A NEW PROOF OF CERTAIN METRIZATION THEOREMS

MARION B. SMITH, JR.

1. Introduction. A well-known problem in topology is the so-called metrization problem. This consists of asking for the topological conditions that are necessary and sufficient in order to guarantee that a topological space be metrizable. The first solution of this problem was given in 1923 by P. Alexandroff and P. Urysohn **(1)**. Their proof relied heavily upon the result of Chittenden **(4)** that the notion of uniformly regular écart is equivalent to that of a metric distance function. In 1937 A. H. Frink **(5)** gave a simplified proof of Chittenden's theorem along with a modification of the Alexandroff-Urysohn condition. Then in 1947 R. H. Bing **(2)** published a paper containing a proof of the Alexander-Urysohn condition which did not use the notion of uniformly regular écart. His method allowed him to prove another modification of the original condition for metrizability.

The condition given by Alexandroff and Urysohn is in some ways not as satisfactory a solution of the topological metrization problem as some of the later ones, such as those given by Smirnov **(9)**, Nagata **(7)**, and Bing **(3)**. However, a very interesting feature of the Alexandroff-Urysohn condition is that it can be viewed as a necessary and sufficient condition for the metrization of uniform spaces. This aspect of the result is developed in Weil **(10)** and Kelley **(6)**.

All the previously mentioned proofs of the Alexandroff-Urysohn theorem and of its modifications have been entirely topological. In this paper a certain combinatorial result concerning triangular arrays of real numbers is established, which serves to remove the major obstacle in proving the metrization theorems of this type. The remaining topological arguments are very brief and elementary.

2. Definitions. Let $A = (a_1, a_2, \ldots, a_n)$ be an ordered n-tuple of real numbers. From A we obtain an m-tuple, $B = (b_1, b_2, \ldots, b_m)$, of real numbers which we call a *reduction* of A. The process of obtaining a reduction can be explained in the following way. We select b_1 to be either a_1 or one less than the minimum of a_1 and a_2. If b_1 is chosen to be a_1, then we may choose b_2 to be either a_2 or one less than the minimum of a_2 and a_3. On the other hand if b_1 is chosen to be one less than the minimum of a_1 and a_2, then b_2 may be

Received April 11, 1961. This paper contains a part of a thesis prepared under the supervision of Professor F. B. Jones and submitted for a Ph.D. degree at the University of North Carolina in 1957.

chosen to be either a_3 or one less than the minimum of a_3 and a_4. In this manner we proceed step-by-step along the n-tuple A. At each step we either select the first "unused" element of A or one less than the minimum of the first pair of elements in the "unused" portion of A. We continue until all of A has been used in producing elements of B. If at each step, except possibly the last one, we use two elements of A, that is, we select each term of B, except possibly the last one, to be one less than the minimum of the appropriate pair in A, then we shall say that the resulting m-tuple B is a *standard* reduction of A. In case n is even such a selection is possible at each step and $m = \frac{1}{2}n$. If n is odd, then there is no such selection possible in the final step, and b_m must be chosen equal to a_n, and $m = \frac{1}{2}(n+1)$. In either case we can write $m = [\frac{1}{2}(n+1)]$, where we use the notation $[x]$ to denote the largest integer not greater than x. A precise definition of reduction can be stated as follows. If (1) $m \leqslant n$, (2) there are k terms $(0 \leqslant k \leqslant m)$, $b_{i(1)}, b_{i(2)}, \ldots, b_{i(k)}$, $i(j) < i(j+1)$, of B such that $b_{i(j)} = \{\min(a_{i(j)+j-1}, a_{i(j)+j})\} - 1$, and (3) the remaining $m - k$ terms of B are such that $b_q = a_{q+j}$, where $i(j) < q < i(j+1)$ and $i(0) = 0$, $i(k+1) = m+1$, then B is said to be a *reduction* of A. If $k = m$, or if $k = m-1$ with $b_m = a_n$, then B is a *standard reduction* of A.

