חשבון אינפיניטסימלי

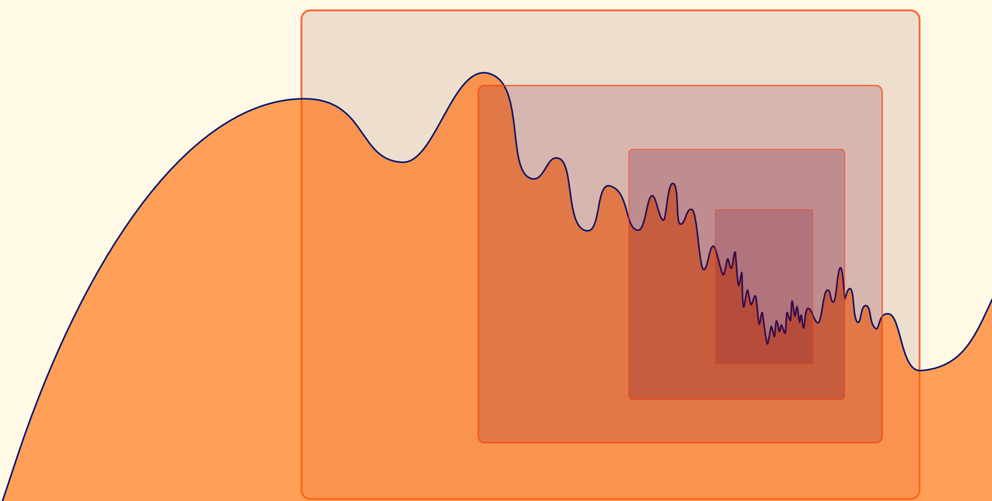
מיכאל הוכמן

עם

יותנן הראל

איתי וייס

אופק שילון



מהדורה שלישית

חשבון אינפיניטסימלי

מיכאל הוכמן

עם יונתן הראל, איתי וייס ואופק שילון

מהדורה שלישית

חשבון אינפיניטסימלי

מיכאל הוכמן

עם יונתן הראל, איתי וייס ואופק שילון

מהדורה שלישית

הפצה : הוצאת מאגנס
טל' 02-6584352 , **shlomi@magnespress.co.il**
www.magnespress.co.il

מסת"ב 978-965-350-150-8
eBook 978-965-350-147-8 **ISBN**

נדפס בישראל

תוכן עניינים

הקדמה

ספר זה נועד ללוות קורס אוניברסיטאי ראשון בחשבון אינפיניטסימלי, ומבוסס על תכנית הלימודים של הקורס כפי שהוא נלמד באוניברסיטה העברית. קיימים בשפה העברית מספר ספרים בחשבון אינפיניטסימלי, אך לדעתנו אין לאף אחד מהם את מכלול התכונות הדרושות מספר כזה כיום. נזכיר במיוחד את הספר "חשבון אינפיניטסימלי" של דוד מיזלר [1], שמלווה באופן מסורתי את הקורס באוניברסיטה העברית. זהו ספר מדויק ומקיף, אך הוא מכוון לקהל יעד השונה במידה ניכרת מקהל הסטודנטים הלומדים כיום את הקורס. הצגת החומר בו מהירה מאד ותמציתית, והוא אינו מרבה בהסברים ובדוגמאות.

הספר שלנו נועד למלא חלל זה. מטרתו היא לא רק לסכם את החומר אלא גם להסביר אותו בצורה הברורה והאינטואיטיבית ביותר האפשרית. הוא נפתח במבוא בנושאים מתמטיים כלליים, ומשם עובר להצגה מלאה אך נינוחה של חומר הקורס: הגדרת המספרים, סדרות וטורי מספרים, רציפות, גזירה ואינטגרציה של פונקציות במשתנה אחד, סדרות וטורי פונקציות, פולינומי טיילור וטורי חזקות. הטקסט כולל דוגמאות ותרגילים רבים בכל הרמות.

כללנו בספר כמה נושאים שאינם שייכים לליבת הקורס, כמו עוצמת המספרים הממשיים, מספרים טרנסצנדנטיים, שיטות נומריות, פונקציית ויירשטראס, נוסחת סטירלינג ועוד. לדעתנו הרחבות אלה חשובות לא רק לתלמידים שימשיכו ללמוד מתמטיקה, שממילא יתקלו בהן בהמשך לימודיהם, אלא במיוחד לתלמידים שעבורם הקורס בחשבון אינפיניטסימלי הוא ההזדמנות האחרונה להכיר את המתמטיקה המודרנית.

נושא אחד שלא כלול בספר הוא החשבון האינפיניטסימלי של פונקציות רבות משתנים. בקורס באוניברסיטה העברית נוגעים בנושא זה לקראת סוף הקורס, ובמקור התכוונו לכלול אותו בספר. החלטנו בסופו של דבר שלא לכלול אותו כיוון שלא נוכל לדון בו בהיקף הראוי. מצד שני, ישנם בשפה העברית ספרים טובים בנושא, כמו הספר של מיזלר [1] או הספר של לינדנשטראוס בחשבון אינפיניטסימלי מתקדם [9].

ארגון הספר

מושגים חדשים מופיעים **בהדגשה** במקום הופעתם הראשון. בסוף כל הוכחה מופיע הסימן ■. הסימן [1] מציין הפניה ביבליוגרפית, ומפנה לרשומה המתאימה ברשימת הספרים שבסוף הספר. הגדרות, סימונים, למות,[1] טענות, משפטים ומסקנות ממוספרים ברצף בכל סעיף. למשל, משפט שמספרו 1.2.3 נמצא בסעיף 1.2, כלומר בסעיף השני של הפרק הראשון, והוא המשפט השלישי באותו סעיף. האיורים, התרגילים והדוגמאות ממוספרים בנפרד. תרגילים קשים מסומנים ב־(*).

כדי להתמצא בספר ניתן להיעזר בתוכן העניינים ובשני המפתחות שבסוף הספר: מפתח סימנים, המסודר לפי סדר הופעת הסימנים ומפנה להופעתם הראשונה, ומפתח נושאים, המסודר לפי סדר א"ב. בסוף הספר תמצאו גם רשימה של הא"ב היווני.

דרישות קדם

כדי לקרוא את הספר אין צורך, עקרונית, בידיעות קודמות במתמטיקה. אנו נפתח כאן מבראשית את כל הנושאים שנעסוק בהם.

אולם בפועל, מומלץ להיות בקיאים במיומנויות חשבון בסיסיות הנלמדות בבית הספר, כגון פתרון של משוואות פשוטות וטיפול באי־שוויונות. במהלך הספר נזכיר גם מושגים מתורת הגאומטריה ומטריגונומטריה, את שיטת ההוכחה באינדוקציה, את מושג הפונקציה ומושגים הקשורים לו. היכרות עם נושאים אלה בוודאי לא תזיק. כדי להשלים את ידיעותיכם בנושאים אלה אפשר להיעזר בספרי הלימוד התיכוניים השונים.

כמה הצעות ללמידה נכונה

החומר שמובא כאן הוא בעל אופי שונה מאד מזה של המתמטיקה התיכונית. כדי להתמודד אתו יידרש מכם שינוי מחשבתי משמעותי. חוויית הלימוד היא כמובן עניין אישי, אך העצות הבאות עשויות לתרום ללימוד יעיל ומהנה יותר.

השתתפו באופן פעיל בלימוד. בזמן הקריאה חפשו דוגמאות למה שמתואר בטקסט, ציירו ציורים, העלו תהיות וענו עליהן. אם במשפט מסוים מופיעה הנחה, בדקו היכן משתמשים בה בהוכחה. שאלו את עצמכם מה קורה אם משנים את הניסוח של משפט מסוים באופן כזה או אחר, למשל על ידי השמטה או שינוי של אחת ההנחות, ונסו להוכיח או להפריך את הטענה החדשה.

