Writing on a computer: some discouraging experiences

Mikhail B. Sevryuk

Institute of Energy Problems of Chemical Physics,
The Russia Academy of Sciences,
Lenin prospect 38, Bldg. 2, Moscow 117829, Russia
E-mail: `rusin@chph.ras.ru`

October 6, 1998

Abstract

Apart from great advantages, preparing a scientific paper on a computer has also undesirable side effects. In the present text, we point out some of down sides of electronic writing and discuss how one can avoid them.

$\mathcal{A}\mathcal{M}\mathcal{S}$-TEX also provides a font as a concession
to perverse notational usage.
Mathematicians used to write things like $\mathbb{R}$ on the blackboard,
or in their manuscripts, to indicate boldface.
Now some of them use similar symbols in print,
so $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX has a "blackboard bold" font `\Bbb`.
— M. D. Spivak, The Joy of TEX (1990)

Blackboard bold, as in $\mathbb{A}$, $\mathbb{B}$, $\mathbb{C}$,
is provided in the `amsfonts` package
with the `\mathbb` command.
— G. Grätzer, Math into LaTEX (1996)

Modern text processors such as `Microsoft`® `Word for Windows`® and markup languages such as TEX and its descendants LaTEX or $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX have revolutionized the process of writing, especially in mathematics and physics. I remember those remote times when mathematical papers were prepared by a typewriter and formulas were filled in by hand, all the special characters, fonts, sub(super)scripts being marked by underlining, encircling, and penciled notes. I disliked filling in the equations in manuscripts, this seemed to me the most unpleasant part of a mathematician's life, and I dreamt of an equipment that would save us from that distressing duty. One of few reveries that have come true completely!

A lot of books, brochures, manuals, and articles are devoted to computerized writing and typesetting, the advantages they offer, and their role in the modern organization of scientific researches. However, my experience as well as the experience of many of my colleagues show that one should use such powerful systems as TeX or `WinWord` rather guardedly. Preparing papers and books electronically may cause (and does cause!) some errors and troubles entirely impossible in the past, in the era of typewriters and hand-written equations. Other misprints (e.g., an unwanted hyphen in the middle of a word) have become much more frequent. It is easy to avoid these troubles, one must just be aware of them. Nevertheless, I have never encountered any mention of those down sides of electronic writing in standard manuals for authors of mathematical papers. I therefore consider it expedient to share my impressions with others.

There is a widespread opinion that TeX is almost infinitely portable. This is perfectly true in the sense that TeX documents are not corrupted in e-mail transmission—just because they are `ASCII` files. But there is no guarantee that the same TeX file would lead to the same printing output after having been processed on two different computers using two different TeX interpreters and font packages. This is probably the case if, say, the file at hand is a LaTeX 2.09 document containing no commands beyond Leslie Lamport's standard 1986 manual. But even in this situation, LaTeX systems installed on different computers break the text into pages rather differently (according to my experience)—without any modifications of the standard LaTeX styles.

On the other hand, the fonts and symbols provided by the standard LaTeX 2.09 packages do not meet all the needs of a mathematical user. Those packages include a vast variety of symbols, dozens of them being whimsical signs hardly ever used in actual works, and miss the blackboard bold font ("open font") and German Fraktur ("Gothic") font, so usual in mathematics. Some people find the blackboard bold symbols unattractive (I disagree with that), but in any case, almost all the users working in mathematics are forced to add additional fonts to the standard LaTeX 2.09 set of fonts to be able to produce the blackboard bold characters. The problem is that those additional fonts and the corresponding TeX control sequences installed on different computers are often incompatible, and this gives rise to many troubles. The absence of the blackboard bold font in the LaTeX 2.09 packages has been a real computer malady.

Several years ago, I prepared a long work in LaTeX 2.09, the blackboard bold symbols occurring in that text very frequently (which is typical of mathematical papers). The local TeX system I used included a beautiful blackboard bold font. The final file was transferred to another computer because the printer attached to the latter was better (the authors were supposed to submit the manuscript in the camera-ready form). It turned out that the

TEX commands invoking the blackboard bold symbols on the first computer corresponded to other symbols within the system on the second computed. Neither my coauthors nor I were aware of that circumstance. Without any warning messages, our LATEX file was compiled and printed, and, of course, nobody read the final manuscript obtained—the text had already been checked very carefully on the first computer! The manuscript was sent to the publisher and was published with incorrect symbols throughout.