If A_i is a reduction of A_{i-1} for $i = 1, 2, \ldots, r$, then $A_0, A_1, \ldots, A_r$ is a *reduction sequence*. The reduction sequence is said to be *standard* if A_i is a standard reduction of A_{i-1} for $i = 1, 2, \ldots, r$. The reduction sequence is said to be *complete* if $A_r = (a_{r1})$ and $A_i \neq A_r$ for $i < r$; we shall call a_{r1} a *residue* of A_0.

For $A = (a_1, a_2, \ldots, a_p, \ldots, a_q, \ldots, a_n)$ we shall use the notation $A(p, q)$ to denote $(a_p, a_{p+1}, \ldots, a_q)$.

A reduction sequence $A_0, A_1, \ldots, A_k$ is said to hold the a_{0j} term of A_0 *stable* provided that for $i = 1, 2, \ldots, k$, there are terms $a_{ii(j)}$ such that: (1) $a_{ii(j)} = a_{0j}$; (2) $A_i(1, i(j) - 1)$ is a standard reduction of $A_{i-1}(1, \overline{i-1}(j) - 1)$; (3) $A_i(i(j)+1, n_i)$ is a standard reduction of $A_{i-1}(\overline{i-1}(j)+1, n_{i-1})$.

A reduction sequence $A_0, A_1, \ldots, A_k$ is said to hold *stable* terms $a_{0s(0,1)}$, $a_{0s(0,2)}, \ldots, a_{0s(0,t)}$ of A_0 (where $s(0, i) < s(0, i+1)$) provided that the $a_{0s(0,i)}$ term, for $i = 1, 2, \ldots, t$, is held stable by the reduction sequence $A_j(s(j, i-1)+1,\ s(j, i+1) - 1)$, where $s(j, 0) + 1 = 0$, and $s(j, t+1) - 1 = n_j$, for $j = 0, 1, \ldots, k$.

3. Examples.

1. An example where B is a reduction of A is:

$$\begin{array}{rcccccccc} A = & (11, & 5, & 2, & 13, & 9, & 3, & 7) \\ & \diagdown & \diagup & | & \diagdown & \diagup & | & | \\ B = & (& 4, & 2, & & 8, & 3, & 7) \end{array}$$

2. An example where B is a standard reduction of A is:

$$\begin{array}{rcccccccc}
A = & (11, & 5, & 2, & 13, & 9, & 3, & 7) \\
 & \diagdown & \diagup & \diagdown & \diagup & \diagdown & \diagup & | \\
B = & (& 4, & & 1, & & 2, & 7)
\end{array}$$

3. An example of a complete reduction sequence is:

$$\begin{array}{rcccccccc}
A_0 = & (8, & 10, & 4, & 6, & 11, & 9, & 6) \\
 & \diagdown & \diagup & | & | & \diagdown & \diagup & | \\
A_1 = & (& 7, & 4, & 6, & & 8, & 6) \\
 & & \diagdown & \diagup & | & & \diagdown & \diagup \\
A_2 = & (& & 3, & 6, & & 5 &) \\
 & & & | & \diagdown & & \diagup & \\
A_3 = & (& & 3, & & 4 & &) \\
 & & & \diagdown & & \diagup & & \\
A_4 = & (& & & 2 & & &)
\end{array}$$

In this example 2 is a residue of A_0.

4. An example of a standard, complete reduction sequence is:

$$\begin{array}{rcccccccc}
A_0 = & (8, & 10, & 4, & 6, & 11, & 9, & 6) \\
 & \diagdown & \diagup & \diagdown & \diagup & \diagdown & \diagup & | \\
A_1 = & (& 7, & & 3, & & 8, & 6) \\
 & & \diagdown & \diagup & & & \diagdown & \diagup \\
A_2 = & (& & 2, & & & 5 &) \\
 & & & \diagdown & & \diagup & & \\
A_3 = & (& & & 1 & & &)
\end{array}$$

5. An example of a reduction sequence which holds the second and sixth terms of A_0 stable is:

$$\begin{array}{rccccccccc}
A_0 = & (5, & 3, & 7, & 8, & 5, & 6, & 3, & 8) \\
 & | & | & \diagdown & \diagup & | & | & \diagdown & \diagup \\
A_1 = & (5, & 3, & & 6, & 5, & 6, & & 2\ ) \\
 & | & | & & \diagdown & \diagup & | & & | \\
A_2 = & (5, & 3, & & & 4, & 6, & & 2\ )
\end{array}$$