פתרו תרגילים, ואחר־כך פתרו עוד תרגילים. אל תתעצלו ואל תוותרו לעצמכם. אי אפשר ללמוד מתמטיקה מבלי לעשות מתמטיקה, כפי שאי אפשר ללמוד לרקוד בלי

[1] למה (lemma) היא טענת עזר, כלומר טענה שמטרתה לסייע בהוכחת טענה חשובה יותר. מקור המילה הוא השפה היוונית.

לקום על הרגליים ולהתאמן. בסוף כל סעיף יש רשימה של תרגילים ברמות שונות. הם שם כדי שתפתרו אותם. אין צורך לפתור את כולם אבל כדאי לעבור על כולם, ולפתור מדגם מייצג שלהם. אפשר למצוא שאלות גם בספרים אחרים.

אל תנסו ללמוד את הספר בעל פה. הלימוד של ספר זה בעל פה הוא משימה בלתי אפשרית, וממילא לא תלמדו כך מתמטיקה. במקום לשנן אינספור פרטים קטנים נסו להבין את המוטיבציה מאחורי הדברים, את מבנה־העל של ההגדרות וההוכחות, וכיצד הן משתלבות זו בזו. הבנת התמונה הגדולה תאפשר לכם להשלים לבד את הפרטים הקטנים.

השתדלו להסתמך על משפטים קודמים במקום להוכיח אותם מחדש. התוצאות שמובאות בטקסט נועדו לשימוש בפתרון בעיות. כשאתם ניגשים להוכיח טענה או לפתור תרגיל חפשו טענות ומשפטים מהטקסט שעשויים להיות רלוונטיים לבעיה שלפניכם.

היעזרו בחברים ובמורים וקיימו דיון בנושאים הנלמדים. תוכלו ללמוד הרבה על ידי הצצה לדרך החשיבה של אחרים. בנוסף לעבודה האינדיבידואלית כדאי לדון עם חברים בחומר ובתרגילים, לשתף זה את זה בפתרונות שמצאתם, וגם לפתור יחד בעיות.

הקפידו על סדר הקריאה ואל תיצרו פערים. סדר הנושאים אינו שרירותי, וכל נושא מבוסס במידה רבה מאד על נושאים שקודמים לו. אם תדלגו על סעיפים חשובים ותרוצו קדימה, מהר מאד תלכו לאיבוד. אם במהלך הקריאה אתם נתקלים בנושא שאינו מוכר לכם, חיזרו אחורה ולמדו אותו, ורק אז המשיכו לנושא החדש.

אל תצפו שההבנה תבוא בן־לילה. תנו לעצמכם זמן לעכל את החומר ואל תצפו לבלוע פרק שלם ברגע. אם אתם "נתקעים" עשו הפסקה, התייעצו, וחזרו לבעיה מאוחר יותר.

יצירת קשר

על אף מאמצים רבים שעשינו לאתר שגיאות בטקסט לפני ההבאה לדפוס, אין ספק שנותרו טעויות שחמקו מעינינו. אם אתם חושבים שנתקלתם בטעות, אנא ספרו לנו עליה בדואל

`infibook@math.huji.ac.il`

רשימת תיקונים ניתן למצוא באתר האינטרנט

`www.math.huji.ac.il/~infibook`

תודות

ראשית כל, תודה גדולה לאופק שילון, שיזם את כתיבת הספר, ולאיתי וייס ויונתן הראל, שיחד עם אופק השתתפו בכתיבה. הפרויקט החל בשנת 2001, כאשר התחלנו ארבעתנו לתרגל את הקורס בחשבון אינפיניטסימלי באוניברסיטה העברית. משיחות עם מרצים עלה שהכול היו שמחים לו היה קיים ספר מעודכן בחשבון אינפיניטסימלי. כך קרה שאופק הציע לכתוב כמה פרקי עזר לתלמידים. למשימת הכתיבה אופק גייס אותי, את איתי ואת יונתן. חלק מהפרקים בספר מבוססים על טיוטות שהוכנו במהלך אותה שנה.

אנו אסירי תודה לחוג ולמכון למתמטיקה באוניברסיטה העברית, שתמכו בפרויקט. תודה לפרופ׳ אנדריי שנקובסקי, פרופ׳ עמנואל פרג׳ון, פרופ׳ גנאדי לוין, פרופ׳ אהוד דה־שליט ופרופ׳ מרדכי ציפין על עזרתם.

רבים סייעו לנו בשלבים השונים של הכתיבה וההגהה. אנו מודים במיוחד לאיתמר צביק, לפרופ׳ עזריאל לוי ולפרופ׳ ברק וייס על ההצעות וההערות הרבות, ולהילה זמר על הסיוע בעריכה. תודה גם לגילי שול, יהודה לוי ואביתר פרוקצ׳ה על עזרתם בהגהה.

הספר הודפס בעזרת תוכנת הדפוס LaTeX ותוכנת LyX, שהן תוכנות חינמיות שפותחו והותאמו לשפה העברית על ידי מתנדבים רבים. גם להם תודה.

מיכאל הוכמן
ירושלים, 2007

פרק 1

מבוא

החשבון האינפיניטסימלי (באנגלית: infinitesimal calculus) הומצא בסוף המאה ה־17. הוא פותח במקביל ובאופן בלתי תלוי על ידי אייזק ניוטון (Isaac Newton, 1643-1727) וגוטפריד לייבניץ (Gottfried Leibnitz, 1647-1716). אמנם העיסוק במתמטיקה התחיל כבר בעת העתיקה,[1] אך המצאת החשבון האינפיניטסימלי היא ללא ספק אחת מפריצות הדרך המשמעותיות ביותר בהיסטוריה שלה. יחד עם תורת הפיזיקה של ניוטון, שפותחה באותן שנים, היא מהווה את קו פרשת המים בין העידן העתיק לעידן המודרני של המדע. גם היום החשבון האינפיניטסימלי משחק תפקיד חשוב במדעים כמו פיזיקה, הנדסה, כלכלה ועוד.

התורה של ניוטון ולייבניץ אפשרה לראשונה לענות על שאלות שונות במתמטיקה ובפיזיקה, שחלקן עמדו זמן רב ללא פתרון, וחלקן אפילו לא נשאלו. למשל,

- איך מחשבים שיפוע של עקומה במישור?
- או מחשבים שטחים ונפחים של גופים במישור ובמרחב?
- האם אפשר לרשום את המספר $\sqrt{2}$ בעזרת "נוסחה"?
- איך מוצאים ביעילות פתרון מקורב למשוואה $x^7+x^4+2x^2+x+1=0$?
- מדוע כוח הכבידה גורם לכוכבי הלכת לנוע סביב השמש במסלולים שצורתם אליפסה?

על אף הצלחתה הרבה, התורה של ניוטון ולייבניץ נשענה על רעיונות מעורפלים למדי. ליקויים אלה הטרידו רק מעטים במאות הראשונות לקיום התורה כיוון שעל אף אי־הבהירות התאורטית, מבחינה מעשית התורה של ניוטון ולייבניץ היא תורה מוצלחת מאד. אולם החל מאמצע המאה ה־18 ובמהלך המאה ה־19, אי־הבהירות הפכה למכשול של ממש להמשך המחקר, והמתמטיקאים הפנו את תשומת לבם למציאת ביסוס תיאורטי מוצק יותר לתורה. הם נדרשו לענות על שאלות כמו:

[1] קיימים עדויות על ניצנים של פעילות מתמטית כבר בממלכה המצרית והבבלית באלף השלישי לפנה"ס.

- כיצד יש לפרש ביטוי מהסוג

$$1+\frac{1}{1+\frac{1}{1+\dots}}$$

האם הוא מייצג מספר ולמה הוא שווה?

- מדוע הנוסחה

$$\frac{1}{1-x}=1+x+x^2+x^3+\dots$$

נכונה כאשר מציבים $x=\frac{1}{2}$, אבל כאשר מציבים $x=2$ מקבלים את השוויון המוזר $-1=1+2+4+8+\dots$, שבצד אחד שלו מספר שלילי ובשני סכום של מספרים חיוביים?

- מהו בעצם סכום של אינסוף מספרים?
- מה זה בכלל מספר?

