $B$ with increasing frequency content with $n$, we have

$$\lim_{n \to \infty} \|T_b b_n\|_A = 0,$$

where $\| \cdot \|_A$ is a norm for square-integrable functions over $A$. \qed

For Lambertian scenes with unbounded $x$ ($T_b$ is non compact), the transport process does not reduce high frequencies in the radiant exitance. Thus, an arbitrarily large number of basis functions is required to faithfully (fixed precision) represent $T_b B$ across all $B$. Unfortunately this represents a practically common situation.

4 CONNECTIONS TO PREVIOUS WORK

The theoretical properties that we identify in Sec. 3 manifest as Theorem 4.1: intuitively imagined to ‘smooth’ out $L$ for successive bounces, the arbitrary locations of discontinuities implies that the use of a fixed basis, regardless of the choice of functions (Fourier, wavelets, radial basis functions), would require an infinite set of basis functions even for simple scenes.

4.1 Finite-rank and finite-bases: cannot bound error

Our theory explains observations that have been made empirically via trial and error. Methods which discretize the transport operator directly (using a matrix), such as radiosity [Goral et al. 1984; Hanrahan et al. 1991] and matrix row-column sampling [Hašan et al. 2007], have struggled with bounding approximation error for arbitrary light distributions. This is often because they represent $T$ using a low- or finite-rank matrix. Although this might be reasonable for local analysis (e.g. see Lessig [2012, p.470] and Mahajan [2007]), the arguments do not hold for the global operator. It is also well known that extending such methods to non-Lambertian transport is challenging. Virtual point lights in instant radiosity [Keller 1997] create singularities at abutting edges (see Section 3.2) thereby necessitating ad hoc strategies including blurring energy contributions over spheres to avoid singularities [Hašan et al. 2009].

Approximating $L$ in multi-bounce transport using a finite set of basis functions is another example of methods that have faced difficulties in scalable extensions to glossy scenes. Although $T$ is often intuitively imagined to ‘smooth’ out $L$ for successive bounces, the arbitrary locations of discontinuities implies that the use of a fixed basis, regardless of the choice of functions (Fourier, wavelets, radial basis functions), would require an infinite set of basis functions even for simple scenes.

4.2 Adaptive methods

High error due to fixed bases motivated the development of adaptive methods where the discretization is directly dependent on the specific distribution of light and therefore enables point-convergence to an arbitrary error threshold. For example, adaptive refinement of surface patches at each Gauss-Seidel iteration while calculating radiosity [Cohen et al. 1988; Heckbert 1990] and irradiance caching [Krivánek et al. 2008] where the decision to add a new point to the cache depends on local variation of the irradiance (spatial as well as angular with respect to the normal). Furthermore adaptation was devised to make caching practical in conjunction with clamping [Krivánek et al. 2006] and for global illumination [Krivánek et al. 2008]. Adaptive sampling enables sparsification of transport matrices for use in real-time applications [Huang and Ramamoorthi 2010]. These works invented targeted schemes for specific applications.

Realizing the importance of adaptive methods, systematic theory and tools were developed to analyze and propagate local variation in the radiance field [Durand et al. 2005]. These methods have been used to inform adaptive sampling strategies for reducing approximation error from optical effects such as depth of field [Soler et al. 2009], shadows [Ramamoorthi et al. 2012], motion blur [Egan et al. 2009] and glossy materials [Bagher et al. 2012]. Some practical methods develop low-dimensional approximations by tailoring data-driven bases to the output of $T$ rather than developing light-agnostic discretizations [Lehtinen et al. 2008]. This approach resembles methods for dimensionality reduction of large matrices [Haliko et al. 2011].

More recently, neural networks were trained on-line for caching transported radiance in real-time [Müller et al. 2021] applications. Neural radiance caches typically map $S \times \Omega$ onto radiance values. Although they do not explicitly adapt the sample positions, the adaptation stems from a combination of two effects. First, rays are traced from the camera towards the scene to determine cache locations. In addition, the radiance cache is not trained across distributions, but on a specific light distribution (the one being rendered).

4.3 Invertibility

The inference of emitting distributions $L_e$ that result in a given transported distribution $L$ involves inverting the transport equation. Although $T$ may not be invertible (see Section 5), two broad strategies have been useful in inferring $L_e$. The first is to consider only direct reflection [Ramamoorthi and Hanrahan 2001, 2004] and therefore only requires approximately inverting the local reflectance operator $K_x$—which is actually not invertible either as explained in Section 5. The second strategy, for multibounce inverse light transport [Bai et al. 2010; Seitz et al. 2005] appears deceptively simple when the operator equation is rewritten as $L_e = (I - T)L$. However since only a sparse sampling of $L$ is available, say from a photograph or a video, and since scene properties (properties of $T$) are usually unknown, the inference problem is ill-posed.

4.4 Volumetric light transport

The transport of light through participating media is governed by a differentio-integral equation from which an operator equation can be derived [Jakob and Marschner 2012; Zhang et al. 2019]. All operators in their equation contain partial integrals, either over direction only or integrating along rays. Hence, just as with the transport in non-Lambertian scenes described in Sec. 3.1, all these operators are not compact in the general case. As with Lambertian scenes with unbounded $x$ ($T_b$ is non compact), the transport process does not reduce high frequencies in the radiant exitance. Thus, an arbitrarily large number of basis functions is required to faithfully (fixed precision) represent $T_b B$ across all $B$. Unfortunately this represents a practically common situation.
surface transport, it is possible that special cases exist for volumetric transport where the operator is compact. We leave this analysis as future work.

