MATHEMATICAL STATEMENTS AND PROOFS

ANDREAS STRÖMBERGSSON

In this note we review some mathematical notation, and try to teach the routine aspects of writing mathematical proofs. The note is meant to serve as a complement for students taking the course in Functional Analysis at Uppsala Univ (2006), using Kreyszig's book [Kreyszig].

We stress that we will *not* try, in this note, to say anything systematic about the much more interesting question of *how to come up with good and relevant ideas* to solve the given problem. (Thus, if we compare mathematics to the game of chess, then in this note we merely aim at teaching the *rules* of chess, i.e. how the various pieces are allowed to move; we do *not* try to say very much about how to become a good chess player.) However, a good mastering of these routines for writing proofs means that one is able to quickly get to the *core* of a problem; to see what the *key questions* are. Hopefully this might be useful, in an indirect way, also to help coming up with the necessary creative ideas.

1. Notation; sets and sequences

We will assume that the reader is familiar with the basic notation connected with *sets*. Thus if A, B are sets then "$a \in A$" means that a is an element in A; "$A \cup B$" is the union of A and B; "$A \cap B$" is the intersection of A and B; and "$A - B$" is the difference set A minus B (i.e. the set containing all the elements in A which do not lie in B). The statement "$A = B$" means that A and B have exactly the same elements; the statement "$A \subset B$" means that A is a subset of B (i.e., every element in A belongs also to B).

As usual $\mathbb{R}$ is the set of all real numbers, $\mathbb{Z}$ is the set of all integers and $\mathbb{C}$ is the set of all complex numbers. We also write $\mathbb{Z}^+$ for the set of *positive* integers and $\mathbb{R}^+$ for the set of positive real numbers. Thus $\mathbb{Z} \subset \mathbb{R} \subset \mathbb{C}$ and $\mathbb{Z}^+ \subset \mathbb{R}^+ \subset \mathbb{R}$. Also, $\mathbb{R}^n$ is the familiar vector space consisting of all n-tuples $(x_1, x_2, ..., x_n)$ of real numbers.

To describe a set we may list all its elements, as in "$\{2, 4, 7\}$" or "$\{5, 6, 7, 8, \cdots\}$." Another way is to form the set of elements with a certain property, for example: "$\{x \mid x + 4 \geqq 7\}$" is the set of all real numbers x such that $x + 4 \geqq 7$; hence $\{x \mid x + 4 \geqq 7\} = [3, \infty)$. The same set could also be written as "$\{x \in \mathbb{R} \mid x + 4 \geqq 7\}$", to stress the fact that we want to consider all *real numbers* x with the stated property. Some other examples:

$$\{b \in \mathbb{Z} \mid 3b \leqq 7 \text{ and } b \text{ is even}\} = \{2, 0, -2, -4, -6, -8, \cdots\};$$

$$\{z \in \mathbb{C} \mid |z - i| = 4\} \quad \text{is a circle in the complex plane with radius 4 and centre} = i.$$

One may also write an expression to the left of the "|" sign; for example

$$\{x^2 + 5 \mid x \in \mathbb{Z} \text{ and } 3 \leqq x \leqq 7\}$$

is the set of all numbers of the form $x^2 + 5$ where x is an arbitrary integer satisfying the condition $3 \leqq x \leqq 7$. Thus, in fact, $\{x^2 + 5 \mid x \in \mathbb{Z} \text{ and } 3 \leqq x \leqq 7\} = \{14, 21, 30, 41, 54\}$.

If A is any set, we write $\mathcal{P}(A)$ for the *power set* of A; this is the set of all subsets of A, including A itself and the empty set. Thus, for example,

$$\mathcal{P}(\{5, 6, 7\}) = \Big\{\emptyset, \{5\}, \{6\}, \{7\}, \{5, 6\}, \{5, 7\}, \{6, 7\}, \{5, 6, 7\}\Big\},$$

and in general, "$M \in \mathcal{P}(N)$" is synonymous with "$M \subset N$".

Sometimes in this note we will also use a less standard notation, $\mathcal{S}(A)$, to denote *the set of all infinite sequences for which all entries are elements in* A. Just as in Kreyszig [Kreyszig, p. 7] we use the notation "$(a_1, a_2, a_3, \cdots)$" to denote an infinite sequence whose elements are $a_1, a_2, a_3, \cdots$. We can also write briefly "(a_j)" to denote the same sequence. (And sometimes we may also write just "$a_1, a_2, a_3, \cdots$" to denote the same sequence.) Note that the definition of $\mathcal{S}(A)$ can also be expressed as follows:

$$\mathcal{S}(A) = \Big\{(\xi_1, \xi_2, ...) \mid \xi_j \in A \text{ for all } j\Big\}.$$

As an example, if $A = \{1, 2\}$, then $\mathcal{S}(A)$ is an infinite set, and here are three examples of elements in $\mathcal{S}(A)$:

$$(1, 2, 1, 2, \cdots); \qquad (1, 1, 2, 1, 2, 2, 1, 1, \cdots); \qquad (1, 1, 1, 1, \cdots).$$

IF THERE IS ANYTHING IN THE ABOVE NOTATION THAT YOU FEEL THAT YOU DO NOT UNDERSTAND AND HAVE TO REVIEW MORE IN DETAIL, THEN PLEASE ASK ME AS SOON AS POSSIBLE!

2. Statements, logical connectives, quantifiers

2.1. **Statements.** We will assume that the reader has some basic understanding of what a *statement* is. A statement can be either *true* or *false*. Examples of statements:

$$e^{\pi i} = -1$$

(this statement happens to be *true*);

$$3 + 4 = 6$$

(this statement happens to be *false*);

$$3 + x = 6$$

(this statement is true or false depending on what x is; for example, if $x = 0$ then the statement is *false*);

$$3 \in \{1, 2, 3, 4\}$$

(this statement is *true*);

$$\{1,2\} \subset \{1,3,4\}$$

(this statement is *false*);

$$A \subset B$$

(this statement is true or false depending on what A and B are; for example, if $A = \{1,2\}$ and $B = \{0,1,2,3,4,\cdots\}$ then the statement is *true*).

We call any letter appearing in a statement, such as "x" in "$3+x=6$", or "A" and "B" in "$A \subset B$", a *variable*. Thus a variable can represent a number, or a vector, or a set; it can also represent a metric space, a vector space, a sequence, and many other things.

Note that if all the variables occurring in the statement are given a specified meaning (as we did above by considering the case $x = 0$ for the statement "$3 + x = 6$"), then it should in principle be possible to determine whether the statement is true or is false. (Although this may be beyond human knowledge of today, for certain complicated statements.)

2.2. **Logical connectives.** In the following we will use boldface letters $\mathbf{P}, \mathbf{Q}, \mathbf{R}, ...$ to denote arbitrary statements.

More complicated statements than those given in the examples above can be formed using the logical connectives "and", "or", "not", "implies": if **P** and **Q** are statements then

$$\mathbf{P} \text{ and } \mathbf{Q}, \qquad \mathbf{P} \text{ or } \mathbf{Q}, \qquad \text{not}(\mathbf{P}), \qquad \mathbf{P} \text{ implies } \mathbf{Q},$$

are also statements. The statement "**P** and **Q**" is true exactly when *both* **P** and **Q** are true. The statement "**P** or **Q**" is true exactly when *at least one* of **P** and **Q** are true. The statement "not(**P**)" is true exactly when **P** is *false*. Finally, the statement "**P** implies **Q**" is true when **Q** is true, and also when **P** is false; but it "**P** implies **Q**" is false in the case when **P** is true and **Q** is false. Thus in tabular form (writing "T" for "True" and "F" for "False"):

	P: T , **Q**: T	**P**: T , **Q**: F	**P**: F , **Q**: T	**P**: F , **Q**: F
P and **Q**	T	F	F	F
P or **Q**	T	T	T	F
not(**P**)	F	F	T	T
P implies **Q**	T	F	T	F

We remark that there are many other symbols for the above four logical connectives: "$\wedge$", ",", "$\vee$", "$\Rightarrow$", "$\neg$". These are collected in the following table; all statements listed on each individual line are completely synonymous:

P and **Q**	$\mathbf{P} \wedge \mathbf{Q}$	$\mathbf{P}, \mathbf{Q}$	
P or **Q**	$\mathbf{P} \vee \mathbf{Q}$		
not(**P**)	$\neg\mathbf{P}$		
P implies **Q**	$\mathbf{P} \Rightarrow \mathbf{Q}$	$\mathbf{Q} \Leftarrow \mathbf{P}$	if **P** then **Q**

Parentheses are used as usual to make it clear in which order to read a statement. However, one usually agrees that "implies" has lower priority than "and,or", and these have lower priority than "not", but these priority rules should always be used in combination with intelligent *spacing* or similar to make the order clear. Thus, for example,

$$\neg\mathbf{P} \text{ and } \mathbf{Q} \implies \neg\mathbf{R} \text{ or } (\mathbf{P} \text{ and } \mathbf{Q})$$

is synonymous with

$$((\neg\mathbf{P}) \text{ and } \mathbf{Q}) \text{ implies } ((\neg\mathbf{R}) \text{ or } (\mathbf{P} \text{ and } \mathbf{Q})).$$

Finally we note one more very common connective: "$\mathbf{P} \Leftrightarrow \mathbf{Q}$", which is also written "**P** and **Q** are equivalent", and also "**P** if and only if **Q**". This can be seen as an abbreva-tion of "($\mathbf{P} \Rightarrow \mathbf{Q}$) and ($\mathbf{P} \Leftarrow \mathbf{Q}$)", and hence the truth table is:

	P: T , **Q**: T	**P**: T , **Q**: F	**P**: F , **Q**: T	**P**: F , **Q**: F
$\mathbf{P} \Leftrightarrow \mathbf{Q}$	T	F	F	T

2.3. **Quantifiers.** There is one more way to form more complicated statements from simpler ones: By using the *quantifiers* $\forall$ ("for all") and $\exists$ ("there exists"). We start with $\forall$. If **P** is any statement in which the variable x occurs, and M is some set, then the following is a new statement:

$$\forall x \in M : \mathbf{P}$$

This statement, "$\forall x \in M : \mathbf{P}$", is true exactly if the statement **P** is *true for every choice of x as an element in M*. We give four examples of such statements:

$$\forall x \in \mathbb{Z} : \ (x + x)^2 = 4x^2$$

(this statement is *true*, since $(x + x)^2 = 4x^2$ holds for every integer x);

$$\forall a \in \mathbb{R} : \ (a - 2)^2 \geqq \tfrac{1}{10}$$

(this statement is *false*, since $(a - 2)^2 \geqq \frac{1}{10}$ it is not true for *every* real number a; for example it fails if $a = 2$);

$$\forall v \in \mathbb{R}^3 : \ 2v + (2, 4, 6) = 2(v + (1, 2, 3))$$

(this statement is *true*, since $2v + (2, 4, 6) = 2(v + (1, 2, 3))$ holds for every vector $v \in \mathbb{R}^3$);

$$\forall \xi \in \mathbb{R} : \ \Big((\xi + 3)^2 \geqq 2 \text{ or } \xi \leqq 10\Big)$$

(this statement is *true*, since $\big((\xi + 3)^2 \geqq 2 \text{ or } \xi \leqq 10\big)$ is true for every real number ξ (why?)).

