

POSTING PHYSICS DOCUMENTS ON THE WEB

**Kevin Zollman
and Dean Zollman**
Department Editors

Many physicists use the World Wide Web to post papers that are being considered for publication. Web posting provides for wide distribution and saves on postage and paper costs in comparison with the mailing of pre-prints. The primary question is how to prepare the papers so that they display properly on the Web.

Most physicists will have prepared the manuscripts for the papers using a word processor, such as Microsoft Word or WordPerfect, or using some variation on TeX. Once the manuscript is prepared for printing, we have three possibilities for putting it on a Web server in a form that others can use:

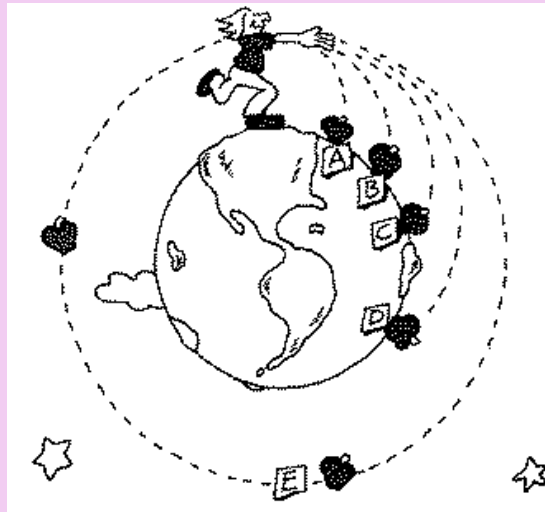
- Convert it to Hypertext Markup Language (HTML).
- Convert it to a document format that is “portable” across different platforms.
- Leave it in the existing file format.

Each of these methods has some advantages and disadvantages, and so we shall treat them one at a time and provide “live” examples that readers can download and view, so as to experience the differences (see “Demonstrations and Source Code,” p. 134).

Kevin Zollman is a senior at Manhattan High School in Kansas and Webmaster for the Physics Department at Kansas State University. E-mail: kzollman@phys.ksu.edu

Dean Zollman is professor of physics at Kansas State University, Manhattan, KS 66506-2601. E-mail: dzollman@phys.ksu.edu

If the surface of the earth were flat, there would be little more to say. The distance the apple travels outward before hitting the ground would simply depend on how hard you threw it. Since the earth is curved, however, the apple falls just a bit further each time it is thrown, as shown in the figure below. The vertical distance it must travel increases because of the earth’s curvature. If we could throw the apple hard enough so that it traveled from North America to, for instance, Africa (path D), it would have an enormous drop before hitting the ground. If we could throw it even harder still, it would eventually miss the earth entirely (path E). The apple now keeps falling and falling, but it never quite catches up with the earth’s surface. Your apple behaves like the moon.



... The law of universal gravitation states that the magnitude of the force is:

$$F = -G \frac{m_1 m_2}{r^2}$$

$$F = -G \frac{\mu M}{r^2}$$

Figure 1. Sample text includes an illustration and two forms of the gravitational equation.

Sample documents

As demonstration examples, we have taken two short excerpts from an out-of-print textbook¹ to represent typical text and a short spreadsheet file to represent data.

The materials adapted from the textbook are a short paragraph followed by a figure and a sentence followed by an equation (see Fig. 1). The equation, Newton’s Law of Gravitation, has been written in two different notations—one involving subscripts; the other, a Greek letter. These different notations will be used to demonstrate some of the problems that we discover as we try to put documents on the Web.

The short spreadsheet file is a summary of simple projectile-motion data, as shown in Fig. 2. The original spreadsheet contained the numerical data and a graph of y versus t .

