A CONTINUITY THEOREM FOR CORES OF RANDOM CLOSED SETS

PEDRO TERÁN

(Communicated by Richard C. Bradley)

Abstract. If a sequence of random closed sets $X_n$ in a separable complete metric space converges in distribution in the Wijsman topology to $X$, then the corresponding sequence of cores (sets of probability measures dominated by the capacity functional of $X_n$) converges to the core of the capacity of $X$. Core convergence is achieved not only in the Wijsman topology, but even in the stronger Vietoris topology. This is a generalization for unbounded random sets of the result proved by Artstein for random compact sets using the Hausdorff metric.

1. Introduction and main result

A selection of a random set $X$ is a random variable almost surely contained in $X$. It is natural to ask about the relationships between the distribution of a random set and its selectable distributions. Typical examples in this spirit are Aumann’s definition of the expectation (or integral) of a random set via its integrable selections [2], and Hess’s result that a set-valued martingale can be represented as the closure of a countable family of martingale selections [10, 24].

In recent years, there has been a growing interest in selectable distributions mainly due to applications in Economics and Statistics; see e.g. [18, 21, 13, 8, 16]. In the latter context, a random set is viewed as an imprecise observation of an underlying random variable (a ‘coarse data’ model), so selectable distributions are the candidates to be the true unobserved distribution. A key question is whether closeness of random set distributions (e.g. the empirical and population distributions) ensures that their selectable distributions are also close in some sense.

Artstein [1] solved that question for random compact sets in a separable complete metric space $E$. He proved that, for any sequence converging in distribution in the Hausdorff metric, the corresponding maximal sets of selectable distributions converge in the Hausdorff metric associated to the Prokhorov metric in the space of probability measures.

Received by the editors December 29, 2006, and, in revised form, October 25, 2007. 2000 Mathematics Subject Classification. Primary 60D05. Key words and phrases. Choquet capacity, core, random closed set, Vietoris topology, Wijsman topology.

This research was partially supported by Spain’s Ministerio de Educación y Ciencia under its research grants MTM2005-02254 and TSI2005-02511, and the Gobierno de Aragón under its research grant PM2004-052.
It is a bit burdensome, however, to state the results in terms of selectionable distributions, because these may not be independent of the underlying probability space. We borrow from game theory and economic theory the slightly different language of cores, which is independent of the probability space.

Recall that a random closed set $X$ induces the mapping $\nu$ given by

$$\nu(A) = P(X \cap A \neq \emptyset)$$

for each compact set $A$, which extends to Borel subsets as an infinitely alternating Choquet capacity called the capacity functional or hitting functional of $X$, e.g. [17, Chapter 1].

The core of $\nu$ is the set of all probability measures dominated by $\nu$. In cooperative game theory, $\nu$ is called the characteristic function of a game and $\nu(A)$ is interpreted as the value of a coalition of players. Then the core is the set of payoff allocations under which no coalition can perform better by leaving the ‘grand coalition’ of all players.

Theorem 1 and Proposition 3.5 in [21] amount to a nonstandard proof that cores are closed. A standard proof for random compact sets is given in [5, Corollary 3.4]. For lack of a fully general standard proof, we adopt the notation core for the closure of the core. Throughout the paper, $\nu_n, \nu$ will consistently denote the capacity functionals of the random sets denoted by $X_n, X$.

Our main result is then as follows. Observe that local compactness of the carrier space, a usual assumption in results about distributions of random closed sets, is not invoked.

**Theorem 1.1.** Let $(E, d)$ be a separable complete metric space and $\{X_n, X\}_n$ be random closed sets in $E$. If $X_n \rightarrow X$ in distribution in the Wijsman topology associated with $d$, then

$$\text{core}(\nu_n) \rightarrow \text{core}(\nu)$$

in the Vietoris topology for closed sets of probability measures in $E$.

The Wijsman topology is quite weaker than the Vietoris topology; e.g. the former is Polish while the latter is metrizable only if the carrier space is compact. Thus it is striking that Wijsman convergence of random sets implies not just Wijsman but Vietoris convergence of cores.

2. Notation and preliminaries

Let $(E, d)$ be a complete separable metric space. The closed ball with center $x$ and radius $\varepsilon$ will be denoted $B(x, \varepsilon)$.

