

# Aprendizaje Activo y Curvas Características al Ítem con $R$ y TikZ

EDVATIO PHYSICORVM



ISSN 1870-9095

**Rubén Sánchez Sánchez**

*Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada del Instituto Politécnico Nacional, Unidad Legaria. Calzada Legaria No. 694, Colonia Irrigación, Alcaldía Miguel Hidalgo, C.P. 11500.*

**E-mail:** rsanchezs@ipn.mx

(Recibido el 2 de abril de 2024, aceptado el 4 de mayo de 2024)

## Resumen

En este documento se detallan los pasos que se pueden seguir para encontrar las Curvas Características al Ítem correspondientes a un trabajo de investigación en Educación de la Física con una metodología activa de enseñanza. Suponemos primero que hemos obtenido los datos de un grupo hipotético que ha seguido los pasos del Aprendizaje Activo de la Física, empleando el ciclo PODS en el aula de clases, con los estudiantes y dirigiendo su proceso, teniendo cuidado de que el estudiante lleve a cabo sus actividades didácticas, en acorde a este esquema didáctico metodológico. Una vez que los datos se recolectan en la etapa del pretest y postest (empleando un cuestionario elaborado para tal fin), empieza el proceso de análisis de datos. En este ejemplo que sigue utilizamos hojas de Microsoft Excel para recabar los datos de cada uno de las etapas antes mencionadas y gracias a la tecnología digital podemos hacer uso del programa estadístico  $R$  para realizar el análisis matemático de sus Curvas Características al Ítem de acuerdo al modelo dicotómico de Rasch. La forma en cómo se manejan los datos en  $R$ , y como se pueden utilizar no sólo para realizar las gráficas de las curvas en  $R$ , sino para reprocesarlas para mejorar su definición en sus gráficas de reporte, utilizando un programa de graficación junto con el de publicación de artículos LaTeX, será el tema que se tratará en este trabajo, donde se detalla el manejo detallado de los datos, su almacenamiento en memoria, su manejo con paquetes de software y su posterior manejo para reformatear los resultados para poder ingresarlos y graficarlos en documentos PDF con el uso del software de graficación TikZ que normalmente se puede emplear junto con el cuerpo del reporte en LaTeX. Esperamos que este ejemplo sea de utilidad a la comunidad de investigadores para la enseñanza de la Física, para poder presentar y publicar sus trabajos de una forma sencilla y limpia.

**Palabras clave:** Aprendizaje Activo de la Física, tecnologías en educación de la Física, reprocesamiento de datos.

## Abstract

This document details the steps that can be followed to find the Item Characteristic Curves corresponding to a research work in Physical Education with an active teaching methodology. We first assume that we have obtained data from a hypothetical group which has followed the steps of Active Learning in Physics, employing the PODS cycle in the classroom, with students and directing their process. The student must be careful to carry out his or her teaching activities in accordance with this methodological teaching scheme. Once the data is collected in the pretest and posttest stage (using a questionnaire developed for this purpose), the process of data analysis begins. In the following example we use Microsoft Excel sheets to collect data from each of the above-mentioned stages and thanks to digital technology we can make use of the statistical program  $R$  to perform the mathematical analysis of its Characteristic Curves According to the dichotomous model of Rasch. How data is handled in  $R$ , and how it can be used not only to make graphs of curves in  $R$ , but also to reprocess them to improve their definition in your report graphs, using a graphics program together with the LaTeX article publishing, will be the topic to be treated in this work, where detailed data handling, storage in memory, its handling with software packages and its subsequent handling to reformat the results in order to be able to enter them and graph them in PDF documents using the TikZ graphics software that can normally be used together with report body, in LaTeX. We hope this example will be useful to the research community for physics education, so that they can present and publish their work in a simple and clean way.

**Keywords:** Active Learning in Physics, technologies in Physical Education, data reprocessing.

## I. INTRODUCCIÓN

El Aprendizaje Activo ha sido una metodología didáctica que ha tenido mucho éxito en los Estados Unidos [1, 2, 3, 4]. Comentamos brevemente sus ventajas así como el proceso principal que se sigue en esta metodología de enseñanza. Una vez que aplicamos la metodología, el trabajo del investigador debe de adaptarse al correcto almacenamiento y el análisis