Note that to see that a statement of the form "$\forall x \in M : \mathbf{P}$" is *false* it suffices to find a single example of a choice of $x \in M$ which makes **P** false. But to see that a statement of the form "$\forall x \in M : \mathbf{P}$" is *true* one would in principle have to check that **P** is true *for every possibly choice of $x \in M$*; of course this is not often possible to do directly, since there are in general infinitely many choices of x; instead one will have to rely on logical reasoning to see that "$\forall x \in M : \mathbf{P}$" is true. (We will say more about logical reasoning below.)

We next turn to $\exists$ ("there exists"). If $\mathbf{P}$ is any statement in which the variable x occurs, and M is some set, then the following is a new statement:

$$\exists x \in M : \mathbf{P}$$

This statement, $\exists x \in M : \mathbf{P}$, is true precisely if *there is some choice of* x *as an element in* M *for which the statement* $\mathbf{P}$ *is true.* Examples:

$$\exists a \in \mathbb{R} :\ a^2 \geqq 9$$

(this statement is *true*, since $a^2 \geqq 9$ holds for the real number $a = 10$; another way to see that the statement is true is to note that $a^2 \geqq 9$ holds for the real number $a = -42$);

$$\exists n \in \mathbb{Z} :\ n^2 = 10$$

(this statement is *false*, since there is no integer n for which $n^2 = 10$);

$$\exists v \in \mathbb{R}^3 :\ v \cdot (1, 2, 3) = 10$$

(this statement is true, since $v \cdot (1, 2, 3) = 10$ holds for the vector $v = (0, 2, 2) \in \mathbb{R}^3$).

We remark that statements containing $\forall$ or $\exists$ may well contain *other* variables as well, and then, just as in some of the examples in §2.1, the statement may be true or false depending on what values these other variables take. For example, the statement

$$\forall a \in \mathbb{R} :\ a^2 \geqq x + y$$

is *true* if $x = 2$, $y = -3$, since $a^2 \geqq 2 + (-3)$ holds for all real numbers a, but it is *false* when $x = 2, y = 1$, since $a^2 \geqq 2 + 1$ does *not* hold for every real number a (for example it fails if $a = 1$). As another example, the statement

$$\exists n \in \mathbb{Z} :\ n^2 + m = 3h$$

is *true* when $m = 2$, $h = 6$, since $n^2 + 2 = 3 \cdot 6$ holds for the integer $n = 4$, but it is *false* when $m = 0$, $h = 1$, since $n^2 + 0 = 3$ does not hold for any integer n.

Combinations: Of course, when forming a statement "$\forall x \in M : \mathbf{P}$" or "$\exists x \in M : \mathbf{P}$", the inner statement $\mathbf{P}$ may very well itself contain quantifiers $\forall$ or $\exists$. We give three examples of such statements:

1. $\forall x \in \mathbb{R} :\ \exists y \in \mathbb{R} :\ x = y^2 - 10.$
This statement is *false*, since the statement "$\exists y \in \mathbb{R} :\ x = y^2 - 10$" is not true for *every* real number x; for example it fails if $x = -11$.

2. $\forall \varepsilon \in \mathbb{R}^+ :\ \exists \delta \in \mathbb{R}^+ :\ \forall x \in [2 - \delta, 2 + \delta] :\ |e^x - e^2| < \varepsilon.$
This statement is *true*, since the statement

$$\exists \delta \in \mathbb{R}^+ :\ \forall x \in [2 - \delta, 2 + \delta] :\ |e^x - e^2| < \varepsilon$$

is true for every $\varepsilon \in \mathbb{R}^+$; however this is probably not obvious to all readers...

3. $\forall x \in \mathbb{R}^3 : \Big(\Big(\forall y \in \mathbb{R}^3 : \ x \cdot y = 0 \Big) \Longrightarrow x = 0 \Big).$

This statement is *true* since the statement

$$\Big(\forall y \in \mathbb{R}^3 : \ x \cdot y = 0 \Big) \Longrightarrow x = 0$$

is true for every vector $x \in \mathbb{R}^3$. (Why?)

2.4. **Grammatical restrictions; free and bound variables.** There are some restrictions when forming statements of the form "$\forall x \in M : \mathbf{P}$" or "$\exists x \in M : \mathbf{P}$". First of all, we must of course require that the statement $\mathbf{P}$ is *makes sense* for every element x in M. Thus for example the following two "statements" are *forbidden* ("nonsense"):

$$\forall v \in \mathbb{R}^3 : \ v + 2 = 3; \qquad \exists n \in \mathbb{Z} : n + (1,2,3) = (4,6,7).$$

(Both these are forbidden since there does not exist any operation of adding an integer and a 3-dimensional vector.)

Secondly, the quantification status of each variable must be clear at each position in the statement. Here are some examples of "statements" which are *forbidden* ("nonsense"):

$$\forall a \in \mathbb{R} : \ \exists b \in \mathbb{R} : \ \exists a \in \mathbb{R} : \ e^a \leqq b$$

$$\forall x \in \mathbb{R} : \ \Big(\Big(\exists x \in \mathbb{R} : \ x^4 + x^2 = y \Big) \text{ and } x + 2 = z \Big).$$

To avoid such unclear quantification we have the following *rules for forming statements:*

- If a variable α occurs in a statement $\mathbf{P}$ and there is no "$\forall \alpha$" or "$\exists \alpha$" in $\mathbf{P}$, then the variable α is called *free* in $\mathbf{P}$.[1]
- If $\mathbf{P}$ is a statement then "$\forall \alpha \in M : \mathbf{P}$" and "$\exists \alpha \in M : \mathbf{P}$" are permitted statements *only if the variable α is free in* $\mathbf{P}$. The variable α is said to be a *bound* in these two statements "$\forall \alpha \in M : \ \mathbf{P}$" and "$\exists \alpha \in M : \ \mathbf{P}$".
- If $\mathbf{P}$ and $\mathbf{Q}$ are statements, then "$\mathbf{P}$ and $\mathbf{Q}$" and "$\mathbf{P}$ or $\mathbf{Q}$" and "$\mathbf{P} \Rightarrow \mathbf{Q}$" are also permitted statements, *as long as there is no variable which is free in one of* $\mathbf{P}, \mathbf{Q}$ *and bound in the other.*

Thus for example

$$\text{“}\Big(\exists n \in \mathbb{Z} : \ n^2 + m = 3h \Big) \text{ and } \Big(\forall n \in \mathbb{Z} : \ \forall a \in \mathbb{Z} : n^2 + a^2 \geqq h - 10 \Big)\text{”}$$

is a permitted statement; and this statement contains two bound variables, a, n and two free variables m, h.

A statement in which all variables are bound is said to be *closed*. Every theorem (or lemma or proposition) in a mathematical theory should be a closed statement!

[1]There is one *exception*: When α occurs as a running variable ("dummy variable") in an expression like "$\sum_{\alpha=3}^{5} \alpha$" or "$\prod_{\alpha=1}^{\infty}(1 + \alpha^{-2})$" or "$\{\alpha \in \mathbb{Z} \mid \alpha^2 + 4 \geqq 100\}$", then α is a *bound variable* in statements containing these expressions.

2.5. **Language conventions.** When we write a mathematical text we often use common language instead of signs like "$\forall$" and "$\exists$". For example, the statement

$$\text{“}M \subset \mathbb{R} \text{ and } \forall v \in M : \forall w \in M : \Big(v \neq w \Rightarrow |v - w| > 1 \Big)\text{”}$$

may be expressed like:

> "M is a set of real numbers and for all numbers v, w in M we have that $\Big(v \neq w \Rightarrow |v - w| > 1 \Big)$" .

The same statement can also be written:

> "M is a set of real numbers such that for all numbers $v \neq w$ in M we have $|v - w| > 1$"

or

> "M is a set of real numbers such that any two distinct numbers in M have distance greater than one"

or

> "M is a set of real numbers whose elements are pairwise separated by a distance greater than one."

Concerning the last formulation: This uses the word *pairwise*, which may possibly cause confusion to a mathematically less experienced reader; the word pairwise is generally used in constructions like "The elements of M are pairwise ****", meaning
"$\forall a \in M : \forall b \in M : \Big(a \neq b \Rightarrow a, b \text{ are } ^{****} \Big)$".

A mathematical text contains a lot of such words which by convention has a very precise meaning to an experienced reader. *IT IS OF BASIC IMPORTANCE THAT WHENEVER YOU READ A STATEMENT IN THE MATHEMATICAL LITERATURE, YOU SHOULD KNOW EXACTLY WHAT IT MEANS,* i.e. exactly how the variables are quantified etc. To test this, if you are hesitant about the meaning of a statement, you may try to write it out in the formal language which we have introduced in §2.2–2.4. *WHENEVER YOU FIND THAT IT IS UNCLEAR HOW TO DO THIS, THEN PLEASE ASK ME!* Also see the exercise in §2.6.

Here are some particular language conventions (all fairly low level):

- One often leaves out "M" in "$\forall x \in M$" or "$\exists x \in M$" if it is clear from the context what M is. For example, in the middle of a long piece of text which deals only with integers, one may perhaps write "...hence there is an x such that $ax = z - by$..." instead of "...hence there is an *integer* x such that $ax = z - by$...".
- The quantifiers $\forall$ and $\exists$ are often used with a more permitting grammar; for example one may put them *after* the relevant statement, and do other obvious abbrevations. Thus "$x^2 + y^2 \geqq 0, \forall x, y \in \mathbb{R}$" is synonymous with "$\forall x \in \mathbb{R} : \forall y \in \mathbb{R} : x^2 + y^2 \geqq 0$."
- If it is clear from the context that x denotes a real number, then one may write e.g. "$\exists x > 5 : \mathbf{P}$" instead of "$\exists x \in (5, \infty) : \mathbf{P}$". Similarly, if it is clear from the context that n denotes an integer, then we may write e.g. "$\forall n \geqq N$" instead of "$\forall n \in \{N, N+1, N+2, \cdots\}$".