The space $F$ is formed by all nonempty closed subsets of $E$. When the need arises to indicate explicitly the carrier space $E$, we will write $F(E)$. The complement, closure and interior of a set $A$ are denoted by $A^c$, $\text{cl} A$, and $\text{int} A$. We also set $d(x, A) = \inf_{y \in A} d(x, y)$.

The $\varepsilon$-enlargement of $A$ is defined to be

$$A^\varepsilon = \{x \in E \mid d(x, A) \leq \varepsilon\} = \bigcup_{x \in A} B(x, \varepsilon).$$

The Hausdorff metric is given by

$$d_H(A, C) = \max\{\sup_{x \in A} d(x, C), \sup_{y \in C} d(y, A)\}$$

$$= \inf\{\varepsilon > 0 \mid A \subset C^\varepsilon, C \subset A^\varepsilon\}.$$
Unless A and C are bounded, the value of the metric can be infinite. The topology generated by $d_H$ will be denoted by $\tau_H$ or $\tau_H(d)$. The definition of $d_H$ extends to arbitrary sets, in which case it becomes a pseudometric.

Also, note that we use the subindex ‘$H$’ consistently whenever it is necessary, so $\rho_H, \delta_H$, etc., are the Hausdorff metrics associated with the metrics denoted by $\rho, \delta$ and so on. Apart from $\tau_H(d)$ we will make use of three other types of topologies on $\mathcal{F}$. There is a wealth of such topologies, and the reader is referred to [3] for a detailed study.

The Victoris topology $\tau_V$ has as a subbase the sets $\{A \in \mathcal{F} \mid A \cap U \neq \emptyset\}$ and $\{A \in \mathcal{F} \mid A \cap F = \emptyset\}$ for $U$ open and $F$ closed. Accordingly, a sequence $\tau_V$-converges when it eventually hits every open set hit by its limit, and it eventually misses every closed set missed by its limit. This is a very old notion, dating back to the work of one of the founders of topology [22]. It is interesting in that there are many ‘if and only if’ links between its topological properties and those of the basic space (e.g. compactness), however often too strong for practical applications in the noncompact case.

The proximal topology $\tau_P(d)$ is a weaker variant of the Victoris topology having the same lower part (i.e. that concerning open sets) but its upper part is generated by the sets $\{A \in \mathcal{F} \mid A \text{ is far from } F\}$ instead. Two sets $A, C$ are far if $\inf_{x \in A, y \in C} d(x, y) > 0$. The difference between Victoris and proximal convergence can be appreciated with the following example. For any real number $a$, define the closed set $A[a] = \{(x, x^{-1} + a) \in \mathbb{R}^2 \mid x > 0\}$. Then, if a sequence $a_n$ converges to 0 but oscillates around it (like, for instance, $a_n = (-1)^n n^{-1}$), then $A[a_n] \rightarrow A[0]$ proximally but not in the Victoris topology, since it hits the open set $\mathbb{R} \times (-\infty, 0)$ infinitely often. The proximal topologies were studied mostly during the nineties, as a part of the effort to systematize hyperspace topology; see e.g. [3].

The Wijsman topology $\tau_W(d)$ [25] is the weak topology generated by the distance functions $\{d(x, \cdot) \mid x \in E\}$. Therefore it appears as the topology of pointwise convergence when each set is identified with the function $d(\cdot, A)$.

The Victoris topology depends only on the topology induced by $d$, whereas Wijsman and proximal topologies, in general, are not invariant under equivalent metrics. In general, $\tau_H(d)$ and $\tau_V$ are incomparable, but both are finer than $\tau_P(d)$, which is finer than $\tau_W(d)$.

Finally, we will be mentioning Painlevé-Kuratowski convergence, a still weaker notion. A sequence $\{A_n\}_n \subset \mathcal{F}$ is said to converge in the Painlevé-Kuratowski sense to $A \in \mathcal{F}$ if $\liminf_n A_n = \limsup_n A_n = A$, where $\liminf_n A_n$ is the set of limits of convergent sequences $a_n \in A_n$ and $\limsup_n A_n$ is the set of limits of convergent subsequences $a_{n_k} \in A_{n_k}$.

