American Mathematical Monthly, Vol. 104, no. 8 (Oct. 1997) 713-719

Metric Spaces in which All Triangles Are Degenerate

Bettina Richmond and Thomas Richmond

In any subspace of the real line **R** with the usual Euclidean metric $d(x, y) = |x - y|$, every triangle is degenerate. In $\mathbf{R}^2$ or $\mathbf{R}^3$ with the usual Euclidean metrics, a triangle is degenerate if and only if its vertices are collinear. With our intuition of a degenerate triangle having "collinear vertices" extended to arbitrary metric spaces, we might expect that a metric space in which every triangle is degenerate must be "linear". It might be reasonable to expect that any "linear" metric space is isometric to a subset of **R** with the usual metric. When classifying all metric spaces that have only degenerate triangles, we find that there are such metric spaces other than (isometric images of) subspaces of **R**. These other spaces, however, all have precisely four points and all are of the same form. In the final section, we illustrate that the usual topology on $\mathbf{R}^2$ can not be generated by any metric in which all triangles are degenerate.

A metric space (M, ρ) is a set of points M with a metric, or distance function, $\rho : M \times M \longrightarrow [0, \infty)$ that satisfies some natural properties we expect of distances:

$\rho(x, y) = \rho(y, x)$ for any $x, y \in M$,

$\rho(x, y) = 0$ if and only if $x = y$, and

$\rho(x, y) + \rho(y, z) \geq \rho(x, z)$ for any $x, y, z \in M$ (triangle inequality).

If (M, ρ) and (N, δ) are two metric spaces, a function $f : M \longrightarrow N$ such that $\rho(x, y) = \delta(f(x), f(y))$ for any $x, y \in M$ is called an *isometry*. Isometries are always one-to-one. The metric spaces (M, ρ) and (N, δ) are *isometric* if there exists an isometry from M onto N. Generally, one does not distinguish between isometric metric spaces.

If (M, ρ) is a given metric space, we say that $\{x_1, x_2, x_3\} \subseteq M$ forms a *degenerate triangle* if $\rho(x_i, x_j) + \rho(x_j, x_k) = \rho(x_i, x_k)$ for some permutation i, j, k of the indices $\{1, 2, 3\}$. Thus, $\{x_1, x_2, x_3\}$ forms a degenerate triangle if the triangle inequality is an equality for some permutation of the points x_1, x_2, and x_3.

We want to study metric spaces in which every triangle is degenerate, that is, metric spaces such that every 3-element subset forms a degenerate triangle. For brevity, such a metric space is called a *degenerate space*, and its metric is called a *degenerate metric.* Triangles with fewer than three distinct vertices are clearly degenerate, and in what follows, we assume that all triangles have three distinct vertices, unless otherwise noted. With this understanding, we can make statements such as "A four-point space has exactly four distinct triangles." The real line with the usual Euclidean metric $d(x, y) = |x - y|$ is denoted by **R**. It is a familiar fact that **R** is a degenerate space. Any subspace of a degenerate space is a degenerate space. Given any degenerate space M, one might suspect that it must be isomorphic to a subspace of **R**, and might attempt to construct an isometry from M into **R**. As we will see, the construction of such an isometry requires that if $\{y, o, p\}$ and $\{o, p, x\}$ are (degenerate) triangles in M with longest sides $\{y, p\}$ and $\{o, x\}$ respectively, then triangle $\{y, o, x\}$ must have longest side $\{y, x\}$. Surprisingly, this need not be true. Thus, the structure of a degenerate space is based upon the structure of its 4-point subspaces. We consider 4-point degenerate spaces in the next section; in Section 2 we find that the problematic 4-point subspaces cannot occur if the space has more than 4 points.

