THE LEAST UPPER BOUND AXIOM

MATH 153, SECTION 55 (VIPUL NAIK)

Corresponding material in the book: Section 11.1.

What students should already know: The intuitive idea of the real line, basic algebra and calculus as done in the course up to this point for motivation/background.

What students should definitely get: The definitions of least upper bound and greatest lower bound, the definition of boundedness, the statements about the existence of these in the real numbers. Computation of these in special cases.

Executive summary

Words ...

(1) The real numbers satisfy the *least upper bound property*: any nonempty subset of the set of real numbers that is bounded from above has a least upper bound. This property does *not* hold if we replace the real numbers by the rational numbers.

(2) The real numbers satisfy the *greatest lower bound property*: any nonempty subset of the set of real numbers that is bounded from below has a greatest lower bound. This property again does *not* hold if we replace the real numbers by the rational numbers.

(3) We can prove the greatest lower bound property using the least upper bound property. There are two proofs of this. One of these proofs involve *reflection*: replacing a set by its set of negatives. The other proof, which is there in the book, is also worth going through. Please go through it. I'll go through it in review session. You will not be asked the proof in the test, but it may be helpful for multiple choice questions and other conceptually based problems.

(4) The natural numbers satisfy a property that is somewhat similar to the greatest lower bound property for the reals, but stronger: any nonempty subset of the set of natural numbers has a least element. This is equivalent to the *principle of mathematical induction*.

(5) If a nonempty subset of the real numbers has a *maximum* element, then that element is also the least upper bound of the set. Conversely, if the least upper bound of a set is in the set, then that is also the maximum element of the set.

(6) If a nonempty subset of the real numbers has a *minimum* element, then that element is also the greatest lower bound of the set. Conversely, if the greatest lower bound of a set is in the set, then that is also the minimum element of the set.

(7) A nonempty finite subset always has a maximum and a minimum element. Thus, its greatest lower bound and least upper bound are both in the set.

(8) For an interval with lower endpoint a and upper endpoint b, the least upper bound is b and the greatest lower bound is a. Note that this holds for all the four possibilities for the interval: $[a, b]$, (a, b), $[a, b)$, and $(a, b]$.

(9) If T is a nonempty subset of a nonempty bounded subset S of $\mathbb{R}$, any lower bound for S remains a lower bound for T and any upper bound for S remains an upper bound for T. However, we may have an upper bound for T that is *not* an upper bound for S. Similarly, we may have a lower bound for T that is *not* a lower bound for S. Thus, the least upper bound for T is $\leq$ the least upper bound for S, and the greatest lower bound for T is $\geq$ the greatest lower bound for S.

(10) A set does *not* have an upper bound if and only if it has arbitrarily large elements. Similarly, a set does *not* have a lower bound if and only if it has arbitrarily small elements (i.e., negative elements of arbitrarily large magnitude).

(11) If M is the least upper bound of a nonempty subset S of $\mathbb{R}$, then, for every $\epsilon > 0$, S has a nonempty intersection with the interval $(M - \epsilon, M]$. In particular, if $M \notin S$, then S has a nonempty intersection

with the interval $(M - \epsilon, M)$. (See also the analogous theorem for greatest lower bounds, which is Theorem 11.1.4 in the book).

Actions ...

(1) To compute the greatest lower bound and least upper bound of a set, we first need to compute the set. Finding the set as a union of intervals is often useful.
(2) Given a set S, we can construct corresponding sets such as $S + \lambda$ (translation), $-S$ (reflection about 0), $f(S)$ (image of S under a function f), and $\text{abs}(S)$ (the set of absolute values of elements of S, i.e., folding about 0). Please review the results that relate bounds for S with bounds on these corresponding sets.

1. Ordering, boundedness, and related properties

1.1. **A total ordering.** The real numbers, as we have come to understand them, are a *totally ordered set.* This means that for any two real numbers a and b, one of three possibilities holds:

(1) $a < b$: Equivalent formulations are $b > a$, $a - b < 0$, and $b - a > 0$. In words, we say that a is *less than* or *smaller than* b.
(2) $a = b$: Equivalent formulations are $b = a$, $a - b = 0$, and $b - a = 0$.
(3) $a > b$: Equivalent formulations are $b < a$, $b - a < 0$, and $a - b > 0$. In words, we say that a is *greater than* or *bigger than* b.

We also know that if $a < b$ and $b < c$, then $a < c$.

