Some properties of Davenport–Schinzel sequences

by

D. P. Roselle* (Baton Rouge, La)
and R. G. Stanton (Winnipeg, Canada)

1. Introduction. Davenport and Schinzel [1] introduced the following problem on sequences. Suppose that \( Z_n = \{1, 2, \ldots, n\} \) and that \( V_n(d) \) denotes the set of all sequences \( a, b, c, \ldots \) of length \( d \), where \( a, b, c, \ldots \) are distinct elements of \( Z_n \). One considers all sequences made up of elements from \( Z_n \) such that no two adjacent elements are equal and no subsequence is an element of \( V_n(j) \) for \( j < d \). If \( N_d(n) \) denotes the maximal length of any such sequence, we call any sequence of length \( N_d(n) \) a Davenport–Schinzel sequence (or a DS sequence). The problem is to determine all DS sequences and, in particular, to determine \( N_d(n) \). Of course, it will be sufficient to determine all normal DS sequences; i.e., DS sequences in which the elements appear in order from left to right.

Davenport and Schinzel consider \( N_d(n) \) for fixed \( d \) and obtain the results \( N_1(n) = 1, N_2(n) = n, N_3(n) = 2n - 1 \). Two proofs of the result for \( N_3(n) \) are given, the second being based on Mrs. Turán's observation that, in a DS sequence of length \( N_3(n) \), there is some element which occurs exactly one time. They remark that \( 1, 2, 1, 3, 1 \) \ldots , \( 1, n, 1 \) and \( 1, 2, \ldots, n - 1, 1, n, n - 1, \ldots, 2, 1 \) are both DS sequences for \( d = 3 \).

Finally, they obtain bounds on \( N_d(n) \) for fixed \( d \), including the result \( N_d(n) \leq 5n - C \), where \( C \) is a constant.

The authors [2] have proved that \( N_d(2) = d \) and, for \( d > 3, N_d(3) = 3d - 4 \) or \( 3d - 5 \), depending upon whether \( d \) is even or odd. It is also shown that, for \( n = 3 \), a DS sequence is unique.

The object of the present paper is to prove

\[
\begin{align*}
(1.1) \quad N_d(d) &= 6d - 14 \quad (d \text{ odd, } d > 4), \\
(1.2) \quad N_d(d) &= 6d - 13 \quad (d \text{ even, } d > 4), \\
(1.3) \quad \left(\frac{d}{2}\right)d - C &\leq N_d(n) \quad (d > n),
\end{align*}
\]

* Partially supported by National Science Foundation grant GP-11397. Part of this work was done during a visit to the University of Manitoba.
where \( C \) is a constant depending only on \( n \). The constant \( C = C(n) \) which appears in (1.3) is given explicitly as a polynomial in \( n \).

2. A lower bound. Given a normal DS sequence, it is clear that each of the sequences \( i, j, i, j, i, \ldots \) (\( i \neq j \)) is of length at most \( d-1 \). For, otherwise, the sequence would contain a subsequence \( j, i, j, i, \ldots \) of length greater than \( d \).

Now suppose that \( d \) is even and \( d > n \) and let \( f \) denote the maximum number of occurrences of any element, say \( 1 \), in a DS sequence of length \( N_d(n) \). There are \( f-1 \) blocks in the sequence; i.e., groups of consecutive elements which contain no 1 but are bounded both on the left and the right by 1's. Each of the remaining \( n-1 \) elements appears in at most \( (d-2)/2 \) of these blocks; for if a symbol occurred in as many as \( d/2 \) of these blocks, there would result a sequence from \( V_n(d+1) \), and this is impossible. On the other hand, since no pair of 1's is to be adjacent, the number of blocks must be no larger than the number of appearances in the blocks. This proves

\[
 f \leq \frac{(n-1)(d-2)}{2} + 1. 
\]

Next, we give a construction of a sequence in which the bound (2.1) is attained. List

\[
 f = \frac{(n-1)(d-2)}{2} + 1
\]