1. Four-Point Degenerate Spaces are of Two Forms. Throughout this section, $X = \{S, T, U, V\}$ denotes a 4-point degenerate space. The four points of X are represented by the vertices and center of a (Euclidean) equilateral triangle as shown in Figure 1. The distances between points are denoted by labels on the corresponding edges of this graph. Often we do not distinguish between an edge and its length. By the longest edge of some set, we mean any edge of maximal length from that set. After a permutation of the vertices of X, we may assume that the edge $\{V, T\}$ is maximal among the six edges, and that $\{V, S\}$ is maximal among the remaining four edges having V or T as a vertex. If we let $\rho(V, S) = a$, $\rho(S, T) = b$, $\rho(U, V) = c$, and $\rho(U, T) = d$, then since $\{V, T\}$ is the longest edge in the triangles $\{V, T, S\}$ and $\{V, T, U\}$, we must have $\rho(V, T) = a + b = c + d$, as shown in Figure 2. By the choice of the edge $\{V, S\}$, we know that $a \geq c$. Hence either $\{V, S\}$ or $\{S, U\}$ is the longest edge of the triangle $\{V, S, U\}$, and thus either $\rho(S, U) = a - c$ (Case 1) or $\rho(S, U) = a + c$ (Case 2). Depending on which edge of triangle $\{S, T, U\}$ is the longest, either $\rho(S, U) = b + d$ (Case A), $\rho(S, U) = b - d$ (Case B), or $\rho(S, U) = d - b$ (Case C). Among the 6 cases that result from considering which are the longest edges of these two triangles, the cases 1A, 1B, 2B, and 2C are easily seen to be impossible. For example, in case 1B, the equations $a - c = b - d$ and $a + b = c + d$ yield $a - c = 0$, contrary to S and U being distinct points with $\rho(S, U) = a - c$. In the case 1C, the equations $a - c = d - b$ and $a + b = c + d$ are equivalent, while in the case 2A, the equations $a + c = b + d$ and $a + b = c + d$ yield $a = d$ and $b = c$. The two configurations corresponding to these two cases 1C and 2A are pictured in Figure 3. Both of these are possible, and any 4-point degenerate space must be of one of these two forms.

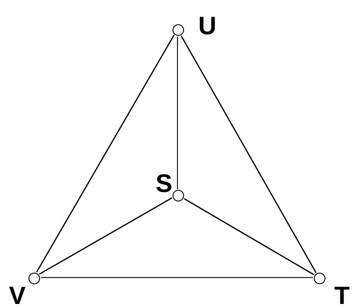

Figure 1

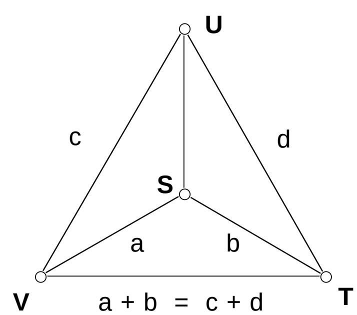

Figure 2

There are familiar models for both of these 4-point metric spaces. The form from Figure 3a can be achieved using Euclidean distance on the real line. If positive distances a, b, and $c < a$ have been chosen, then letting V be any real number, $T = V + a + b$, $S = V + a$, and $U = V + c$ gives a 4-point subspace of $\mathbf{R}$ in which the Euclidean distances between the points agree with those in the form. Because of this model, we call this configuration, shown in Figure 4, the *linear form* of a 4-point degenerate space. The other 4-point degenerate space, shown in Figure 3b, can be redrawn as a square with

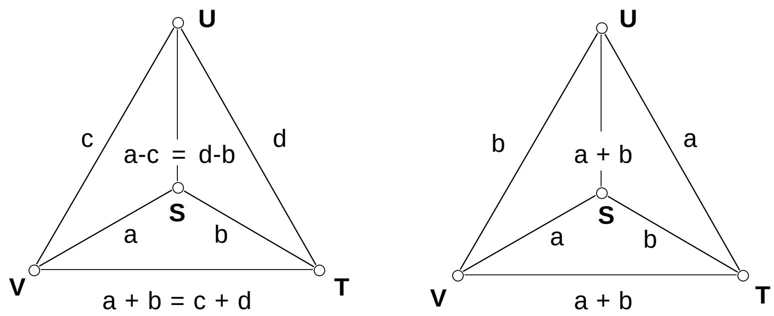

Figure 3a

Figure 3b

Figure 3

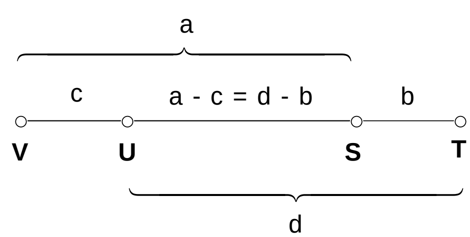

Figure 4. Linear Form.

its diagonals as shown in Figure 5a. It should be clear that no such configuration can be realized with positive Euclidean distances, but the square suggests a familiar model, shown in Figure 5b. Let $\{S, U\}$ and $\{T, V\}$ be distinct pairs of diametrically opposite points on a circle with circumference $2a + 2b$ such that the points V and S determine a central angle of $\frac{\pi a}{a+b}$ radians. Let the distance between two points on the circle be the length of the shortest arc of the circle connecting the two points. This familiar arc-length metric restricted to the set $\{S, T, U, V\}$ is a realization of the 4-point degenerate space of Figure 3b. Because of this model, we call this configuration the *circular form* of a 4-point degenerate space.