Conversion to HTML

A simple way to “convert” documents is to copy a

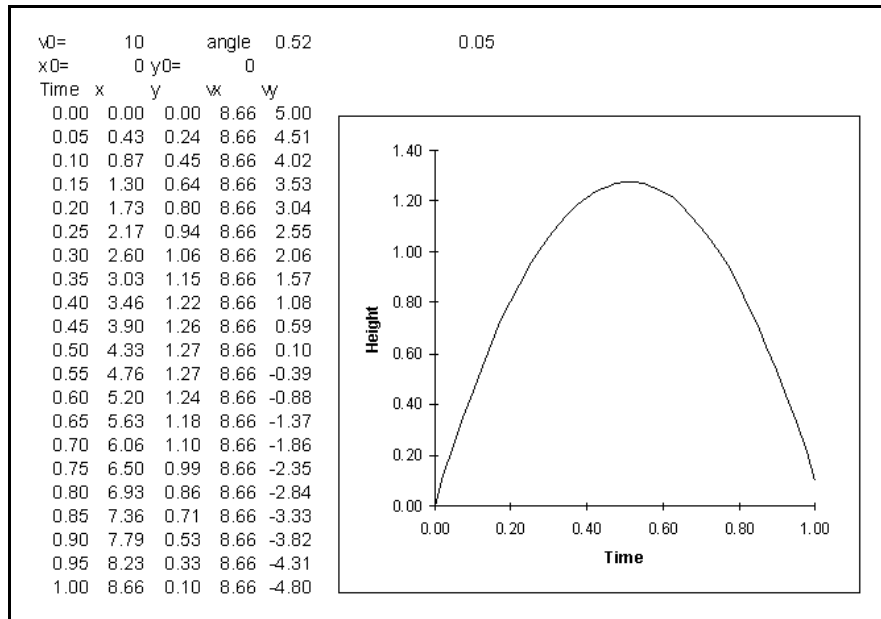


Figure 2. Sample spreadsheet summarizes projectile motion in data and includes a simple graph.

document file and paste it into an HTML editor such as Netscape Gold 3.0. For files that contain only text, this approach works rather well. However, special formatting such as indenting, tabs, and tables disappears with this simple conversion. Such formatting must be recreated in the HTML editor. Equations and graphics are lost completely, and so they need to be inserted by hand. For Netscape Gold, each graphic needs to be placed in its own file, and a statement that calls that file must be inserted into the document.

In converting the document in Fig. 1, we saved the figure as a GIF file. Then we used a screen-capture program to make an image of the equations, and we also converted it to a GIF file. Inserting the two GIF files at the appropriate places created a document that looks quite similar to the original. The HTML editor generated the source code shown in Box 1.

The HTML editor provides the capability of creating some equations directly. For example, if we are satisfied with the Law of Gravitation as

$$F = -G(m_1 m_2)/r^2$$

we can use the following line of HTML:

```
<P>F=-G(m<SUB>1</SUB>m<SUB>2</SUB>)/r<SUP>2</SUP></P>
```

The current version of 3.2 HTML cannot produce Greek letters. This capability would be an important element in scientific communication. For the present, when dealing with full-blown equations and Greek letters, we are stuck with inserting them as a GIF file, unless we make the document available in some other, non-HTML format. We discuss some examples below.

The cut-and-paste method does not work for converting the spreadsheet. The formatting of the table disappears, and all the numbers appear in one line on the screen. Even using the Insert Table option in the HTML editor produces similar, unsatisfactory results.

Most word-processing programs now have an add-on utility that allows automatic conversion of any file to an HTML document. WordPerfect 7.0 (<http://www.corel.com>) even has an HTML converter built into it, whereas Microsoft Word 6.0 or 7.0 users must obtain a free add-on to complete the conversion (<http://www.microsoft.com>). Microsoft is promising built-in functionality in Word 97, which is to be released by the time this article is published. If you use a word-processing program for creation of the original document, then you will be able to create an HTML document rather easily.

To convert a document to HTML in MS Word, you just need to "Save As" an .htm document. Just as with cutting and pasting, the conversion of text goes very smoothly, although some formatting can get lost. Unfortunately, our version of the Word Internet Assistant was unable to save graphics correctly. Both the figure and the equations in Fig. 1 were treated as pictures, and clicking on a message box that asked us if we wished to save the pictures always resulted in an error message.

To obtain the final document we inserted the figures into the document in the same way as we did with the HTML editor. Then the word processor generated HTML source code that looked essentially identical to the code in Box 1.

From our spreadsheet of projectile data, Internet Assistant in Excel produced a usable table for viewing on the Web. Again, however, the utility did not convert the graphics, which had to be inserted into the file with the help of our

Demonstrations and Source Code

The Department Editors are making available on the Internet live Web Mechanics demonstrations to illustrate the procedures described in this article and in other articles in this series. These demonstrations will be available during one year following publication in the magazine, and they can be accessed at the URL <http://bluegiant.phys.ksu.edu/cip/>. For present and future reference purposes, all source code from this series of articles will be maintained indefinitely by *Computers in Physics* and made available for downloading at the URL ftp://ftp.aip.org/cip/cip_sourcecode.



HTML editor. Part of the HTML for the spreadsheet is shown in Box 2. (We have not reproduced the entire HTML source because it repeats the same code for each row of the table.)

Portable file formats

Even prior to the popularity of the Web, file formats that would give identical displays on many different machines were being developed. One of the most popular is Adobe's Portable Document Format (PDF—see <http://www.adobe.com>). The software for creating and displaying PDF files is Adobe Acrobat. It lets us all view and comment on the same fully formatted version of a document even though you, for example, may be using a Macintosh, one of us has Windows, and another person is running a Unix machine. Using Acrobat Exchange, you can convert a document to a PDF file that looks identical to the printed version when displayed. We can view that file with Acrobat Reader, which is available free from Adobe, or, if we also have Acrobat Exchange, we can leave "sticky notes" on the document for your consideration.