estadístico de sus datos, para poder obtener de este trabajo de investigación educativa, el mayor provecho, posible correspondiente a su análisis. Una de las técnicas para el análisis, la podemos hallar con el modelo dicotómico de Rasch [5, 6, 7]. Aquí queremos mostrar con un ejemplo de datos hipotéticos y cómo podemos utilizar el software  $R$  [8], para poder obtener las *Curvas Características al Ítem* (o en inglés: *Item Characteristic Curves, ICC*) [9], y posteriormente

poder pasar estos datos a un programa graficador para su preparación en un artículo que se pueda publicar en una revista de investigación educativa. La forma en cómo manejar los datos de un trabajo de investigación educativa, merece especial cuidado y tener tanto los hábitos como los recursos para realizar un buen análisis estadístico, así como manejar una buena presentación de los resultados de la investigación, ayuda a un mejor entendimiento de los resultados, que nos ayuden a comprender mejor la ganancia en el aprendizaje de los estudiantes de Física.

## II. ENTRADA DE DATOS

Suponemos, para este ejemplo, que hemos realizado un estudio sobre la eficiencia, de un grupo que ha aplicado el *Aprendizaje Activo de la Física* junto al ciclo PODS (por las etapas básicas de aprendizaje: Predicción, Observación, Discusión y Síntesis), para el entendimiento de la *ley de reflexión de la luz sobre una superficie plana* y se les ha evaluado con un test validado para tal fin a semejanza por ejemplo del FTGOT [10], del sitio de PhysPort [11]. Aplicamos un pretest antes de nuestra metodología didáctica y posteriormente, ya que han aprendido el principio de la reflexión luminosa con una Clase Demostrativa Interactiva, les aplicamos un respectivo postest.

Nosotros queremos obtener las curvas características al ítem o ICC (*Item Characteristic Curves*) para poder comparar nuestros resultados de la probabilidad de respuesta correcta ante la misma pregunta o ítem de la prueba o test. Sin embargo, por razones técnicas, nuestras gráficas no son lo suficientemente claras al pasarlas al formato para publicación en LaTeX y quisiéramos mejorarlas empleando un formato más nativo, para los PDF generados vía LaTeX. Aquí es donde el manejo de los resultados de uno de los programas que manejamos a otro es necesario.

Suponiendo que las ICC las hemos obtenido usando *R*, y las queremos pasar a TikZ, entonces hay que saber hasta cierto grado, como las curvas se representan internamente en *R*, y que hay que hacer para pasarlas en TikZ.

Para simplificar supongamos que tenemos del postest unos pocos datos como se muestra en la figura 1.

Luego procesamos los datos utilizando el paquete *ltm* (*Latent Trait Models under IRT*) [12] de *R*, y obtenemos los resultados de los parámetros de cada curva ICC. En las siguientes secciones mostramos un ejemplo simplificado del proceso completo para obtener en un formato más propio al PDF las gráficas ICC.

Como el proceso y el criterio va a ser el mismo tanto para el pretest como para los resultados del postest, vamos a simplificar los pasos mostrando sólo una parte ilustrativa del mismo.

## III. HALLAR LOS RESULTADOS DE ICC EN R

Supongamos que tenemos una entrada de datos, para ejercer una prueba de conocimientos del grupo, antes de aplicar la

metodología activa de enseñanza. Entonces esto representa el pretest para un grupo de 20 estudiantes dados por una hoja de Excel mostrada en la figura 1.

	A	B	C	D	E
1	Q1	Q2	Q3	Q4	Q5
2	0	0	0	0	0
3	0	0	1	1	0
4	0	1	0	0	0
5	0	1	0	0	1
6	0	1	0	1	0
7	1	0	1	0	1
8	0	1	0	0	1
9	1	0	0	1	0
10	0	0	1	1	0
11	1	1	0	0	0
12	1	0	1	1	1
13	0	1	0	0	0
14	1	0	1	0	1
15	0	1	0	1	0
16	0	0	1	0	1
17	1	1	1	0	0
18	0	0	1	0	0
19	0	1	0	0	1
20	0	1	1	0	0
21	0	0	0	1	1

**FIGURA 1.** Ejemplo de una muestra ficticia de un grupo de 20 estudiantes que contestan un test de 5 preguntas (Q), los resultados de la prueba están en Excel. Un “1” indica que el alumno contestó correctamente a la pregunta (Quest) y un “0” indica que el estudiante contestó incorrectamente. Suponemos que esto está en una hoja Excel o “sheet” llamada “pretest”. Se observa que tenemos nombres de columna como “Q1”, “Q2”, . . . , “Q5”, marcados en amarillo.

