Proc. of the 15th International Colloquium on Automata, Languages and Programming, Lecture Notes in Computer Science vol. 317, Springer (Berlin, 1988), 188-200

THE COMPLEXITY OF MATRIX TRANSPOSITION ON ONE-TAPE OFF-LINE TURING MACHINES WITH OUTPUT TAPE

(Extended Abstract)

by

Martin Dietzfelbinger *,**
Lehrstuhl Informatik II
Universität Dortmund
Postfach 50 05 00
D - 4600 Dortmund 50, F.R.G.

and

Wolfgang Maass *
Dept. of Mathematics, Statistics, and
Computer Science
University of Illinois at Chicago
Box 4348
Chicago, IL 60680, U.S.A.

---

* Written under partial support by NSF-grant DCR-8504247

** This work is based on a part of the first author's Ph.D.-thesis at the University of Illinois at Chicago, Chicago, Illinois, U.S.A.

Abstract.

A series of existing lower bound results for one-tape Turing machines (TM's) is extended to the strongest such model for the computation of functions: one-tape off-line TM's with a write-only output tape. ("Off-line" means: having a two-way input tape.)

The following optimal lower bound is shown:
Computing the transpose of Boolean $\ell \times \ell$-matrices takes $\Omega(\ell^{5/2}) = \Omega(n^{5/4})$ steps on such TM's. ($n = \ell^2$ is the length of the input.)

§1 Introduction.

During the last few years one has developed lower bound arguments for a sequence of restricted Turing machines (TM's) of increasing power. Techniques have been devised that allow one to prove optimal superlinear lower bounds on the computation time for several concrete computational problems on one-tape TM's without input tape [2], on one-tape TM's with a one-way input tape ("on-line one-tape TM's") [4], and finally on one-tape TM's with a two-way input tape ("off-line one-tape TM's"; this is the standard model for the definition of space-complexity classes). For this model one has proved an optimal lower bound of $\Omega(n^{3/2} / (\log n)^{1/2})$ for the matrix transposition function [5], and a barely superlinear lower bound of $\Omega(n \cdot \log n / \log\log n)$ for a related decision problem [6].

In this paper we consider the next more powerful type of restricted TM's (for which the preceding lower bound arguments do not suffice): off-line one-tape TM's with an additional one-way output tape. Whereas the addition of the output tape obviously makes no difference for solving decision problems, it was already noted in [5] that these machines can do matrix transposition in $O(n^{5/4})$ steps (as opposed to $\Omega(n^{3/2} / (\log n)^{1/2})$ steps for the previously considered version without output tape, where the output has to appear on the work tape).

This stronger model is also of some interest from a technical point of view, because it exhibits a feature that is characteristic for TM's with several work tapes (which are so far intractable for lower bound arguments): the extensive use of the work tape as an intermediate storage device. This feature played only a minor role in the analysis of matrix transposition on one-tape off-line TM's without output tape, because one could easily show that any use of the work tape as an intermediate storage device is inefficient for this model. (Once some bits have been written on the work tape, they can be moved later only by time-consuming sweeps of the work tape head: during each sweep only $\leq$ log n bits can be moved, where n is the length of the input. The number of bits that can be moved is $\approx$ log n rather than constant since the input tape can be used as a unary counter, thus can store log n bits.)*

In this paper, we prove an optimal lower bound of $\Omega(n^{5/4})$ for the transposition of Boolean matrices on one-tape off-line TM's with output tape. The lower bound argument employs Kolmogorov complexity to enable us to analyze the possible flow of information during the transposition of a suitably chosen matrix on such a machine.

This analysis differs from previous lower bound arguments with Kolmogorov complexity by its emphasis on the time-dimension of the computation: it is not enough to watch which information <u>ever</u> reaches a certain interval on the work tape, rather it is essential to note which information may be present in such an interval at specific time points. In particular, the argument exploits the fact that in certain situations the same information may have to be brought into the same tape area several times (because after is was first brought there, it had to be overwritten to make space for some other information).

Moreover, the Kolmogorov complexity Lemmata (Lemma 2 and 5) employ a new trick (from [1]), which allows us to prove optimal lower bounds for matrix tranposition even in the case where the matrix entries are single bits. (The technique of [5] could only handle the case with entries of bitlength $\geq$ log n.)