1's and insert \( n \) between the first \( (d-2)/2 \) 1's, insert \( n-1 \) between the next \( (d-2)/2 \) 1's, and so forth. Finally, list a normal DS sequence of length \( N_{d-1}(n-1) \) made up of the entries 2, 3, \ldots, \( n \) to the right of the last 1. Using the remark that the sequence of length \( N_{d-1}(n-1) \) contains no subsequence \( i, j, i, \ldots \) of length greater than \( d-2 \), we see that the sequence we have constructed contains no subsequence from \( V_n(j) \) for \( j > d \). This proves the inequality

\[
 N_d(n) \geq (n-1)(d-2) + 1 + N_{d-1}(n-1) \quad (d \text{ even}).
\]

For example,

\[
 N_d(4) \geq 3(d-2) + 1 + 3(d-1) - 5 = 6d - 13 \quad (d \text{ even}).
\]

We now consider the case \( d \) odd and \( d > n \). A similar argument shows that, if \( f \) denotes the maximum frequency of any element, then

\[
 f \leq \frac{(n-1)(d-1)}{2} + 1.
\]

In practice, the inequality (2.4) seems to be less useful than (2.1). In fact, to obtain a lower bound for \( N_d(n) \), \( d \) odd and \( d > n \), we begin by listing only \( \frac{(n-1)(d-1)}{2} - n + 1 \) 1's. Next, we insert \( n \) between the first \( (d-3)/2 \) 1's, \( n-1 \) between the second \( (d-3)/2 \) 1's, and so forth. Finally, we insert a normal DS sequence of length \( N_{d-1}(n-1) \) made up of 2, 3, \ldots, \( n \) between the last two 1's. It is easy to verify that the resulting sequence has no adjacent terms equal and has no subsequences from \( V_n(j) \) for \( j > d \). It is now only necessary to compute the length of this sequence in order to prove

\[
 N_d(n) \geq (d-3)(n-1) + 2 + N_{d-1}(n-1) \quad (d \text{ odd}).
\]

For example,

\[
 N_d(4) \geq 3(d-3) + 2 + 3(d-1) - 4 = 6d - 14 \quad (d \text{ odd}).
\]

It is interesting to note that, as we shall show in § 4 and § 5, (2.3) and (2.6) are exact values for \( N_d(4) \).

3. The constant \( C \). Let \( d > n \) and put

\[
 C_d(n) = \begin{cases} 
 \left(\begin{array}{c}
 n/2 \\
 j 
\end{array}\right) d - C_{d-1}(n-1) & (d \text{ even}), \\
 \left(\begin{array}{c}
 n/2 \\
 j 
\end{array}\right) d - C_{d-1}(n) & (d \text{ odd}). 
\end{cases}
\]

It follows from (2.2) that

\[
 \left(\begin{array}{c}
 n/2 \\
 j 
\end{array}\right) d - C_{d-1}(n) \geq (n-1)(d-2) + 1 + \left(\begin{array}{c}
 n-1 \\
 j 
\end{array}\right) (d-1) - C_{d-1}(n-1) = \left(\begin{array}{c}
 n/2 \\
 j 
\end{array}\right) d - \left(\begin{array}{c}
 n+1 \\
 j 
\end{array}\right) + 2 - C_{d-1}(n-1),
\]

which implies

\[
 C_d(n) \leq C_{d-1}(n-1) + \left(\begin{array}{c}
 n/2 \\
 j 
\end{array}\right) - 2.
\]

In the same way, it follows from (2.5) that

\[
 C_d(n) \leq C_{d-1}(n-1) + \left(\begin{array}{c}
 n/2 \\
 j 
\end{array}\right) - 5.
\]