Any two real numbers are *comparable.* We can say for sure that one real number is bigger or smaller than another. In other words, the real numbers can be arranged in increasing order, since they are arranged on a line. This *total ordering* is something very specific to the real numbers and subsets thereof. It breaks down for more complicated mathematical structures, such as the complex numbers, vector spaces over the real numbers, or polynomials. In many of those cases, we can artificially impose an ordering, but it does not behave in a nice and familiar way.

1.2. **The reals, the rationals, and the integers.** There are many subsets of the reals with qualitatively different behavior, but we concentrate on four important subsets that, when studied together, shed light on many of the issues of interest.

(1) The *positive integers* or *natural numbers*, denoted $\mathbb{N}$. These start at 1, and go like $1, 2, 3, \ldots$. A very important feature of the positive integers is that *every nonempty subset of the positive integers has a smallest element.* This can be reformulated as the *principle of mathematical induction*: if a statement is true for 1 and its truth for k implies its truth for $k + 1$, then it is true for all positive integers.
(2) The *integers*, denoted $\mathbb{Z}$. Unlike the positive integers, they do not satisfy the property that every nonempty subset has a least element. The integers stretch to infinity in both directions. However, just like the positive integers, they are *discrete* – for every integer, there is a unique *successor* (a unique smallest integer that's the *next* integer) and a unique *predecessor* (a unique largest integer that's the *previous* integer).
(3) The *rational numbers*, denoted $\mathbb{Q}$. These differ from the integers in the following important respect: they have a *density property.* Between any two rational numbers, there is yet another rational number. And this process can be repeated *ad infinitum*, so between any two rational numbers, there are *infinitely many* rational numbers. Although the rational numbers are *dense*, they are not *complete* – there are holes. We can have a bunch of rational numbers that seem to be heading to some specific number that turns out to be irrational.
(4) The *real numbers*, denoted $\mathbb{R}$. These include the rational numbers and more numbers that can be arrived at by taking limits of rational numbers. The set of real numbers is *complete* in the sense that it has no holes. If a sequence of real numbers is headed somewhere finite, it is headed towards a real number.

We would like to make precise the notion of completeness of the real numbers, i.e., the idea that the real numbers don't have holes. There are many different (and equivalent) ways of doing this. The one we will follow is the *least upper bound axiom*, which we discuss next.

Aside: Sequences and sizes of infinite sets. A *sequence* of real numbers is defined as a function from $\mathbb{N}$ to $\mathbb{R}$. The sequence can be written as a *list*:

$$f(1), f(2), \dots, f(n), \dots$$

Conversely, any list can be thought of as a function:

$$a_1, a_2, \dots, a_n, \dots$$

where the function is defined as $f(n) = a_n$.

An infinite subset of the reals is *countable* if there is a sequence of real numbers that includes all elements in the subset. In other words, we can count, or list, all elements of the subset. It turns out that:

(1) The set of natural numbers (positive integers) is countable. We can just use the sequence $1, 2, 3, \dots$, which corresponds to the identity function on $\mathbb{N}$.
(2) The set of integers is countable. We can use the sequence $0, 1, -1, 2, -2, 3, -3, \dots$. Note that this sequence involves *alternating* between the positive and negative numbers, so it does not preserve the usual ordering.
(3) The set of rational numbers is countable. The way we list the rational numbers does not preserve the usual ordering, and involves a snake-like path.
(4) The set of real numbers is *not* countable. In other words, there is no way to list *all* real numbers.

Thus, in this sense, even though all the four sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ are infinite, $\mathbb{R}$ is a bigger set than the other three. Note that although $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$, the ostensibly bigger and much more dense set $\mathbb{Q}$ is *not* bigger in terms of number of elements than the ostensibly smaller $\mathbb{N}$.

1.3. **Boundedness, upper bounds, and lower bounds.** Suppose S is a nonempty subset of the real numbers. An element $a \in \mathbb{R}$ is termed an *upper bound* for S if $x \le a$ for all $x \in S$. In particular, we allow the upper bound to be in the set. A nonempty subset of $\mathbb{R}$ that has an upper bound is said to be *bounded from above*.

An element $b \in \mathbb{R}$ is a *lower bound* for S if $b \le x$ for all $x \in S$. In particular, we allow the lower bound to be in the set. A nonempty subset of $\mathbb{R}$ that has a lower bound is said to be *bounded from below*.