4. Lemmas.

LEMMA 1. *If A_i is an ordered n_i-tuple and $A_0, A_1, \ldots, A_r$ is a standard reduction sequence, then $n_i = [(n_0 + 2^i - 1)2^{-i}]$ for $i = 0, 1, \ldots, r$.*

Proof. By an earlier remark $n_i = [(n_{i-1} + 1)2^{-1}]$ for $i = 1, 2, \ldots, r$. The lemma follows by mathematical induction.

LEMMA 2. *If $A_0, A_1, \ldots, A_k$ is a reduction sequence holding stable t terms, and A_i is an ordered n_i-tuple, then $n_k \leqslant [\{n_0 + (2^k - 1)(2t + 1)\}2^{-k}]$.*

Proof. Let $a_{0s(0,q)}$, where $q = 1, 2, \ldots, t$, be the terms held stable, and consider A_i and A_{i+1} for some i, $0 \leqslant i < k$. Let $m_0 = s(i, 1) - 1$, $m_t = n_i - s(i, t)$, and $m_p = s(i, p + 1) - s(i, p) - 1$ for $p = 1, 2, \ldots, t - 1$. Let $m_0' = s(i + 1, 1) - 1$, $m_t' = n_{i+1} - s(i + 1, t)$, and $m_p' = s(i+1, p+1) - s(i + 1, p) - 1$ for $p = 1, 2, \ldots, t - 1$. Then

https://doi.org/10.4153/CJM-1963-003-0 Published online by Cambridge University Press

$$n_i = t + \sum_{p=0}^{t} m_p \qquad \text{and} \qquad n_{i+1} = t + \sum_{p=0}^{t} m_p'.$$

Since $m_p' = [(m_p + 1)2^{-1}]$ for $p = 0, 1, \ldots, t$, we have

$$n_{i+1} = t + \sum_{p=0}^{t} [(m_p + 1)2^{-1}] \leqslant \left[t + \sum_{p=0}^{t} (m_p + 1)2^{-1} \right] = [(n_i + 2t + 1)2^{-1}].$$

The lemma follows by mathematical induction.

LEMMA 3. *Let $A_0, A_1, \ldots, A_r$ be a complete reduction sequence, and let $v(0), v(1), \ldots, v(m)$ be a non-decreasing sequence of integers with $v(0) = 0$ and $v(m) \leqslant r$. If $A_{v(i)}, A_{v(i)+1}, \ldots, A_{v(i+1)}$ is a reduction sequence holding stable k_i terms for $i = 0, 1, \ldots, m - 1$, and if $A_{v(m)}, A_{v(m)+1}, \ldots, A_r$ is a standard reduction sequence, then*

$$2^{r-1} < n_0 + \sum_{i=1}^{m} 2(2^{v(i)} - 2^{v(i-1)})k_{i-1}.$$

Proof. From Lemma 2 we have, for $i = 1, 2, \ldots, m$,

$$n_{v(i)} \leqslant [\{n_{v(i-1)} + (2^{v(i)-v(i-1)} - 1)(2k_{i-1} + 1)\}2^{v(i-1)-v(i)}].$$

It follows, by finite induction, that we have

$$n_{v(m)} \leqslant \left[\left\{ n_0 + \sum_{i=1}^{m} (2^{v(i)} - 2^{v(i-1)})(2k_{i-1} + 1) \right\} 2^{-v(m)} \right].$$