The following notions and definitions are used in this paper. The definition of Turing machines that we use is standard (see e.g. [3]). A k-tape TM is a TM with k (read/write) work tapes. The work tape alphabet is assumed to be {0,1,B}. (If larger work tape alphabets

---

* This feature of one-tape TM's with two-way input tape can be used to show that such machines can simulate f(n)-time-bounded k-tape TM's in $O(f(n)^2/\log n)$ steps.

Γ were used, the lower bound in this paper would change by the constant factor $1/\log(|\Gamma|)$.) The function MATRIX TRANSPOSITION is induced by the operation of transposing a matrix: given an input $x \in \{0,1\}^n$, $n = \ell^2$, regard x as the representation of a Boolean matrix $A \in \{0,1\}^{\ell \times \ell}$ in row-major order, and output the transpose A^t in row-major order (or, equivalently, A in column-major order). That means, if the input is $x = b_1 b_2 \ldots b_n$ with $b_m \in \{0,1\}$, then the output is

$y = b_{\pi(1)} b_{\pi(2)} \ldots b_{\pi(n)}$, where the permutation π of $\{1,2,\ldots,n\}$ is defined

by $\pi((i-1)\cdot \ell + j) = (j-1)\cdot \ell + i$, for $1 \le i,j \le \ell$. (A variation of this function was used in [5,6] for separating two-tape TM's from one-tape off-line TM's without output tape; before that, it had occured in [7] as an example of a permutation that is hard to realize on devices similar to Turing machines.)

The Kolmogorov complexity of a binary string is defined as follows. Let an effective coding of all deterministic Turing machines (with any number of tapes) as binary strings be given and assume that no code is a prefix of any other code. Denote the code of a TM M by $\ulcorner M \urcorner$. Then the Kolmogorov complexity of $x \in \{0,1\}^*$ (w.r.t. this fixed coding) is

$K(x) := \min\{|\ulcorner M \urcorner u| \mid u \in \{0,1\}^*, \text{ M on input u prints x}\}$.

A string $x \in \{0,1\}^*$ is called incompressible if $K(x) \ge |x|$. (A trivial counting argument shows that for all n there is an $x \in \{0,1\}^n$ with $K(x) \ge n = |x|$.)

The paper is organized as follows: in §2, we state the theorem and sketch the proof of the upper bound; in §§3-5 we indicate the main steps of the proof of the lower bound; in §6 we discuss related open problems. (The proof of the Kolmogorov complexity Lemmata is given in the Appendix.)

§2 Main Result.

Theorem. The time complexity of MATRIX TRANSPOSITION on one-tape off-line Turing machines with a one-way output tape is $\Theta(\ell^{5/2}) = \Theta(n^{5/4})$. (Here $n = \ell^2$ is the length of the input, which is a Boolean $\ell \times \ell$-matrix given in row-major order.)

Remark. The lower bound still holds if the output tape is two-way, and even if the TM can print several times to each cell on the output tape, as long as the same symbol is printed each time. The changes needed in the proof are contained in [1] and will be given in the full paper.

Sketch of the proof of the upper bound (This was already noted in [5]): Let the input $x \in \{0,1\}^{\ell^2}$ represent the Boolean $\ell \times \ell$-matrix $A = (a_{ij})_{1 \le i,j \le \ell}$, i.e. $x = a_{11} \dots a_{1\ell} a_{21} \dots a_{2\ell} \dots a_{\ell 1} \dots a_{\ell\ell}$.

Split A into submatrices A_k, $1 \le k \le \ell^{1/2}$, where A_k consists of the columns $(k-1)\cdot\ell^{1/2} + 1, \dots, k\cdot\ell^{1/2}$ of A. For $k = 1,2,\dots,\ell^{1/2}$ compute and output the transpose A_k^t of A_k as follows: first, write A_k in row-wise order on the work tape (this takes one sweep over the input, hence $O(\ell^2)$ steps); then, output A_k column by column (this takes $\ell^{1/2}$ sweeps over the representation of A_k on the work tape, which consists of $\ell^{3/2}$ bits, hence $O(\ell^2)$ steps). Altogether, A is printed column by column, and $O(\ell^{5/2}) = O(n^{5/4})$ steps are made.

Of course, there are many variations of this method that all lead to the desired effect, in particular if the output tape is two-way. In the lower bound proof it is shown that any method for computing A^t in $c \cdot \ell^{5/2}$ steps, for a sufficiently small $c > 0$, must act on a constant fraction of the input bits in a way similar as the method just described. (The head movements must allow for copying these bits to a certain worktape area, and later using the information in this area to print them on the output tape.)