For the purposes of this paper, a random element $\xi$ is a Borel $E$-valued function defined on a probability space $(\Omega, A, P)$. Its law or distribution $\mathcal{L}(\xi)$ is the probability measure $P \circ \xi^{-1}$. A sequence $\{X_n\}_n$ of random elements of $E$ is said to converge in distribution to a random element $X$ if

$$E[f(X_n)] \rightarrow E[f(X)]$$

for every bounded continuous function $f : E \rightarrow \mathbb{R}$.

In that situation, the sequence of laws $\{\mathcal{L}(\xi_n)\}_n$ is said to converge weakly to $\mathcal{L}(\xi)$. Weak convergence gives rise to a metrizable topology, the weak topology, on
probability measures. For instance, the Prokhorov metric

\[ \rho(P, Q) = \inf\{\varepsilon > 0 \mid \forall A \in \mathcal{F}, P(A) \leq Q(A^\varepsilon) + \varepsilon\} \]

induces the weak topology.

Let \((\Omega, \mathcal{A}, P)\) be a probability space. A mapping \(X : \Omega \to \mathcal{F}\) is a random closed set if it is Effros measurable; namely, the events

\[ \{\omega \in \Omega \mid X(\omega) \cap U \neq \emptyset\} \]

are measurable for every open set \(U \subset E\). This is equivalent to \(X\) being \(\tau_W(d)\)-Borel, so \(X\) is a random element of \(\mathcal{F}\) in the sense above and the definitions in the last paragraphs apply also to random closed sets.

We denote by \(\text{Sel}(X)\) the family of all distributions of selections of \(X\). It is well known that \(\text{Sel}(X)\) depends not only on \(L(X)\) but also on the atomic structure of \((\Omega, \mathcal{A})\). The measurable selection theorem ensures that \(\text{Sel}(X)\) is nonempty. If \(\xi\) is a selection of \(X\), we write \(\xi \in X\).

A separable topological space is called Polish if it admits a complete metric. The continuous image of a Polish space, if it is \(T_2\), is called Suslin. Any \(T_2\) topology coarser than a Polish topology is Suslin, since it makes the identity mapping continuous.

### 3. Proof of Theorem 1

**Proof.** We begin by observing the following facts:

1. A separable metric \(d\) admits an equivalent metric \(d'\) which is totally bounded and majorized by \(d\).
2. For a totally bounded metric, the Wijsman and Hausdorff topologies coincide [11].
3. The Wijsman topology of a separable metric admitting an equivalent complete metric is Polish [3].
4. Since \(d' \leq d\), by definition \(\tau_W(d')\) is coarser than \(\tau_W(d)\).
5. The Borel \(\sigma\)-algebras of two comparable Suslin topologies are equal [9].
6. An \(\mathcal{F}\)-valued mapping is a random closed set if and only if it is \(\tau_W(d)\)-Borel (Hess’s measurability theorem, e.g. [3]).
7. For any weakly convergent sequence of probability measures \(P_n \to P\) in a complete separable metric space, there exist random elements \(\xi_n, \xi\) defined on \([0, 1]\), with laws \(L(\xi_n) = P_n\) and \(L(\xi) = P\), such that \(\xi_n \to \xi\) pointwise (Skorokhod’s representation theorem [20]).

The consequence is that \(d'_H\) metrizes the Polish topology \(\tau_W(d')\), with respect to which \(X_n, X\) are Borel and \(X_n \to X\) in distribution. Thus, the Skorokhod representation theorem yields random closed sets \(Y_n, Y\) defined on \([0, 1]\) such that \(L(Y_n) = L(X_n), L(Y) = L(X)\) and \(Y_n(t) \to Y(t)\) in \(\tau_W(d')\) for every \(t \in [0, 1]\).

From the eighth fact:

8. If \(X\) is a random closed set defined on a nonatomic space, then \(\text{core}(\nu) = \text{cl Sel}(X)\) (from [16 Corollary 4.8, part (3)]),

we have \(\text{core}(\nu) = \text{cl Sel}(Y), \text{core}(\nu_n) = \text{cl Sel}(Y_n)\).

In order to prove the ‘lower’ part of the definition of \(\tau_V\)-convergence, it suffices to show that every element of \(\text{cl Sel}(Y)\) is the limit of a sequence of distributions taken from \(\text{cl Sel}(Y_n)\). Since the weak topology is metrizable, every element of \(\text{cl Sel}(Y)\)
is the limit of a sequence in Sel(Y), and we will be done if we prove that every $P \in \text{Sel}(Y)$ is the weak limit of a sequence $P_n \in \text{Sel}(Y_n)$.