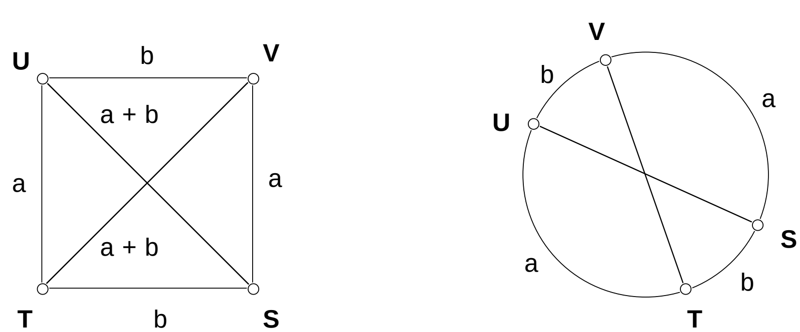

Figure 5a

Figure 5b

Figure 5. Circular Form.

2. Five-Point Degenerate Spaces are of Linear Form. The structure of any degenerate space is based on its 4-point degenerate subspaces, which must be of linear or circular form. In this section, we show that in any 5-point degenerate space all 4-point subspaces are of linear form. Thus, the problem encountered in defining an isometry from any degenerate metric space M into $\mathbf{R}$ can only arise in 4-point spaces (of circular form).

Before considering general 5-point degenerate spaces, we make a few observations on the specific metrics presented above as models of 4-point degenerate spaces. It is clear that extending our Euclidean model of a 4-point degenerate space of linear form to any larger subset of $\mathbf{R}$ results in a degenerate space. However, a degenerate space $\{S, T, U, V\}$ of circular form cannot be extended to any larger degenerate space consisting of points on a circle with the arc-length metric. If R were another point on the circle, relabel the points of $\{S, T, U, V\}$ so that R, S, T, U, and V appear in that order along the circle (traced in any specified direction). Now R, T, and U do not lie in any semicircle, and thus, the sum of the distances from any one of these points to the other two is strictly greater than the arc-length of a semicircle. In our "length of shortest arc" metric, however, no distance can exceed the length of a semicircle. Thus, the triangle formed by R, T, and U cannot be degenerate.

We now show that there is no extension of the 4-point degenerate space of circular form to any 5-point degenerate metric space. We start with the 4-point circular form $\{S, T, U, V\}$ in its square representation and add a fifth point R, as shown in Figure 6, and assume that this space is degenerate. Two opposite edges of the square have lengths x and the other pair of opposite edges of the square have lengths y. Let a be the maximal length of the four edges having R as a vertex. Without loss of generality, we may label the 5-point degenerate space as shown in Figure 7.