Posteriormente se escribe un archivo en lenguaje *R* [13], para leer los datos del pretest mostrados en la figura 1, con una hoja de Excel, y convertirlos a un formato nativo que *R*, puede entender. La salida de este archivo lo nombramos “RaschEntry.RData”.

Entonces, supongamos que tenemos un archivo de *R*, llamado “leer.R” con el siguiente código sencillo:

```
# Archivo “leer.R”
library(readxl)
# Lectura del archivo en Excel “pretest.xlsx” mostrado en la
# figura 1.
x1 <- read_excel("pretest.xlsx", sheet = "pretest", col_names
= TRUE, col_types = "numeric")
x1
save(x1, file="RaschEntry.RData")
q()
```

Después escribimos otro archivo en *R*, donde se releen los datos de entrada para el postest, al cual llamamos “ICCpre.R” con el siguiente código:

```
# Archivo “ICCpre.R”
# Carga de datos de entrada en un formato que R entiende
```

```
load ("RaschEntry.RData")
# Carga de la librería ltm, para graficar las ICC
library(ltm)
x1
pd1 <- rasch (x1, constraint = cbind(ncol(x1) + 1, 1),
IRT.param = TRUE)
pdf("plot.pdf") #Véase la figura 2 de las curvas ICC
rasch1 <-plot (pd1, type="ICC")
dev.off()
write.table(rasch1, "pretest.dat")
q()
```

La salida del código ICCpre.R es un archivo de datos donde están los puntos de las ICC para graficar en TikZ. Pero antes, hay que reformatear la salida del archivo "pretest.dat", con un programa externo que traslade la salida, en un formato que entienda TikZ, y lo podamos usar como archivo de entrada para un programa graficador en el código TikZ.

#### IV. TRANSFORMAR LOS RESULTADOS DE LAS CURVAS ICC DE R PARA TikZ

Ahora debemos usar un programa que trate la salida de este código en R. Ese código puede estar escrito en varios lenguajes *script* como lo pueden ser Python [14], Ruby [15], o Perl [16].

En este ejemplo, utilizaremos un programa escrito en el lenguaje de programación Perl para formatear la salida de R, que recoge los datos a graficar recogidos en el archivo "pretest.dat" y tenga de salida unos datos de graficación que llamaremos "fig4.dat". En este tipo de archivo, todos los datos de graficación usan espacios entre sí, y el primer renglón sirve como guía a los otros renglones, del archivo.

```
#!/usr/bin/perl -w
# Archivo reformateador en Perl "sinpi4.pl"

use strict;

# Abriendo archivos de entrada y salida
open(FILE, "<pretest.dat");
open(OUT, ">fig4.dat");

# Leyendo archivo de entrada
# y realizando cambios con expresiones regulares
while($_ = <FILE>) {
    $_ =~ s/^"//g;
    $_ =~ s/^\d\d?\d?\s//;
    $_ =~ s/(Q|P)//g;
    print OUT $_; #Imprimiendo el archivo de salida
}

# Cerrando el archivo de entrada y el de salida
close(OUT);
```

```
Aprendizaje Activo y Curvas Características al Ítem con R y TikZ
close(FILE);
```

Nótese, que por ejemplo, la línea con (Q|P) es para quitar la letra "Q" (de *question* en inglés o la letra "P" (de *pregunta* en español), y por eso están separadas por un "or" lógico representado por la barra vertical.

Abriendo una terminal en la computadora, este archivo se puede correr como se corre un archivo en perl, (suponiendo que el archivo sinpi4.pl tiene permisos para correr o ser ejecutado. Aquí el símbolo \$ es el "prompt" del sistema operativo de nuestro computador):

```
$ perl sinpi4.pl
```

La salida de este comando, será el archivo "fig4.dat", el cual esta formateado de tal forma que sea legible para TikZ y el código del programa "fig4pi.tex" lo pueda leer.

A continuación mostramos el archivo auxiliar "colors.tex", donde se muestran las paletas de colores, que se utilizarán para darles color y diferenciar a las 5 curvas ICC, y correspondientes a los datos mostrados en la hoja de Excel mostrada en la figura 1. Este archivo utiliza el paquete de LaTeX conocido como "xcolor", para definir los 5 colores que se usarán en las curvas ICC.

```
% Archivo auxiliar de paleta de colores "colors.tex"
% usados en el programa "fig4pi.tex"
\usepackage{xcolor}
\definecolor{marine}{HTML}{044599}
\definecolor{apricot}{rgb}{0.98,0.81,0.69}
\definecolor{antiquefuchsia}{rgb}{0.57,0.36,0.51}
\definecolor{amaranth}{rgb}{0.9,0.17,0.31}
\definecolor{amber}{rgb}{1.0,0.75,0.0}
```

Mostramos a continuación el código del archivo "fig4pi.tex" en TikZ [17], para poder graficar las curvas ICC correspondientes a los datos mostrados en la figura 1.