Acrobat Exchange is a commercial product with a suggested retail price of about \$295. Once it is installed on a machine, you create a PDF file as simply as printing. When you select "print" from the menu, one of the choices for "printer" is an Acrobat PDF file. The printed file may be viewed by anyone with the free reader. For example, see the PDF versions of Figs. 1 and 2 at the URL <http://BlueGiant.phys.ksu.edu/cip/marapr/>.

Once the document has been "printed" to an Acrobat file, you can use Exchange to add more features. Bookmarks can be created, and thumbnails of each page can be placed in the left margin. These options allow users to move around in the document. Although such features are superfluous in our half-page example, we have put them in anyway. Within the document we can also have links to other Web pages, and so we have added two links to other, longer Acrobat files at our Web site.

For our spreadsheet data file, Acrobat works quite well. We selected the text and graphs to be printed to the PDF file. The result was a document that is easily read and looks just like our spreadsheet.

For use with the Web, you need to install the Acrobat Reader Plug-in, also available free from the Adobe Web site. When the Web browser encounters a PDF file, it automatically launches Acrobat Reader. You are then looking at the document just as it would have printed. Equations, figures, and any special formatting are as

they should be. You can choose to print the document, if its creator has not disabled this function.

Acrobat has some disadvantages. You need to purchase a program to be able to create the files—only reading is free. The file-creation process is relatively slow. The single page that appears in Fig. 1 required 30 s on a 133-MHz Pentium with 32 Mbytes of RAM. Graphics slow the process down considerably. The biggest disadvantage is related to the creation of links, bookmarks, and other features added to the document after it becomes a PDF file. Because Acrobat Exchange has no editing features, we must return to our word processor to make any text and graphics changes. When we "print" again, all the bookmark and link information is lost and must be tediously recreated.

The advantages are related to ease of use, to getting a Web document that is identical to the printed version, and to being able to add enhancements such as links and bookmarks. Furthermore, Acrobat files are being adapted by organizations of importance to physicists. For example, many organizations including the American Institute of Physics (<http://www.aip.org/ojs/service.html>), the American Physical Society (<http://publish.aps.org/indexjrnls.html>), and the Institute of Physics (<http://www.ioppublishing.com/Journals/>) are making their journals available online. For these organizations the articles are stored in two formats, one of which is PDF. (The other format varies with the organization.) Thus, you may view individual articles in

Box 1. Cut and paste to an HTML editor created this source code. The illustration and equations were dropped in separately as GIF files.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2//EN">
<HTML>
<HEAD>
  <TITLE></TITLE>
  <META NAME="Author" CONTENT="DEAN ZOLLMAN">
  <META NAME="GENERATOR" CONTENT="Mozilla/3.0Gold (Win95; I) [Netscape]">
</HEAD>
<BODY>

<P>If the surface of the earth were flat, there would be little more to say. The distance the apple travels outward before hitting the ground would simply depend on how hard you threw it. Since the earth is curved, however, the apple falls just a bit further each time it is thrown, as shown in the figure below. The vertical distance it must travel increases because of the earth's curvature. If we could throw the apple hard enough so that it traveled from North America to, for instance, Africa (path D), it would have an enormous drop before hitting the ground. If we could throw it even harder still, it would eventually miss the earth entirely (path E). The apple now keeps falling and falling, but it never quite catches up with the earth's surface. Your apple behaves like the moon. </P>

<P><IMG SRC="z2fig1.gif" HEIGHT=317 WIDTH=358></P>

<P>... The law of universal gravitation states that the magnitude of the force is:</P>