5 DISCUSSION

Monte Carlo methods. Literature from the era of Monte Carlo (MC) dominance in rendering is notably sparse in our literature review. The main reason for this is that MC methods do not rely on finite-rank approximations, fixed bases or discretizations of the transport operator. However we believe that the results in this paper will gain relevance with emerging trends to revisit older strategies for applicability towards real-time rendering or the use of modern representations such as neural networks.

Relevance. There are two strands of emerging research which could benefit from this analysis. First, precomputed global illumination will play a significant role at least until gaming consoles are equipped with hardware raytracing capability. These methods rely on encoding transport using textures (lightmaps) which involves important choices surrounding discretization and representation using bases. Recent trends in realtime global illumination also show the implicit usage of low rank approximations of T in the form of basis functions [Koskela et al. 2019] or probes [Hu et al. 2021; Majercik et al. 2019]. Another recent area is the use of neural networks to encode radiance in a scene [Mildenhall et al. 2020], for which the impact of discretization (sampling) is not well studied. The development of these representations in tandem with light transport capability is still in its early stages [Boss et al. 2021; Srinivasan et al. 2021; Zhang et al. 2021]. Adapting solutions to visibility is a common heuristic. Although discontinuities due to visibility do pose problems they are not related to boundedness or compactness as discussed in this paper. We believe that a better understanding of T could spur principled choices in such applications.

Global vs. local reflectance. The (global) reflectance operator K acts on an incident radiance field over $S \times \Omega$ to produce an exitant radiance field. Its local counterpart $K_x : O \rightarrow \hat{O}$ can be defined as an integral operator that turns incident radiance at a point $x$ into exitant radiance at $x$. It can be shown that $K_x$ is compact while $K$ is not (See supplemental). The latter can be viewed as a tensor product of the identity operator (in space) and $K_x$ in directions, hence its non compactness. For this reason an infinite set of local operators $K_x$ must share eigenvalues with each of them for global eigenfunctions of $K$ to exist, while the tensor product nature of $K$ causes infinite multiplicity of these eigenvalues. Applications that analyze, approximate or invert $K_x$ do so using a finite rank approximation for $K_x$ [Ramamoorthi and Hanrahan 2004], which amounts to projecting radiance onto a subspace via low-pass filtering (in addition to the low-pass effect driven by the kernel of $K$), although $K_x$—as a Hilbert-Schmidt operator in infinite dimensions—is not invertible.

Difficulty computing eigenvalues. Most numerical methods for estimating eigenvalues and eigenfunctions of operators assume that their eigenvalues are decreasing. However, since T is not compact, its eigenvalues may not even be countable, and if so, will certainly not tend to 0. Thus an empirical demonstration on real scenes is challenging. In particular, approximating eigenvalues of $T$ (or $T_p$ when not compact) using an eigensolver over a transport matrix is absolutely not guaranteed to provide an approximation of the actual eigenvalues. Another practical problem is that in some scenes (non closed scenes where G is not defined everywhere), the series given by theorem 3.5 does not stand for a proper singular value decomposition since $\{\phi_i^j\}$ is not orthogonal.

Open problems and future work. In the special cases where $T_p$ is compact, we know that its eigenvalues are countable. When $G$ is defined everywhere (closed scene), $T$ has countable singular values. Other than these situations, we do not know whether $T$ or $T_p$ have a countable set of eigenvalues. We do expect infinite multiplicity for the latter as previously explained. Since it is not compact, it remains to be proven whether $T$ is invertible. Since light transport operators do not satisfy key properties (continuous kernel, self-adjointness, etc) required by existing tools in operator analysis, further work is necessary to enable their analyses.

6 CONCLUSION

In this paper, we prove theoretical properties pertaining to the compactness of the light transport operator. Our results provide mathematical reassurance of intuitions gleaned from decades of empirical experience. For example, the lack of compactness confirms that the light transport operator does not lend itself to low-dimensional approximations except in specific cases away from edges and diffuse surfaces. The absence of uniform convergence implies that estimation errors will depend on specific light emission radiance and radiance analytics exxists the difficulties in extending algorithms that perform well on scenes with diffuse materials to scenes with glossy materials. While these properties may be the consequence of the mathematical model considering light transport as a non-oscillatory simplification of Maxwell theory, their study remains relevant in the context of geometric optics. We hope that future work will leverage properties such as the existence of an SVD of $T$ to design adaptive algorithms where the discretization of $T$ could sidestep problems caused by its lack of compactness.

ACKNOWLEDGMENTS

Cyril Soler and Ronak Molazem were supported by the French National Research Agency through the ANR-16-CE33-0026 project (CalLiTrOp – “Comprehensive Analysis of Light Transport Operators for image synthesis”). Kartic Subh was supported by a Royal Society University Research Fellowship.

REFERENCES


A Theoretical Analysis of Compactness of the Light Transport Operator