Let \( D_d(n) \) and \( D_d(n) \) denote the maximum values of \( C_d(n) \) and \( C_d(n) \) as indicated by (3.2) and (3.3). That is, \( D_d(2) = D_d(3) = 0 \) and

\[
 D_d(n) = D_{d-1}(n-1) + \left(\begin{array}{c}
 n/2 \\
 j 
\end{array}\right) - 2,
\]

\[
 D_d(n) = D_{d-1}(n-1) + \left(\begin{array}{c}
 n/2 \\
 j 
\end{array}\right) - 5.
\]
It is an easy induction to prove that
\[
D(n) = D(n) + \left\lfloor \frac{n-1}{2} \right\rfloor, 
\]
so it will be sufficient to determine \(D(n)\). Combining (3.4) and (3.5), we find that
\[
D(n) = D(n-2) + 2 \left( \frac{n+1}{2} \right) - 7. 
\]
Iterating (3.7) and using the initial conditions, we are able to show
\[
D(n) = \begin{cases} 
\frac{9n^2 + 9n^2 - 32n + 9}{12} & (n \text{ odd}, n \geq 3), \\
\frac{2n^2 + 9n^2 - 32n + 12}{12} & (n \text{ even}). 
\end{cases} 
\]

The first few values of \(D(n)\) and \(D(n)\) are given by the following table

<table>
<thead>
<tr>
<th>(n)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>(D(n))</td>
<td>0</td>
<td>4</td>
<td>13</td>
<td>27</td>
<td>48</td>
<td>76</td>
<td>113</td>
<td>159</td>
<td>216</td>
</tr>
<tr>
<td>(D(n))</td>
<td>1</td>
<td>5</td>
<td>14</td>
<td>29</td>
<td>50</td>
<td>79</td>
<td>116</td>
<td>163</td>
<td>220</td>
</tr>
</tbody>
</table>

Combining the results (3.8), (2.2), and (2.5), we get that, for \(d > n\),
\[
N_d(n) \geq \begin{cases} 
\left( \frac{n}{2} \right) d - D(n) & (d \text{ even}), \\
\left( \frac{n}{2} \right) d - D(n) - \left\lfloor \frac{n-1}{2} \right\rfloor & (d \text{ odd}). 
\end{cases} 
\]

where \(D(n)\) is given by (3.8).

4. The case \(n = 4\) and \(d\) even. If \(d\) is even and \(d > 4\) then, according to (2.1), \(f = 3d/2 \geq 2\). On the other hand, it follows from (2.3) that \(f \geq N_d(4) \geq 6d-13\), so that \(f \geq 3d/2 - 3\). We will show that in a DS sequence of length \(N_d(4), f = 3d/2 - 2\).

If one of the elements, say 1, occurs \(f = 3d/2 - 3\) times, then there are \(f - 1\) blocks between 1's. Since each of 2, 3, and 4 can occur in at most \((d-2)/2\) of these blocks, we see that the appearances of 2, 3, and 4 can overlap in at most one block. We consider two cases.

Case 1. There is no overlapping. Here the number of interior blocks occupied by 2, 3, and 4 is \((d-2)/2, (d-2)/2, \text{ and } (d-4)/2\), respectively. Since this sequence contains subsequences 1, 2, 1, ... and 1, 3, 1, ... of length
\[
2 \left( \frac{d-2}{2} \right) + 1 = d-1, 
\]
it is clear that 2 and 3 can occur at only one of the two ends of the sequence. However, 4 can occur at both ends.

If 2 and 4 occur at one end of the sequence and 3 and 4 occur at the opposite end, then the sequence is of length at most
\[
f + f - 1 + 2N_d(4) = 5d - 9. 
\]
This is not a DS sequence since, for \(d > 4, N_d(4) \geq 6d - 13\). Next, if 2, 3, and 4 appear at one end of the sequence, the maximum length of the sequence is
\[
1 + f + f - 1 + N_d(3) = 6d - 14. 
\]
And, again, this is not a DS sequence.