By the way, contrary to the general opinion, the blackboard bold problem does not confine itself with the five letters $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$. In addition, one should consider at least $\mathbb{H}$ for quaternions, $\mathbb{O}$ for Cayley numbers (octonions), and $\mathbb{A}$ for algebraic numbers, as well as $\mathbb{F}$, $\mathbb{K}$, or $\mathbb{D}$ for an arbitrary basic field (the notation $\mathbb{F}_q$ is also used very often for the finite field of q elements and the notation $\mathbb{D}_q$, for the dihedral group of order $2q$). The symbols $\mathbb{S}^n$ and $\mathbb{T}^n$ for the n-dimensional sphere and torus, as well as $\mathbb{A}^n$ and $\mathbb{P}^n$ for the n-dimensional affine and projective spaces (over an arbitrary field) are sometimes utilized. On the other hand, $\mathbb{A}_n$, $\mathbb{B}_n$, $\mathbb{C}_n$, $\mathbb{D}_n$, $\mathbb{E}_6$, $\mathbb{E}_7$, $\mathbb{E}_8$, $\mathbb{F}_4$, $\mathbb{G}_2$ denote, in some works, the Weyl groups of the root systems of types A_n, B_n, C_n, D_n, E_6, E_7, E_8, F_4, G_2, respectively. The notation $\mathbb{H}^n$ can mean the n-dimensional hyperbolic (Lobachevskiĭ) space. One occasionally uses the blackboard bold letter $\mathbb{P}$ for the probability measure and the notation $\mathbb{L}^p$ for the space L^p of functions f such that $|f|^p$ is integrable, as well as the notation $\mathbb{H}$ for the space $L^2(\mathbb{T}^n, \mathbb{R}^n)$. I have encountered the symbol $\mathbb{D}^n$ for $\mathbb{R}^n \cup \mathbb{S}^{n-1}$, the compactification of $\mathbb{R}^n$ by an $(n-1)$-dimensional sphere of "infinite" radius. In a PhD thesis I read recently, the author denoted respectively by $\mathbb{T}$, $\mathbb{O}$, and $\mathbb{I}$ the tetrahedral, octahedral, and icosahedral subgroups of $SO(3)$. In a preprint I read very recently, the author used the blackboard bold characters $\mathbb{E}$ and $\mathbb{G}$ for certain subsets of $\mathbb{C}^N$ and $\mathbb{Z}^N$, respectively, as well as the blackboard bold $\mathbb{E}$ for any Banach space over $\mathbb{R}$ or $\mathbb{C}$, etc. etc.

Another difficulty is that in many particular installations of the blackboard bold font, its letters do not reduce in size properly while appearing as sub(super)scripts or in footnotes (the same refers to the German Fraktur characters). Of course, the wide availability of the $\mathcal{AMS}$ fonts would easy and eventually remove all these troubles.

The reason for my two-computer disaster described above is obvious. But I have also encountered other TEX calamities (not connected with blackboard bold symbols) which I cannot explain. Once, I refereed a paper prepared in TEX (judging from the fonts used). I pointed out in my report that some equations in the paper were wrong. The author agreed with my criticism and in the revised version, he corrected those equations, but other several equations turned out to be totally corrupted: they became senseless combinations of symbols! The latter equations were OK in the initial version. I had and still have no idea of how this mysterious phenomenon could happen. It was clear that the author had not read the revised version throughout, but it is so natural not to read the sections that have not been modified... (Of course, I eventually made the author produce a version where all the equations were correct.) Such a case had been impossible in the pre-computer times.

One sad event more, this time without any enigmas. A paper prepared in $\mathcal{A}\mathcal{M}\mathcal{S}$-TeX contained arrows created by the `@>>>` and `@<<<` commands. To obtain the final camera-ready output, one processed the file by LaTeX. Instead of arrows, @ >>> and @ <<< themselves were gotten, and nobody paid attention to these meaningless symbols.