פתרון לשאלות אלה ניתן במהלך המאה ה־19. מרכזיים בפתרונן היו מושגי הגבול והרציפות, שנעסוק בהם רבות במהלך הספר.

במאה ה־19 חלו גם כמה תמורות במסגרת שבה מתבצע המחקר המתמטי. השינוי בא בעקבות גילוים של מספר פרדוקסים[2] ביסודות המתמטיקה, והחשש שללא אמצעי זהירות חמורים אי אפשר יהיה למנוע חדירה של שגיאות למתמטיקה. בניסיון להציב את המתמטיקה על קרקע מוצקה יותר עסקו רבים בחקר הלוגיקה ובנושאים יסודיים אחרים, מחקר שהגיע לשיאו בסוף המאה ה־19 ובתחילת המאה ה־20. כתוצאה ממאמצים אלו יש לנו כיום הבנה תאורטית טובה של יסודות המתמטיקה, ולצדה מסגרת מוסכמת של כללים שלפיה מתנהל המחקר. מטבע הדברים גם החשבון האינפיניטסימלי מתקיים כיום באותה מסגרת.

בספר זה נכיר את החשבון האינפיניטסימלי בגרסה המודרנית שלו, אך נפתח אותו בסדר ההפוך מהסדר ההיסטורי אותה תיארנו לעיל. אמנם המטרה המרכזית היא ללמוד את התורה של ניוטון ובני זמנו (וגם כמה תוספות מאוחרות יותר), אך לפני שנוכל לדון בה נצטרך ללמוד את השפה של המתמטיקה המודרנית, להגדיר את המספרים, ללמוד את תורת הגבולות והרציפות, ורק אז נוכל להתחיל לטפל בתורה של ניוטון ולייבניץ. זו אמנם לא הדרך הקצרה ביותר אל המטרה, אך יש לצדה גם שכר, שכן התחנות השונות שנעבור בדרך חשובים ומעניינים בפני עצמם.

ההיסטוריה של החשבון האינפיניטסימלי, ושל המתמטיקה בכלל, היא נושא מרתק. מי שמעוניין יוכל לקרוא עליה בפירוט רב יותר בספר [18].

---

[2] במתמטיקה פרדוקס הוא תופעה המובילה למסקנה אשר סותרת את ההיגיון. איננו מאמינים שמצב כזה ייתכן, ואמנם, עד כה הפרדוקסים שהתגלו במתמטיקה הם פרדוקסים רק לכאורה, המעידים על כשל בניתוח או בפירוש של תופעה ולא על סתירה אמתית.

פרק 2

יסודות

פרק זה עדיין אינו עוסק בחשבון אינפיניטסימלי, אלא בכמה נושאים מתמטיים כלליים. נציג כאן בקצרה את המסגרת הלוגית שבה פועלת המתמטיקה המודרנית, הכוללת את השפה המתמטית ואת מושג ההוכחה, וננסה לשכנע אתכם שיש סיבה טובה לקיומה. נדון בקצרה גם בקבוצות, שהן מהאובייקטים הבסיסיים ביותר במתמטיקה.

החומר בפרק זה מוצג באופן לא פורמלי, והדיון אינו עומד בסטנדרטים שהוא עצמו מציג. ניתן היה לנהל דיון מדויק יותר, אך הדבר היה גוזל זמן רב מדי, ואין הוא קשור ישירות לנושא אליו אנו מכוונים. תוכלו למצוא דיון מפורט יותר בספרים אחרים. הפניות מתאימות מופיעות במהלך הפרק.

2.1 סמנטיקה ותחביר

בשפת הדיבור מקובל, מטעמי נוחות, להתבטא באופן שאינו תמיד מדויק לחלוטין. במתמטיקה, לעומת זאת, מקפידים הקפדה יתירה על דיוק. אפשר ללמוד על הצורך בכך מהדוגמאות הבאות. הראשונה מביניהן היא דוגמה קלסית המיוחסת לפילוסוף והמדען היווני אריסטו:[1]

כל אדם הוא בן תמותה
סוקראטס הוא אדם
לכן סוקראטס הוא בן תמותה.

לכאורה הכול תקין, ואם מקבלים את ההנחות אז המסקנה נובעת. לעומת זאת אותו היגיון מוביל לכאורה לטיעון הבא:

[1] Aristotle, 384-322 לפנה"ס.

בני האדם הם חלק מבעלי החיים,
חלק מבעלי החיים הם בעלי כנף,
לכן בני האדם הם בעלי כנף.

המסקנה שגויה, אף שלכאורה הפעלנו את אותו ההיגיון. במקרה זה לא תתקשו לשים את האצבע על השגיאה, אך אם לא ננקוט אמצעי זהירות ניצור לעצמנו בהמשך סבך נוראי של אי־דיוקים וחצאי אמיתות. כדי למנוע זאת עלינו ראשית כל להבהיר אחת ולתמיד את המשמעות של מרכיבי השפה השונים.

טענות

כל טענה במתמטיקה היא אמתית או שקרית. הערך "אמת" או "שקר" המיוחס לטענה נקרא **ערך האמת** של הטענה. יוצאים מן הכלל הם טענות המכילות משתנים. במקרה זה ערך האמת נקבע רק ברגע שמציבים איברים במקום המשתנים. למשל, "$1+1=2$" היא טענה נכונה, "$1+2=2$" היא טענה שקרית, ואילו "$1+x=2$" היא טענה נכונה אם מציבים $x=1$ וטענה שקרית לכל הצבה אחרת.

מובן שלעתים אנו לא יודעים לקבוע מהו ערך האמת של טענה מסוימת. אין בכך לגרוע מהעובדה שיש לטענה ערך אמת אחד ויחיד.

קשרים

דרך אחת ליצור טענות היא לבנות אותן מטענות פשוטות יותר בעזרת **קשרים** (connectives), שהנפוצים ביניהם הם הביטויים לא, גם, או, אם־אז, ואם־ורק־אם. משמעותם לרוב כמו בשפת הדיבור אך יש כמה הבדלים. אם P,Q הן טענות אז הטענות שניתן ליצור מהן בעזרת הקשרים הן הטענות הבאות:

- "לא P" (בסימנים: $\neg P$). זו הטענה ההפוכה מ־P, כלומר ערך האמת שלה הפוך משל P. הטענה "לא P" נקראת **השלילה** של P.
- "P וגם Q" (בסימנים: $P \wedge Q$) היא הטענה ששתי הטענות P,Q נכונות.
- "P או Q" (בסימנים: $P \vee Q$) היא הטענה שלפחות אחת הטענות P,Q נכונה. שימו לב שהטענה "P או Q" נכונה גם כששתי הטענות נכונות. למשל, הטענה "$1+1=2$ או $1+2=3$" היא טענה נכונה. כאן יש שוני לעומת שפת הדיבור, שבה לפעמים המילה "או" פירושה שאחד מהשניים נכון אך לא שניהם, כמו במשפט "אתה רוצה מרק גזר או מרק שעועית?".
- "אם P אז Q" (בסימנים: $P \implies Q$). טענה זו מקבל ערך "אמת" בכל מקרה למעט כאשר P אמיתית ו־Q שקרית. לכן, אם הטענה "אם P אז Q" אמיתית, ואם P אמיתית, אז בהכרח Q אמיתית; אבל אם הטענה "אם P אז Q" אמיתית ו־P שקרית אי אפשר ללמוד דבר על ערך האמת של Q, שכן גם

הערך "אמת" וגם הערך "שקר" עולים בקנה אחד עם כך ש־"אם P אז Q" אמיתית. להמחשה, הטענה "אם יורד גשם אז יש עננים בשמיים" היא טענה נכונה, וכאשר יורד גשם אפשר להסיק ממנה שמעונן. לעומת זאת, מכך שלא יורד גשם אי אפשר להסיק ממנה דבר, כי לפעמים מעונן למרות שלא יורד גשם. דוגמה מתמטית יותר היא הטענה "אם $2=1$ אז $3=1$". זו טענה נכונה! במקרים כאלה אומרים שהטענה **מתקיימת באופן ריק**. דוגמה נוספת מהשפה המדוברת היא המשפט "אם לסבתא שלי היו גלגלים אז היא הייתה אוטובוס".