La salida de este programa es un archivo PDF llamado para este ejercicio "fig4pi.pdf". El archivo comienza declarando el tipo de archivo LaTeX, que se va escribir para formar un programa en TikZ [18]. En este caso, como sólo se va a ocupar para generar una gráfica de varias curvas, es declarado como "standalone", y se le rodeará con un borde en blanco de 10 pt (10 unidades de "punto", siendo 1 pt = 0.35278 mm) [19, 20, 21].

```
% Archivo graficador en TikZ "fig4pi.tex"
% LaTeX solo produce una imagen en PDF
\documentclass[border=10pt]{standalone}
%%<
\usepackage{verbatim}
%%> Comentarios
\begin{comment}
:Title: A plot of TeX
:Tags: plots
:Author: Anonymous
A diagram showing a common TeX workflow, using the
pgfplots package.
```

```

\end{comment} %Fin de comentarios

%Incluimos los colores definidos en el archivo "colors.tex"
\include{colors}

% Paquete para realizar gráficas
\usepackage{pgfplots}

\begin{document}

% Colocando estilos y prefigurando el
% archivo de salida fig4pi.pdf
\pgfplotsset{scale only axis,
xlabel near ticks, ylabel near ticks,
axis on top,
every axis title shift=3pt,
width=10cm, height=6cm,
compat=1.18,
every axis legend/.style={
cells={ anchor=center} ,%Centered entries
inner xsep=3pt, inner ysep=2pt,
nodes={ inner sep=2pt,text depth=0.15em},
anchor=north west,
shape=rectangle,
fill=white,
draw=black,
at={(0.98,0.98)}
}
}

\begin{tikzpicture}
% Ajustes a los ejes de la gráfica
\begin{axis}[
font=\small,
title=\textbf{Item Characteristic Curves in pretest},
xlabel=\emph{Ability (\$\theta_s\$)},
ylabel=\emph{Probability (\$P\$)},
ymin=-0.05, ymax=1.05,
ymajorgrids=true,
yminorticks=false,
xmin=-4, xmax=4,
xmajorgrids=false,
enlarge x limits=0.05,
]
#Iterando en cada una de las 5 curvas ICC para graficarlas
#con un color determinado
\foreach \lval\clr in {1/marine, 2/apricot, 3/antiquefuchsia,
4/amaranth, 5/amber}
{
\edef\temp{
\noexpand\addplot[\clr] table[x=z,y=\lval]{ fig4.dat};
\noexpand\addlegendentry {\$Q\$\lval}
}\temp
}
\end{axis}
\end{tikzpicture}

\end{document}

```

El código anterior en TikZ, redibuja la salida gráfica del programa codificado en R, llamado en este ejemplo "ICCpre.R", una muestra de las curvas ICC del modelo de Rasch en R, se puede ver en la figura 2.

En la figura 3, mostramos de manera esquemática, y con un diagrama de flujo, como primero recolectamos los datos en hoja de Excel de Microsoft, luego realizamos las gráficas de las curvas ICC del modelo de Rasch en R, utilizando el paquete *ltm*. Posteriormente tratamos los datos de R, con un programa *script* escrito en Perl, y finalmente graficamos estas mismas curvas pero utilizando un programa escrito en TikZ. Así, podemos controlar la presentación de estas curvas ICC del modelo de Rasch en TikZ. Por ejemplo, podemos escoger los colores que queramos utilizar para las curvas ICC, a través del archivo auxiliar de colores, que aquí llamamos "colors.tex". También el formato de salida por defecto es PDF, por lo que la integración en un documento LaTeX, es sencillo, y las figuras gráficas son de muy buena resolución y por lo tanto, son muy claras, y conservan fielmente las características de las gráficas generadas por el paquete *ltm* del lenguaje de programación R.