A nonempty subset of $\mathbb{R}$ is said to be *bounded* if it is both bounded from above and bounded from below. Some comments:

(1) a is an upper bound for S if and only if $S \subseteq (-\infty, a]$.
(2) b is a lower bound for S if and only if $S \subseteq [b, \infty)$.
(3) a is an upper bound for S and b is a lower bound for S if and only if $S \subseteq [b, a]$.

1.4. **Least upper bound and greatest lower bound.** Suppose S is a nonempty subset. An element $a \in \mathbb{R}$ is termed a *least upper bound* for S if a is an upper bound for S and, for any upper bound a' of S, $a \le a'$. In other words, a is the *least* possible element we can choose as an upper bound.

We can make two petty observations without much thought:

(1) For a least upper bound to exist, the set must be bounded from above.
(2) If a and a' are both least upper bounds for a nonempty subset S, then $a = a'$. This is because each one is less than or equal to the other.

Now, whatever we have said so far applies inside the real numbers, but it also applies inside the rational numbers, inside the integers, and inside the positive integers. In fact, it applies more abstractly inside any totally ordered set. However, what we are going to say *now* is something very specific to the real numbers, in so far as it captures a *completeness* property of the real numbers. This says that:

> Any subset of the real numbers that is bounded from above has a least upper bound *which is also a real number*.

As already mentioned in the two points made above, the "bounded from above" condition is clearly *necessary*, and also, the least upper bound must be *unique*. Thus, for a subset S of $\mathbb{R}$ that is bounded from above, we denote the least upper bound of S by $\text{lub}(S)$.

In a similar vein, we can define the notion of *greatest lower bound*, and we have the following:

> Any subset of the real numbers that is bounded from below has a greatest lower bound.

Similar to the previous case, the "bounded from below" condition is *necessary*, and the greatest lower bound must be *unique*. We denote this by $\text{glb}(S)$.

1.5. **Pedestrian observations.**

(1) The following are equivalent for a nonempty subset S of $\mathbb{R}$:
  (a) S has a *maximum* element – an element that is larger than every other element of S.
  (b) The least upper bound of S is in S.

  Moreover, if these equivalent conditions hold, the maximum element equals the least upper bound.

(2) The following are equivalent for a nonempty subset S of $\mathbb{R}$:
  (a) S has a *minimum* element – an element that is smaller than every other element of S.
  (b) The greatest lower bound of S is in S.

  Moreover, if these equivalent conditions hold, the minimum element equals the greatest lower bound.

(3) A finite set has a maximum element and a minimum element. Thus, any finite set contains its least upper bound and greatest lower bound.

(4) For an interval with lower endpoint a and upper endpoint b, the least upper bound is b and the greatest lower bound is a. Note that this holds for all the four possibilities for the interval: $[a, b]$, (a, b), $[a, b)$, and $(a, b]$. The key distinction between the closed and open situation is not in the value of the upper/lower bound but in *whether that value is contained in the set we start with*.

(5) If T is a nonempty subset of a nonempty bounded subset S of $\mathbb{R}$, any lower bound for S remains a lower bound for T and any upper bound for S remains an upper bound for T. However, we may have an upper bound for T that is *not* an upper bound for S. Similarly, we may have a lower bound for T that is *not* a lower bound for S. Thus, the least upper bound for T is $\leq$ the least upper bound for S, and the greatest lower bound for T is $\geq$ the greatest lower bound for S.

(6) A (set does *not* have an upper bound) if and only if (it has arbitrarily large elements). Similarly, a set (does *not* have a lower bound) if and only if (it has arbitrarily small elements (i.e., negative elements of arbitrarily large magnitude)).

1.6. **Boundedness of a set and of related sets.** Given a nonempty set $S \subseteq \mathbb{R}$, we can construct the following sets:

(1) *Translation*: $S + \lambda$, which is defined as $\{s + \lambda : s \in S\}$: a is an upper (respectively lower) bound for S if and only if $a + \lambda$ is an upper (respectively lower) bound for $S + \lambda$. S is bounded from above (respectively below) if and only if $S + \lambda$ is. The least upper bound and greatest lower bound both get translated to the right by λ. In other words, the least upper bound of $S + \lambda$ is λ plus the least upper bound of S, and the greatest lower bound of $S + \lambda$ is λ plus the greatest lower bound of S.