- "P אם ורק אם Q" (בסימנים: $P \iff Q$). טענה זו אומרת שערך האמת של P, Q זהה: או ששניהם נכונים או ששניהם שקריים. את הביטוי "אם ורק אם" מקצרים בדרך כלל לראשי תיבות: אמ"מ.

כאשר מצרפים יותר משתי טענות יחד חשוב להבהיר באיזה סדר הם מצורפים. לשם כך אפשר להשתמש בפיסוק או בסוגריים. על כך שההקפדה חשובה אפשר ללמוד מהטענה $(x<0)\wedge((x>0)\vee(x=1))$, אשר שקרית כשמציבים $x=1$, ולעומתה הטענה $((x<0)\wedge(x>0))\vee(x=1)$, הנבדלת מהטענה הראשונה רק במיקום הסוגריים, ואשר אמיתית עבור $x=1$. הנה דוגמה נוספת: הטענה $(\neg(x<1))\wedge(x>2)$ והטענה $\neg((x<1)\wedge(x>2))$ נבדלות במיקום הסוגריים אך הראשונה נכונה כאשר $x=1$ (למעשה היא נכונה תמיד), ואילו באותם תנאים הטענה השנייה שקרית.

כמתים

דרך נוספת לקבל טענות חדשות מישנות היא בעזרת **כמתים** (quantifiers). אם $P(x)$ טענה התלויה במשתנה x אפשר ליצור ממנה את הטענות הבאות:

- "לכל x מתקיים $P(x)$" (בסימנים: $\forall x\, P(x)$). טענה זו נכונה רק אם $P(x)$ היא טענה נכונה לכל איבר a שנציב במקום x.
- "קיים x כך שמתקיים $P(x)$" (בסימנים: $\exists x\, P(x)$). טענה זו נכונה אם יש לפחות איבר אחד שהצבתו במקום x הופכת את $P(x)$ לטענה נכונה, אחרת הטענה המורכבת שקרית.

כאשר טענה נפתחת ב"לכל x..." או ב"קיים x..." ואחריה מופיעה טענה $P(x)$ התלויה ב־x, המשתנה x נקרא **המשתנה המכומת**; משתנה שאינו מכומת נקרא **משתנה חופשי**. למשל בטענה "לכל x מתקיים $x=y$" המשתנה x הוא מכומת ואילו y הוא חופשי. אין לכמת על אותו משתנה יותר מפעם אחת בכל הקשר. למשל, קשה לתת פרוש לטענה "לכל x קיים x כך ש־ $x=1$", ולעולם לא ניצור טענות כאלה. לעומת זאת, הטענה "(קיים x כך ש־ $x=1$), וגם (קיים x כך ש־ $x=2$)" היא טענה תקינה, אם כי מבחינה סגנונית היה עדיף לבחור במשתנים שונים בשני חלקיה, כמו למשל בטענה "(קיים x כך ש־ $x=1$) וגם (קיים y כך ש־ $y=2$)". נציין

שאפשר ליצור טענה מהסוג $\forall x P(x)$ גם כאשר הטענה $P(x)$ אינה באמת תלויה ב־x. למשל, הטענה "לכל x מתקיים $1+1=2$" היא טענה כזו (היא נכונה, כמובן).

ישנו קשר הדוק בין הכמתים "קיים" ו"לכל". הטענה "לכל x מתקיים $P(x)$" והטענה "לא קיים x כך ש־$P(x)$ שקרי" אומרות בדיוק את אותו הדבר. באופן דומה הטענה "קיים x כך ש־$P(x)$" והטענה "לא לכל x מתקיים ש־$P(x)$ שקרי" אומרות את אותו הדבר. בסימנים:

$$\begin{aligned} \forall x\, P(x) &\iff \neg(\exists x\, \neg P(x)) \\ \exists x\, P(x) &\iff \neg(\forall x\, \neg P(x)) \end{aligned}$$

חשוב להקפיד על סדר הופעת הכמתים בטענה. התבוננו למשל בטענות

$$\exists x \forall y\, (y \text{ מכיר את } x) \qquad , \qquad \forall x \exists y\, (y \text{ מכיר את } x)$$

בשמאלית כתובה הטענה שכל אדם מכיר מישהו, ובימנית כתוב שיש מישהו שמכיר את כולם. אלה טענות שונות לחלוטין.

לעומת זאת, סדר הכמתים אינו משמעותי כשמדובר בכמתים מאותו סוג. אם $P(x,y)$ היא טענה התלויה במשתנים x,y אז הטענות

$$\forall y \forall x\, P(x,y) \qquad , \qquad \forall x \forall y\, P(x,y)$$

שקולות, כי שתיהן אומרות ש־$P(a,b)$ נכונה לכל זוג איברים a,b שנציב במקום x,y. לכן לפעמים מקצרים וכותבים $\forall x,y\, P(x,y)$. באופן דומה, הטענות

$$\exists y \exists x\, P(x,y) \qquad , \qquad \exists x \exists y\, P(x,y)$$

שקולות ושתיהן אומרות שקיים זוג איברים a,b כך ש־ $P(a,b)$ נכונה. לכן לפעמים מקצרים ורושמים $\exists x,y\, P(x,y)$.

המתמטיקה כשפה מדוברת

בפועל מתמטיקה אינה נכתבת רק בסימנים. מתמטיקאים הם בכל זאת בני אדם, לכן ברוב המקרים כותבים מתמטיקה בשפת דיבור אשר רק מקרבת את הלוגיקה היבשה, ודואגים לדייק במקומות החשובים. גם אנו ננהג כך. על הקורא (כלומר עליכם!) מוטלת המשימה לפרש נכונה את הכתוב. השימוש העיקרי שנעשה בסימני הלוגיקה יהיה במקרים שבהם אנו ניצבים בפני טענה מורכבת במיוחד. דווקא אז פירוקה למרכיבים וכתיבתה בכתיב פורמלי עשויים לשפוך אור על המצב.

במהלך הספר נגדיר מושגים חדשים רבים. הפירוש של מושג נקבע אך ורק לפי מה שכתוב בהגדרה שלו. כלל זה חל גם על מילים מהשפה המדוברת המשמשות

לעתים לציין מושגים מתמטיים, וחשוב לזכור שהפירוש הרגיל של מילה אינו בהכרח הפירוש המתמטי שלה. למשל המילה "גבול" היא מילה עברית מוכרת שפירושה נמצא בכל מילון, אך אנו נשתמש בו במובן אחר שאותו נגדיר בבוא העת, ואינו דומה לפירושים הרגילים של המילה.

בשפות כמו עברית וערבית, שבהן כיוון הקריאה הוא מימין לשמאל, קיימת אי בהירות לגבי אופן הקריאה של משפטים המכילים סימנים לועזיים. אנו נאמץ את הכלל שקוראים קטע רצוף של לועזית משמאל לימין. אם קטע לועזי נמצא בתוך משפט עברי אז קוראים את העברית בסדר הרגיל עד לקטע הלועזי, אז קוראים את הלועזית משמאל לימין, ואחר־כך ממשיכים מימין לשמאל. למשל את המשפט "אם $x > 1$ אז $x^2 > 1$" קוראים כך: "אם איקס גדול מאחת אז איקס בריבוע גדול מאחת". אפשר להשתמש בסוגריים כדי לכוון את הקריאה: למשל הטענה "(x אוהב שוקולד)$\forall x$" היא מהצורה $\forall x\, P(x)$ כאשר $P(x)$ היא הטענה "x אוהב שוקולד", ולכן קוראים אותה "לכל איקס, איקס אוהב שוקולד". נדגיש שכלל זה שונה מהמוסכמה בשפה הערבית, שבה כותבים גם מתמטיקה מימין לשמאל.

הקשר בין השפה למתמטיקה שייך לתחום במתמטיקה ובפילוסופיה הנקרא לוגיקה (או לוגיקה מתמטית). ניתן לקרוא עליו ביתר פירוט בספרים [11],[14].