§3 Printing phases, work tape intervals, and visibility.

In this and the following two sections we outline the proof of the lower bound.

Fix a one-tape off-line TM M that computes MATRIX TRANSPOSITION. Choose ℓ large enough (how large, can be seen from the proofs of the Kolmogorov complexity Lemmata), and fix an incompressible string $x \in \{0,1\}^n$, where $n = \ell^2$. (Assume for simplicity that $\ell^{1/2}/2^{20} \in \mathbb{N}$.) Consider the computation of M on x as input, consisting of, say, T steps. We want to show that $T \geq C \cdot \ell^{5/2}$, for some fixed C (e.g., $C = 2^{-20}$). The input $x = b_1 b_2 \ldots b_n$ represents $A = (a_{ij})_{1 \leq i,j \leq \ell} \in \{0,1\}^{\ell \times \ell}$, where $a_{ij} = b_{(j-1) \cdot \ell + i}$. The output $y = b_{\pi(1)} b_{\pi(2)} \ldots b_{\pi(n)}$ represents A^t.

Definition. (Printing times, printing phases.)
For $1 \leq m \leq n$, let

$t_{pr}(m)$:= the timestep at which the copy of b_m is printed to the $\pi(m)$-th output cell.

(Then $t_{pr}(\pi(1)) < t_{pr}(\pi(2)) < \ldots < t_{pr}(\pi(n))$; note $\pi = \pi^{-1}$.)

Split $\{1,2,\ldots,T\}$ into $\ell^{3/2}$ disjoint intervals P_γ (the printing phases), so that each P_γ contains exactly $\ell^{1/2}$ of the printing times $t_{pr}(m)$.

Informally, we talk of γ as the "color" of printing phase P_γ. The bits b_m whose printing time belongs to P_γ inherit the color: if $t_{pr}(m) \in P_\gamma$, both copies of b_m (the m-th input bit and the $\pi(m)$-th output bit) are said to have color γ.

First, we observe a trivial fact: if T is to be $< C \cdot \ell^{5/2}$, then M cannot print too slowly. More precisely, we can assume w.l.o.g.

that $\geq \ell^{3/2}/2$ of the P_γ consist of $< \ell$ steps. (Otherwise, M makes

$\geq (\ell^{3/2}/2) \cdot \ell = \ell^{5/2}/2$ steps, and we are done.) We choose a set

$G_1 \subseteq \{1,2,\ldots,\ell^{3/2}\}$ of colors γ such that P_γ has length $< \ell$ for all

$\gamma \in G_1$, and $|G_1| = \ell^{3/2}/2$. Further, we let
$B_1 := \{m \mid t_{pr}(m) \in P_\gamma \text{ for some } \gamma \in G_1\}$ (the set of bits with color in G_1). We will focus on these bits in the following (and regard the other bits as "uncolored"). We list some simple observations.

<u>Lemma 1.</u> Let $\gamma \in G_1$. Then

(a) on the output tape, the bits of color γ occupy an interval of $\ell^{1/2}$ cells, corresponding to a part of a column of A;

(b) on the input tape, all bits of color γ have distance $\geq \ell$ from one another;

(c) during P_γ, the work tape head visits $< \ell$ cells.

(Proof omitted.)

We now turn to the area on the work tape that the work tape head scans during P_γ. Intuitively, at the beginning of P_γ this area must contain all the information necessary to print the $\ell^{1/2}$ bits of color γ. (By Lemma 1 (b),(c), at most one bit of color γ on the input tape can be inspected during P_γ. The incompressibility of x entails that the other bits of the input inspected during P_γ will not contain any information that can be used to weaken this requirement.) For technical reasons, we need these areas to be either identical or disjoint for different γ. This can be achieved as follows, reducing the number of useful (colored) bits only by a constant fraction.

<u>Definition.</u> (Work tape intervals.)
Split the work tape into blocks of ℓ cells each. For $\gamma \in G_1$, let V_γ be an interval consisting of two adjacent blocks such that during P_γ the work tape head is always in V_γ (possible by Lemma 1(c)). Let W_γ be V_γ augmented by the block to the left and to the right of V_γ. (W_γ has 4ℓ cells.)
It is easy to see that we can fix a set $G_2 \subseteq G_1$ with $|G_2| = |G_1|/4 = \ell^{3/2}/8$ such that for $\gamma, \gamma' \in G_2$ the intervals W_γ and $W_{\gamma'}$ are either disjoint or identical. Let

$B_2 := \{m \in B_1 \mid t_{pr}(m) \in P_\gamma$ for some $\gamma \in G_2\}$, the set of bits with color in G_2, and focus on these colors and bits from here on.