- In the present text we will sometimes use the notation "$\forall X$ [metric space]", meaning "for all metric spaces X", as well as "$\forall H$ [Hilbert space]", meaning "for all Hilbert spaces H", and some similar things. This is because we prefer not to introduce a special symbol for the set of all metric spaces, or the set of all Hilbert spaces, etc.
- One common mathematical expression is "there exists a *unique* x in M such that $\mathbf{P}$"; this is often written "$\exists!\, x \in M : \mathbf{P}$". The meaning of the this statement is that there is *exactly one* choice of x in the set M such that $\mathbf{P}$ is true. To give an example, the statement "there is a unique integer x such that $x^3 = 27$" can be written as "$\exists!\, x \in \mathbb{Z} :\ x^3 = 27$", and using only our basic symbols the same statement can be written out as follows: "$\exists x \in \mathbb{Z} :\ \Big(x^3 = 27 \text{ and } \forall y \in \mathbb{Z} :\ \Big(y^3 = 27 \Rightarrow y = x\Big)\Big)$." [2]

Finally we remark that any good mathematical text contains a lot of words and formulations which, technically, are actually irrelevant for the statement or proof at hand. For example any remarks on what *we want to prove* are irrelevant at a technical level, but such remarks often make the proof much easier to read.

2.6. **An exercise.** Write out every lemma and theorem in Kreyszig, chapter 1, as closed statements using the symbols introduced in this sections! Also write out every definition as a closed statement! (Recall from §2.4 that a closed statement is a statement in which every variable is bound.)

Answers. Below we give answers to the exercise for some cases, and give some comments. We do not keep strictly the same order as in Kreyszig's book.

Definition 1.3-2 The definition of "open" looks as follows as a closed statement. Note that in this form the statement in itself contains all information regarding which M and which X the definition is concerned with:

$$\forall X \text{ [metric space]} :\ \forall M \in \mathcal{P}(X) :\ \Big[\underline{M \text{ is open}} \Longleftrightarrow \Big(\forall x \in M :\ \exists r > 0 :\ B(x,r) \subset M\Big)\Big].$$

(The notation "$\mathcal{P}(X)$" was explained in §1; note that "$M \in \mathcal{P}(X)$" is the same as saying that M is an arbitrary subset of X.) Furthermore, the definition of "closed" looks as follows as a closed statement:

$$\forall X \text{ [metric space]} :\ \forall K \in \mathcal{P}(X) :\ \Big[\underline{K \text{ is closed}} \Longleftrightarrow \Big(X - K \text{ is open}\Big)\Big].$$

[2] A general procedure is as follows: The statement "$\exists!\, x \in M :\ \mathbf{P}$" can be written in basic notation as "$\exists x \in M :\ \Big(\mathbf{P} \text{ and } \forall y \in M :\ \Big(\mathbf{P}_{|y \mapsto x} \Rightarrow y = x\Big)\Big)$", where "$\mathbf{P}_{|y \mapsto x}$" is the statement $\mathbf{P}$ but with the variable x everywhere replaced by y (and the variable y is chosen so that it does not occur anywhere in the original statement $\mathbf{P}$).

Definition of bounded (p. 16, problem 6)

$$\forall X \text{ [metric space]} : \forall A \in \mathcal{P}(X) - \{\emptyset\} : \Big[\underline{A \text{ is bounded}} \Longleftrightarrow \sup_{x,y \in A} d(x,y) < \infty\Big].$$

Note here that "$A \in \mathcal{P}(X) - \{\emptyset\}$" is a short way of writing "A is a nonempty subset of X". Further, note that "$\sup_{x,y\in A} d(x,y)$" is just an alternative notation for "$\sup\{d(x,y) \mid x,y \in A\}$", and this is a concept which we define in §4 below. As we explain below Theorem 1, "$\sup\{d(x,y) \mid x,y \in A\} < \infty$" is the same thing as saying "$\{d(x,y) \mid x,y \in A\}$ is bounded above", and this can be expanded using Definition 4.3 in §4. Hence our definition can be expressed as follows:

$$\forall X \text{ [metric space]} : \forall A \in \mathcal{P}(X) - \{\emptyset\} :$$
$$\Big[\underline{A \text{ is bounded}} \Longleftrightarrow \exists b \in \mathbb{R} : \forall t \in \{d(x,y) \mid x,y \in A\} : t \leqq b\Big].$$

This can also be simplified slightly to read:

$$\forall X \text{ [metric space]} : \forall A \in \mathcal{P}(X) - \{\emptyset\} :$$
$$\Big[\underline{A \text{ is bounded}} \Longleftrightarrow \exists b \in \mathbb{R} : \forall x \in A : \forall y \in A : d(x,y) \leqq b\Big].$$

Definition 1.3-3 Definition of "continuous at a point":

$$\forall (X,d) \text{ [metric space]} : \forall (Y,\overline{d}) \text{ [metric space]} : \forall T \text{ [map } X \to Y] : \forall x_0 \in X :$$
$$\Big[\underline{T \text{ is continuous at } x_0 \in X}$$
$$\Longleftrightarrow \Big(\forall \varepsilon > 0 : \exists \delta > 0 : \forall x \in X : d(x,x_0) < \delta \Longrightarrow \overline{d}(Tx,Tx_0) < \delta\Big)\Big].$$

Definition of "continuous":

$$\forall (X,d) \text{ [metric space]} : \forall (Y,\overline{d}) \text{ [metric space]} : \forall T \text{ [map } X \to Y] :$$
$$\Big[\underline{T \text{ is continuous}} \Longleftrightarrow \Big(\forall x_0 \in X : T \text{ is continuous at } x_0\Big)\Big]$$

Definition 1.4-1 Definition of "$x_n \to x$" (note that we have a lot of other notations for exactly the same concept, e.g. "$\lim_{n\to\infty} x_n = x$"):

$$\forall X \text{ [metric space]} : \forall (x_n) \in \mathcal{S}(X) : \forall x \in X : \Big[\underline{x_n \to x} \Longleftrightarrow \lim_{n\to\infty} d(x_n,x) = 0\Big].$$

Note that Kreyszig here assumes that the reader is already familiar with the concept of limit of a sequence of real numbers (in this case the sequence $d(x_1,x), d(x_2,x), d(x_3,x), \cdots$).

If we expand also on this definition (cf. Definition 4.1 in §4 below) then the above definition would look as follows:

$$\forall X \text{ [metric space]} : \forall (x_n) \in \mathcal{S}(X) : \forall x \in X :$$
$$\Big[\underline{x_n \to x} \Longleftrightarrow \Big(\forall \varepsilon > 0 : \exists N \in \mathbb{Z}^+ : \forall n \geqq N : |d(x_n, x) - 0| < \varepsilon\Big)\Big].$$

If we use the fact that $d(x_n, x) - 0 = d(x_n, x)$, and also $|d(x_n, x)| = d(x_n, x)$ since $d(x_n, x) \geqq 0$, then the above definition can also be written in a slightly shorter way:

$$\forall X \text{ [metric space]} : \forall (x_n) \in \mathcal{S}(X) : \forall x \in X :$$
$$\Big[\underline{x_n \to x} \Longleftrightarrow \Big(\forall \varepsilon > 0 : \exists N \in \mathbb{Z}^+ : \forall n \geqq N : d(x_n, x) < \varepsilon\Big)\Big].$$

Note that in Definition 1.4-1 (the first three lines) Kreyszig also defines what we mean by saying that "the sequence (x_n) converges". This is defined as follows:

$$\forall X \text{ [metric space]} : \forall (x_n) \in \mathcal{S}(X) : \Big[\underline{(x_n) \text{ converges}} \Longleftrightarrow \exists x \in X : x_n \to x\Big]$$

(And in here, the "$x_n \to x$" can of course be expanded via the definition which we discussed above.)

Lemma 1.4-2 (a):

$$\forall X \text{ [metric space]} : \forall (x_n) \in \mathcal{S}(X) :$$
$$\Big[(x_n) \text{ is convergent} \Longrightarrow \Big(\Big((x_n) \text{ is bounded}\Big) \text{ and } \Big(\lim_{n\to\infty} x_n \text{ is unique}\Big)\Big)\Big]$$

Here note that the meaning of "(x_n) is convergent" was defined in Definition 1.4-1 (see our discussion above; note that "is convergent" is synonymous with "converges"; hence "(x_n) is convergent" can be written out as "$\exists x \in X : \lim_{n\to\infty} x_n = x$").
Furthermore, regarding the statement "(x_n) is bounded": See our discussion of p.16, problem 16 above! Note that saying that a *sequence* $(x_n) \in \mathcal{S}(X)$ should of course be interpreted as saying that the corresponing set of points, $\{x_1, x_2, x_3, \cdots\} \subset X$, is bounded. (But this ought to have been written out in the book!) Hence from the third format in our discussion of p.16, problem 16 above, we see that "(x_n) is bounded" can be written out as

$$\text{“}\exists b \in \mathbb{R} : \forall n \in \mathbb{Z}^+ : \forall m \in \mathbb{Z}^+ : d(x_n, x_m) \leqq b.\text{”}$$

Finally, the statement "$\lim_{n\to\infty} x_n$ is unique" can be written out as

$$\text{“}\forall x \in X : \forall z \in X : \Big(\Big(\lim_{n\to\infty} x_n = x \text{ and } \lim_{n\to\infty} x_n = z\Big) \Rightarrow x = z\Big).\text{”}$$

In fact this is a general pattern for writing out the claim that "something is unique". (Compare our discussion about "∃!" above in §2.5. The difference now is that right here we do *not* wish to make the claim that "$\lim_{n\to\infty} x_n$" exists, but *only* that it is unique.)