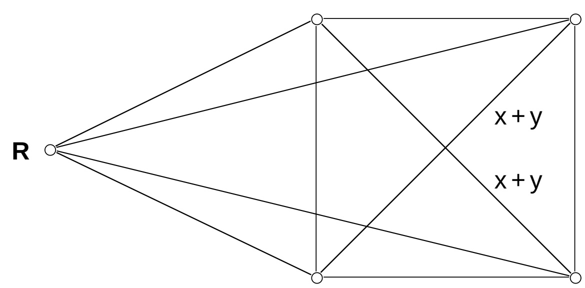

Figure 6

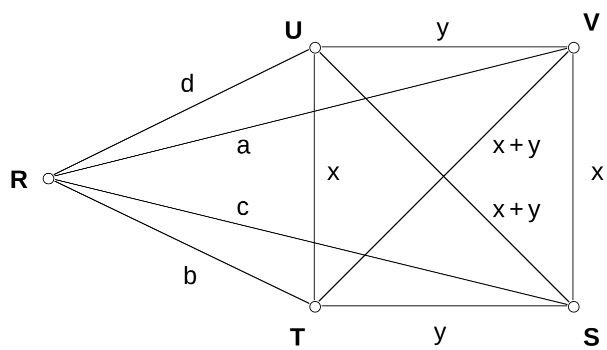

Figure 7

We first show that $a \le x+y$, that is, $x+y$ is the longest edge of any in Figure 7. Suppose to the contrary that $a > x+y$. Let j be the largest integer such that $a > x+jy$ and let k be the largest integer such that $a > kx+y$. Since a is the length of the longest edge in triangles $\{R,S,V\}$ and $\{R,U,V\}$, we have $a = x+c = d+y$, and thus $c > jy$ and $d > kx$. Adding these inequalities gives $c+d > x+y$, so the longest edge in triangle $\{R,S,U\}$ has length c or d. If c were the longest, then $c = d+x+y > (k+1)x+y \ge a$, contrary to a being the longest edge incident on R. Similarly, d cannot be the length of the longest edge of $\{R,S,U\}$. Thus, $a \le x+y$.

Now we show that the assumption that Figure 7 represents a degenerate space leads to the contradiction that the set of real numbers $\{x,y,a\}$ has no maximum. By the preceding paragraph, the longest edges of triangles $\{R,S,U\}$ and $\{R,T,V\}$ have length $x+y$, and thus $a+b = c+d = x+y$. Suppose $x = \max\{x,y,a\}$. The choice of a implies that $x = \max\{x,y,a,b,c,d\}$. From triangles $\{R,S,V\}$ and $\{R,T,U\}$, we have $x = a+c = d+b$. Since $a+b = c+d$, it follows that $a = d$ and $b = c$. Thus, $x = a+c = a+b = x+y$, giving the contradiction that $y = 0$. By the symmetry of x and y, the case $y = \max\{x,y,a\}$ is also impossible. Finally, suppose $a = \max\{x,y,a\}$. From triangles $\{R,V,S\}$ and $\{R,V,U\}$, we have $a = c+x = d+y$. It follows that $2a = c+d+x+y = 2(a+b)$, contrary to $b \ne 0$.

Thus, the 4-point degenerate space of circular form cannot be extended to a 5-point degenerate space, and therefore cannot be extended to any degenerate space with more than 4 points.

3. Classification of All Degenerate Spaces. Eliminating circular form subsets from spaces with 5 or more points clears the way for the classification of not only 5-point degenerate spaces, but of all degenerate spaces.

It is clear that every metric space with fewer than three points is degenerate and isometric to a subspace of $\mathbf{R}$. Any degenerate 3-point space consists of one degenerate triangle with edges of length, say, a, b, and $a+b$. Such a space is clearly isometric to the subspace $\{-a, 0, b\}$ of $\mathbf{R}$. Four-point degenerate spaces are of one of the two forms previously described. In any degenerate space with more than four points, every 4-point subspace is degenerate, of linear form.

Suppose M is a degenerate space with more than four points. We will construct an isometry between M and a subspace of $\mathbf{R}$. Pick any two distinct points o and p of M. The point o plays the roll of the "origin" and the point p plays the roll of a "positive point". Define a function $f : M \longrightarrow \mathbf{R}$ by

$$f(x) = \begin{cases} -\rho(o,x) & \text{if } \{x,p\} \text{ is the longest edge of triangle } \{x,o,p\} \\ \rho(o,x) & \text{if } \{x,o\} \text{ or } \{o,p\} \text{ is the longest edge of triangle } \{x,o,p\}. \end{cases}$$

We first show that f is well defined. Observe that $\{x,p\}$ and $\{x,o\}$ cannot both be longest edges of triangle $\{x,o,p\}$, for then $\rho(x,p) = \rho(x,o)+\rho(o,p)$ and $\rho(x,o) = \rho(x,p)+\rho(p,o)$ lead to the contradiction that $\rho(o,p) = 0$. If $\{x,p\}$ and $\{o,p\}$ are both longest edges of triangle $\{x,o,p\}$, then $\rho(x,p) = \rho(x,o)+\rho(o,p)$ and $\rho(o,p) = \rho(o,x)+\rho(x,p)$, and it follows that $\rho(o,x) = 0 = -\rho(o,x)$, so that both definitions of the value of $f(x)$ agree.