Let $\xi$ be a selection of $Y$ with $L(\xi) = P$, and define

$$Z_n = Y_n \cap B(\xi, d'(\xi, Y_n) + n^{-1}),$$

which are random closed sets. By the measurable selection theorem, there exist selections $\xi_n \in Z_n \subset Y_n$. By construction, we have

$$d'(\xi_n, \xi) \leq d'(\xi, Y_n) + n^{-1}.$$  

For each $t \in [0, 1]$, we know

$$d'(x, Y_n(t)) \to d'(x, Y(t))$$

for every $x \in E$;

in particular, taking $x = \xi(t)$,

$$d'((\xi(t)), Y_n(t)) \to d'((\xi(t)), Y(t)) = 0$$

for almost every $t$. We conclude that $\xi_n \to \xi$ a.s., whence the sequence $P_n = L(\xi_n)$ converges weakly to $P$.

For the proof of the ‘upper’ part of $\tau_V$-convergence, consider the following facts:

(9) $\tau_V$ is the supremum of all Wijsman topologies over equivalent metrics [4].
(10) For any metric $\delta$, the proximal topology $\tau_P(\delta)$ is the supremum of all Wijsman topologies over equivalent metrics [3].

Hence, fix a metrization $\delta$ of weak convergence in the space of probability measures in $E$. Due to the arbitrariness of $\delta$, it is simpler for us to prove $\tau_P(\delta)$-convergence than $\tau_V(\delta)$-convergence (the latter involving the exact value of $\delta$ and not just whether the infimum is zero or nonzero).

To achieve proximal convergence, we need to ‘upgrade’ the convergence $Y_n \to Y$ from $\tau_V(d')$ to $\tau_P(d')$. But the Wijsman topology is coarser than the Hausdorff topology, which is invariant under uniformly equivalent metrizations; in view of fact (10), in any metric space the proximal topology is coarser than the Hausdorff topology. Then the identity $\tau_V(d') = \tau_H(d')$ yields $\tau_P(d') = \tau_V(d')$.

Let $Q$ be a weakly closed family of probability measures in $E$. We have to prove that $\text{cl}({\nu_n})$ is far (in the metric $\delta$) from $Q$ whenever $\text{cl}({\nu})$ is so and $n$ is large enough. Assume the contrary, namely that there exists a sequence $\{n_k\}$ such that $\text{cl}({\nu_{n_k}})$ is not far from $Q$, and let us prove that $\text{cl}({\nu})$ is not far from $Q$.

Since two sets are far if and only if their closures are far, $\text{Sel}(Y_{n_k})$ is not far from $Q$. Therefore, there exists a sequence of probability measures $P_{n_k} \in \text{Sel}(Y_{n_k})$ such that $\delta(P_{n_k}, Q) \to 0$. Take selections $\xi_{n_k} \in Y_{n_k}$ with law $P_{n_k}$, and define

$$Z_{n_k} = Y \cap B(\xi_{n_k}, d'(\xi_{n_k}, Y) + n_k^{-1}).$$

Using the measurable selection theorem again, we obtain selections $\eta_{n_k} \in Z_{n_k} \subset Y$, and

$$d'((\xi_{n_k}, \eta_{n_k}) \leq d'(\xi_{n_k}, Y) + n_k^{-1}.$$  

Observe that for any fixed $t \in [0, 1]$ and for a fixed $\epsilon > 0$, the set $Y(t)$ is far from the set

$$\text{cl}(Y(t)^c) = \text{cl}\{x \in E \mid d'(x, Y(t)) > \epsilon\}.$$  

Then, the $\tau_P(d')$-convergence $Y_{n_k}(t) \to Y(t)$ implies that $Y_{n_k}(t)$ is eventually far from $\text{cl}(Y(t)^c)$. Consequently,

$$\xi_{n_k}(t) \in Y_{n_k}(t) \subset \text{cl}(Y(t)^c)^c = \text{int} Y(t)^c,$$
and then

\[ d'(\xi_n(t), Y(t)) \leq \varepsilon \]

for almost every \( t \). But this proves \( d'(\xi_n, Y) \rightarrow 0 \) a.s., whence also \( d'(\xi_n, \eta_n) \rightarrow 0 \) a.s.