Lemma 1.4-2(b):

$$\forall X \text{ [metric space]} : \forall (x_n) \in \mathcal{S}(X) : \forall x \in X : \forall (y_n) \in X : \forall y \in X : \\ \Big[\Big(x_n \to x \text{ and } y_n \to y \Big) \Rightarrow d(x_n, y_n) \to d(x,y) \Big].$$

Theorem 1.4-4: This statement has no variables; it says simply "$\mathbb{R}$ is a complete metric space". (And similarly for $\mathbb{C}$.) However, if we write out what this means in view of Def 1.4-3, then Theorem 1.4-4 (for $\mathbb{R}$) reads:

$$\forall (x_n) \in \mathcal{S}(\mathbb{R}) : \Big((x_n) \text{ is Cauchy} \Big) \Rightarrow \Big((x_n) \text{ is convergent} \Big).$$

Using also the definitions of "Cauchy" and "convergent" the same statement reads:

$$\forall (x_n) \in \mathcal{S}(\mathbb{R}) : \bigg(\Big(\forall \varepsilon > 0 : \exists N \in \mathbb{Z}^+ : \forall m > N : \forall n > N : d(x_m, x_n) < \varepsilon \Big) \\ \Rightarrow \Big(\exists x \in X : \lim_{n \to \infty} d(x_n, x) = 0 \Big) \bigg).$$

Theorem 1.4-5.

$$\forall X \text{ [metric space]} : \forall (x_n) \in \mathcal{S}(X) : \Big((x_n) \text{ is convergent} \Longrightarrow (x_n) \text{ is Cauchy} \Big).$$

Here the inner statements "(x_n) is convergent" and "(x_n) is Cauchy" can be written out using the definitions, but we skip these details here since we have done very similar things above.

Theorem 1.4-6(a).

$$\forall X \text{ [metric space]} : \forall M \in \mathcal{P}(X) - \{\emptyset\} : \forall x \in X : \Big(x \in \overline{M} \Longleftrightarrow \exists (x_n) \in \mathcal{S}(M) : x_n \to x \Big)$$

Theorem 1.4-6(b).

$$\forall X \text{ [metric space]} : \forall M \in \mathcal{P}(X) - \{\emptyset\} : \\ \bigg[\bigg(\forall x \in X : \Big(\Big(\exists (x_n) \in \mathcal{S}(M) : x_n \to x \Big) \Longrightarrow x \in M \Big) \bigg) \Longleftrightarrow M \text{ is closed} \bigg].$$

3. How to write proofs

NOTE: THE FOLLOWING SECTION §3.1 IS STILL PRELIMINARY! It is very difficult – perhaps impossible – to find a good balance between a description of proof rules that is formally correct, and still permits, without long detours, most of the standard methods used in "everyday mathematics" (which you will need to solve problems in the Functional Analysis course)! There are several details in the following presentation that I am not happy with!

3.1. **What is a proof?** A *proof* is a list of statements. (Most often we express each statement as a sentence in common English.) The basic requirement is that each statement in the proof should be a logical consequence of statements made earlier in the proof, *or* a consequence of a known theorem. Also, the *very last statement* in the proof should be the desired statement which we wanted to prove.

However, there is one exception to the first rule: We are allowed to make an *assumption*, whenever we want, in the proof. An *assumption* is a completely arbitrary statement; we can choose it to be anything we want![3] However, at each stage in the proof we must remember *which assumptions are in force*; a statement in the middle of a proof is most often *not* an "unconditional truth"; instead, it is only true *conditionally* on all the assumptions which are presently in force! In the system which we present here, there are *four* ways to *close* an assumption, that is, to make an assumption *stop* being in force after a certain point. The assumption which we close must always be the *most recently* made assumption.

We now describe the four ways to close an assumption:

1. ⇒-Introduction: If the most recent assumption is a statement $\mathbf{P}$, and we have, after this, arrived at a statement $\mathbf{Q}$, then we may *close the assumption* $\mathbf{P}$ and after this conclude "$\mathbf{P} \Rightarrow \mathbf{Q}$".

2. ∀-Introduction: If the most recent assumption is of the form "Let x be an arbitrary element in M" and we have after this arrived at a statement $\mathbf{Q}$ (which contains x as a free variable), then we may *close the assumption* and after this conclude "$\forall x \in M : \mathbf{Q}$". (Restriction: There must be no other assumption in force which contains x as a free variable!)

3. ∃-Elimination: Suppose that $\mathbf{Q}$ is a statement which contains a free variable x and that $\mathbf{P}$ is a statement which does *not* contain x as a free variable. In our proof, suppose that we have arrived at the statement $\mathbf{P}$, and that the most recent assumption was "$x \in M$ and $\mathbf{Q}$". Also assume that *before* this assumption in the proof we have the statement "$\exists x \in M : \mathbf{Q}$". Then we may *close the assumption* and after this conclude "$\mathbf{P}$".

4. Proof by contradiction: Suppose that we have arrived at a contradiction in our proof, i.e. we have arrived at a statement of the form $\mathbf{Q}$ and $\neg\mathbf{Q}$ (where $\mathbf{Q}$ is some arbitrary

[3] But if we do not choose it intelligently, with planning ahead, the assumption will be useless for us!

statement). Assume that the most recent assumption is a statement **P**. Then we may *close the assumption* **P**, and after this conclude "¬**P**".

We now list the other basic ways (which do not involve making or closing an assumption) to obtain new statements in a proof. Note that there is one restriction regarding which earlier statements in the proof we are allowed to use: We are not allowed to make use of any earlier statement which depends on some assumption which has already been closed!

5. Elementary logic: In this type of deduction we only use properties of the logical connectives " and ", " or ", "not", "⇒", "⇔"; this type of conclusion can always be checked using a *truth table*. We give an example: Suppose that somewhere earlier in the proof we have the statement "$(\exists x \in \mathbb{Z} : y = x^2)$ or $(\exists x \in \mathbb{Z} : y = 3x)$" and at another place we have the statement "not$(\exists x \in \mathbb{Z} : y = 3x)$"; then we may deduce the new statement "$\exists x \in \mathbb{Z} : y = x^2$". If we want to check this type of deduction using a truth table we rewrite each involved statement using only the symbols " and ", " or ", "not", "⇒", and unspecified symbols $\mathbf{P}, \mathbf{Q}, \mathbf{R}, ...$ for the inner statements involved: Thus the two statements which we already had in our proof are "**P** or **Q**" and "¬**Q**", and the statement which we deduce is: "**P**". Note that we are allowed to use the same letter only for *exactly identical inner statements*. Truth table of the statements involved:

	P: T , **Q**: T	**P**: T , **Q**: F	**P**: F , **Q**: T	**P**: F , **Q**: F
earlier in proof: "**P** or **Q**"	T	T	T	F
earlier in proof: "¬**Q**"	F	T	F	T
to conclude: "**P**"	T	T	F	F

The point is that in each column where all the earlier statments are true (there is only one such column above: the second one), the statement which we want to conclude is *also* true!

Another example: If we have a statement "$y = 5 \Rightarrow (x = 3$ and $x \neq 3)$" (that is, "$y = 5 \Rightarrow (x = 3$ and not$(x = 3))$") earlier in the proof, we may conclude "not$(y = 5)$". Truth table:

	P: T , **Q**: T	**P**: T , **Q**: F	**P**: F , **Q**: T	**P**: F , **Q**: F
earlier in proof: "**P** ⇒ (**Q** and not(**Q**))"	F	F	T	T
to conclude: "not(**P**)"	F	F	T	T

With some experience these elementary logic deductions can be made fairly quickly, and certainly without actually writing out the truth table.

Another example, slightly more complicated: Suppose that somewhere earlier in the proof we have the statement "$x \neq 4$ or $y = 3$ or $y = 4$" and at another place we have the statement "$x = 4$ or $y = 3$"; then we may conclude the new statement

$$\text{“}y = 3 \text{ or } (x = 4 \text{ and } y = 4)\text{”}.$$

(Note that the truth table in this case would have 8 columns, but hopefully you can convince yourself that the conclusion is correct without actually drawing the full table.)

6. Arbitrary manipulations. You are allowed to use all your acquired knowledge about how to manipulate various mathematical objects! For example, you know how to simplify an expression like "$(a+2b)^2(a-b)-a^3+4b^3$" (if a,b are real or complex numbers); namely:

$$(a+2b)^2(a-b)-a^3+4b^3 = (a^2+4ab+4b^2)(a-b)-a^3+4b^3$$
$$= (a^3+3a^2b+0ab^2-4b^3)-a^3+4b^3 = 3a^2b.$$

This means that whenever you see "$(a+2b)^2(a-b)-a^3+4b^3$" in a statement you are allowed to replace it by "$3a^2b$" (if you also present the above computation, to justify your step). For example, if somewhere earlier in the proof we have the statement "$4+c \leqq (a+2b)^2(a-b)-a^3+4b^3$", then we may conclude the new statement "$4+c \leqq 3a^2b$".

Similarly you may use any techniques you know for how to manipulate functions (differentiation, integration) and vectors (linearity rules, etc).

7. Exchange $\neg\forall \leftrightarrow \exists\neg$ and $\forall\neg \leftrightarrow \neg\exists$. If α is a variable and M is a set, then any occurence of the string "$\neg\forall\alpha \in M$:" may be replaced by "$\exists\alpha \in M : \neg$". Also any occurence of the string "$\neg\exists\alpha \in M$:" may be replaced by "$\forall\alpha \in M : \neg$".

An example: If somewhere earlier in the proof we have the statement

$$\text{“}(\exists z \in \mathbb{Z} :\ y = 4z+3) \Rightarrow (\forall x \in \mathbb{Z} : \neg(y = x^2))\text{”},$$

then we may deduce the new statement

$$\text{“}(\exists z \in \mathbb{Z} :\ y = 4z+3) \Rightarrow \neg(\exists x \in \mathbb{Z} : y = x^2)\text{”}.$$

8. $\forall$-Elimination. Suppose that somewhere earlier in the proof we have the two statements "$* \in M$" and "$\forall\alpha \in M$: **P**", where $*$ is some expression, and **P** is a statement containing α as a free variable, and no variable which occurs in the expression $*$ occurs as a bound variable in **P**. Then we may conclude as a new statement "[**P** with every occurrence of α replaced with $*$]."

We give three examples: If somewhere earlier in the proof we have the two statements "$v \in \mathbb{R}^3$" and "$\forall w \in \mathbb{R}^3 : \exists u \in \mathbb{R}^3 - \{\mathbf{0}\} : w \cdot u = 0$" then we may conclude the new statement "$\exists u \in \mathbb{R}^3 - \{\mathbf{0}\} : v \cdot u = 0$".

If somewhere earlier in the proof we have the two statements "$r - r_0 \in \mathbb{R}^+$" and "$\forall\varepsilon > 0 : \exists\delta > 0 : \forall x \in [a-\delta, a+\delta] : |e^x - e^a| < \varepsilon$", then we may conclude the new statement "$\exists\delta > 0 : \forall x \in [a-\delta, a+\delta] : |e^x - e^a| < r - r_0$".