There are several ways for M to get the information about the bits of color γ into V_γ, before P_γ. The most natural one gives rise to the following definition.

<u>Definition.</u> (Visibility.)
Let W be any interval on the work tape, and b_m, $1 \leq m \leq n$, any input bit. We say that b_m <u>is visible from W at step t</u> if at step t the input tape head scans b_m and the work tape head scans a cell in W.

§4 The case of many "underinformed" intervals.

For b_m a bit of color γ, we know that M prints b_m to cell $\pi(m)$ "from V_γ", that is, while the work tape head is in V_γ. Intuitively, it seems reasonable for M to make such bits b_m visible at least from W_γ at some step before $t_{pr}(m)$, to allow for a "direct" transfer of b_m from the input tape to W_γ and from there to the output tape. (Otherwise, these bits would have to be moved by the work tape head over $\geq \ell$ cells, and this takes too much time.) By a Kolmogorov complexity argument we show in the following lemma that in fact it costs M too much time to deviate too far from this obvious strategy.

Lemma 2. (The "underinformed" interval.)
Let M, ℓ,n,x be as above, ℓ large enough. Assume that W is an interval of length 4ℓ on the work tape, and that V consists of the 2ℓ cells in the center of W. Let $r \geq \ell^{1/2}/2$, and assume that there are $\geq r$ bits b_m that are printed "from V" (i.e. at a time step when the work tape head is in V) but are never visible from W before being printed. Then the work tape head spends $\geq r \cdot \ell/(16 \cdot \log n)$ steps in W.

Proof: See Appendix.

This lemma allows us to show that if M is to make $< C \cdot \ell^{5/2}$ steps, then it cannot be the case that too many "colored" bits b_m are never visible from the associated work tape interval W_γ before being printed from V_γ.

Namely, assume that there is a set $G_3 \subseteq G_2$, with $|G_3| = |G_2|/2 = \ell^{3/2}/16$, and a set $B_3 \subseteq B_2$, with $|B_3| = |B_2|/4 = \ell^2/32$, such that for each $\gamma \in G_3$ there are exactly $\ell^{1/2}/2$ indices $m \in B_3$ so that b_m has color γ and b_m is never visible from W_γ before $t_{pr}(m)$. Then for each $\gamma \in G_3$ there are

$$\geq r_\gamma := |\{\gamma' \in G_3 | W_\gamma = W_{\gamma'}\}| \cdot \ell^{1/2}/2$$

bits b_m that satisfy the hypothesis of Lemma 2 with $W = W_\gamma$, $V = V_\gamma$, namely all bits with a color γ' such that $W_{\gamma'} = W_\gamma$. From Lemma 2 it follows that M spends $\geq r_\gamma \cdot \ell/(16 \cdot \log n)$ steps with the work tape head in W_γ. By summing up these bounds for a family of $\gamma \in G_3$ that form a set of representatives for the equivalence relation over G_3 defined by $W_\gamma = W_{\gamma'}$, we see that M makes

$$\geq |G_3| \cdot (\ell^{1/2}/2) \cdot \ell/(16 \cdot \log n) = \ell^3/(512 \cdot \log n)$$

steps altogether, which is $> C \cdot \ell^{5/2}$, for ℓ large enough. Thus if M is to make $< C \cdot \ell^{5/2}$ steps, the above assumption cannot be true. This case is considered in the following section.

§5 The case of many "overburdened" intervals.

We have seen in the preceding section that w.l.o.g. we can assume that many bits b_m are visible from the associated interval before they are printed. More precisely, we will assume that there is a set $G_4 \subseteq G_2$, with $|G_4| = \ell^{3/2}/16$, and a set

$B_4 \subseteq B_2$, with $|B_4| = |B_2|/4 = \ell^2/32$ such that for each $\gamma \in G_4$

there are $\ell^{1/2}/2$ indices $m \in B_4$ so that b_m has color γ and b_m is visible from W_γ at some step before $t_{pr}(m)$. We focus on the $\ell^2/32$ bits in B_4 and "uncolor" all the other bits (and printing phases).