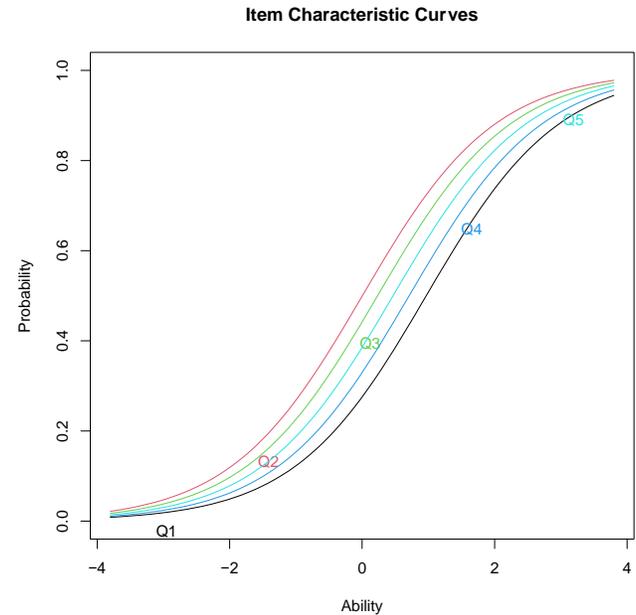


FIGURA 2. Item Characteristic Curves o ICC del modelo de Rasch (plot.pdf), tal y como las calcula R, con el paquete *ltm*.

En la figura 4, mostramos la gráfica que se obtiene con el código del programa en TikZ "fig4pi.tex", de las curvas ICC del modelo de Rasch dicotómico, para los datos de ejemplo, mostrados en la figura 1.

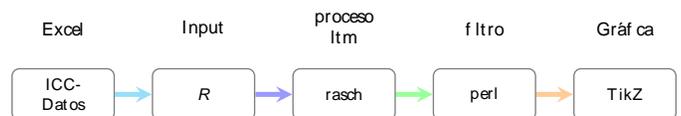
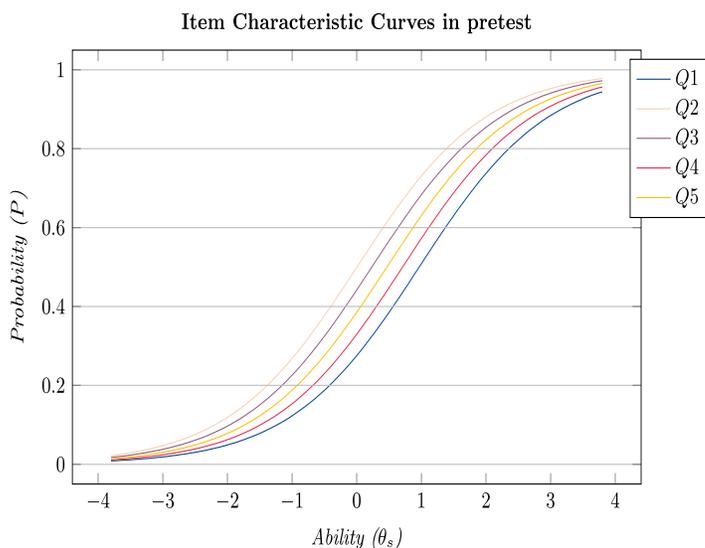


FIGURA 3. Esquema del proceso del tratamiento de datos para obtener las curvas Item Characteristic Curves o ICC del modelo de Rasch y trasladarlas de R, a TikZ.

Las curvas ICC de la figura 4 con TikZ, siguen fielmente las características dibujadas en la figura 2 con R.

El software de TikZ/PGF y PGFPLOTS, nos da la opción de presentar estos resultados, en un formato que es compatible con documentos, que se publican en formato PDF. Así que el tratamiento de los datos con TikZ, nos abre otra puerta para mostrar nuestros resultados gráficos, siguiendo un formato ampliamente conocido y utilizado en la comunidad científica.



**FIGURA 4.** Gráfica de las curvas ICC del modelo de Rasch en TikZ (fig4pi.pdf), utilizando los datos de salida de R, después de formatear esos datos con el programa en Perl que aquí nombramos “sinpi4.pl”. Se nota como la probabilidad ( $P$ ) de contestar correctamente a cada pregunta ( $Q_1, Q_2, \dots, Q_5$ ) aumenta, conforme al nivel de habilidad de los estudiantes.

## V. CONCLUSIONES

Como se puede observar, el graficar nuestras curvas ICC del modelo de Rasch en TikZ, tiene ciertas ventajas. Por ejemplo, os sirve para tener un buen control de los rasgos y detalles de las gráficas, y nos ofrece otra alternativa de presentar resultados estadísticos. Las curvas, mantienen fielmente las características de las gráficas hechas por R, y también la salida se puede integrar fácilmente con LaTeX, que es un programa que se utiliza ampliamente en la publicación de resultados científicos.