Thus we are led to consider

Case 2. Two of the elements, say 3 and 4, appear in the same block. The resulting sequence contains subsequences 1, a, 1, ... of length \(d-1\) for a = 2, 3, and 4, so that 2, 3, and 4 can occur at any one of the end positions.

If 4 is at one end and 2 and 3 are at the opposite end, then the sequence has length at most
\[
f + f - 2 + 1 + 2N_d(4) = 5d - 9 < 6d - 13. 
\]
Next, if 2 is on one end and 3 and 4 are at the opposite end, the maximum length is
\[
f + f - 2 + 1 + N_d(4) + 2 = 4d - 8, 
\]
a contradiction.

The remaining possibility is when 2, 3, and 4 appear at the same end. Here the length is at most
\[
f + f - 2 + N_a(4) + N_d(3), \quad \text{where } a \geq 1. 
\]
This is the same as \(6d - 2a - 10\) if \(a\) is even and \(6d - 2a - 11\) if \(a\) is odd. But then the requirement \(N_d(4) \geq 6d - 13\) implies that \(a = 1\). However, we are able to rule out this case by making the observation that the unique DS sequence of length \(N_d(3) \text{ (d odd)}\) contains a subsequence \(a, b, a, \ldots \text{ of length at least } d-2\), for all \(a\) and \(b\). Thus the sequence constructed above with \(a = 1\) would contain a subsequence of length at least \(1 + 3 + d-1 - 2 = d+1\), a contradiction. This was the last case to be considered and therefore we are able to conclude that \(f \neq 3d/2 - 3\).

We remark that the construction of §2 provides a sequence for which \(f = 3d/2 - 2\). We will now show that this is the unique DS sequence for the case \(d > n = 4\).
Suppose \( f = 3d/2 - 2 \), and that 1 occurs \( f \) times. Since there are \( f-1 \) blocks between adjacent 1's and each of symbols 2, 3, and 4 can occur in at most \((d-2)/2\) of these, we conclude that no overlapping occurs. Since this sequence contains subsequences 1, a, 1, ..., of length \( d-1 \) for \( a = 2, 3, \) and 4, each of 2, 3, and 4 can occur at only one of the end positions.

If 2 occurs at one end and 3 and 4 occur at the opposite end, then the sequence is of length at most
\[
f + f - 1 + 1 + N_{d-1}(2) = 4d - 5.\]
However, since \( 3d - 5 < 6d - 13 \) for \( d > 4 \), this is not a DS sequence. On the other hand, if 2, 3, and 4 occur at the same end, we obtain the sequence constructed in § 2. It follows that, for \( d \) even and \( d > 4 \), \( N_d(4) = 6d - 13 \) and the DS sequence is unique.

5. The case \( u = 4 \) and \( d \) odd. If \( d \) is odd and \( d > 4 \) then, according to (2.4), \( f \leq 3(d-1)/2 + 1 \) and, using (2.9), we get \( df \geq N_d(4) \geq 6d - 14 \), so that \( f \geq 3(d-1)/2 - 2 \). Notice that if \( f = 3(d-1)/2 - 2 \), then \( N_d(4) = 6d - 14 \). Since we can construct a sequence of this length with \( f = 3(d-1)/2 - 1 \) (see § 2), we need only consider
\[
\frac{3(d-1)}{2} - 1 \leq f \leq \frac{3(d-1)}{2} + 1
\]
in order to prove \( N_d(4) = 6d - 14 \). However, this has the disadvantage of leaving the question of uniqueness unanswered.

If \( f = \frac{3}{2}(d-1) + 1 \), then there is exactly one of the elements 2, 3, and 4 between each pair of 1's. Since this sequence contains all subsequences 1, a, 1, ..., of length \( 2(d-1)/2 + 1 = d \), there are no other occurrences of 2, 3, or 4. Thus, the length of this sequence is \( 2f - 1 \). But since \( 2f - 1 < 6d - 14 \), the sequence is not a DS sequence.