Definition. ("Visibility times".)
For $\gamma \in G_4$ and $m \in B_4$, where b_m has color γ, we let

$t_{vis}(m)$:= the largest $t \leq t_{pr}(m)$ such that b_m is visible from W_γ at step t.

To motivate the following definitions, we sketch the idea behind the rest of the proof. We ask at what frequency the time steps $t_{vis}(m)$, $m \in B_4$, occur. Clearly, if the average distance between two consecutive such steps is $> 32 \cdot C \cdot \ell^{1/2}$, we are done: M makes $> C \cdot \ell^{5/2}$ steps. Thus we may assume tht on the average the visibility times are closer together, hence that the following situation occurs frequently: many ($> C' \cdot \ell^{1/2}$ for a big constant C') bits b_m all have their visibility times $t_{vis}(m)$ within a time interval of length ℓ.

From the definitions it follows immediately that all these bits have different colors and that the work tape areas associated with these bits are identical or adjacent. Taking into account the other $\ell^{1/2}/2$ bits of each of these colors, we can show (Lemma 4) that for some work tape interval the following is true: there is a time step t_δ and there are many colors γ with $W_\gamma = W$, so that for altogether $> C'' \cdot \ell$ bits b_m we have

$$t_{vis}(m) < t_\delta < t_{pr}(m).$$

Here $C'' \gg 4 \cdot \log 3$, so that $C'' \cdot \ell$ far exceeds the "capacity" of W. This means that on the one hand all these bits could have been stored in W before t_δ, and that at some step after t_δ they are needed in W, but it is impossible that at t_δ all the required information is contained in W. By a Kolmogorov complexity argument (Lemma 5) it can be shown that in this situation M must spend a lot of time carrying the missing information into W.

<u>Definition.</u>

$$B_4^E := \{m \in B_4 \mid t_{vis}(m) < t_{vis}(m') \text{ for } \geq 3 \cdot \ell^{1/2}/8$$
$$\text{bits } b_{m'} \text{ of the same color } \gamma \text{ as } b_m\}$$

(the bits with "early" visibility times w.r.t. the other bits of the same color),

$$B_4^L := \{m \in B_4 \mid$$
$$t_{vis}(m') < t_{vis}(m) \text{ for } \geq \ell^{1/2}/4 \text{ bits } b_{m'} \text{ of the same color } \gamma \text{ as } b_m \text{ and}$$
$$t_{vis}(m) < t_{vis}(m') \text{ for } \geq \ell^{1/8}/8 \text{ bits } b_{m'} \text{ of the same color } \gamma \text{ as } b_m\}$$

(the bits with "late" visibility times).

Clearly, $|B_4^E| = |B_4^L| = |B_4|/4 = \ell^2/128$.

Partition $\{1,2,\ldots,T\}$ into $\ell^{3/2} \cdot 2^{-16}$ disjoint intervals P'_δ, $1 \leq \delta \leq \ell^{3/2} \cdot 2^{-16}$, so that each P'_δ contains exactly $128 \cdot \ell^{1/2}$ timesteps $t_{vis}(m)$ with $m \in B_4^L$. The P'_δ are called the <u>visibility phases</u>.

We may assume w.l.o.g. that there is a set $D \subseteq \{1,2,\ldots,\ell^{3/2} \cdot 2^{-16}\}$, $|D| = \ell^{3/2} \cdot 2^{-17}$ so that for $\delta \in D$ the visibility phase P'_δ consists of $< \ell/2$ steps. (Otherwise M makes $\geq C \cdot \ell^{5/2}$ steps, and we are done.) The following is immediate from the definitions.

Lemma 3.

(a) If $m \neq m'$ $m,m' \in B_4^L$, and $t_{vis}(m)$, $t_{vis}(m') \in P'_\delta$ for some $\delta \in D$, then $b_m, b_{m'}$ have different colors.

(b) If $t_{vis}(m)$, $t_{vis}(m') \in P'_\delta$ for some $\delta \in D$, $m,m' \in B_4^L$, and γ,γ' are the colors of b_m, $b_{m'}$, respectively, then W_γ and $W_{\gamma'}$ are either adjacent or identical.

(Proof omitted.)

For each $\delta \in D$, we choose $64 \cdot \ell^{1/2}$ bits b_m with $m \in B_4^L$ and $t_{vis}(m) \in P'_\delta$ so that all the bits chosen have the same W_γ. (This is possible by Lemma 3(b).) We then call this interval W'_δ , its central 2ℓ cells V'_δ. The subset of B_4^L consisting of all the bits chosen is called B_5^L.