Accordingly,

\[
\inf \{ \delta(P, Q) \mid P \in \text{Sel}(Y), Q \in Q \} \leq \delta(\mathcal{L}(\eta_n), Q) \\
\leq \delta(\mathcal{L}(\eta_n), P_n) + \delta(P_n, Q) \rightarrow 0,
\]

implying that \( \text{Sel}(Y) \), and hence \( \overline{\text{core}(\nu)} \), is not far from \( Q \).

Since this is valid for any metric \( \delta \) generating the weak topology, the proof is complete in view of facts (9) and (10).

\[ \square \]

4. Comparison to former work

The former literature on this topic is widely heterogeneous as regards the modes of convergence and the assumptions placed on the carrier space or the random sets. Due to this variation, we find it convenient to dedicate this section to making a careful comparison. Whenever conditions on the space or the random sets are not specified, it should be understood that \( \mathbb{E} \) is meant to be a complete separable metric space and \( X_n, X \) random closed sets.

As mentioned in the introduction, Artstein [1, Theorem 5.4] was the first to prove a result of this kind. He did so for random compact sets, assuming \( X_n \rightarrow X \) in distribution in the Hausdorff metric and obtaining convergence of the cores in the Hausdorff metric associated to the Prokhorov metric.

Let us show that our result reduces to Artstein’s when it is assumed that the random sets are compact. Since \( \tau_H = \tau_V \) on compact sets [15], the conclusions of both theorems are the same. Incidentally, the fact that the Vietoris topology is independent of the metric implies that the Prokhorov metric could be replaced by any other metrization of weak convergence. Our assumption is formally but not actually weaker than Artstein’s: since all Hausdorff metrics are equivalent on compact sets, from facts (1), (2) and (4) in our proof we see that

\[ \tau_W(d) \supset \tau_W(d') = \tau_H(d') = \tau_H(d) \supset \tau_W(d) \]

on compact sets, so both assumptions are actually the same.

The paper [21] by Sun uses nonstandard models and the theory of Loeb measures to study the selectionable distributions of multifunctions (or correspondences). On the one hand, many of his results do not assume closedness of the values, which we need in view of the importance of topologies on closed sets to our approach (the discussion below is restricted to the closed-values case). But on the other hand, the use of ZFC models with specific properties and the use of Loeb measure spaces implies that we ignore whether the results also hold for other models or other measure spaces. (In fact, Sun provides examples that mappings on \([0,1]\) with its Lebesgue measurable structure may fail the theorems obtained for Loeb spaces!)

The combination of Propositions 3.12 and 3.14 in [21] with fact (8) in our proof yields the following: If \( X_n, X \) are random compact sets defined on a Loeb measure space and \( X_n \rightarrow X \) almost surely in the sense of Painlevé-Kuratowski, then \( \text{core}(\nu_n) \rightarrow \text{core}(\nu) \) in the sense of Painlevé-Kuratowski. Almost sure convergence cannot be weakened to convergence in distribution using the Skorokhod representation, because in general Painlevé-Kuratowski is not topological. That can be
A CONTINUITY THEOREM FOR RANDOM CLOSED SETS 4423

done if $E$ is additionally assumed to be locally compact, in which case the space of probability measures is so also. Then, Painlevé-Kuratowski convergence in both the hypothesis and the conclusion coincides with Wijsman convergence (see e.g. [14]) and a (weakened) version of Artstein’s result is obtained.

It is worth commenting that [21] Proposition 3.14 is valid in an arbitrary probability space and for random closed sets. It claims that, if $X = \liminf_n X_n$, then $\text{core}(\nu) \subset \liminf_n \text{core}(\nu_n)$. Note that $X \subset \liminf_n X_n$ is the same as $X_n \to X$ (pointwise) in the lower Vietoris topology $\tau_V$. Thus an easy extension of [21] Proposition 3.14 says that

$$X_n \to X \text{ a.s. in } \tau_V \Rightarrow \text{core}(\nu) \subset \liminf_n \text{core}(\nu_n)$$

(the right-hand side being the same thing as $\text{core}(\nu_n) \to \text{core}(\nu)$ in $\tau_V$) whereas what was proved in the ‘lower’ part of the proof of Theorem [1.1] is tantamount to

$$X_n \to X \text{ in distribution in } \tau_W(d) \Rightarrow \text{core}(\nu) \subset \liminf_n \text{core}(\nu_n),$$

where the comparison is easier if we regard $\tau_W(d)$ as the supremum of the lower Vietoris and upper Wijsman topologies (see e.g. [12]). Again Skorokhod techniques are not useful to weaken a.s. convergence in the first statement because the topology $\tau_V$ fails the $T_2$ axiom.