<CENTER><P><IMG SRC="z2eq1.gif" HEIGHT=98 WIDTH=120></P></CENTER>
</BODY>
</HTML>
```

Box 2. Excel Internet Assistant generated this HTML source code for our sample spreadsheet (excerpt only). Graph had to be dropped in separately as a GIF file.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2//EN">
<HTML>
<HEAD>
  <TITLE>Fig. 2
</TITLE>
  <META NAME="GENERATOR" CONTENT="Mozilla/3.0Gold (Win95; I) [Netscape]">
</HEAD>
<BODY>
<H1>Projectile data</H1>
<P>
<HR><IMG SRC="z2proj.gif" HEIGHT=165 WIDTH=315 ALIGN=ABSCENTER><!-- The
following table has been generated by the Internet Assistant Wizard for Microsoft Excel. You can
find this add-in on "http://www.microsoft.com/msoffice/freestuf/msexcel/index.htm" -->
<!-- START OF CONVERTED OUTPUT -->
<TABLE>
<TR VALIGN="bottom">
<TD ALIGN="left">Time </TD>
<TD ALIGN="left">x</TD>
<TD ALIGN="left">y</TD>
<TD ALIGN="left">v<SUB>x</SUB></TD>
<TD ALIGN="left">v<SUB>y</SUB></TD>
</TR>
<TR VALIGN="bottom">
<TD ALIGN="right">0.00</TD>
<TD ALIGN="right">0.00</TD>
<TD ALIGN="right">0.00</TD>
<TD ALIGN="right">8.66</TD>
<TD ALIGN="right">5.00</TD>
</TR>
<TR VALIGN="bottom">
<TD ALIGN="right">0.05</TD>
<TD ALIGN="right">0.43</TD>
<TD ALIGN="right">0.24</TD>
<TD ALIGN="right">8.66</TD>
<TD ALIGN="right">4.51</TD>
</TR>
</TR>
</TABLE>
<P><!-- END OF CONVERTED OUTPUT -->
<HR></P>
<P>Last Update: 12/27/96 </P>
<P>Name: Dean Zollman </P>
</BODY>
</HTML>
```

has significant advantages over straight text.

Leave it alone

Your third option for posting a document on the Web is to leave the file in its original format. In this case, the users should take advantage of the “Helpers” facility on their Web browser. Under “Helpers,” users indicate the software that they would like to have launched when their browser encounters files of a particular type. For example, under general preferences in Netscape Navigator we can specify that any time Navigator encounters a filename with the extension .doc, it should open the copy of Microsoft Word that is located on our machine. If everyone who wishes to view our files uses the same helpers as we do, or some other software that is compatible with our helpers, posting of documents can be quite easy.

For posting, the document is simply saved to file space on the Web server. For example, the file of the document described in Fig. 1 is available on our server under the title box2-1.doc. If you have instructed your Web browser to use a word-processing program every time it sees a .doc file, you will be able to open the file with no problem. Your machine will automatically launch a copy of your word processor and load the document into it.

For this approach to work satisfactorily on the Web, the users must have configured their browsers ap-

a format that looks identical to the printed version using your Acrobat plug-in to your Web Browser.

You may also obtain guidelines and submit proposals to the National Science Foundation in electronic form. NSF allows only two formats—ASCII text and Acrobat—for submission. Moreover, some NSF application forms and information are available in Acrobat format (see <https://www.fastlane.nsf.gov/>). If you are submitting an electronic proposal with equations or other special formats, Acrobat

appropriately. They should choose “General Preferences,” then “Helpers,” and specify the location of the programs. The process is rather user-friendly, and so most people can do it. However, they may not know that they need to do anything and thus never get to your file. We suggest that you place a message on the link to the file that says, for example, “WordPerfect or compatible word processor needed as a browser helper to view this file.”

Word-processor vendors are cooperating by making



available free viewers for the document files. These viewers can be used for displaying, not editing, documents. Moreover, all vendors provide some capability for conversion of other vendors' documents for use within their own programs. Our MS Word software can read WordPerfect documents and vice versa. However, a time lag always occurs between the introduction of a new version of a software package and the competition being able to read from that version. The solution for most documents is to save them on the Web in an older format.

WordPerfect 5.1 and MS Word 2.0 can be displayed by essentially all word-processing programs. If you save documents in one of these formats, anyone who has set up an appropriate helper application will be able to read it. On our Web site, we have posted our sample document in both formats.

A similar approach works for spreadsheets. We told our browser to open Excel when it encounters a .xls file, and we are therefore now able to download and view such files without any problem. We can also work interactively with the data, because we receive all the spreadsheet formulae. We can change values and see how other variables are affected. This approach can be particularly useful when collaborators wish to share data over the Web. You can place the file on a server and tell your colleagues what software to use to open it. You should not care, in this case,

if the rest of the world receives an error message.

This approach is not limited to word-processing and spreadsheet programs. You can display PostScript files with a viewer such as GhostView and TeX files with one of the many TeX viewers available. Files from a variety of other programs, including algebraic manipulators, could be treated similarly.

Making the best choice

For physicists, placing documents on the Web is not difficult, but it is not yet automatic. Our need to communicate with equations and diagrams makes straight HTML coding unsatisfactory. With a little work, however, an HTML document can include graphics. Other options, such as converting to portable file formats or specifying helper applications for existing file formats, are probably easier in terms of preparation, but require a little more computer savvy on the part of the surfers who will view your documents. The choice depends on the time you wish to devote to posting your documents and the sophistication of your users.

Reference

1. Jacqueline D. Spears and Dean Zollman, *The Fascination of Physics* (Addison-Wesley, Reading, MA, 1985), pp. 155–156.