תרגילים

בשאלות הבאות הסימנים P, Q, R מייצגים טענות לא ידועות או משתנים המקבלים את הערכים "אמת" ו"שקר". אם φ, ψ טענות התלויות ב־P, Q, R נאמר שהן שקולות אם ערך האמת שלהן זהה לכל בחירה של P, Q, R. למשל אם $\varphi = \neg(P \wedge Q)$ ואם $\psi = (\neg P) \vee (\neg Q)$, אז φ, ψ שקולות, אבל φ אינה שקולה לטענה $(\neg P) \vee Q$.

1. מצאו את כל הזוגות השקולים מבין הטענות הבאות:

(א) $P \wedge (\neg Q)$.

(ב) $P \wedge Q$.

(ג) $(\neg P) \wedge Q$.

(ד) $(\neg P) \vee Q$.

(ה) $P \vee (\neg Q)$.

(ו) $\neg(P \wedge (\neg Q))$.

(ז) $\neg((\neg P) \vee Q)$.

(ח) $\neg((\neg P) \vee (\neg Q))$.

2. לכל טענה בעמודה הימנית מצאו טענה שקולה בעמודה השמאלית:

$P \vee (Q \wedge R)$	$(P \wedge (\neg Q)) \vee ((\neg P) \wedge Q)$
$P \implies Q$	$Q \vee (\neg P)$
$P \wedge (Q \vee R)$	$(P \wedge Q) \vee (P \wedge R)$
$\neg(P \vee Q)$	$(\neg P) \wedge (\neg Q)$
$\neg(P \iff Q)$	$(P \vee Q) \wedge (P \vee R)$

3. כתבו את הטענות הבאות במילים, וקבעו אילו מהן נכונות תמיד:

(א) $\forall x \forall y (x \neq y)$.

(ב) $\forall x \forall y \exists z ((x \neq y) \implies (x \neq z))$.

(ג) $\neg \exists x \forall y \forall z ((x = y) \implies (x = z))$.

4. תהי $P(x,y)$ טענה התלויה במשתנים x,y. מצאו את כל הזוגות השקולים מבין הטענות הבאות:

(א) $\neg \forall x \, \forall y \, P(x,y)$.

(ב) $\forall x \, \forall y \, \neg P(x,y)$.

(ג) $\exists x \, \neg \forall y \, P(x,y)$.

(ד) $\neg \exists x \, \forall y \, P(x,y)$.

(ה) $\forall x \, \neg \exists y \, P(x,y)$.

(ו) $\forall x \, \forall y \, P(x,y)$.

(ז) $\neg \exists x \, \exists y \, P(x,y)$.

(ח) $\forall x \, \neg \forall y \, P(x,y)$.

5. יהיו $P(x), Q(x)$ טענות התלויות במשתנה x. מצאו את כל הזוגות השקולים מבין הטענות הבאות:

(א) $\forall x \, (P(x) \wedge Q(x))$.

(ב) $(\forall x \, P(x)) \wedge (\forall x \, Q(x))$.

(ג) $\exists x \, (P(x) \wedge Q(x))$.

(ד) $(\exists x \, P(x)) \wedge (\exists x \, Q(x))$.

(ה) $\forall x \, (P(x) \implies Q(x))$.

(ו) $(\forall x \, P(x)) \implies (\forall x \, Q(x))$.

(ז) $\exists x \, (P(x) \implies Q(x))$.

(ח) $(\exists x \, P(x)) \implies (\exists x \, Q(x))$.

2.2 הוכחות

הניסיון מלמד שלבני האדם אין כשרון טבעי בכל הקשור לאבחנה בין אמת לשקר, לא כל שכן כשמדובר בטענות מתמטיות. לדוגמה, מהו לדעתכם ערך האמת של שתי הטענות הבאות:[2]

- בהינתן דלי מלא במים אפשר לערבב את המים כך שאף חלק ממנו אינו שב למקומו המקורי.
- אפשר לפרק כדור מלא (כלומר כדור הכולל את הפנים ולא רק את הקליפה) למאה חלקים, להזיז אותם מבלי למתוח או לכווץ, ולהרכיב מהחלקים שני כדורים שכל אחד מהם זהה בנפחו לכדור המקורי.

האינטואיציה בוודאי אומרת שהטענה הראשונה נכונה והשנייה שקרית, אך ההפך הוא הנכון![3]

מכיוון שהאינטואיציה שלנו מוגבלת, במתמטיקה טענה מתקבלת כנכונה רק אם יודעים להוכיח אותה, והיא מתקבלת כשקרית רק אם יודעים להפריך אותה (כלומר להוכיח את השלילה שלה). בהעדר הוכחות כאלה ערך האמת שלה נחשב לא ידוע.

הוכחה של טענה מתמטית היא סדרה של נימוקים המובילה אל המסקנה המבוקשת. לנימוקים בכל שלב מותר להסתמך רק על הנחות הטענה, על טענות שהוכחו קודם לכן, ועל שיקולים לוגיים. לעומת זאת הוכחה אינה יכולה להתבסס על אינטואיציה, ניסיוננו היומיומי או הקבלות לא מדויקות עם תחומים אחרים (אם כי כל אלה הם מקורות טובים להשראה).

למילה "הוכחה" פירושים מעט שונים במקצועות אחרים. המובן המתמטי שלה הוא התובעני ביותר מביניהם. כדאי במיוחד להשוות את המושג המתמטי של הוכחה עם המושגים המקבילים במדעים המדויקים ובמערכת המשפטית.

במדעים מדויקים כמו פיזיקה וביולוגיה, טענה נחשבת לנכונה כל עוד איש לא צפה בתופעה הסותרת אותה. אין זה כך במתמטיקה. העובדה שאין דוגמה זמינה המפריכה טענה מסוימת מחזקת אולי את ההשערה שהטענה נכונה, אך אינה מהווה הוכחה לכך.

לעומת זאת, כדי להוכיח בבית משפט שאדם אשם בעבירה כלשהי יש להראות שהאדם אשם מעבר לספק סביר. לכן הבאת עדויות ממקורות אמינים מקרב אותנו אל ההרשעה, גם אם אפשר להעלות על הדעת תרחיש קיצוני שמסביר את הראיות אך מזכה את הנאשם. מנגד, במתמטיקה הוכחה חייבת להיות משכנעת מעבר לכל

---

[2] אין אלה טענות על העולם הפיזיקלי, שבו החומר מורכב מאטומים, אלא על המודל המתמטי של העולם שבו החומר רציף וניתן לחלוקה אינסופית. כאן שוב מתגלה החשיבות של ניסוח מדויק ומלא של טענה.

[3] שקריות הטענה הראשונה היא תוצאה של משפט נקודת השבת של בראואר (Brouwer, 1912), הטענה השנייה היא תוכנו של משפט הנקרא פרדוקס טרסקי־בנך (Tarski-Banach, 1924). לא נוכיח משפטים אלה.

ספק ותקפה גם במקרי קיצון. ניתן (ורצוי!) להיעזר בעדויות ודוגמאות על מנת לגבש דעה, אך זה רק שלב מקדים לחיפוש אחר הוכחה.

כאשר אפשר להוכיח טענה Q על ידי הנחת טענה P אומרים ש־P **גוררת** את Q, או ש־Q **נובעת** מ־P. אומרים גם ש־P **היא תנאי מספיק** ל־Q וש־Q היא **תנאי הכרחי** ל־P. שימו לב שבמקרה כזה הטענה "אם P אז Q" נכונה. אם בנוסף Q גוררת את P (כלומר, כל אחת מהטענות P,Q גוררת את השנייה) אומרים שהטענות **שקולות**, וש־P היא **תנאי הכרחי ומספיק** ל־Q (וכמובן, גם להפך). במקרה זה הטענה "P אם ורק אם Q" היא טענה נכונה.

כמה הוכחות לדוגמה

אין זה אפשרי למנות את כל צורות ההוכחה השונות, וגם אין בכך טעם. במקום זאת נסתפק בכמה דוגמאות, ונסמוך על השכל הישר שלכם.