If somewhere earlier in the proof we have the two statements "$y + t \in \mathbb{Z}$" and "$\forall x \in \mathbb{Z} :\ (x+y)^2 = x^2 + 2xy + y^2$", then we may conclude the new statement: "$((y+t)+y)^2 = (y+t)^2 + 2(y+t)y + y^2$."

9. Substitution. Suppose that somewhere earlier in the proof we have the statement "$(\alpha = *) \Rightarrow$ **P**", where $*$ is some expression and **P** is a statement containing α as a free variable, and no variable which occurs in the expression $*$ occurs as a bound variable in **P**.[4] Then we may conclude as a new statement "[**P** with every occurrence of α replaced with $*$]."

Example: If somewhere earlier in the proof we have the statement "$x = y - z \Longrightarrow a = x^2 + (y+z)^2$" then we may conclude the new statement "$a = (y-z)^2 + (y+z)^2$".

10. $\exists$-Introduction. Let $\mathbf{P}$ be a statement containing a free variable α, and let $*$ be an expression such that no variable which occurs in $*$ occurs as a bound variable in $\mathbf{P}$. Let $\mathbf{Q}$ be the statement "[$\mathbf{P}$ with every occurrence of α replaced with $*$]." The rule for $\exists$-introduction is as follows: Suppose that somewhere earlier in the proof we have the two statements "$* \in M$" and "$\mathbf{Q}$"; then we may conclude the new statement "$\exists \alpha \in M : \mathbf{P}$".

Examples: If somewhere earlier in the proof we have the two statements "$2x - 1 \in \mathbb{R}$" and "$\frac{(2x-1)+x}{(2x-1)-x+1} = 2$" then we may conclude the new statement "$\exists y \in \mathbb{R} : \frac{y+x}{y-x+1} = 2$".

If somewhere earlier in the proof we have the two statements "$\log\left(1 + \frac{\varepsilon}{e^a}\right) \in \mathbb{R}^+$" and "$\forall x \in \left[a - \log\left(1 + \frac{\varepsilon}{e^a}\right), a + \log\left(1 + \frac{\varepsilon}{e^a}\right)\right] : |e^x - e^a| < \varepsilon$", then we may conclude the new statement $\exists \delta > 0 : \forall x \in [a - \delta, a + \delta] : |e^x - e^a| < \varepsilon$".

11. Recall a theorem or a definition. You may anywhere in your proof write down the statement of a definition or an already known theorem (or a lemma or a proposition), *written in closed form* (cf. all our examples in §2.6).

We end by several remarks, which may be skipped on a first reading:

Remark 1. Many of the rules above are actually redundant, meaning that we can always achieve the same thing *without* ever using that particular rule. However, we do not worry about this; our aim here is not to give a set of rules which is as small as possible, but instead we are trying to cover as many as possible of the methods used in "everyday mathematics". The following proof schema illustrates how to do the derivation mentioned in rule 4, proof by contradiction, by only using rule 1 and rule 5 ($\Rightarrow$-Introduction and Elementary logic):

(1)	$\mathbf{P}$	(A new assumption!)
(2)	$\cdots$	(Some argument)
(3)	$\mathbf{Q}$ and $\neg\mathbf{Q}$	
(4)	$\mathbf{P} \Rightarrow (\mathbf{Q}$ and $\neg\mathbf{Q})$	($\Rightarrow$-Introduction, closing assumption (1))
(5)	$\neg\mathbf{P}$	(Elementary logic, from (4))

Here using rule 4, proof by contradiction, we could have gone directly from (3) to (5) (again closing assumption (1)), but we see that we can achieve the same thing in just one more step using $\Rightarrow$-Introduction and Elementary logic.

Here is another proof schema which shows how the $\exists$-introduction rule can be avoided by using the $\forall$-elimination and the exchange $\neg\exists \leftrightarrow \forall\neg$ rule. As in the $\exists$-introduction rule we assume that $\mathbf{P}$ is a statement containing a free variable α, that $*$ is an expression such that no variable which occurs in $*$ occurs as a bound variable in $\mathbf{P}$, and that $\mathbf{Q}$ is the statement "[$\mathbf{P}$ with every occurrence of α replaced with $*$]." Note that using the $\exists$-introduction rule we may go directly from (1), (2) to (7) below; however the proof shows that the same thing can be achieved without using the $\exists$-introduction rule:

(1)	$* \in M$	
(2)	$\mathbf{Q}$	
(3)	$\neg\exists\alpha \in M : \mathbf{P}$	A new assumption!
(4)	$\forall\alpha \in M : \neg\mathbf{P}$	Exchange $\neg\exists \leftrightarrow \forall\neg$, from (3)
(5)	$\neg\mathbf{Q}$	$\forall$-elimination using (1) and (4)
(6)	$(\neg\exists\alpha \in M : \mathbf{P}) \Rightarrow \neg\mathbf{Q}$	$\Rightarrow$-introduction, closing assumption (3)
(7)	$\exists\alpha \in M : \mathbf{P}$	Elementary logic from (2) and (6)

Remark 2. The $\exists$-introduction rule (rule number 10) is often used together with the substitution rule. For example, in the last example above, we assume that we have already succeeded to prove the statement "$\forall x \in \left[a - \log\left(1 + \frac{\varepsilon}{e^a}\right), a + \log\left(1 + \frac{\varepsilon}{e^a}\right)\right] : |e^x - e^a| < \varepsilon$". However, the part of the proof leading to this probably looks much simpler if we introduce some notation for the special number $\log\left(1 + \frac{\varepsilon}{e^a}\right)$; let's call it δ! Then the relevant part of the proof might schematically look as follows:

(1) $\delta = \log\left(1 + \frac{\varepsilon}{e^a}\right)$ (A new assumption!)

(2) $\cdots$ (some clever argument)

(3) $\delta \in \mathbb{R}^+$ and $\forall x \in [a-\delta, a+\delta] : |e^x - e^a| < \varepsilon$ (conclusion from the clever argument)

(4) $\delta = \log\left(1 + \frac{\varepsilon}{e^a}\right) \Rightarrow \Big(\delta \in \mathbb{R}^+$ and $\forall x \in [a-\delta, a+\delta] : |e^x - e^a| < \varepsilon\Big)$

($\Rightarrow$-introduction, closing assumption (1))

(5) $\log\left(1 + \frac{\varepsilon}{e^a}\right) \in \mathbb{R}^+$ and $\forall x \in \left[a - \log\left(1 + \frac{\varepsilon}{e^a}\right), a + \log\left(1 + \frac{\varepsilon}{e^a}\right)\right] : |e^x - e^a| < \varepsilon$

(substitution, from (4))

(6) $\exists\delta > 0 : \forall x \in [a-\delta, a+\delta] : |e^x - e^a| < \varepsilon.$ ($\exists$-introduction, from (5))

Since this type of substitution following an $\Rightarrow$-introduction is very common and standard, one usually does not write it out. Thus the above proof may look as follows when written in common words: "Set $\delta = \log\left(1 + \frac{\varepsilon}{e^a}\right)$. Then [here give a clever argument....] we conclude $\delta \in \mathbb{R}^+$ and $\forall x \in [a-\delta, a+\delta] : |e^x - e^a| < \varepsilon$. Hence we have proved that $\exists\delta > 0 : \forall x \in [a-\delta, a+\delta] : |e^x - e^a| < \varepsilon$."

Remark 3. The rule about recalling a theorem or definition (rule number 11) is almost always used together with the $\forall$-Elimination rule (repeatedly used). We give an example: Suppose that in our proof we have arrived at the statement "$\forall x \in U : \exists\varepsilon > 0 : B(x, \varepsilon) \subset U$", where U is some subset of a metric space D. That is, we have the following lines somewhere in our proof:

(1)	D is a metric space
(2)	$U \subset D$
(3)	$\forall x \in U : \exists\varepsilon > 0 : B(x, \varepsilon) \subset U$

We may then continue as follows:

(4) $\forall X$ [metric space] : $\forall M \in \mathcal{P}(X) : \Big[M$ is open $\Longleftrightarrow \Big(\forall x \in M : \exists r > 0 : B(x, r) \subset M\Big)\Big]$

(Recalling Definition 1.3-2)

(5) $\forall M \in \mathcal{P}(D) : \Big[M$ is open $\Longleftrightarrow \Big(\forall x \in M : \exists r > 0 : B(x, r) \subset M\Big)\Big]$ ($\forall$-Elimination using (4) and (1))

(6) $\Big[U$ is open $\Longleftrightarrow \Big(\forall x \in U : \exists r > 0 : B(x, r) \subset U\Big)\Big]$ ($\forall$-Elimination using (5) and (2))

(7) U is open (Elementary logic on (3) and (6))

Usually these steps (4)–(7) are considered obvious, so that in a common mathematical text, the proof following after statement number (3) may look something like: "Hence by Definition 1.3-2 we see that U *is open*."

3.2. **Proof strategies.** When trying to find a proof, one finds oneself in the following situation: We are *assuming* that certain statements $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n$ hold, and we *want to prove* a certain statement $\mathbf{P}$. We propose the following (very naive!) strategy for solving this task: Find some rule among (1)–(15) below that seems to fit well to your situation. Applying your chosen rule means that you have to write some lines of text into your proposed proof, and after this your task has been modified to a *new* task: Assuming some *new* statements $\mathbf{A}'_1, \mathbf{A}'_2, ..., \mathbf{A}'_m$ (possibly same as before), and we want to prove some *new* statement $\mathbf{P}'$. Now *repeat from start!*

> NOTE: THE FOLLOWING LIST IS PROBABLY STILL *VERY* INCOMPLETE. I HOPE TO ADD MORE OF THE MOST USUAL AND BASIC APPROACHES.

(1) Try to remember some known result (a theorem or a lemma) which says directly that $\mathbf{P}$ follows from (some of) $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n$.
[The formal proof will use rules 11 and 8 in §3.1; cf. Remark 3.]

(2) If $\mathbf{P}$ is of the form "$\forall x \in M : \mathbf{Q}$" then start the proof by "Let x be an arbitrary element in M". Then try to find an argument which, from this assumption (together with the old assumptions $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n$) proves the statement $\mathbf{Q}$. Finally write the last sentence: "This is true for every element x in M, hence we have proved $\forall x \in M : \mathbf{Q}$". This completes the task.
[The formal proof will use rule 2 in §3.1; cf. Remark 3.]