Next, if \( f = 3(d-1)/2 \), there will be overlapping in at most one of the interior blocks. If there is overlapping in no interior block then the length of the sequence is at most \( f + f - 1 + 1 \), and, again, we see that it is not a DS sequence. On the other hand, if there is overlapping in one of the interior blocks, then the length is at most
\[
f + f - 2 + N_{d-1}(2) = 4d - 5 < 6d - 14.\]

Finally, if \( f = 3(d-1)/2 - 1 \), then there will be overlapping in 0, 1, or 2 of the interior blocks. If there is overlapping in none of the interior blocks, the sequence is of length at most \( f + f - 1 + N_{d-1}(2) + 2 \) or \( f + f - 1 + N_{d-1}(2) + 2 \).

Neither of these is as large as \( 6d - 14 \). If there is overlapping in 2 of the interior cells, either 2 and 3 or 2 and 4 occur in one block and 2 and 4 occur together in a second block. In the former case the sequence has length at most
\[
f + f - 3 + N_{d-2}(2) + N_{d-1}(2) = 4d - 7 < 6d - 14.\]

In the latter case, the sequence is of length at most
\[
f + f - 3 + 2N_{d-2}(2) = 5d - 10 < 6d - 14.\]

Finally, if there is overlapping in one of the interior blocks then the sequence is seen to be of length at most
\[
f + f - 2 + N_{d-1}(2) + 2 \text{ or } f + f - 2 + N_{d-1}(3)\]

The first of these is less than the second which is equal to \( 6d - 14 \). It follows that, for \( d \) odd and \( d > 4 \), we have proved \( N_d(4) = 6d - 14 \).

6. Tabulation of \( N_d(n) \). The table of this paragraph gives the known values of \( N_d(n) \). The value \( N_5(5) \) was obtained by computer. We have established the inequality:

\[(6.1) \quad 5a - 8 \leq N_d(n) \leq \frac{n}{n-1} N_d(n-1) + 2\]

and the values of \( N_d(n) \), for \( n \leq 8 \), follow from it; however, we omit the proof of (6.1) since Professor Schönzel informs us that Davenport and J. H. Conway have established a stronger result. We also owe to Professor Schönzel the information that Conway has shown that \( N_d(9) = 37 \), \( N_d(10) = 42 \), but that the equality \( N_d(n) = 5n - 8 \) suggested by the table is false, and in particular that \( N_5(13) \geq 58 \).

<table>
<thead>
<tr>
<th>( n )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>10</td>
<td>14</td>
<td>16</td>
<td>20</td>
<td>22</td>
<td>26</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>12</td>
<td>16</td>
<td>23</td>
<td>28</td>
<td>35</td>
<td>40</td>
<td>47</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>5</td>
<td>9</td>
<td>17</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>7</td>
<td>13</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>8</td>
<td>15</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>9</td>
<td>17</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>10</td>
<td>19</td>
<td>42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A combinatorial problem connected with differential equations II

by

H. Davenport

edited by A. Schinzel

WARSZAWA

1. Let us call a sequence admissible if it has no immediate repetition and contains no subsequence of the form $a, b, a, b$ with $a 
eq b$. Let $N(n)$ be the greatest length (that is, greatest number of terms) of an admissible sequence formed from $n$ distinct elements.

The problem of estimating $N(n)$ has been investigated in [1] and it has been proved there that

$$5n - C < N(n) = O(n \log n).$$

$(N(n)$ was denoted by $N_4(n)$.) The aim of this paper is to improve the above result in both directions. We prove

\textbf{Theorem 1.} We have

$$N(n) = O\left(\frac{n \log n}{\log \log n}\right).$$

\textbf{Theorem 2.} We have

$$\lim_{n \to \infty} \frac{N(n)}{n} \geq 8.$$

\textbf{Theorem 3.} For positive integers $l, m$, the following inequality holds

$$N(lm + 1) \geq 6lm - m - 5l + 2.$$