כדוגמה ראשונה נבחן את הדוגמה שפתחנו בה את הסעיף הקודם:

> כל אדם הוא בן תמותה, סוקראטס הוא אדם, לכן סוקראטס הוא בן תמותה.

כאן מופיעות שתי הנחות ומסקנה. נתרגם את הטענות לשפה פורמלית יותר: ההנחה הראשונה היא "לכל x, אם x הוא אדם אז x הוא בן תמותה". ההנחה השנייה היא "סוקראטס הוא אדם". מכאן נובעת משיקולים לוגיים טהורים הטענה "סוקראטס הוא בן תמותה". אין מה להוסיף ואין צורך בהוכחה נוספת: זהו צעד לוגי בסיסי מהסוג המשמש אבן בניין להוכחות מורכבות יותר.

הנה דוגמה קצת יותר מורכבת מעולם המספרים:

טענה: אין מספר חיובי מינימלי.

נציג לטענה שתי הוכחות: אחת על דרך החיוב, כלומר נראה את הטענה בדיוק, והשנייה על דרך השלילה, כלומר נראה שהשלילה של הטענה שקרית (זה גורר כמובן שהטענה אמתית). **הוכחה** (על דרך החיוב) הטענה שקולה לטענה שאם x מספר חיובי אז יש מספר חיובי קטן ממנו. נוכיח טענה זו: יהי x מספר חיובי. אז גם $x/2$ הוא מספר חיובי, ומתקיים $x/2 < x$. לכן יש מספר חיובי קטן יותר מ־x, כפי שרצינו להראות. ■

הוכחה (על דרך השלילה): נצא מתוך הנחה שהטענה אינה נכונה ונראה שהנחה זו מובילה לסתירה. מכך נסיק שההנחה שלנו הייתה שגויה, ולכן הטענה כן נכונה.

אם כן, נניח שקיים מספר חיובי קטן ביותר (זו השלילה של הטענה), ונסמן מספר זה ב־x. המספר x חיובי, ולכן המספר $x/2$ גם־כן חיובי. אבל $x/2$ קטן יותר מ־x, בסתירה לכך ש־x הוא המספר החיובי הקטן ביותר. ■

כל טענה הניתנת להוכחה על דרך השלילה ניתנת להוכחה גם על דרך החיוב. אין יתרון לשיטה זו או אחרת, ואפשר לבחור את השיטה הנוחה יותר בכל מקרה ומקרה.

לפעמים אפשר להוכיח טענה על־ידי חלוקתה לכמה תת־טענות משלימות, והוכחת כל אחת מהן בנפרד. למשל,

טענה 2.2.1 אם n מספר שלם אז כשמחלקים את n^2 ב־4, השארית לעולם אינה 3.

הוכחה נוכיח את הטענה בשני מקרים: כאשר n זוגי וכאשר n אי־זוגי. מאחר שאחת האפשרויות תמיד מתקיימת הרי שאם בכל אחת מהאפשרויות המסקנה נכונה, אז היא נכונה בכל מקרה.

נניח ש־n זוגי. אז יש k שלם כך ש־ $n = 2k$ ולכן

$$n^2 = (2k)^2 = 4k^2$$

k^2 הוא מספר שלם, ולכן n^2 מתחלק ב־4, דהיינו השארית היא 0, וממילא השארית אינה 3.

נניח ש־n אי־זוגי. אז יש k שלם כך ש־ $n = 2k + 1$ ולכן

$$n^2 = (2k+1)^2 = (2k)^2 + 2 \cdot (2k) \cdot 1 + 1^2 = 4(k^2 + k) + 1$$

זהו סכום של 1 ומספר שלם המתחלק ב־4, לכן שארית חלוקתו ב־4 היא 1, וממילא אינה 3. ■

קורה לעתים שכדי להוכיח טענה יש לפרק אותה למספר רב של מקרים, ואת המקרים לתת־מקרים. אין בכך פסול, אך כמובן שיש להקפיד שלא לשכוח אף מקרה. מספיק שלא נבדוק מקרה אחד כדי לפסול את ההוכחה.

ביטוי שתפגשו מפעם לפעם בהוכחות הוא הביטוי **בלי הגבלת הכלליות** (without loss of generality), המשמש כדי לציין שלפנינו כמה מקרים שאין טעם להבדיל ביניהם, ולכן נבדוק מקרה אחד מייצג. למשל:

טענה 2.2.2 אם x, y מספרים שונים אז המספר הגדול ביניהם הוא גדול יותר מהממוצע שלהם.

הוכחה בלי הגבלת הכלליות נניח ש־ $y < x$. הממוצע של x, y הוא $\frac{1}{2}(x+y)$, ומתקיים

$$\frac{1}{2}(x+y) < \frac{1}{2}(x+x) = x$$

כפי שרצינו. ■

כאן בדקנו רק אחד משני מקרים אפשריים. ההצדקה לכך היא שהוכחת המקרה $x < y$ זהה למקרה $x > y$ אלא שהתפקידים שמשחקים x ו־y התחלפו (אם נחליף בהוכחה כל הופעה של x ב־y וכל הופעה של y ב־x נקבל הוכחה למקרה השני).

קיצור ההוכחה בדרך זו היא לגיטימית, אך יש להיזהר שלא לדלג בטעות על מקרה חשוב מתוך אמונה שגויה שהוא דומה למקרה שבדקנו. על הכותב, וגם על הקורא, מוטלת האחריות לוודא שקיצור דרך כזה הוא מוצדק, ואם יש ספק כדאי להוכיח את שני המקרים במלואם.

כמה שגיאות לדוגמה

כשלא מקפידים על כללי הלוגיקה אפשר "להוכיח" טענות שגויות ולתת הוכחות שגויות לטענות נכונות. הסכנה כאן היא במידה רבה פסיכולוגית, מכיוון שתמיד אפשר לבדוק האם הוכחה היא נכונה או לא: כל שיש לעשות הוא לעבור על שלבי ההוכחה ולוודא שכל אחד מהשלבים כשר. אולם הלהיטות להוכיח דברים מובילה פעמים רבות לכך שאיננו בודקים הוכחות בזהירות מספקת. להלן כמה דוגמאות כאלה.

"**טענה**" : $1 = -1$.

"**הוכחה**" עלינו להראות ש־ $1 = -1$. נעלה את שני האגפים של שוויון זה בריבוע. מכיוון ש־ $(-1)^2 = 1$, נקבל את השוויון $1 = 1$, שהוא שוויון נכון. לכן הטענה $1 = -1$ נכונה.

כאן הראינו שהטענה שאותה אנו מבקשים להוכיח גוררת טענה אחרת, שהיא במקרה טענה נכונה. אין זה אומר דבר על ערך האמת של הטענה המקורית. הנה ניסיון הוכחה נוסף של אותה טענה:

"**הוכחה**" נניח תחילה ש־ $4 = 0$. את הטענה הזו נוכיח מיד, אך תחילה נראה שהיא גוררת את מה שרצינו. ואמנם אם $4 = 0$ נוכל להחסיר 2 משני האגפים ולקבל את השוויון $2 = -2$. נחלק בשניים ונקבל $1 = -1$, כנדרש.

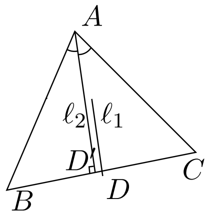
נוכיח כעת ש־ $4 = 0$. נשים לב ש־ $4 = 2 \cdot 2 = 2 \cdot (1 + 1)$. אבל הראינו כבר ש־ $1 = -1$ ולכן $1 + 1 = 1 + (-1) = 0$. מכאן ש־ $4 = 2 \cdot 0 = 0$, כפי שרצינו, והשלמנו את ההוכחה.

כאן יצרנו נימוק מעגלי: רצינו להוכיח טענה P, ולשם כך הראינו שיש טענה אחרת, Q, שממנה נובעת P. בשלב זה אם היינו מראים ש־Q נכונה בעזרת נימוק שמסתמך רק על עובדות ידועות היינו יכולים להסיק ש־P נכונה. אולם בהוכחת הטענה Q הסתמכנו על P, שהיא טענה שעדיין לא הוכחה (הוכחתה מותנית בנכונות של Q). ההוכחה לעיל היא למעשה הוכחה לטענה "$1 = -1$ אמ"מ $4 = 0$".