(3) If $\mathbf{P}$ is of the form "$\exists x \in M : \mathbf{Q}$" then try to find some *construction* of a specific element x in M for which you can prove that $\mathbf{Q}$ holds. This often requires creativity, and this is often where the "core" of a proof lies! *If* you succeed, then the proof will look something like: "Now set $x =$[your formula/construction]. Note that this x indeed lies in the set M. Furthermore,........; hence $\mathbf{Q}$ holds for our specific choice of x. Hence we have proved $\exists x \in M : \mathbf{Q}$."
[The formal proof will use rules 10 and 9 in §3.1; cf. Remark 2.]

(4) Alternatively, again if $\mathbf{P}$ is of the form "$\exists x \in M : \mathbf{Q}$", then you may make the *assumption* $\forall x \in M : \neg\mathbf{Q}$, and try to see if you can derive a *contradiction* from this. The proof will look as follows: "Assume $\forall x \in M : \neg\mathbf{Q}$. Then .... But this is a contradiction. Hence the assumption $\forall x \in M : \neg\mathbf{Q}$ cannot hold. Hence we have proved $\exists x \in M : \mathbf{Q}$".
[The formal proof will use rules 4 and 7 in §3.1.]

(5) If $\mathbf{P}$ is of the form "$\neg\forall x \in M : \mathbf{Q}$": Note that $\mathbf{P}$ is synonymous to "$\exists x \in M : \neg\mathbf{Q}$"; now try to use rule (3) or (4) to prove this statement!

(6) If **P** is of the form "$\neg\exists x \in M$: **Q**": Note that **P** is synonymous to "$\forall x \in M$: $\neg$**Q**"; now try to use rule (2) to prove this statement!

(7) If **P** is of the form "**Q** $\Rightarrow$ **R**" then start the proof by "Assume that **Q** holds." Then try to find an argument which, from this assumption (together with the old assumptions $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n$), proves the statement **R**. Finally write the last sentence: "Hence **Q** $\Rightarrow$ **R**".

(8) Alternatively, again if **P** is of the form "**Q** $\Rightarrow$ **R**": Note that this is logically equivalent to $\neg$**R** $\Rightarrow$ $\neg$**Q**; hence try to apply (8) to this format instead!

(9) If **P** is of the form "**Q** and **R**" then you first have to find a proof of **Q** from the given assumptions $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n$, and then find a proof of **Q** from the given assumptions $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n$. Finally write the last sentence: "Since both **Q** and **R** hold, it follows that [**Q** and **R**] hold."

(10) If **P** is of the form "$\neg$(**Q** and **R**)": Note that by elementary logic **P** is equivalent to "$\neg$**Q** or $\neg$**R**". Hence try to apply (11) or (12) below.

(11) If **P** is of the form "**Q** or **R**" then try to find an assumption which from the assumption $\neg$**Q** (together with the old assumptions $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n$) proves the statement **R**. If you succeed, the proof will look something like: "Assume $\neg$**Q**. Then .... Hence **R** holds. This was under the assumption $\neg$**Q**; hence we have proved in general that [**Q** or **R**] holds." (Note that this strategy is unsymmetric in **Q** and **R**, so you can choose yourself which inner part of **P** to call "**Q**" and which part to call "**R**".)

(12) Alternatively, again if **P** is of the form "**Q** or **R**" then you may actually be in a lucky situation where you can prove "**Q**" from the given assumptions! If this happens, the proof will look something like: "............ Hence **Q** holds. Hence also [**Q** or **R**] holds."

(13) If **P** is of the form "x is blob" or "x, y are blob" or "x blob y" or "$\text{blob}(x) = y$" or "$\text{blob}(x) \leqq y$" or similar, where "blob" is some concept or relation or function that you know the definition of: Use the definition of "blob" (alternatively, use a known result (a lemma or a theorem) which characterizes "blob") to translate **P** into a synonymous statement **Q** which does not contain "blob". Then try to find a proof of **Q** (using the existing assumptions $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n$). Finally write the last statement: "By the definition of "blob", this means that **P** holds."

(14) Alternatively, if one of our *assumptions* $\mathbf{A}_j$ is a statement of the form "x is blob" or "x, y are blob" or "x blob y" or "blob$(x) = y$" or "blob$(x) \leqq y$" or similar, where "blob" is some concept or relation or function that you know the definition of: Use the definition of "blob" (alternatively, use a known result (a lemma or a theorem) which characterizes "blob") to translate $\mathbf{A}_j$ into a synonymous statement $\mathbf{A}'_j$ which does not contain "blob". Now you can use $\mathbf{A}'_j$ as assumption!

(15) If $\mathbf{P}$ is of the form "$x \in M$" (where M is some set) then look into the definition of M to see what we have to prove. For example, if $\mathbf{P}$ is "$x \in \{y : \mathbf{Q}\}$", where $\mathbf{Q}$ is some statement involving the variable y, then our task is to prove that the statement [$\mathbf{Q}$ with y everywhere replaced by x] is true. If we can prove this, then we can write as last sentence: "Hence $x \in \{y : \mathbf{Q}\}$ holds."

Note that often there are many choices of which rule to use! It is a matter of experience and skill to be able to make as fruitful choices as possible!

Examples of more specialized strategies

When you become fluent in using the strategies in §3.2, the following strategies are things which you would *easily find out (or reconstruct) yourself!* We stress again that *all the stuff that requires true creativity is still left unsolved in the proof schemes below, as "..........."!*

(16) To prove the statement "$M \subset N$" (where M, N are two sets)

The standard proof looks like this:

"Let x be an arbitrary element in M. Then .... Hence $x \in N$. This is true for all elements x in M, hence we have proved $M \subset N$."

Here one has to fill in the ".........." with a proof of "$x \in N$", and in this proof we may use the assumption "$x \in M$".

Remark on how you could have thought of this yourself, using (1)–(15) *above:* The above proof scheme comes from using the definition of "$M \subset N$" (thus, following rule (13) above); this definition says that "$M \subset N$" is synonymous with "$\forall x \in M : x \in N$. In order to prove this statement, we use rule (2) above. This gives the proof scheme above. (To make it clear, let's write out the above proof scheme in slightly more detail – more detail than is needed in a usual mathematical text:

"Let x be an arbitrary element in M. Then .... Hence $x \in N$. This is true for all elements x in M, hence we have proved $\forall x \in M : x \in N$. Hence, by definition, we have $M \subset N$."

(17) To prove the statement "$M \not\subset N$" (i.e., "$\neg(M \subset N)$").

One standard proof is to construct an element which belongs to M but not to N. The proof will look something like:

> "Set $x =$ .... Note that $x \in M$. Furthermore ...., hence $x \notin N$. Since $x \in M$ and $x \notin N$, it follows that $M \not\subset N$.

Here one has to fill in the first "........." with an appropriate (clever) choice (or construction) of x (which must belong to M). This choice should be such that we can replace the second "........." with a proof of the statement $x \notin N$.

Remark on how you could have thought of this yourself, using (1)–(15) *above:* The above proof scheme again comes from using the definition of "$M \subset N$"; since "$M \subset N$" is synonymous with "$\forall x \in M : x \in N$ we see that "$M \not\subset N$" is synonymous with "$\neg\forall x \in M : x \in N$". In order to prove this statement, we use rule (5), which says that the statement which we want to prove is "$\exists x \in M : \neg(x \in N)$". Finally we apply rule (3). This leads to the proof schema above.

(18) To prove the statement "$M = N$" (where M, N are two sets)

Try to prove $M \subset N$ and $N \subset M$ (for this: see (16) above). If you can prove both of these then $M = N$ holds!

(19) To prove that a subset M of a metric space X is open

The standard proof looks like this:

> "Let x be an arbitrary point in M. Now set $\varepsilon =$ .... Note that $\varepsilon > 0$. Furthermore.................. Hence $B(x, \varepsilon) \subset M$. Hence we have proved $\exists \varepsilon > 0 : B(x, \varepsilon) \subset M$. This is true for every x in M. Hence by definition 1.3-2, M *is an open set.*"

Here one has to fill in the first "........" with an appropriate choice of ε (must be a positive real number!), which is chosen cleverly so that we can replace the second ".........." with a proof of the statement "$B(x, \varepsilon) \subset M$".

For a more detailed example of this, and motivation how one sees the above proof scheme using only the more general rules from §3.2; see Problem 1 in §3.3 below (where the above is done for the case $M = B(x_0, r)$).

(20) To prove that a subset M of a metric space X is *not* open

One standard proof looks like this:

> "Let x be the point .... in M. Let $\varepsilon > 0$ be an arbitrary number. Let y be the point .... in $B(x, \varepsilon)$. We now have...............; hence $y \notin M$! Hence $B(x, \varepsilon) \not\subset M$. But $\varepsilon > 0$ is arbitrary; hence there does *not* exist any $\varepsilon > 0$ such that $B(x, \varepsilon) \subset M$. Hence by definition 1.3-2, M *is not an open set.*"

Here one has to fill in the first "........" with an appropriate choice (or construction) of a point x in M. Then one has to fill in the second "..........." with an appropriate choice (or construction) of a point y, *which may depend on* ε, and which must belong to the set $B(x,\varepsilon)$. These choices should be cleverly made so that we are able to fill in the third "............" with a proof of the statement "$y \notin M$"!

Remark on how you could have thought of this yourself, using (1)–(15) *above:* First (via rule (13)) use Kreyszig's definition 1.3-2 to reformulate the statement which we wish to prove, "M is not an open set", into: "$\neg\forall x \in M : \exists\varepsilon > 0 : B(x,\varepsilon) \subset M$". We next apply rule (5) and (3) to see that our task is to construct some $x \in M$ such that "$\neg\exists\varepsilon > 0 : B(x,\varepsilon) \subset M$". To this we apply rule (6) and (2) to see that our task is to prove, for each $\varepsilon > 0$, that "$\neg(B(x,\varepsilon) \subset M)$". For this we apply the rule (17) above. Then we arrive at the above proof scheme!

(21) To prove that a subset M of a metric space X is closed

One standard proof, using Theorem 1.4-6, looks like this:

> "Let (x_n) be an arbitrary sequence of points in M. Assume that $x_n \to x$ for some $x \in X$. Then......... Hence $x \in M$. This is true whenever x is the limit of a sequence (x_n) of points in M. Hence by Theorem 1.4-6(b), M is closed.

Here it remains to fill in the "........" with a proof of "$x \in M$", and in this proof we may use the two assumptions "$(x_n) \in \mathcal{S}(M)$" and "$x_n \to x$". (Note: The first step to find out such a proof may be to apply rule (15) to the statement "$x \in M$".)

Remark: A *different* approach to proving that M is closed is to try to use Definition 1.4-1 directly.