We now show that f is an isometry. If $x = y$, then clearly $\rho(x,y) = 0 = |f(x) - f(y)| = d(f(x), f(y))$. If one of x or y is o, say $y = o$, then $f(x) = \pm\rho(o,x)$, so $\rho(x,o) = |f(x)-0| = |f(x)-f(o)| = d(f(x),f(y))$. We may thus assume that x, y, and o are distinct. If $\{x,y,o,p\}$ is not already a 4-point subset of M it may be extended to a 4-point subset of M, which, after our unexpected combinatorial detour, we now know must be of linear form as shown in Figure 4. The linear form of Figure 4 can be linearly ordered in two natural ways: left to right or right to left. Give $\{x, y, o, p\}$ the natural linear order from its linear representation as in Figure 4 in which $o < p$. Without loss of generality, let us assume $x < y$ in this order. There are three cases to consider: either $x < y < o$, $x < o < y$, or $o < x < y$. If $x < y < o$, then $\rho(x, o) = \rho(x, y) + \rho(y, o)$, so $\rho(x, y) = \rho(x, o) - \rho(y, o)$. Since $x < o < p$, $f(x) = -\rho(o, x)$, and similarly $f(y) = -\rho(o, y)$. Thus, $\rho(x, y) = f(y) - f(x)$. Since the distance from x to y is positive, we have $\rho(x, y) = f(y) - f(x) = |f(x) - f(y)|$. If $x < o < y$, then $\rho(x, y) = \rho(x, o) + \rho(o, y)$. Since $o < y$, either $\{o, p\}$ or $\{o, y\}$ is the longest edge of $\{y, o, p\}$, so $f(y) = \rho(o, y)$. As before, $x < o < p$ implies $f(x) = -\rho(o, x)$. Thus, $\rho(x, y) = f(y) - f(x) = |f(x) - f(y)|$. Finally, if $o < x < y$, then $\rho(x, y) = \rho(o, y) - \rho(o, x) = f(y) - f(x) = |f(x) - f(y)|$. Thus, for any choice of $x, y \in M$, we have shown that $\rho(x, y) = |f(x) - f(y)| = d(f(x), f(y))$. This completes the proof that any degenerate space with more than four points is isometric to a subspace of $\mathbf{R}$.

Note that if M is a 4-point degenerate space of circular form, the function f defined in the preceding paragraph is still well defined but is not an isometry.

4. Topological Considerations. Our motivation for considering degenerate spaces originated with a topological question. Denote the topology generated by the Euclidean metric on $\mathbf{R}^n$ by τ_n. With the usual Euclidean metrics, $\mathbf{R}^2$ has non-degenerate triangles but $\mathbf{R}$ does not. Does this alone provide another verification that $(\mathbf{R}, \tau_1)$ and $(\mathbf{R}^2, \tau_2)$ are not homeomorphic?

There are two immediate observations we should make. First, metrizable topological spaces can be generated by several different metrics which may not share common properties such as boundedness or "degenerateness". In particular, observe that though $(\mathbf{R}, \tau_1)$ is generated by a degenerate metric $d(x, y) = |x - y|$, it is also generated by non-degenerate metrics. If we embed $\mathbf{R}$ into $\mathbf{R}^2$ as the graph of $y = x^2$ and give this parabola $P = \{(x, y) \in \mathbf{R}^2 : y = x^2\}$ the usual Euclidean distances from $\mathbf{R}^2$, then P has no degenerate triangles, yet P is homeomorphic to $(\mathbf{R}, \tau_1)$. Secondly, it is easily seen that "metrizable with a degenerate metric" is a topological property, for if (M, ρ) is a degenerate space homeomorphic to a topological space X, then a degenerate metric can be defined on X by making the homeomorphism from X to M an isometry. Thus, though the Euclidean metric on $\mathbf{R}$ is degenerate and the Euclidean metric on $\mathbf{R}^2$ is not, this is not sufficient to conclude that $(\mathbf{R}, \tau_1)$ is not homeomorphic to $(\mathbf{R}^2, \tau_2)$. For this conclusion, we must show that $(\mathbf{R}^2, \tau_2)$ admits no degenerate metric.

If our goal were to show only that $(\mathbf{R}^2, \tau_2)$ is not generated by any degenerate metric, then we could simply note that this is an immediate consequence of the fact that $(\mathbf{R}^2, \tau_2)$ is not homeomorphic to a subspace of $(\mathbf{R}, \tau_1)$ and the classification of infinite degenerate spaces. However, since we want to use the fact that $(\mathbf{R}^2, \tau_2)$ admits no degenerate metric to show that $(\mathbf{R}^2, \tau_2)$ is not homeomorphic to $(\mathbf{R}, \tau_1)$, we now outline a proof not dependent upon this latter fact.

