The Nonstraddling Lemma and a new look at the Heine-Borel Theorem

Philip G Spain

Abstract

Any family of intervals in the real line determines a *linking equivalence* relation on its union. The equivalence classes are *order convex* so are therefore themselves intervals, partitioning the union of the original family.

If one starts from a cover of a bounded closed interval by open intervals one can then apply the *Nonstraddling Lemma*, a result of the utmost simplicity, to clinch the proof of the *Heine-Borel Theorem*.

The *Structure Theorem for Open Sets* emerges naturally from this discussion.

The natural setting for this scheme of proof is in the wider context of *complete linearly ordered spaces* where the generalisation of the *Heine-Borel Theorem* is known as the *Haar-König Theorem*.

Linearly Ordered Spaces

Although the title of this note refers to the *Heine-Borel Theorem*, whose setting is the real line $\mathbf{R}$, the discussion does not rely on the fact that the real line is *gapless* (or *dense-in-itself*). I therefore present the material in the context of *complete linearly ordered spaces*.

> *The reader who prefers to abstain from this generality may in what follows: replace* $\mathbf{E}$ *by* $\mathbf{R}$*, 'join' by 'intersect', and 'separate' by 'disjoint'; and may then skip over the gaps.*

A *linearly ordered* space is a set $\mathbf{E}$ with a *partial order* $<$ such that any two elements of $\mathbf{E}$ are *comparable*.

Intervals & Gaps

I employ the usual terminology for intervals in a linearly ordered space $(\mathbf{E}, <)$. *Left* (or *right*) *open* [or *closed*] signifies the *omission* [or *inclusion*] of the *left* (or *right*) endpoint; and use (,), [, and] conventionally to make these distinctions explicit. (It might have been helpful if standard nomenclature referred to *defining* endpoints, as I shall do.)

Whether $\mathbf{E}$ has *extreme elements*, or not, I will follow the convention for $\mathbf{R}$ and write $-\infty$ and $+\infty$ to denote *points at infinity*, placeholders to allow one to describe open and closed half-lines in interval notation: $\genfrac{}{}{0pt}{}{(}{[} a, +\infty)$ denotes a left $\genfrac{}{}{0pt}{}{\textit{open}}{\textit{closed}}$ interval, unbounded above: and similarly for intervals unbounded below.

Convention. *I reserve the term* open interval *for an interval presented in 'open bracket' form: that is, an interval defined by endpoints that it* does not contain.
Similarly, a closed interval *is an interval presented in 'closed bracket' form: that is, an interval defined by endpoints that it* does contain. *Closed intervals, in this sense, are automatically bounded.*

Gaps

Definition 1. *If r and t in a linearly ordered space $(\mathbf{E}, <)$ are such that $r < t$ and $(r, t) = \emptyset$, then (r, t) is a* gap.
Then an interval that is *right closed* at r is simultaneously *right open* at t.
Indeed, an interval is simultaneously open and closed at an end precisely when there is a gap at this end.

Also, the union of two disjoint open intervals can again be an interval. In $\mathbf{Z}$, for example, $[1, 2] = (0, 3) = \{1, 2\} = (0, 2) \cup (1, 3)$, these last two open intervals being disjoint.

Definition 2. *The space* $\mathbf{E}$ *is* gapless *(or* dense-in-itself*) when between any two distinct elements there lies yet another.*

In a space that may have gaps the appropriate dichotomy is *joining* $\longleftrightarrow$ *separated* rather than *intersecting* $\longleftrightarrow$ *disjoint* .

Definition 3. *Two intervals I and J in the linearly ordered space* $(\mathbf{E}, <)$

$$\begin{matrix} \textit{join} \\ \textit{are separated} \end{matrix} \qquad \textit{if} \quad I \cup J \quad \begin{matrix} \textit{is} \\ \textit{is not} \end{matrix} \quad \textit{an interval.}$$

In any gapless space, as in $\mathbf{R}$, two open intervals *join* precisely when they *intersect*, and are *separated* precisely when they are *disjoint*.

Remark 1. *If two intervals intersect, or, less restrictively, if a defining endpoint of one belongs to the other, then these intervals join.*
If (r, t) is a gap and if I and J are intervals containing r and t respectively, then I and J join.

How can two disjoint intervals join?