(22) To prove that a subset M of a metric space X is *not* closed

One standard proof, using Theorem 1.4-6, looks like this:

> "Set $x_n =$ .... for each $n \in \mathbb{Z}^+$. Note that each x_n lies in M, i.e. (x_n) is a sequence of points in M. Now .... Hence $x_n \to x$ (where $x \in X$). Now .... Hence $x \notin M$. Hence there is a sequence of points in M which converges to a point in X outside M. This implies, by Theorem 1.4-6(b), that M is not closed.

Here the first "..........." should be filled in by a clever choice (construction) of a sequence of points in x_n. This sequence should be such that we can fill in the second ".........." with a proof that $x_n \to x$, where x is a specific point in X (which you will probably have to compute). Your sequence should also be such that we can fill in the third ".........." with a proof that $x \notin M$.

Remark: A *different* approach to proving that M is not closed is to try to use Definition 1.4-1 directly.

(23) To prove that a subset M of a metric space X is bounded

One standard proof looks like this:

"Let x and y be an arbitrary elements in M. Then .... Hence $d(x,y) \leqq C$. (Note that C is independent of x, y!) This is true for arbitrary $x, y \in M$. Hence, by definition, M is bounded."

Here it remains to replace ".........." with a proof of the statement "$d(x,y) \leqq C$", where we may use the assumptions $x \in M$ and $y \in M$. Note that the constant "C" which we achieve here must be *independent of* x, y*!* Remark: Regarding the definition of "bounded", see the exercise which we treated above on p. 9.

(24) To prove that a subset M of a metric space X is *not* bounded

One standard proof looks like this:

"Let C be an arbitrary positive number. Then set $x =$ .... and $y =$ ....; note that x and y are elements in M. Now ...., hence $d(x,y) > C$. But here C was an arbitrarily large number. Hence, by definition, M is not bounded.

Here it remains to replace the first two ".........." with clever choices (constructions) of the points x and y (which must belong to M). Note that your choice of these two points may (actually: must) *depend on the constant* C*!* Finally, the third "........" should be replaced with a proof of the statement "$d(x,y) > C$".

(25) To prove that a sequence (x_n) is convergent

One approach is to somehow compute or guess the limit x, and then write out a proof of "$x_n \to x$" using the definition 1.4-1. (You may work out a more precise proof scheme yourself using the rules (1)–(15) and our discussion of Definition 1.4-1 in §2.6.)

Another approach is as follows: If (x_n) is a sequence in a metric space X which we know to be *complete*, then try to prove that "(x_n) is Cauchy", via Definition 1.4-3. It then follows from Definition 1.4-3 (the second part) that (x_n) is actually convergent.

You may try to collect many more examples of specific proof strategies yourself, by practicing, and also by studying other people's solutions!

3.3. **Examples.** In this section we show very carefully how the rules given in §3.2 are used to find the proofs solving some problems. (Actually, the rules only help us do the "routine" part of the solution.)

1. Prove that any open ball is an open set. (Kreyszig; §1.3, problem 1 (p.23)):

Discussion leading to solution. First we need to understand exactly what we are asked to prove. It is clear from the context of Kreyszig, §1.3 that the statement should be proved *for an arbitrary metric space* X. Also note that "open ball" is defined in Def

1.3-1 (p.18); it is written by the symbol $B(x_0, r)$ it depends on two parameters x_0, r, where $x_0 \in X$ and r is a real number $r > 0$. We have to prove that *any* open ball is an open set, i.e. that every possible open ball in the metric space X is an open set. To consider every possible open ball is the same thing as considering $B(x_0, r)$ for every $x_0 \in X$ and every $r > 0$. Hence the precise statement which we are asked to prove is the following:

$$(*) \qquad \forall X \text{ [metric space]} : \forall x_0 \in X : \forall r > 0 : \Big(B(x_0; r) \text{ is an open set}\Big).$$

We now apply our basic strategy rules in §3.2. Since the statement (*) is of the form "$\forall X$ [metric space] : **Q**", we use rule number (2). We may as well use this rule for all three "$\forall$" at once. The rule says that the proof should look as follows:

> Let X be an arbitrary metric space. Let x_0 be an arbitrary point in X and let r be an arbitrary positive real number. Then ....
> Hence $B(x_0, r)$ is an open set. This is true for all $x_0 \in X$ and $r > 0$, and for every metric space X; hence we have proved "$\forall X$ [metric space] : $\forall x_0 \in X : \forall r > 0 : \Big(B(x_0; r)$ is an open set$\Big)$", i.e. the problem is solved.

It now remains to fill in ".........." above; we must replace ".........." with a proof of "$B(x_0, r)$ is an open set" (and in this proof we may use the given assumptions about X, x_0, r). We again use the rules in §3.2. There isn't really any other rule to use than (13), i.e. we should use definitions to translate into a new (more "basic") statement. There are *two* concepts which we may translate: "$B(x_0, r)$" and "is an open set". If we start by translating "is an open set" (using Def 1.3-2) then the statement which we want to prove looks as follows: "$B(x_0, r)$ contains a ball about each of its points", i.e. in symbols:

$$\forall x \in B(x_0, r) : \exists \varepsilon > 0 : B(x, \varepsilon) \subset B(x_0, r).$$

(Note that we have to use a different letter than "r" for the radius of the new, since r is already occupied.) With this, our proof should look as follows:

> Let X be an arbitrary metric space. Let x_0 be an arbitrary point in X and let r be an arbitrary positive real number. Then ....
> Hence $\forall x \in B(x_0, r) : \exists \varepsilon > 0 : B(x, \varepsilon) \subset B(x_0, r)$. Hence by definition 1.3-2, $B(x_0, r)$ is an open set. This is true for all $x_0 \in X$ and $r > 0$, and for every metric space X; hence we have proved "$\forall X$ [metric space] : $\forall x_0 \in X : \forall r > 0 : \Big(B(x_0; r)$ is an open set$\Big)$", i.e. the problem is solved.

It remains to replace ".........." by a proof of

$$(**) \qquad \forall x \in B(x_0, r) : \exists \varepsilon > 0 : B(x, \varepsilon) \subset B(x_0, r).$$

(and in this proof we may use the given assumptions about X, x_0, r). There are now no less than three occurrences of the notation $B(x, r)$, and each of these we could translate using definition 1.3-1; however it may be better to just remember in our head this definition 1.3-1, and instead continue applying the "logic reducing" parts of rules in §3.2. Since the desired statement (**) is of the form $\forall x \in B(x_0, r) : \exists \varepsilon > 0 : \mathbf{Q}$, we use rule number

(2), and then rule number (3) on the inner statement. With this, our proof should look as follows:

Let X be an arbitrary metric space. Let x_0 be an arbitrary point in X and let r be an arbitrary positive real number. Let x be an arbitrary point in $B(x_0, r)$. Now set $\varepsilon =$ .... Note that $\varepsilon > 0$. .... Hence $B(x, \varepsilon) \subset B(x_0, r)$. Hence we have proved $\exists \varepsilon > 0 : B(x, \varepsilon) \subset B(x_0, r)$. This is true for every x in $B(x_0, r)$. Hence $\forall x \in B(x_0, r) : \exists \varepsilon > 0 : B(x, \varepsilon) \subset B(x_0, r)$. Hence by definition 1.3-2, $B(x_0, r)$ *is an open set*. This is true for all $x_0 \in X$ and $r > 0$, and for every metric space X; hence we have proved "$\forall X$ [metric space] : $\forall x_0 \in X : \forall r > 0 : \Big(B(x_0; r)$ is an open set$\Big)$", i.e. the problem is solved.

It now remains to fill in the two ".........." above; we must replace the first ".........." by an appropriate choice of ε, which is chosen cleverly so that we can replace the second "............" with a proof of

$$B(x, \varepsilon) \subset B(x_0, r).$$

For this we use rule (16) in §3.2; with this our proof should look as follows:

Let X be an arbitrary metric space. Let x_0 be an arbitrary point in X and let r be an arbitrary positive real number. Let x be an arbitrary point in $B(x_0, r)$. Now set $\varepsilon =$ .... Note that $\varepsilon > 0$. Now let p be an arbitrary point in $B(x, \varepsilon)$. .... Hence $p \in B(x_0, r)$. This is true for every $p \in B(x, \varepsilon)$, hence we have proved $B(x, \varepsilon) \subset B(x_0, r)$. Hence we have proved $\exists \varepsilon > 0 : B(x, \varepsilon) \subset B(x_0, r)$. This is true for every x in $B(x_0, r)$. Hence $\forall x \in B(x_0, r) : \exists \varepsilon > 0 : B(x, \varepsilon) \subset B(x_0, r)$. Hence by definition 1.3-2, $B(x_0, r)$ *is an open set*. This is true for all $x_0 \in X$ and $r > 0$, and for every metric space X; hence we have proved "$\forall X$ [metric space] : $\forall x_0 \in X : \forall r > 0 : \Big(B(x_0; r)$ is an open set$\Big)$", i.e. the problem is solved.

The remaining task is to replace the second "..........." with a proof of "$p \in B(x_0, r)$", and in this proof we may use all the assumptions made about X, x_0, r, x, p, and we must also fill in a clever choice of ε (in the first ".........") that makes this proof work. Since the desired statement is "$p \in B(x_0, r)$" we can only use rule (15) and/or (13) in §3.2; both these rules tell us to use definition 1.3-1 to see that $B(x_0, r)$ is the same thing as[5] $\{y \in X : d(y, x_0) < r\}$ Hence "$p \in B(x_0, r)$" is synonymous with "$p \in \{y \in X : d(y, x_0) < r\}$", and this is synonymous with "$d(p, x_0) < r$". Similarly, we may also translate the two assumptions which involve $B(...)$. With this our proof should look as follows:

[5]Here, for safety, we use y instaed of x as the "running set variable", since x is already occupied in our proof.

> Let X be an arbitrary metric space. Let x_0 be an arbitrary point in X and let r be an arbitrary positive real number. Let x be an arbitrary point in $B(x_0, r)$, i.e. (by def 1.3-1) $x \in X$ and $d(x, x_0) < r$. Now set $\varepsilon =$ .... Note that $\varepsilon > 0$. Now let p be an arbitrary point in $B(x, \varepsilon)$, i.e. (by def 1.3-1) $p \in X$ and $d(x, p) < \varepsilon$. .... Hence $d(p, x_0) < r$. By def 1.3-1 this implies that $p \in B(x_0, r)$. This is true for every $p \in B(x, \varepsilon)$, hence we have proved $B(x, \varepsilon) \subset B(x_0, r)$. Hence we have proved $\exists \varepsilon > 0 : B(x, \varepsilon) \subset B(x_0, r)$. This is true for every x in $B(x_0, r)$. Hence $\forall x \in B(x_0, r) : \exists \varepsilon > 0 : B(x, \varepsilon) \subset B(x_0, r)$. Hence by definition 1.3-2, $B(x_0, r)$ *is an open set.* This is true for all $x_0 \in X$ and $r > 0$, and for every metric space X; hence we have proved "$\forall X$ [metric space] : $\forall x_0 \in X : \forall r > 0 : \Big(B(x_0; r)$ is an open set$\Big)$", i.e. the problem is solved.