This establishes the lemma if $v(m) = r$, and the lemma is trivially true if $r = 0$. Therefore, suppose $r > 0$ and $v(m) < r$, then let $q = r - 1 - v(m)$. Since $A_{v(m)}, A_{v(m)+1}, \ldots, A_{r-1}$ is a standard reduction sequence we have, by Lemma 1, $n_{r-1} = [(n_{v(m)} + 2^q - 1)2^{-q}]$. Therefore,

$$\begin{aligned} n_{r-1} &\leqslant \left[\left\{ n_0 + \sum_{i=1}^{m} (2^{v(i)} - 2^{v(i-1)})(2k_{i-1} + 1) + 2^{r-1} - 2^{v(m)} \right\} 2^{-(r-1)} \right] \\ &\leqslant \left[\left\{ n_0 + \sum_{i=1}^{m} 2(2^{v(i)} - 2^{v(i-1)})k_{i-1} + 2^{r-1} - 1 \right\} 2^{-(r-1)} \right]. \end{aligned}$$

Since $r > 0$, $n_{r-1} = 2$, and the lemma's inequality follows immediately.

LEMMA 4. *If $A_0 = (a_{01}, a_{02}, \ldots, a_{0n})$ and M is the supremum of the set of all residues of A_0, then*

$$\sum_{i=1}^{n} 2^{-a_{0i}} \geqslant 2^{-(M+2)}.$$

Proof. Let p be an integer such that $2(p - 1) < n \leqslant 2p$. Define $B_0 = \{a_{0i} : a_{0i} \in A_0$ and $a_{0i} \leqslant M\}$, and, for $j = 1, 2, \ldots, p$, define $B_j = \{a_{0i} : a_{0i} \in A_0$ and $a_{0i} = M + j\}$. Let b_j be the number of elements in B_j, for $j = 0, 1, \ldots, p$, and assume that:

$$\sum_{j=0}^{p} b_j 2^{-(M+j)} < 2^{-(M+2)}. \tag{1}$$

This implies that $b_0 = b_1 = b_2 = 0$. Now assume that $p < 3$, and let A_0, $A_1, \ldots, A_r$ be a standard, complete reduction sequence. Since $n \leqslant 2p \leqslant 4$, we have $r \leqslant 2$. No term of A_0 is less than $M + 3$, so $a_{r1} \geqslant M + 1$, contradicting that M is the supremum of the residues of A_0. Thus $p \geqslant 3$.

If q is an integer such that $0 \leqslant q \leqslant p - 3$, and if we multiply both sides of (1) by $2^{M+p-q+1}$, we get

$$\sum_{j=0}^{p} 2^{p-q-j+1} b_j < 2^{p-q-1}.$$

Thus,

$$\sum_{j=3}^{p-q} 2^{p-q-j+1} b_j \leqslant 2^{p-q-1} - 2.$$

Summing over q, we obtain

$$\sum_{q=0}^{p-3} \sum_{j=3}^{p-q} 2^{p-q-j+1} b_j \leqslant 2^p - 2p.$$

Changing the order of the summation and simplifying, we find

$$\sum_{j=3}^{p} (2^{p-j+2} - 2) b_j \leqslant 2^p - 2p.$$

Since $n \leqslant 2p$ we have:

$$n + \sum_{j=3}^{p} (2^{p-j+2} - 2) b_j \leqslant 2^p. \tag{2}$$

Let $A_0, A_1, \ldots, A_r$ be a reduction sequence such that: (a) $r \geqslant p$, (b) if, for i and j integers ($0 \leqslant i \leqslant r$, $0 \leqslant j \leqslant p$), we define ${}_iC_j = \{a_{ik} : a_{ik} \in A_i$ and $a_{ik} \leqslant M + p - j\}$, then A_i, A_{i+1} holds stable all the terms in ${}_iC_i$ for $i = 0, 1, \ldots, p - 1$ and (c) $A_p, A_{p+1}, \ldots, A_r$ is a standard, complete reduction sequence. Next, we define ${}_id_j$ to be the number of elements in ${}_iC_j$ and, from the preceding, we know ${}_{i+1}d_j = {}_id_j$ if $i + 1 \leqslant j \leqslant p$. By induction it is easy to see that for all integers t ($1 \leqslant t \leqslant p$), ${}_0d_j = {}_td_j$ if $t \leqslant j \leqslant p$. Since ${}_0d_p = 0$, it follows that ${}_id_p = 0$ for all $i \leqslant p$. So, if $0 \leqslant i \leqslant p$, A_i does not consist of only a single term. This means that $A_0, A_1, \ldots, A_r$ satisfies our definition of a complete reduction sequence, and also that $r > p$.