This is impossible for two closed intervals. Two *abutting* intervals [where a single point is a defining endpoint for intervals on either side of it, with one open, the other closed, there] are disjoint yet join.
Otherwise, the critical case, consider a pair of open intervals $\kappa = (r, t)$ and $\lambda = (u, w)$, neither of which is empty.
There are 3 possibilities with $r \leq u$, that is, when '*κ is to the left of λ*':

1. $t \leq u$. Here κ and λ are separated: for $\kappa \cup \lambda = (r, w) \setminus [t, u]$.
2. $r \leq u < w \leq t$. Here $\kappa \cap \lambda = \lambda$.
3. $r \leq u < t < w$. Here $\kappa \cap \lambda = (u, t)$.

There are another similar 3 possibilities when $w \leq r$. Putting these together we find:

Remark 2. *Two open intervals $\kappa = (r, t)$ and $\lambda = (u, w)$, neither empty, are simultaneously* disjoint *and* joined *when they* intertwine minimally across a gap*. That is,* either *when $r < u < t < w$ and (u, t) is a gap* or *when $u < r < w < t$ and (r, w) is a gap.*

Endpoints and Straddling

An interval I *straddles* two disjoint sets E, F if both $I \cap E \neq \emptyset$ and $I \cap F \neq \emptyset$.

The Nonstraddling Lemma, which follows easily from the No Endpoints Lemma, tells us that an interval cannot straddle [*informally*: split itself across] a family of pairwise separated open intervals if wholly contained in their union.

Lemma 1 (No Endpoints Lemma)**.** *The union of a family $\mathcal{O}$ of pairwise separated open intervals in a linearly ordered space $\mathbf{E}$ does not contain the defining endpoints of any of these intervals.*

Proof. Immediate from Remark 1. □

Lemma 2 (Nonstraddling Lemma)**.** *If $\mathcal{O}$ is a family of pairwise separated open intervals in a linearly ordered space $\mathbf{E}$, and I is an interval in $\mathbf{E}$ such that $I \subseteq \bigcup \mathcal{O}$, then $I \subseteq O$ for some $O \in \mathcal{O}$.*

Proof. Consider an $O \in \mathcal{O}$ that intersects I. All points of I belong to $\bigcup \mathcal{O}$ and therefore cannot be endpoints of any interval in $\mathcal{O}$ (by Lemma 1). In particular, the defining endpoints of O do not lie inside I, which means that $I \subseteq O$. □

Corollary 1 (Uniqueness of pairwise separated open partitions)**.** *If $\mathcal{O}$ and $\mathcal{Q}$ are families of pairwise separated open intervals in a linearly ordered space $\mathbf{E}$, and $\bigcup \mathcal{O} = \bigcup \mathcal{Q}$, then $\mathcal{O} = \mathcal{Q}$.* □

Order Convexity & Completeness

Definition 4 (Order convexity)**.** *A subset E of of a linearly ordered space $\mathbf{E}$ is* order convex *if $[r,\ t] \subseteq E$ whenever $r,\ t \in E,\ r \leq t$.* Intervals are clearly order convex.

A linearly ordered space $\mathbf{E}$ is *complete* if every set with an upper bound has a *supremum*, and, equivalently, every set with a lower bound has an *infimum.* This form of completeness is sometimes referred to as *conditional* (or *order* or *Dedekind*) *completeness.* Recall that M is the *supremum* of a set E if (**i**) $e \leq M$ ($\forall e \in E$), and (**ii**) $M \leq N$ whenever $e \leq N$ ($\forall e \in E$): and similarly for an *infimum*.

Intervals are the only order convex sets in complete linearly ordered spaces, a fact not often mentioned in the literature: though see, for instance, [Bou, IV.2.4 Proposition 1] and [SS, Counterexample 39].

Proposition 1 (Characterization of Order Convex Sets)**.** *A subset of a complete linearly ordered space* $\mathbf{E}$ *is order convex* (*if and*) *only if it is an interval.*

Proof. Consider a nonempty order convex subset E of $\mathbf{E}$: define $\substack{M\\m} = \substack{sup\\inf}\ E$ $\big($with the usual conventions governing $\substack{M\\m} = \substack{+\\-}\infty\,\big)$: the existence of m and M is guaranteed by the completeness of $\mathbf{E}$. Then $m = M$ only when E is a singleton, and so an interval: thus, without loss of generality, we may assume $m < M$. If $s \in (m,\ M)$ then $\exists\ \substack{t\\r} \in \substack{[s,\,M]\\[m,\,s]} \bigcap E$ $\big($by the definition of $\substack{sup\\inf}\big)$. Hence $s \in [r, t] \subseteq E$ and it follows that $E\backslash\{m, M\} = (m, M)$. $\square$

This is the only place where we use completeness. Proposition 1 does not extend to spaces that are not complete. For example, in $\mathbf{Q}$ (the rationals) the order convex set $\{q \in \mathbf{Q} : q^2 < 2\}$ is not an interval (even though it is the intersection of $\mathbf{Q}$ with an interval in $\mathbf{R}$).