Now there is not much more reduction that can be done; we have to face the task of trying to come up with the creative definition of ε and then fill in the second ".........." with the remaining proof. In this proof we are allowed to make use of all assumptions which were made before this second "..........". Let us carefully organize these assumptions: We are assuming that X *is a metric space,* x_0, x, p *are points in* X, r *and* ε *are positive real numbers, and we know that* $d(x, x_0) < r$ *and* $d(x, p) < \varepsilon$. Also, we are actually allowed to *choose* ε, in a way that may depend on X, x_0, r, x but *not* on p (since in our proof X, x_0, r, x have been chosen *before* we make the choice of ε). Under these assumptions, we want to prove the statement "$d(p, x_0) < r$."

Up to this point, the work has been *routine.* The explanation of how to come up with a good definition of ε lies partly outside the purpose of the present text; however we attempt an explanation in this particular case.

Let us spell out our task in an even denser form: We wish to choose $\varepsilon > 0$ in such a way that we can prove $d(p, x_0) < r$ from $d(x, x_0) < r$ and $d(x, p) < \varepsilon$. [One may also draw a picture with the points p, x_0, x, etc!] Hopefully now we think of *the triangle inequality!* (Kreyszig p.3, (M4).) In fact, in an arbitrary metric space the triangle inequality is the *only* known relation which connects the three distances $d(p, x_0)$, $d(x, x_0)$, $d(x, p)$. Since we wish to prove that $d(p, x_0)$ is small it seems reasonable to use the triangle inequality in the form

$$d(p, x_0) < d(p, x) + d(x, x_0) = d(x, p) + d(x, x_0).$$

(We here also used the symmetry, Kreyszig p.3 (M3).) If we use this inequality, and also the assumptions $d(x, p) < \varepsilon$ and $d(x, x_0) < r$ we obtain

$$d(p, x_0) < d(x, p) + d(x, x_0) < \varepsilon + r.$$

This is *worse* than the statement which we wanted to prove. However, we don't really have to use both inequalities $d(x, p) < \varepsilon$ and $d(x, x_0) < r$ in that way. Let us first use *only* $d(x, p) < \varepsilon$: [6]

$$d(p, x_0) < d(x, p) + d(x, x_0) < \varepsilon + d(x, x_0).$$

[6] We cannot really lose anything by using this inequality, since p is chosen after ε and thus we cannot have any more precise information about $d(x, p)$ than just "$d(x, p) < \varepsilon$".

We want to make the right hand side $\leqq r$, and we are free to choose ε depending on X, x_0, r, x! Hence let us set $\varepsilon = r - d(x, x_0)$![7] It is important to check that this ε belongs to the permitted set, in this case $\mathbb{R}^+$, the set of positive real numbers; this is true because of the assumption $d(x, x_0) < r$! (Note how nicely all our assumptions now come into play.) And the choice $\varepsilon = r - d(x, x_0)$ also clearly gives us the desired inequality:

$$d(p, x_0) < d(x, p) + d(x, x_0) < \varepsilon + d(x, x_0) = (r - d(x, x_0)) + d(x, x_0) = r, \qquad \text{i.e. } d(p, x_0) < r.$$

Hence we finally obtain a complete solution to our problem:

Solution.

Let X be an arbitrary metric space. Let x_0 be an arbitrary point in X and let r be an arbitrary positive real number. Let x be an arbitrary point in $B(x_0, r)$, i.e. (by def 1.3-1) $x \in X$ and $d(x, x_0) < r$. Now set $\varepsilon = r - d(x, x_0)$. Note that $\varepsilon > 0$ (since $d(x, x_0) < r$). Now let p be an arbitrary point in $B(x, \varepsilon)$, i.e. (by def 1.3-1) $p \in X$ and $d(x, p) < \varepsilon$. By the triangle inequality we have:

$$d(p, x_0) < d(x, p) + d(x, x_0) < \varepsilon + d(x, x_0) = \big(r - d(x, x_0)\big) + d(x, x_0) = r.$$

Hence $d(p, x_0) < r$. By def 1.3-1 this implies that $p \in B(x_0, r)$. This is true for every $p \in B(x, \varepsilon)$, hence we have proved $B(x, \varepsilon) \subset B(x_0, r)$. Hence we have proved $\exists \varepsilon > 0 : B(x, \varepsilon) \subset B(x_0, r)$. This is true for every x in $B(x_0, r)$. Hence $\forall x \in B(x_0, r) : \exists \varepsilon > 0 : B(x, \varepsilon) \subset B(x_0, r)$. Hence by definition 1.3-2, $B(x_0, r)$ *is an open set.* This is true for all $x_0 \in X$ and $r > 0$, and for every metric space X; hence we have proved "$\forall X$ [metric space] $: \forall x_0 \in X : \forall r > 0 : \Big(B(x_0; r) \text{ is an open set}\Big)$", i.e. the problem is solved.

4. Review of limit, limsup, liminf, sup and inf

In this section we review some basic asymptotic concepts for sets of real numbers. This also give some further examples of the formalism and routines introduced in earlier sections. We start with defining "$\lim_{n\to\infty} a_n$", although we expect that the reader already has some familiarity with this concept.

Definition 4.1. If (a_n) is a sequence of real numbers (i.e., $(a_n) \in \mathcal{S}(\mathbb{R})$) and $b \in \mathbb{R}$, then we define "$\lim_{n\to\infty} a_n = b$" to mean

$$\forall \varepsilon > 0 : \exists N \in \mathbb{Z}^+ : \forall n \geqq N : |a_n - b| < \varepsilon.$$

We also review the concepts "$\lim_{n\to\infty} a_n = \pm\infty$":

[7] We could also choose $\varepsilon = \frac{1}{10}(r - d(x, x_0))$; this would give "some room to spare" in our inequalities.

Definition 4.2. If (a_n) is a sequence of real numbers (i.e., $(a_n) \in \mathcal{S}(\mathbb{R})$), then we define "$\lim_{n\to\infty} a_n = +\infty$" to mean

$$\forall C \in \mathbb{R} : \exists N \in \mathbb{Z}^+ : \forall n \geqq N : a_n > C.$$

Similarly, we define "$\lim_{n\to\infty} a_n = -\infty$" to mean

$$\forall C \in \mathbb{R} : \exists N \in \mathbb{Z}^+ : \forall n \geqq N : a_n < C.$$

We next define what supremum and infimum of a set of real numbers mean (compare Kreyszig [Kreyszig, p. 619: §A1.6]).

Definition 4.3. Let E be a subset of $\mathbb{R}$ and let $b \in \mathbb{R}$. We say that b is an *upper bound* of E if $x \leqq b$ for all $x \in E$. We say that E is *bounded above* if E has an upper bound. Finally, we say that $a \in \mathbb{R}$ is a *supremum* of E (also called *least upper bound* of E) written $a = \sup E$, if a is an upper bound of E and $a \leqq b$ holds for every upper bound b of E.

Written out as closed statements (similarly to the exercises we did in §2.6), the three definitions above look as follows:

$$\forall E \in \mathcal{P}(\mathbb{R}) : \forall b \in \mathbb{R} : \Big[\underline{b \text{ is an upper bound of } E} \Longleftrightarrow \forall x \in E : x \leqq b\Big];$$

$$\forall E \in \mathcal{P}(\mathbb{R}) : \Big[\underline{E \text{ is bounded above}} \Longleftrightarrow \exists b \in \mathbb{R} : b \text{ is an upper bound of } E\Big];$$

and

$$\forall E \in \mathcal{P}(\mathbb{R}) : \forall a \in \mathbb{R} :$$
$$\Big[\underline{a = \sup E} \Longleftrightarrow \Big[\Big(a \text{ is an upper bound of } E\Big) \text{ and } \forall b \in \mathbb{R} : \Big(\Big(b \text{ is an upper bound of } E\Big) \Rightarrow a \leqq b\Big)\Big]\Big].$$

We also write out the definition of "$a = \sup E$" with the concept "upper bound" expanded:

$$\forall E \in \mathcal{P}(\mathbb{R}) : \forall a \in \mathbb{R} :$$
$$\Big[\underline{a = \sup E} \Longleftrightarrow \Big[\Big(\forall x \in E : x \leqq a\Big) \text{ and } \forall b \in \mathbb{R} : \Big(\Big(\forall x \in E : x \leqq b\Big) \Rightarrow a \leqq b\Big)\Big]\Big].$$

Theorem 1. *If E is a nonempty set of real numbers which is bounded above, then E has a supremum.*

This is a basic theorem about real numbers which we will not prove, since it depends on how we *define* the set of real numbers and we do not wish to discuss these matters here. We will simply take the theorem as part of our background knowledge about real numbers.

Note that it is clear from Definition 4.3 that if E is a set of real numbers which *does not* have an upper bound, then E does not have a supremum either. In this situation we will write $\sup E = \infty$. Thus, in view of Theorem 1, a nonempty subset $E \subset \mathbb{R}$ *has a (real) supremum if and only if E is bounded above.*

Next we prove that the supremum is always *unique*.

Theorem 2. *If E is a nonempty set of real numbers which is bounded above, then E has a* ***unique*** *supremum, i.e.* $\sup E$ *is a uniquely defined real number!*

Discussion leading to proof. ......

A very analogous concept to "sup" is "inf":

Definition 4.4. Let E be a subset of $\mathbb{R}$ and let $b \in \mathbb{R}$. We say that b is a *lower bound* of E if $x \geqq b$ for all $x \in E$. We say that E is *bounded below* if E has a lower bound. Finally, we say that $a \in \mathbb{R}$ is an *infimum* of E (also called *greatest lower bound* of E) written $a = \inf E$, if a is a lower bound of E and $a \geqq b$ holds for every lower bound b of E.

STILL TO WRITE: lim inf, lim sup, and more details!

References

[Kreyszig] E. Kreyszig, *Introductory Functional Analysis with Applications*, John Wiley & Sons, 1989.