Obviously,

$${}_0C_j = \sum_{i=0}^{p-j} B_i, \quad \text{so} \quad {}_0d_j = \sum_{i=0}^{p-j} b_i \quad \text{and} \quad b_j = {}_0d_{p-j} - {}_0d_{p-j+1}.$$

Using this in (2), along with the fact that ${}_0d_{p-2} = 0$, we obtain:

$$n + \sum_{i=0}^{p-3} 2(2^{i+1} - 2^i)\, {}_0d_i \leqslant 2^p. \tag{3}$$

However, $A_0, A_1, \ldots, A_r$ is a complete reduction sequence and from Lemma 3 we have:

$$(4) \qquad 2^{r-1} < n + \sum_{i=0}^{p-1} 2(2^{i+1} - 2^i)\ {}_0d_i.$$

So, since $d_{p-2} = d_{p-1} = d_p = 0$, we have $2^{r-1} < 2^p$, contradicting $r > p$, which was obtained earlier. Therefore, inequality (1) is false, and the lemma is proved.

It may be noted that the inequality we actually established is a good deal stronger than the one stated in Lemma 4, as one can see by comparing the negation of (1) with the stated inequality. However, the form in the lemma is the more convenient form when using it in proving certain metrization theorems. Also, the above proof is somewhat constructive in that, given M, the supremum of the set of all the residues of a certain n-tuple, the proof displays a complete reduction sequence resulting in M as the residue.

5. Alexandroff-Urysohn metrization theorem.

THEOREM. *A topological space S is metrizable if and only if S is a T_1-space such that there exists a sequence $G_1, G_2, \ldots,$ having the following properties:* (1) *for each natural number i, G_i is an open covering of S,* (2) *if g_1 and g_2 are intersecting elements of G_{i+1}, then $g_1 + g_2$ is a subset of some element of G_i, and* (3) *if p is a point of an open set R, then there is a natural number i such that if $j \geqslant i$ and g contains p and belongs to G_j, then g is contained in R.*

Proof. The necessity of the condition is easily established. We shall use the notation $U(p, 2^{-i})$ to represent the 2^{-i}-neighbourhood of p, that is the set $\{x | x \in S, d(p, x) < 2^{-i}\}$. Assuming the space S to be metric, we define G_i to be the set of all 2^{-i}-neighbourhoods. Since each such neighbourhood is open, condition (1) is satisfied. If g_1 and g_2 belong to G_{i+1} and $p \in g_1 \cdot g_2$, then $U(p, 2^{-i})$ belongs to G_i and contains $g_1 + g_2$. Consequently, condition (2) is true. If p is a point of an open set R, there is a natural number i such that $U(p, 2^{-i}) \subset R$. Then, if $j \geqslant i + 1$ it is clear that each element of G_j which contains p must lie within $U(p, 2^{-i})$ and hence in R.

To prove the sufficiency of the condition we begin defining $H_i = G_i$ if $i > 1$, $H_1 = G_1 + \{S\}$, and $H_i = \{S\}$ if $i \leqslant 0$. Obviously, the sequence $H_1, H_2, \ldots$ will satisfy the conditions stated in the hypothesis of the theorem. If x and y are points in S, let N be the largest integer i such that there is an element of H_i containing x and y. Define D to be the set of all positive numbers d such that there are open sets $g_1, g_2, \ldots, g_k$, where x is in g_1, y in g_k, $g_i \cdot g_{i+1} \neq 0$ $(i = 1, 2, \ldots, k-1)$, g_i belongs to $H_{a(i)}$ $(i = 1, 2, \ldots, k)$, and