Moreover, if $\mathbf{E}$ is a linearly ordered space with the property that every order convex subset is an interval, then $\mathbf{E}$ is complete. For if B is bounded above in $\mathbf{E}$ then its set of upper bounds $\mathcal{M}$ is order convex (and extends to $+\infty$): the left endpoint of $\mathcal{M}$ is the supremum of B.

$\mathcal{I}$-linking

Any family of intervals determines an equivalence relation on its union.

Definition 5. *Given a family* $\mathcal{I}$ *of intervals in linearly ordered space* $\mathbf{E}$ *a (n* $\mathcal{I}-$*)stretch is an interval which is the union of a finite number of intervals from* $\mathcal{I}$*.*
Two points $\substack{r\\t}$ *are* $\mathcal{I}$*-linked* $\big[$*write* $r \sim_{\mathcal{I}} t\big]$ *if there is an* $\mathcal{I}$*-stretch that contains them both.*

Lemma 3. *Let* $\mathcal{I}$ *be a family of intervals in a linearly ordered space* $\mathbf{E}$*. Then* $\sim_{\mathcal{I}}$ *is an equivalence relation on* $\bigcup \mathcal{I}$*, and the equivalence classes are order convex.*

Proof. Reflexivity and symmetry are clear. As to transitivity: if $r \sim_{\mathcal{I}} s \sim_{\mathcal{I}} t$ then there are $\mathcal{I}$-stretches $P = \bigcup \mathcal{P}$ and $Q = \bigcup \mathcal{Q}$ such that $\substack{r\\s} \in P$ and $\substack{s\\t} \in Q$. Then $P \cup Q = \bigcup \mathcal{P} \cup \bigcup \mathcal{Q}$ is a stretch that contains both $\substack{r\\t}$.

Two points in the same equivalence class belong to a common stretch, and stretches, being intervals, are order convex. $\square$

Remark 3. *If an* I ($\in \mathcal{I}$) intersects *an equivalence class* κ *then* $I \subseteq \kappa$*:* for then every point in I is linked to every point in κ.

Theorem 1 (Linking equivalence decomposition)**.** *Let* $\mathcal{I}$ *be a family of intervals in a complete linearly ordered space* $\mathbf{E}$*. Then the* $\mathcal{I}$*-equivalence classes are intervals.*
If, further, $\mathcal{I}$ *is a family of* open *intervals, then all the* $\mathcal{I}$*-equivalence classes are open intervals; and they are pairwise separated.*

Proof. *Disjoint intervals*:
The equivalence classes are order convex, by Lemma 3, so, by Proposition 1, they are intervals. And, automatically, being equivalence classes, they are pairwise disjoint.

Openness:
Suppose, further, that all the intervals in $\mathcal{I}$ are open. Consider a defining endpoint s of a class κ. If $s \in \kappa$ then there is an $I \in \mathcal{I}$ such that $s \in I$. Then $I \subseteq \kappa$, by Remark 3, and κ is open and closed at its s-end, which can happen only if $\mathbf{E}$ has a gap at s. *When* $\mathbf{E}$ *is gapless this is impossible and the proof is complete.*

Separatedness [when $\mathbf{E}$ *may have gaps]*:
Consider two distinct equivalence classes $\kappa = (r, t)$, and $\lambda = (u, w)$, and suppose that they join. Then, following Remark 2, without loss of generality we may assume that $r < u < t < w$ and that (u, t) is a gap. Now $u \in \kappa$ so there is an $I \in \mathcal{I}$ such that $u \in I$. Similarly $t \in \lambda$ so there is a $J \in \mathcal{I}$ such that $t \in J$. Then I and J join (Remark 1). Thus $\kappa \ni u \sim_{\mathcal{I}} t \in \lambda$ and therefore $\kappa = \lambda$ — a contradiction. □

The Heine-Borel Theorem

The extension of the Heine-Borel Theorem to complete linearly ordered spaces is usually ascribed to Haar and König.