Clearly, $|B_5^L| = |B_4^L| / 2 = \ell^2/256$.

Definition.

For $\delta \in D$, let t_δ be the first step of P'_δ , and let

$$B_\delta := \{m \in B_4^E \mid \text{for some } m' \in B_5^L,\ t_{vis}(m') \in P'_\delta \text{ and } b_m, b_{m'} \text{ have the same color } \gamma\}$$

(the set of "early" bits whose color "occurs" in P'_δ).

Lemma 4.

(a) b_m is never visible from W'_δ after t_δ, and b_m is printed from V'_δ after t_δ, for all $\delta \in D$ and $b_m \in B_\delta$.

(b) $|B_\delta| = 32 \cdot \ell$, for all $\delta \in D$.

(c) m occurs in at most $\ell^{1/8}/8$ of the B_δ, for all $m \in B_4^E$.

(Proof omitted; use Lemma 1 (b) and the definitions.)

Lemma 5. (The "overburdened" interval.)
Let M, ℓ, n, x be as above, ℓ large enough, and let W be an interval of 4ℓ cells on the work tape, V the 2ℓ cells in the center of W. Let $t_0 < t_1$ be time steps. Let $r \geq 16 \cdot \ell$, and assume that there are $\geq r$ bits b_m that are printed "from V" during $(t_0, t_1]$ but are never visible from W after t_0 before being printed from V. Then the work tape head spends $\geq r \cdot \ell/(8 \cdot \log n)$ steps in W during the interval $(t_0, t_1]$.

Proof:
By a Kolmogorov complexity argument, see Appendix.

One last complication arises when we try to put together the lower bounds for the time that the work tape head spends in various W'_δ during various time intervals. Lower bounds for different W'_δ can be added up, since they are disjoint. But there may be many $\delta \in D$ with the same W'_δ. The dynamics of such an "overburdened" interval may be quite complex, because the set of bits that "belong" to this interval whose visibility time is over but that have not been printed yet changes constantly. The following technical lemma, whose proof is based on Lemma 4, resolves this problem. It shows that $\{1,2,\dots,T\}$ can be split into sufficiently many disjoint time intervals to which Lemma 5 can be applied.

Lemma 6.
Let $D_0 \subseteq D$, and let W be such that $W = W'_\delta$ for all $\delta \in D_0$. Let V be the 2ℓ cells in the center of W. Then there is a $q \in N$ and there are timesteps $T = t^*_0 > t^*_1 > \dots > t^*_q$ with $t^*_1, \dots, t^*_q \in \{t_\delta \mid \delta \in D_0\}$ such that for the sets $B^*_1, \dots, B^*_q$ defined by

$$B^*_s := \{m \in \bigcup_{\delta \in D_0} B_\delta \mid b_m \text{ is printed from } V \text{ in } (t^*_s, t^*_{s-1}] \text{ and is not visible from } W \text{ in } (t^*_s, t_{pr}(m)]\}$$

holds:
(a) $|B_s^*| \geq 4\ell$ for $1 \leq s \leq q$, and

(b) $\sum_{s=1}^{q} |B_s^*| \geq 32 \cdot |D_o| \cdot \ell^{1/2}$.

(Proof omitted; use Lemma 4.)

By applying Lemma 5 in this situation, we get the result we need.

<u>Corollary</u>.
If M, ℓ, n, x are as in Lemma 5, and D_o and W are as in Lemma 6, then M spends

$$\geq 4 \cdot |D_o| \cdot \ell^{3/2} / \log n$$

steps with the work tape head in W.
(Proof omitted.)

Using the corollary, we can finish the proof of the Theorem. Let $\delta \epsilon D$ be arbitrary. Let $D_o := \{d' \epsilon D | W'_{\delta'} = W'_\delta\}$, and apply the corollary to conclude that the work tape head spends

$$\geq |D_o| \cdot 4 \cdot \ell^{3/2} / \log n$$

steps in W'_δ. Summing up these lower bounds for a family of $\delta \epsilon D$ that form a system of class representatives for the equivalence relation on D defined by $W'_\delta = W'_{\delta'}$, we conclude that M makes

$$\geq |D| \cdot 4 \cdot \ell^{3/2} / \log n = \ell^3 / (2^{13} \cdot \log n)$$

steps altogether, and this is $\geq C \cdot \ell^{5/2}$ for ℓ large enough. This finishes the proof of the Theorem.