Suppose ρ is a degenerate metric on $\mathbf{R}^2$ that generates the Euclidean topology τ_2. Pick distinct points a, b, and x from $\mathbf{R}^2$ such that $\{a, b\}$ is the longest side of triangle $\{a, x, b\}$. Thus $\rho(a, x) < \rho(a, b)$ and $\rho(x, b) < \rho(a, b)$. Using the fact that $\rho(a, x)$ is a continuous function of x, if the "middle point" x is moved slightly to $x(t)$, $\rho(a, b)$ remains the longest side of the resulting triangle $\{a, x(t), b\}$. These perturbations of the middle point are actually not restricted to slight movements: in fact, every point z of $\mathbf{R}^2 \setminus \{a, b\}$ is a middle point of $\{a, z, b\}$! Suppose there exists some point z of $\mathbf{R}^2 \setminus \{a, b\}$

such that $\{a, b\}$ is not the longest side of $\{a, z, b\}$. Let the points $x(t)$ slide along a path in $\mathbf{R}^2 \setminus \{a, b\}$ from x to z. There must exist a first point y on the path for which $\rho(a, y) = \rho(a, b)$ or $\rho(y, b) = \rho(a, b)$. Since $\rho(a, x(t)) + \rho(x(t), b) = \rho(a, b)$ for all the points $x(t)$ on the path before y, it follows that $\rho(a, y) + \rho(y, b) = \rho(a, b)$, and hence $y \in \{a, b\}$. This contradicts the fact that the path avoids the points a and b. Thus, if we select any triangle $\{a, x, b\}$ with x "between" a and b (that is, with longest side $\{a, b\}$), then every point of $\mathbf{R}^2 \setminus \{a, b\}$ must be "between" a and b. In particular, $\rho(a, y) \leq \rho(a, b)$ for every $y \in \mathbf{R}^2$. This implies that our degenerate metric on $\mathbf{R}^2$ is bounded by $2\rho(a, b)$ since $\rho(x, y) \leq \rho(x, a) + \rho(a, y) \leq \rho(a, b) + \rho(a, b)$ for any x and y. Our contradiction will come from the fact that this should hold for any points a and b that happen to form the longest side of some triangle. Pick any two distinct points $z, w \in \mathbf{R}^2$ and let $\epsilon = \rho(z, w)$. Let $\{x_n\}$ be any sequence converging to the origin 0 in $(\mathbf{R}^2, \tau_2)$. Since $\{\rho(x_n, 0)\}$ converges to zero, we may pick points x_j and x_k such that the longest edge of triangle $\{x_j, x_k, 0\}$ has length less than $\frac{\epsilon}{2}$. But ρ is bounded by twice the length of the longest edge of this triangle, contrary to the choice of $\epsilon = \rho(z, w)$ for some $z, w \in \mathbf{R}^2$. This shows that $(\mathbf{R}^2, \tau_2)$ does not admit any degenerate metric, and thus is not homeomorphic to $(\mathbf{R}, \tau_1)$.

Using the topological property "metrizable by a degenerate metric" to distinguish between $\mathbf{R}$ and $\mathbf{R}^2$ is a good exercise in manipulating metrics, but one should recognize the dependence of our proof upon another more common topological property. The argument that if x is "between" a and b, then every $y \in \mathbf{R}^2 \setminus \{a, b\}$ is "between" a and b utilized the fact that $\mathbf{R}^2 \setminus \{a, b\}$ is path connected (relative to τ_2). Removing two points from $(\mathbf{R}, \tau_1)$ never gives a path connected space, providing a more standard argument that $(\mathbf{R}, \tau_1)$ is not homeomorphic to $(\mathbf{R}^2, \tau_2)$. On the other hand, proving that $(\mathbf{R}^2, \tau_2)$ is not metrizable by a degenerate metric shows not only that $(\mathbf{R}^2, \tau_2)$ is not homeomorphic to $(\mathbf{R}, \tau_1)$, but also that $(\mathbf{R}^2, \tau_2)$ is not homeomorphic to any subspace of $(\mathbf{R}, \tau_1)$.

REFERENCES

1. Kaplansky, I., *Set Theory and Metric Spaces,* Chelsea Publishing Co., New York, 1977.
2. Munkres, J. R., *Topology, A First Course,* Prentice-Hall Inc., Englewood Cliffs, NJ, 1975.

Department of Mathematics
Western Kentucky University
1 Big Red Way
Bowling Green, KY 42101

Bettina.Richmond@wku.edu
Tom.Richmond@wku.edu