הנה דוגמה מסוג אחר. היא משתמשת במושגים מתורת הגאומטריה של המישור, ובעיקר במשפטי החפיפה של משולשים.

"**טענה**" : כל משולש הוא שווה־שוקיים.

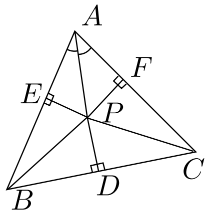
"**הוכחה**" : יהי ABC משולש במישור. עלינו להראות ש־ $AB = AC$. תהי D הנקודה על הצלע BC המחלקת את BC לשני קטעים שווי־אורך, כלומר $BD = DC$. נעביר ישר ℓ_1 דרך D הניצב ל־BC. כמו־כן נעביר ישר ℓ_2 דרך הקדקוד A באופן שחוצה את הזווית $\angle BAC$.



איור 2.2.1

אם שני הישרים מקבילים אז ℓ_2 ניצב לישר BC וחותך אותו בנקודה שנסמן אותה ב־D'. ראו איור 2.2.1. נשים לב שאז המשולשים $\Delta ABD'$ ו־$\Delta ACD'$ חופפים כי (א) יש להם צלע משותפת AD', (ב) $\angle BAD' = \angle CAD'$ כי AD' חוצה את הזווית $\angle BAC$, ו־(ג) AD' ניצב ל־BC ולכן הזוויות $\angle AD'B$ ו־$\angle AD'C$ ישרות, ובפרט שוות. לכן $AB = AC$, והמשולש הוא שווה שוקיים.

אחרת הישרים ℓ_1, ℓ_2 אינם מקבילים, ולכן הם נפגשים בנקודה P. נוריד מ־P אנכים לצלעות AB, AC ונקרא לנקודות המפגש עם הצלעות E, F בהתאמה. כמו־כן נחבר את P עם הקדקודים B, C. הבנייה מתוארת באיור 2.2.2.



איור 2.2.2

נשים לב שיש כאן כמה זוגות של משולשים חופפים. (*) המשולשים BDP, CDP חופפים, מכיוון שלפי בחירת D מתקיים $BD = DC$, הצלע PD משותפת, ו־ $\angle BDP = \angle CDP$ כי שתיהן ישרות. (*) המשולשים APE, APF חופפים, כי לפי הבניה $\angle AEP = \angle AFP$ (שתיהן זוויות ישרות), $\angle EAP = \angle FAP$ כי AP נבחר להיות חוצה זווית, והצלע AP משותפת לשני המשולשים. (*) המשולשים BEP, CFP חופפים, כי מתקיים $\angle BEP = \angle CFP$ (שתיהן ישרות), $EP = FP$ לפי העובדה שהמשולשים APE, APF חופפים, ו־ $BP = CP$ לפי העובדה שהמשולשים BDP, CDP חופפים.

מחפיפת המשולשים AFP, AEP נובע השוויון $AE = AF$, ומחפיפת המשולשים BEP, CFP נובע השוויון $BE = CF$. לכן

$$AB = AE + EB = AF + FC = AC$$

כפי שרצינו להראות.

הטענה כמובן אינה נכונה. השגיאה בהוכחה זו עדינה יותר מאשר בדוגמאות הקודמות, ומקורה אינה כשל לוגי, אלא הנחה סמויה שמקורה באיור שליווה את ההוכחה. ראשית, ציירנו את הנקודה P כאילו שייכת לפנים של המשולש, בעוד שיש מקרים בהם היא נמצאת מחוץ לו. שנית, השתמשנו בשוויונות $AB = AE + EB$ ו־ $AC = AF + FC$ המבוססים על ההנחה השגויה שהנקודה E נמצאת בין הנקודות A, B ושהנקודה F נמצאת בין הנקודות A, C.

הלקח מדוגמה זו הוא שאין להסתמך בהוכחה על אמצעי המחשה לא פורמליים. כדאי להיעזר בציורים, אך יש לעשות זאת בזהירות, ולהשתמש בהם כהמחשה ולא כהצדקה לשלב כלשהו בהוכחה.

החיפוש אחר ההוכחה

עד כאן הבאנו דוגמאות שהן בגדר "עשה" ו"אל תעשה". כפי שציינו למעלה אפשר תמיד לגלות הוכחה שגוייה על ידי בדיקה מדוקדקת. מצד שני, לא הסברנו כיצד למצוא הוכחות נכונות. הסיבה היא שאין שיטה כזאת. אמנם יש דפוסים שחוזרים על עצמם, ולאחר שתראו מספיק הוכחות ותפתרו מספיק תרגילים תפתחו לכך חוש מסוים, אך מעבר לכך הדרך היחידה היא ניסיון וטעייה.

לא תמיד השלימו המתמטיקאים עם מצב זה, ובסוף המאה ה־19 ובתחילת המאה ה־20 עוד קיוו רבים שהבנה טובה יותר של הלוגיקה תביא לגילוי של שיטה להוכחה או הפרכה של כל טענה. התקווה הייתה שאפשר לבנות מכונה (היום היינו קוראים לה תכנית מחשב) שמקבלת טענה כקלט ומייצרת הוכחה או הפרכה שלה. תקווה זו שיקפה את ההתפתחויות המרשימות בלוגיקה באותן שנים, אבל גם את מגמת האוטומציה בתחומים אחרים: בתקופה זו הומצאו בזו אחר זו מכונות חדשות שהחליפו בהצלחה את בני האדם בכל מיני משימות, החל ממשימות מכאניות שונות וכלה בחישובים מספריים מסובכים.

ואולם הסתבר שאין סיכוי לכך. המתמטיקאי קורט גדל[4] הוכיח ב־1931 שלא קיימת ולא יכולה להיות קיימת שיטה שתוכל לספק הוכחה או הפרכה לכל טענה מתמטית. יתרה מזאת, גם שיטות חלקיות שהומצאו הן בעלות יעילות נמוכה ביותר, ואינן מתקרבות עדיין ליכולות של מתמטיקאים בשר ודם.

תורת ההוכחה שייכת לתחום הלוגיקה המתמטית, וניתן לקרוא עליה בספרים [14],[15].

2.3 תורת הקבוצות

העולם המתמטי מורכב מ"איברים" (שמות נרדפים ל"איבר" הם "עצם", "ישות", "אובייקט" ועוד). חלק מאותם איברים הם אוספים של איברים אחרים. אוסף של איברים נקרא **קבוצה** (set).

אם A קבוצה, אז לכל איבר x מתקיימת בדיוק אחת משתי האפשרויות: או ש־x **שייך** ל־A, או ש־x **לא שייך** ל־A. אם x שייך ל־A נסמן $x \in A$, אחרת נסמן $x \notin A$. אם x שייך ל־A נאמר גם ש־A **מכילה** את x.

קבוצה אינה כוללת כל מידע נוסף על איבריה פרט לעצם השתייכותם לקבוצה. בפרט, אין משמעות לשאלה כמה פעמים x מופיע ב־A, או לשאלה מהו המיקום של x ב־A, או לכל שאלה אחרת מסוג זה.

שתי קבוצות הן שוות אם יש להן אותם איברים, דהיינו אם A, B קבוצות אז $A = B$ אמ"מ מתקיים $\forall x (x \in A \iff x \in B)$.

---

[4] Kurt Gödel, 1906־1978.