§6 Open problems.

It seems that the methods presently available do not suffice for proving superlinear lower bounds on the computation time for concrete problems from P on TM's with two (or more) work tapes. Therefore one has started to develop lower bound arguments for a sequence of more restricted types of TM's where various abilities of two-tape TM's are "phased in" in a stepwise process. In this sequence of models that approximate the power of two-tape TM's from below we have, with the model considered in this paper, reached the "end of the rope" as far as already existing models are concerned.

A TM with two work tapes is substantially more powerful since it can move information repeatedly forth and back between the two work tapes. This suggest that one should consider the following new hierachy of restricted two-tape TM's: at the k-th level of this hierarchy consider TM's with two tapes where at any time point one of the two tapes is a two-way read-only tape, the other one is a read/write tape, and the two tapes may switch their roles up to k times. (One may also allow that k grows with n.) For $k = 1$ such a TM can perform matrix transposition as fast as the model considered in this paper, for $k = 2$ it can do it substantially faster, and for $k = O(\log n)$ one can implement on this model the standard $O(n \cdot \log n)$-algorithm for matrix transposition on two-tape TM's (via the "butterfly-graph").

Already for $k = 2$ the known arguments do not suffice for proving lower bounds for TM's from this hierarchy. In particular, it would be desirable to show that the power of these machines strictly increases with k (at least for $k \leq \log n$).

References.

[1] Dietzfelbinger, M.: Lower bounds on computation time for various models in computational complexity theory. Ph.D.-Thesis, University of Illinois at Chicago, 1987.

[2] Hennie,F.C.: One-tape off-line Turing machine computation. Inf. and Control 8: 553-578, 1965.

[3] Hopcroft,J.E., and Ullman, J.D.: Introduction to Automata Theory, Languages, and Computation. Reading (Mass.), Addison-Wesley,1979.

[4] Maass,W.: Combinatorial lower bound arguments for deterministic and nondeterministic Turing machines, Trans. Amer. Math.Soc. 292: 675-693, 1985.

[5] Maass,W., and Schnitger,G.: An optimal lower bound for Turing machines with one work tape and a two-way input tape. In Structure in Complexity Theory, ed. A.Selman, pp. 249-264. Lecture Notes in Computer Science 223. Berlin, Springer-Verlag, 1986.

[6] Maass, W., Schnitger,G., and Szemerédi, E.: Two tapes are better than one for off-line Turing machines. In Proc. 19th ACM Sympos. on Theory of Computing, pp. 94-100, 1987.

[7] Paul,W., and Stoss, H.-J.: Zur Komplexität von Sortierproblemen. Acta Informatica 3: 217-225, 1974.

Appendix: Proofs of the Kolmogorov complexity Lemmata

We begin with Lemma 5. The proof of Lemma 2 will be a slight variation of that of Lemma 5.

Proof of Lemma 5: (This is an extension of an argument in [5].) Let L (R) be the leftmost (rightmost) ℓ cells of W. (So, W is the union of L, V, R.) Choose a cell boundary c_L to the right of a cell in L so as to minimize the number of times in $(t_0,t_1]$ the work tape head crosses this boundary from left to right, and let the number of crossings be $\#C_L$. Similarly, choose a cell boundary c_R to the left of a cell in R that minimizes the number of times in $(t_0,t_1]$ the work tape head crosses this boundary from right to left, and let the number of crossings be $\#C_R$.

Clearly, the work tape head spends $\geq \ell\cdot(\#C_L+\#C_R)$ steps in L and R taken together (by minimality), and hence in W. Thus, it suffices to show that the work tape head enters $[c_L,c_R]$ (the interval between c_L and c_R) at least $r/(8\cdot\log n)$ times in $(t_0,t_1]$. For this, we describe a method for producing the input x as output of some Turing machine.