Proposition 3.15 in [21] contains two one-sided statements whose conjunction is as follows: If $d_H(X_n, X) \to 0$ in probability, then $\text{core}(\nu_n) \to \text{core}(\nu)$ in the metric $\delta_H$, where $\delta$ is the metric given by

$$\delta(P, Q) = \sum_{m=1}^{\infty} 2^{-m} | \int \phi_m dP - \int \phi_m dQ |$$

and $\{ \phi_m \}_m$ is the sequence of continuous functions given in [21] p. 87 (this sequence is rather special and possibly cannot be replaced by an arbitrary dense sequence).

Although that result applies in the unbounded case, it seems that the assumption is too strong unless $X_n, X$ are taken to be at least bounded. As an exception to this, it may happen that one is interested in applying only one of Sun’s one-sided statements. That is actually the case in some situations which appear in stochastic programming; see e.g. [23].

The paper [8] by Feng and Feng assumes that $E$ is a compact metric space. According to its Corollary 3.4, if $\nu_n \to \nu$ in the pseudometric $\Lambda$ defined in [8] p.92], then $\text{core}(\nu_n) \to \text{core}(\nu)$ in $\Delta_H$, where $\Delta$ is the metrization of weak convergence defined by

$$\Delta(P, Q) = \sum_{m=1}^{\infty} 2^{-m} ||f_m||^2 \int | f_m dP - f_m dQ |$$

and $\{ f_m \}_m$ is a sup-norm dense sequence of continuous functions on $E$. As a matter of fact, $\Lambda$ is actually a metric: two capacities at $\Lambda$-distance 0 are equal on compact sets, and capacitability arguments show that then they must be equal on Borel sets.

But, as shown by the authors, for probability measures (seen as the capacity functionals of random singletons) $\Lambda$-convergence is stronger than weak convergence, whence the hypothesis that $\Lambda(\nu_n, \nu) \to 0$ is not weaker than $X_n \to X$ in distribution for any of the topologies $\tau_W, \tau_P, \tau_H, \tau_V$ we have considered (all of which agree with ordinary convergence when restricted to singletons). The conclusion is
that [8, Corollary 3.4], even in the case of $E$ compact, does not improve [1, Theorem 5.4]. However, notice that [8, Theorem 3.3] provides an explicit upper estimate of $\Delta_H(\text{core}(\nu_n), \text{core}(\nu))$.

Regarding the paper [7] by Feng and Nguyen, its Theorem 1 can be rephrased, using its proof’s first lines, [18, Theorem 4.6] and the identity of the hit-and-miss and Hausdorff topologies in compact spaces (e.g. [14]) as follows: Let $E$ be a compact subset of $\mathbb{R}^d$ and $X_n \rightarrow X$ in distribution in $\tau_H$. Then $\rho_H(\text{core}(\nu_n), \text{core}(\nu)) \rightarrow 0$. Thus it follows from Artstein’s result.

Feng and Nguyen also prove the converse (convergence of cores implies weak convergence). Our result suggests that the situation in the general case might be more complicated.

5. CONCLUDING REMARKS

If the random sets $X_n, X$ are compact, the completeness assumption on $E$ can be dropped. That is done by using Dudley’s version of the Skorokhod representation in a separable metric space [6] and Sion’s version of the measurable selection theorem [19], valid for random compact sets in a hereditarily Lindelöf regular $T_2$ space with weight at most $\aleph_1$ (in particular, a separable metric space).

REFERENCES


Facultad de Ciencias Económicas y Empresariales, Grupo Decisión Multicriterio Zaragoza, Universidad de Zaragoza, Gran Vía 2, E-50005 Zaragoza, Spain

Current address: Miguel Servet 2, 2ºD, E-33213, Gijón, España

E-mail address: teran@unizar.es

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use