$$\sum_{i=1}^{k} 2^{-a(i)} = d.$$

From condition (2) of the hypothesis we see that for every reduction of $(a(1), a(2), \ldots, a(k))$ there is a corresponding coherent collection containing x and y of open sets from the sequence of H's. Hence, N is an upper bound for the residues of this k-tuple, and by Lemma 4, $d \geqslant 2^{-(N+2)}$. So, the set D is bounded below, and we proceed to define $d(x, y)$ as the infimum of D if $x \neq y$, and for $x = y$ we define $d(x, y) = 0$. We now prove that this distance function satisfies the requirements of a metric. From our definition it follows that $d(x, y) = 0$ if and only if $x = y$. And it also follows that $d(x, y) = d(y, x)$. To establish the triangle inequality for the function, we let x, y, and z be points of S and assume $d(x, z) - d(x, y) - d(y, z) = \epsilon > 0$. From the definition of the distance function it follows that there is a coherent collection containing x and y such that the corresponding sum of dyadic fractions is less than $d(x, y) + \frac{1}{2}\epsilon$. Likewise, there is a coherent collection containing y and z whose corresponding sum of dyadic fractions is less than $d(y, z) + \frac{1}{2}\epsilon$. The union of these two coherent collections gives us a coherent collection containing x and z such that the corresponding sum of dyadic fractions is less than $d(x, y) + d(y, z) + \epsilon = d(x, z)$. This contradicts the definition of $d(x, z)$ as the infimum of the set of all such sums; hence the triangle inequality is true. Finally we must show that the limit point relation defined by the original topology is equivalent to that given by the distance function. Let p be a point of an open set R. By part (3) of the given condition on the G's, there is a natural number k such that if q is not in R, then no element of G_k contains p and q. Thus, $d(p, q) \geqslant 2^{-k-1}$ and so $U(p, 2^{-k-1})$ is contained in R. On the other hand, for any p in S and natural number j, consider the union of all elements of G_{j+1} which contains p. This set is usually called a star of p and will be denoted by $\mathrm{st}(p, j + 1)$. If q is in $\mathrm{st}(p, j + 1)$ then $d(p, q) \leqslant 2^{-j-1} < 2^{-j}$. Hence, $\mathrm{st}(p, j + 1)$ is contained in $U(p, 2^{-j})$. This completes the argument proving that the distance function defined above preserves the limit point relation and, in fact, is a metric for S.

6. Concluding remarks. The manner in which the inequality obtained in Lemma 4 can be applied in proving certain other metrization theorems is almost identical with the way in which it was used above. Several other metrization theorems, such as Frink's and Bing's, are stated in terms of conditions similar to the condition (3) in the Alexandroff-Urysohn theorem. Whenever such a condition is present, the inequality in Lemma 4 can be used to establish a positive lower bound for a set of certain dyadic fractions, and then the distance between points can be defined by means of the infimum of this set. The same method can also be applied to obtain a simple proof of a theorem of H. Ribeiro **(4)** giving a necessary and sufficient condition for a topological space to be weakly metric.

References

1. P. Alexandroff and P. Urysohn, *Une condition nécessaire et suffisante pour qu'une classe (L) soit une classe (B)*, C. R. Acad. Sc., *177* (1923), 1274–1276.

2. R. H. Bing, *Extending a Metric*, Duke Math. J., *14* (1947), 511–519.

3. ——— *Metrization of topological spaces*, Can. J. Math., *3* (1951), 175–186.

4. E. W. Chittenden, *On the equivalence of écart and voisinage,* Trans. Amer. Math. Soc., *18* (1917), 161–166.

5. A. H. Frink, *Distance functions and the metrization problem,* Bull. Amer. Math. Soc., *43* (1937), 133–142.

6. J. L. Kelley, *General topology* (Princeton, 1955).

7. J. Nagata, *On a necessary and sufficient condition of metrizability,* J. Inst. Polytech. Osaka City Univ., *1* (1950), 93–100.

8. H. Ribeiro, *Sur les espaces à métrique faible,* Portugaliae Mathematica, *4* (1943), 21–40.

9. Yu. M. Smirnov, *A necessary and sufficient condition for metrizability of a topological space,* Doklady Akad. Nauk. S.S.S.R., N.S., *77* (1951), 197–200.

10. A. Weil, *Sur les espaces à structure uniforme et sur la topologie générale,* Actualités Sci. Ind., *551* (Paris, 1937).

University of Utah
University of Wisconsin