Recall: *A family of sets* $\mathcal{E}$ *covers a set* E *if* $E \subseteq \bigcup \mathcal{E}$.

Theorem 2 (Heine-Borel Haar-König Theorem — Intervals)**.** *Let* K *be a closed interval in a complete linearly ordered space* $\mathbf{E}$*, and let* $\mathcal{I}$ *be a cover of* K *by open intervals. Then* K *is covered by a finite subfamily of* $\mathcal{I}$*.*

Proof. By Theorem 1, we can express $\bigcup \mathcal{I}$ in the form $\bigcup \kappa$, where the κ, the $\sim_{\mathcal{I}}$-equivalence classes, are pairwise separated open intervals. Then $K \subseteq \kappa$ for some κ, by the Nonstraddling Lemma. Thus the defining endpoints of K are linked by an $\mathcal{I}$-stretch, a finite subfamily of $\mathcal{I}$: hence K is covered by a finite subfamily of $\mathcal{I}$. □
There is no useful conclusion to be drawn concerning open coverings of an interval that does not contain its defining endpoints.

Order topology

Up to now our discussion has not required any concepts from general topology — just openness and closedness for intervals.

Open sets

The natural *order topology* on a linearly ordered space $(\mathbf{E}, <)$ is the topology generated by the intervals $(-\infty, e)$ and (e, ∞) as e ranges through $\mathbf{E}$ (that is, the topology for which these sets are a subbasis): a set is open in $\mathbf{E}$ if and only if it is the union of all the open intervals it contains.

The characterisation of open sets is a direct consequence of Theorem 1.

Theorem 3 (Structure Theorem for Open Sets)**.** *Let U be an open subset in a complete linearly ordered space* $\mathbf{E}$*. Then U is the union of a uniquely determined family of pairwise separated open intervals. This family is countable if* $\mathbf{E}$ *is* separable*.*

Proof. Let $\mathcal{I}$ be some family of open intervals such that $U = \bigcup \mathcal{I}$. By Theorem 1, U is the union of a family of pairwise separated open intervals. By Corollary 1 such an expression for U is unique, so independent of the family $\mathcal{I}$. Further, any separated family of intervals with nonempty interior is countable if $\mathbf{E}$ has a countable dense subset. □

Closed sets & Compact sets

Recall: *A set is* closed *when its complement is open.*
A set is compact *when every open cover of it can be reduced to a finite subcover.*

The generalisation of Theorem 2 to open covers of closed bounded sets is routine.

Theorem 4 (Heine-Borel Haar-König Theorem)**.** *Let K be a closed bounded set in a complete linearly ordered space* $\mathbf{E}$*. Then K is compact.*

Proof. First, K has an infimum and a supremum in $\mathbf{E}$ (because K is bounded), and both of these belong to K (because K is closed).
Consider a cover $\mathcal{U}$ of K by open sets. Express each $V \in \mathcal{U} \cup \{\mathbf{E} \setminus K\}$ as a union of open intervals. These together form a family $\mathcal{I}$ that covers $\mathbf{E}$. Then, by the Nonstraddling Lemma, the closed interval $[\inf K, \sup K]$ lies in some $\mathcal{I}$-equivalence class. Thus K is covered by a finite subfamily of $\mathcal{I}$, so, after discarding any $\mathcal{I}$-subintervals of $\mathbf{E} \setminus K$, and choosing a superset in $\mathcal{U}$ for each of the remaining $\mathcal{I}$-sets, we have a cover of K by a finite subfamily of $\mathcal{U}$. □

Connectedness

Recall that a topological space is *disconnected* if it is the union of two nonempty disjoint open subsets: otherwise it is *connected.*

$\mathbf{R}$ *is connected.* For, consider two nonempty disjoint open subsets U and V of $\mathbf{R}$. Then $U = \bigcup \mathcal{I}$ and $V = \bigcup \mathcal{J}$ for two families $\mathcal{I}$ and $\mathcal{J}$ of pairwise disjoint open intervals. Now, by Lemma 1, the family of pairwise disjoint open intervals $\mathcal{I} \bigcup \mathcal{J}$ cannot contain the defining endpoints of its intervals — so $U \cup V \neq \mathbf{R}$.