Many TeX users often forget typing a backslash before the names of Greek letters and then do not check the printed output carefully. A lot of manuscripts I refereed were choked with the English names of Greek letters in place of the actual Greek letters (and it was in fact almost impossible to examine such papers on their mathematical merits). This "misprint", extremely popular in the present age of TeX, did not occur at all, of course, in the good old days of typewriters and handwritten formulas. As a rule, it cannot filter through the refereeing and copy-editing system of the best journals, but I observed the names of Greek letters in the formulas even in Springer books. I also encountered TeX-prepared manuscripts with blank spaces in place of some formulas, the origin of this misfortune being unclear to me.

Up to now, I discussed TeX, but similar effects happen to `Word for Windows`® documents as well. As a rule, the rich variety of `Windows` fonts the `Word` exploits does not include the blackboard bold font. I experienced the cases where the `WYSIWYG` principle failed for `WinWord` (the printed text exhibited blank spaces instead of characters perfectly visible on the screen—even in the preview mode). Transferring a `WinWord` file (as well as a graphic `SigmaPlot`® file) to another computer may destroy the document almost completely. I myself encountered the following effects:

a) several symbols in equations were corrupted in a horrible and entirely irregular way (e.g., $\hbar$ and $\widehat{H}$ turned into η and $\vec{H}$, respectively, while 1 and W were replaced with blank spaces), of course, this refers to the `Equation Editor` rather than to the `WinWord` itself;

b) additional symbols appeared;

c) some Greek letters were replaced with the "corresponding" Latin ones (e.g., α with a or γ with g).

Needless to say, all this occurred without any warning messages. Such symbols as $\vec{H}$ (with overlapping H and $\exists$) were unknown before sophisticated text processors conquered the world.

Curiously, in all the (not numerous) `WinWord` installations with the blackboard bold font I have seen, the blackboard bold characters do not reduce in size automatically while appearing in footnotes—like in some TEX installations.

How can one overcome all the above-mentioned unpleasant side effects of computerized typesetting? The following two very simple rules seem to be sufficient:

1) *Print the whole document and read carefully through it after every change in the environment you use (installing a new version of the system, adding a new font package, transferring the file to another computer, etc.).*

2) *Read carefully once more through the final manuscript before submitting it to the journal or sending it to the publisher.*

Nowadays, many journals ask their contributors to send a diskette with the file of the paper along with the hard copy. In the instructions for authors, it is often pointed out that electronic submission enables one to avoid introducing errors while processing the paper. This contradicts heavily the experience of many contributors, including my own. A lot of my papers were submitted together with the electronic version (either in TEX or in `WinWord`). I strictly followed all the instructions of the journals. The result has been always the same—numerous misprints in the proofs, as if the text were rekeyed. I do not know the origin of this fact.

Moreover, using electronic versions of the papers leads sometimes to rather specific errors. Beginning users of LATEX often ignore the LATEX cross-reference system for numbering equations and other elements of the text. They create numbered displayed equations by the `equation` environment, but in a text reference to any particular equation, type the explicit number of the equation instead of using the `\ref` command. Journal copy-editors sometimes change the equation numbering (e.g., via breaking a numbered equation into two parts occupying two lines, in which case each part gets its own number provided that the new environment is `eqnarray`, this would shift the numbers of all the subsequent equations). Since the manuscript is formally prepared in LATEX, the copy-editor does not bother about the necessity to change the text references accordingly. The result is highly regrettable—the actual equation numbers and the references to the equations within the text do not match. I experienced this twice.

One should therefore follow the third simple rule:

3) *Do not trust electronic submission. Check the proof as carefully as you can even if the printed final version of the paper was accompanied with the file.*

On the other hand, when the publisher uses the file supplied by the author(s) without carefully copy-editing (not to mention when a camera-ready output is utilized), the result is often unsatisfactory as well. The authors' files seldom meet all the standard printing conventions—e.g., a uniform style of the references or (in the case of TEX) the right size of the "delimiters" which is ensured by typing `\left` and `\right` commands—and a certain intervention is almost always needed to produce a professionally designed paper.