(2) *Reflection about* 0: $-S$, which is the set $\{-s : s \in S\}$. a is an upper (respectively lower) bound for S if and only if $-a$ is a lower (respectively upper) bound for S. S is bounded from above (respectively below) if and only if $-S$ is bounded from below (respectively above). The least upper bound for $-S$ is the negative of the greatest lower bound for S, and the greatest lower bound for $-S$ is the negative of the least upper bound for S (these statements are subject to existence).

(3) Folding about 0: The set $\text{abs}(S) = \{|s| : s \in S\}$. Although this could be denoted $|S|$, that notation is often used for the *size* (or number of elements) of S, so we will refrain from that notation. S is bounded (i.e., bounded from both above and below) if and only if $\text{abs}(S)$ is bounded from above. Note that $\text{abs}(S)$ is *always* bounded from below by 0. Moreover, the least upper bound of $\text{abs}(S)$ is the maximum of the absolute values of the greatest lower bound and least upper bound of S.

(4) *An increasing function*: Suppose $f : \mathbb{R} \to \mathbb{R}$ is an increasing function and $f(S)$ is the image of S under f. Then, the least upper bound of the image $f(S)$ is the image under f of the least upper bound of S, while the greatest lower bound of the image $f(S)$ is the image under f of the greatest lower bound of S.

(5) *A decreasing function*: Suppose $f : \mathbb{R} \to \mathbb{R}$ is a decreasing function and $f(S)$ is the image of S under f. Then, the greatest lower bound of the image $f(S)$ is the image under f of the greatest lower bound of S, while the least upper bound of the image $f(S)$ is the image under f of the least upper bound of S.

(6) The least upper bound (respectively greatest lower bound) of a set of rational numbers may be rational or irrational. Similarly, the least upper bound (respectively greatest lower bound) of a set of irrational numbers may be rational or irrational.

2. Some important results

2.1. **There are numbers arbitrarily close up to the least upper bound.** The statement of the theorem is:

> If M is the least upper bound of a nonempty subset S of $\mathbb{R}$, then, for every $\epsilon > 0$, S has a nonempty intersection with the interval $(M - \epsilon, M]$. In particular, if $M \notin S$, then S has a nonempty intersection with the interval $(M - \epsilon, M)$.

The proof is straightforward. Suppse there exists $\epsilon > 0$ such that S has empty intersection with the interval. Then, we can easily check that $M - \epsilon$ is *also* an upper bound for S, which contradicts the minimality of M as an upper bound for S.

A similar statement holds for the greatest lower bound.

Basically, what these results are saying is that for something to qualify as the least upper bound or greatest lower bound, it either must be in the set or there must be stuff arbitrarily close to it that is in the set. It should not be possible to isolate it away from the set that it claims to be the best bound on.

2.2. **Deriving the least upper bound and greatest lower bound results from each other.** In the book's presentation, the statement about the existence of a least upper bound is stated as an *axiom* while the statement about the existence of a greatest lower bound is stated as a *theorem*, proved using the least upper bound axiom.

In reality, there is an elaborate construction procedure for the real numbers, and the details of that construction procedure guarantee both the least upper bound and the greatest lower bound properties. However, since we are not going through the construction procedure, we take a shortcut by treating the least upper bound existence statement as a *given*. It turns out, though, that the statement about the greatest lower bound can be deduced from it. There are two ways of deducing it:

(1) Reflection about 0: This interchanges the role of upper and lower bounds. We can use the least upper bound axiom on $-S$ to deduce the greatest lower bound property for S.
(2) The method as given in the book: The crux of this method is the following observations: (a) the greatest lower bound of a set is the least upper bound of the set of lower bounds of the set, (b) the set of lower bounds on a nonempty set S is bounded from above by an element of S, hence is bounded from above. We use the least upper bound axiom to argue that the set of lower bounds of S has a least upper bound, and then use (a) to show that this least upper bound *of the set of lower bounds* of S is also the greatest lower bound for S. For a formal proof, refer to the book.

The second approach may seem a little dense (in the non-mathematical sense) – why go for this kind of twisted logic when we can simply use the reflection about 0 proof? There is a deep reason.

The proof that involves the use of reflection about 0 is drawing on the additive structure of the real numbers and the behavior of the negation operation. Thus, this proof *does not generalize* to arbitrary totally ordered sets. On the other hand, the second proof uses *nothing special* about the real numbers and is equally valid for any totally ordered set. In other words, the second proof is valid over any totally ordered set, and says that a totally ordered set satisfies the least upper bound property if and only if it satisfies the greatest lower bound property.