Suppose we are given:

(i) the program of M, coded as a bitstring in some standard form;

(ii) the contents of $[c_L,c_R]$ at time t_0;

(iii) the number ℓ and the positions of all three tape heads at the first time in $(t_0,t_1]$ at which the work tape head visits $[c_L,c_R]$;

(iv) the position of the input and output tape heads, and the state of M at each time the work tape head crosses c_L or c_R towards V;

(v) the bits b_m that are visible from $[c_L,c_R]$ during $(t_0,t_1]$ but are not printed from $[c_L,c_R]$ during $(t_0,t_1]$ before being visited on the input tape (these bits are given as a single string in the order they are visited by the input tape head);

(vi) the bits b_m of the input that are neither visited by the input tape head during $(t_0,t_1]$ while the work tape head is in $[c_L,c_R]$ nor printed from $[c_L,c_R]$ during $(t_0,t_1]$ (these bits are given in one consecutive string in the order they appear in x);

(vii) the code $\ulcorner M'\urcorner$ for a Turing machine that works as follows. First, simulate the computation of M on input x during all time periods that the work tape head spends in $[c_L,c_R]$: Starting with an empty input tape, using the information given by (ii) and (iii), start simulating M at the first timestep in $(t_0,t_1]$ at which the work tape head is in $[c_L,c_R]$. Whenever the input tape head visits a cell that has no bit written to it as yet, copy the next bit given by the string described in (v) to this cell, and continue the simulation. Whenever M prints a bit b_m to the $\pi(m)$-th cell on the output tape, immediately copy this bit to the corresponding m-th cell on the input tape. Whenever the work tape head leaves $[c_L,c_R]$, interrupt the simulation and resume it with the step the work tape head enters $[c_L,c_R]$ again, using the information given by (iv). The simulation is finished when the work tape head leaves $[c_L,c_R]$ for the last time, or when M halts. After this happens, fill in the bits still missing on the input tape, using the string described in (vi). Output the contents of the input tape.

It is clear that the procedure just described outputs x. So if we estimate the number of bits needed to code the information described in (i)-(vii), in the form required by the definition of Kolmogorov complexity, we obtain an upper bound for $K(x)$. For the different parts of the string, we get the following esimates:

(i): c_M bits, for some constant c_M;

(ii): $\leq 4\ell\cdot\log 3 \leq 8\cdot\ell$ bits;

(iii): $\leq 4\cdot\log n$ bits;

(iv): $\leq (\#C_L+\#C_R)\cdot(2\cdot\log(n)+c_M)$ bits;

(v), (vi):

$\leq \ell^2 - r$ bits

(recall that, by the hypothesis, $\geq r$ bits are printed from V before they are visible from W ; at least these bits are printed form $[c_L, c_R]$ during $(t_0, t_1]$ before being visited by the input tape head);

(vii): c_0 bits, for some constant c_0 .

Furthermore, $O(\log n)$ bits are needed to separate the substrings that belong to (i),...,(vii), when concatenated to a single string. (For example, we can precede each of these substrings by its length in binary, with each bit doubled.) We get

$$K(x) \leq c_M + 8\ell + (\#C_L + \#C_R)\cdot(2\cdot\log(n) + c_M) + \ell^2 - r + c_0 + c_1\cdot\log n \ .$$

Since x is incompressible, $\ell^2 \leq K(x)$. We get, for ℓ_M so large that $\log \ell_M \geq c_M$, and for some constant c'_M :

$$r - 8\ell - c'_M \cdot\log n \ \leq \ (\#C_L + \#C_R)\cdot 3\cdot\log n \ .$$

By assumption, $r \geq 16\ell$, hence $r - 8\ell \geq r/2$. This, together with a trivial transformation, yields:

$$r\cdot(\ 4 - 8\cdot c'_M\cdot\log(n)/r\ )/(24\cdot\log n) \ \leq \ \#C_L + \#C_R \ .$$

For ℓ_M so large that $8\cdot c'_M\cdot\log n < 16\ell \leq r$ we get

$$r/(8\cdot\log n) \leq \#C_L + \#C_R \ ,$$

as desired.

Proof of Lemma 2: This proof is essentially the same as the previous one, excepting that we simulate the computation of M for all time periods in $\{1,2,\dots,T\}$ the work tape head spends in $[c_L,c_R]$, and that in (ii) we just need the number of cells between c_L, c_R – the work tape being initially blank. We get the following estimate for $K(x)$:

$$\ell^2 \le K(x) \le$$
$$c_M + 2\cdot\log n + (\#C_L+\#C_R)\cdot(2\cdot\log n + c_M) + \ell^2 - r + c_0 + c_1\cdot\log n ,$$

hence, for ℓ_M large enough,

$$r - c'_M\cdot\log n \le (\#C_L+\#C_R)\cdot 3\cdot\log n .$$

As before, we conclude that

$$r/(4\cdot\log n) \le \#C_L+\#C_R ,$$

for ℓ_M large enough; hence M spends

$$\ge \ell\cdot r/(4\cdot\log n)$$

steps in M.