In fact, a linearly ordered space is connected if and only if it is complete and has no gaps ([Ke, Chapter 1, Problem I]). The essential point: if B is bounded above and has no supremum then both $\bigcup\{(-\infty, b) : b \in B\}$ and $\mathcal{M}$, the set of upper bounds of B, are open and complementary in $\mathbf{E}$.

A little history

It is now more than 50 years since Edwin Hewitt published his magisterial outline of the rôle of compactness in analysis [H], which still repays reading. The unifying theme was that many propositions of analysis are:

> trivial for finite sets;
>
> true and reasonably simple for infinite compact sets (subject to obviously necessary additional hypotheses);
>
> either false or extremely difficult to prove for noncompact sets.

Although Hewitt discussed many examples, more and less elementary, emphasizing the sweeping out of a local property to a finite union of intervals, and then to a compact interval, he did not present a proof of the Heine-Borel Theorem. Its evolution is outlined in [Dugac], as reported in [MacTutor].

> The first proof of this theorem was given by Dirichlet in his lectures of 1862 (published 1904) before Heine proved it in 1872. Dugac shows that Dirichlet used the idea of a covering and a finite subcovering more explicitly than Heine. This idea was also used by Weierstrass and Pincherle. Borel formulated his theorem for countable coverings in 1895 and Schönflies and Lebesgue generalized it to any type of covering in 1900 and 1898 (published 1904), respectively. Dugac shows that the story is in fact much more complicated and includes names such as Cousin, Thomae, Young, Vieillefond, Lindelöf. The priority questions are nicely illustrated with quotes from the correspondence between Lebesgue and Borel and other letters.

The Heine-Borel Theorem did not appear in the First Edition [GHH1] of Hardy's influential *Pure Mathematics*, but was included in the Second Edition [GHH2]. Indeed, the First Edition does not even include a description of the construction of the real numbers.

Another treatment of the Heine-Borel Theorem, based on a combinatorial result, König's Infinity Lemma (the same Dénes König as of [HK]), and the standard ultrametric on the space of binary sequences, has appeared recently [MRR].

There is a faint hint of the present treatment (the use of a *transitive* relation) in the 'creeping lemma' of [MR]. See also [Ka] for another approach. (I am indebted to the referee of an earlier version of this note for this last reference.)

Acknowledgment

I thank Professor Earl Berkson for his perspicacious comments on an earlier version of this note.

References

[Bou] N. Bourbaki, *General Topology. Part 1*, Addison-Wesley 1996.

[Dugac] P. Dugac, *Sur la correspondance de Borel et le théorème de Dirichlet–Heine–Weierstrass–Borel–Schoenflies–Lebesgue*, Arch. Internat. Hist. Sci. **39** (122) (1989) 69–110. MR 1092039.

[HK] A. Haar and D. König, *Über einfach geordnete Mengen*, Crelle's J. **139** (1910) 16–28.

[GHH1] G.H. Hardy, *A Course of Pure Mathematics*, First Edition, Cambridge University Press 1908.

[GHH2] G.H. Hardy, *A Course of Pure Mathematics*, Second Edition, Cambridge University Press 1914.

[H] E. Hewitt, *The Rôle of Compactness in Analysis*, American Math. Monthly **67** (1960) 499–516. MR 0230853.

[Ka] I. Kalantari, *Induction over the continuum*, in *Induction, Algorithmic Learning Theory, and Philosophy*, M. Friend, N.B. Goethe and V.S. Harizanov (editors), 145–154. Springer, Dordrecht 2007. MR 2377000.

[Ke] J.L. Kelley, *General Topology*, Van Nostrand 1955. MR 0070144.

[MacTutor] //www-history.mcs.st-andrews.ac.uk .

[MRR] M. MacAuley, B. Rabern and L. Rabern, *A Novel Proof of the Heine-Borel Theorem*, arXiv:0808.0844v1 [math.HO] 6 Aug 2008.

[MR] R.M.F. Moss and G.T. Roberts, *A creeping lemma*, American Math. Monthly **75** (1968) 649–652. MR 120617.

[SS] L.A. Steen and J.A. Seebach, Jr., *Counterexamples in Topology*, Holt, Rinehart and Winston 1970. MR 0266131.

Department of Mathematics and Statistics, University of Glasgow, Glasgow G12 8QW
Philip.Spain@glasgow.ac.uk