## AGRADECIMIENTOS

El autor quiere agradecer el apoyo recibido por el CONAHCYT, de México y también por el apoyo otorgado por el Instituto Politécnico Nacional a través del proyecto SIP 20242173 titulado: *La enseñanza de la Física utilizando el Aprendizaje Activo*.

## REFERENCIAS

- [1] Sokoloff, D. R., Thornton, R. K., Laws, P. W., *RealTime Physics: Active Learning Laboratories, Module 1: Mechanics* (Wiley, 3<sup>rd</sup> Edition, USA, 2011).
- [2] Sokoloff, D. R., Laws, P. W., Thornton, R. K., *RealTime Physics: Active Learning Laboratories, Module 2: Heat and Thermodynamics* (Wiley, 3<sup>rd</sup> Edition, USA, 2011).
- [3] Sokoloff, D. R., Laws, P. W., *RealTime Physics: Active Learning Laboratories, Module 3: Electricity and Magnetism* (Wiley, 3<sup>rd</sup> Edition, USA, 2012).
- [4] Sokoloff, D. R., *RealTime Physics: Active Learning Laboratories, Module 4: Light and Optics* (Wiley, 3<sup>rd</sup> Edition, USA, 2012).
- [5] Fischer, G. H., Molenaar, I. W., *Rasch Models, Foundations, Recent Developments, and Applications* (Springer Verlag, Nueva York, USA, 1995).
- [6] Boone, W. J., Staver, J. R., Yale, M. S., *Rasch Analysis in the Human Sciences* (Springer, New York, USA, 2014).
- [7] Boomsma, A., van Duijn, M. A. J., Snijders, T. A. B., *Essays on Item Response Theory, Lecture Notes in Statistics 157* (Springer, The Netherlands, 2001).
- [8] CRAN, *The R Project for Statistical Computing*, <<https://www.r-project.org>>, Consultado el 18 de febrero de 2024.
- [9] Baker, F. B., Kim S.-H., *The Basics of Item Response Theory Using R* (Springer, USA, 2017).
- [10] Kaltakci, D., *Four-tier Geometrical Optics Test (FTGOT)*, <<https://www.physport.org/assessments/assessment.cfm?A=FTGOT>>, Consultado el 20 de febrero de 2024.
- [11] PhysPort, <<https://www.physport.org>>, Consultado el 20 de febrero de 2024.
- [12] Rizopoulos, D., *Package 'ltm' Latent Trait Models under IRT*, página web (An R package for latent variable modeling and item response theory) <<https://cran.r-project.org/web/packages/ltm/ltm.pdf>>, Consultado el 21 de febrero de 2024.
- [13] Davies, T. M., *The book of R: A First Course in Programming and Statistics* (No Starch Press, 1<sup>st</sup> edition, San Francisco, 2016).
- [14] Lopez, F., Romero, V., *Mastering Python Regular Expressions* (Pack Publishing, Birmingham, England, 2014).
- [15] Agarwal, S., *Understanding Ruby Regexp* (LeanPub, British Columbia, Canada, 2024).
- [16] Stubblebine, T., *Regular Expression Pocket Reference: Regular Expressions for Perl, Ruby, Php, Python, C, Java and .NET* (O'Reilly Media, 2<sup>nd</sup> Edition, USA, 2007).
- [17] Kolda, T. G., *Unlocking LaTeX Graphics, A Concise Guide to TikZ/PGF and PGFPLOTS* (MathSci.ai, 1<sup>st</sup> Edition, California, USA, 2024).
- [18] Johnson, H. C., *Mastering LaTeX: Advanced Drawing with Tikz, Including 3D Drawing* (The Telephasic Workshop, USA, 2024).
- [19] Knuth, D. E., *The TeXbook* (Addison Wesley Publishing Company, Inc., Spi Edition, Massachusetts, 1984).
- [20] Mittelbach, F., Goossens, M., with Braams, J., Carlisle, D., and Rowley, C., and contributions by Detig, C., and Schrod, J., *The LaTeX Companion* (Addison Wesley Series on

*Rubén Sánchez Sánchez*

Tools and Techniques for Computer Typesetting, Second Edition, San Francisco, 2006).

[21] Lamport, L., *LaTeX: A Document Preparation System* (Addison Wesley Professional, Second Edition, San Francisco, 1994).