כדי לתאר קבוצה במפורש נרשום את האיברים שלה בין סוגריים מסולסלים. דרך אחת היא פירוט ישיר של רשימת האיברים: למשל

$$U = \{1, 2, 3\}$$

היא הקבוצה שאיבריה הם המספרים $1, 2, 3$. לאור העובדה שסדר האיברים וריבוים אינם בעלי חשיבות, מתקיימים גם השוויונות

$$U = \{1, 3, 2\} = \{2, 1, 3\} = \{1, 1, 1, 2, 2, 3\} = \{1, 1, 2, 2, 3, 3\}$$

וכן הלאה. מתקיימים בין השאר יחסי השייכות הבאים:

$$1 \in U \quad , \quad 2 \in U \quad , \quad 3 \in U \quad , \quad 4 \notin U$$

קבוצה יכולה להכיל בין איבריה גם קבוצות אחרות. כך אפשר לדבר על קבוצה של קבוצות, כמו

$$U' = \{\{1\}, \{2\}, \{3\}\}$$

זו קבוצה שונה מ־U, למשל כי אינה מכילה את 1 (היא מכילה כמה קבוצות, שאחת מהן היא הקבוצה $\{1\}$, אבל היא אינה מכילה את האיבר 1). כמו כן מתקיים

$$\{1\} \in U' \qquad , \qquad \{1\} \notin U$$

לעתים נגדיר קבוצה על ידי ציון תכונה מסוימת של איבריה. אז נרשום קבוצה בצורה $A = \{x : ...\}$, כאשר במקום שלוש הנקודות מופיע תנאי כלשהו.[5] יש לקרוא סימון זה כאילו כתוב "A היא קבוצת כל האיברים שמקיימים ...". באופן דומה נרשום $A = \{x \in B : ...\}$ כדי לציין את קבוצת האיברים מ־B עם תכונה מסוימת. יש לקרוא זאת כך: "A היא קבוצת כל האיברים ב־B שמקיימים ...".

לדוגמה, תהי $U = \{1, 2, 3\}$ כמו קודם, ונגדיר

$$\begin{aligned} V &= \{x : x + 1 \in U\} \\ W &= \{x \in U : x + 1 \in U\} \end{aligned}$$

אז $V = \{0, 1, 2\}$ ו־ $W = \{1, 2\}$.

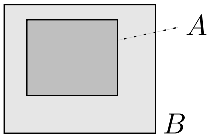
[5] לפעמים משתמשים בסימון $A = \{x \mid \ldots\}$ במקום בסימון $A = \{x : \ldots\}$.

תת־קבוצות והקבוצה הריקה

יהיו A, B קבוצות. נאמר ש־A **תת־קבוצה** (subset) של B, ונסמן זאת $A \subseteq B$,[6] אם כל איבר של A הוא גם איבר של B, כלומר

$$A \subseteq B \qquad \Longleftrightarrow \qquad \forall x\,((x \in A) \implies (x \in B))$$

במקרה זה נאמר גם ש־B **מכילה** את A או ש־A **מוכלת**[7] ב־B. אם A היא תת־קבוצה של B אבל אינה שווה ל־B אומרים ש־A **תת־קבוצה ממש** של B. אם A אינה תת־קבוצה של B נסמן $A \not\subseteq B$.



$A \subseteq B$

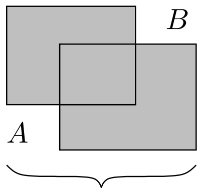
איור 2.3.1 A היא תת־קבוצה של B

דרך נוחה לתאר את יחס ההכלה היא בעזרת תרשים ון (Venn), שהוא ציור מהסוג שבאיור 2.3.1. בתרשים אנו מתארים כל קבוצה כצורה שבה האיברים הם אותן נקודות במישור הנמצאות בתוך הצורה. יחס ההכלה $A \subseteq B$ מתבטא בכך שהצורה A מוכלת כולה בצורה B.

למשל, בדוגמאות מהעמוד הקודם מתקיימים היחסים הבאים:

$$\{1\} \subseteq U \qquad , \qquad \{1\} \not\subseteq U'$$

כל קבוצה היא תת־קבוצה של עצמה, כלומר, לכל קבוצה A מתקיים בבירור $A \subseteq A$. כמו־כן, אם A, B קבוצות אז $A = B$ אמ"מ ($A \subseteq B$ וגם $B \subseteq A$). אולם שימו לב שלא לכל שתי קבוצות A, B מתקיימת אחת האפשרויות $A \subseteq B$ או $B \subseteq A$. למשל אם U, U' כמו בעמוד הקודם אז $1 \in U$ אבל $1 \notin U'$ ולכן $U \not\subseteq U'$. באותו אופן, $\{1\} \in U'$ אבל $\{1\} \notin U$ ולכן $U' \not\subseteq U$.

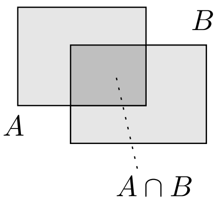


$A \cup B$

איור 2.3.2 האיחוד של הקבוצות A ו־B

הקבוצה שאינה מכילה אף איבר נקראת **הקבוצה הריקה** (emptyset), ומסומנת ב־$\emptyset$. הסיבה שהשתמשנו בהא הידיעה בקשר לקבוצה הריקה היא שיש רק קבוצה אחת כזו: שכן אם A, B קבוצות ריקות אז כל איבר באחת מהן שייך לשנייה (כי אין בהן בכלל איברים. זו דוגמה לטענה המתקיימת באופן ריק. עיינו בדיון בקשר "אם־אז" בעמוד 5). מסיבות דומות הקבוצה הריקה היא תת־קבוצה של כל קבוצה אחרת.

פעולות בין קבוצות



$A \cap B$

איור 2.3.3 החיתוך של הקבוצות A ו־B

הגדרה 2.3.1 יהיו A, B קבוצות. **האיחוד** (union) של A ו־B הוא קבוצה שאיבריה הם האיברים השייכים ל־A או ל־B (או לשניהם). קבוצה זו מסומנת $A \cup B$. **החיתוך**

---

[6] ישנם ספרים בהם כותבים $\subset$ במקום $\subseteq$.

[7] למילה "מכיל" יש כעת שני פירושים שונים: אפשר לפרש את הטענה "A מכיל את B" בתור $B \in A$ או בתור $B \subseteq A$. למרבה הצער זהו אי דיוק מושרש היטב. לרוב הכוונה ברורה מההקשר, ואם יש ספק אפשר להשתמש בסימנים $\in$ ו־$\subseteq$.

(intersection) של A ו־B הוא קבוצה שאיבריה הם האיברים המשותפים ל־A ול־B. קבוצה זו מסומנת על ידי $A\cap B$. בסימנים,

$$\begin{aligned} A\cup B &= \{x\,:\,x\in A\vee x\in B\}\\ A\cap B &= \{x\,:\,x\in A\wedge x\in B\}\end{aligned}$$

שימו לב לקשר ולדמיון בין הסימנים $\cap,\cup$ והקשרים $\wedge,\vee$. ניתן להיעזר בהם כדי לזכור את המשמעות שלהם (חשבו איך!).

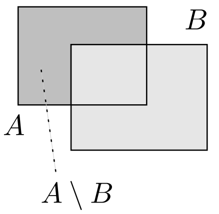
למשל, אם U,U' כמו קודם אז

$$\begin{aligned} U\cup U' &= \{1,2,3,\{1\},\{2\},\{3\}\}\\ U\cap\{2,3,4\} &= \{2,3\}\\ U\cap U' &= \emptyset\end{aligned}$$

היחס $A\cap B=\emptyset$ בין קבוצות אומר שלקבוצות A,B אין איברים משותפים. קבוצות כאלה נקראות קבוצות **זרות** (disjoint).

לעתים יש צורך לתאר חיתוך ואיחוד של יותר משתי קבוצות. נניח ש־I קבוצה שנחשוב עליה כקבוצה של פרמטרים, ולכל $i\in I$ תהי A_i קבוצה. האיחוד של כל ה־A_i־ים הוא הקבוצה המכילה את כל האיברים x שעבורם קיים $i\in I$ עם $x\in A_i$, ומסומנת ב־$\bigcup_{i\in I}A_i$. באופן דומה החיתוך של כל ה־A_i־ים הוא הקבוצה המכילה את כל האיברים x כך שלכל $i\in I$ מתקיים $x\in A_i$, ומסומנת ב־$\bigcap_{i\in I}A_i$.

לדוגמה, תהי I קבוצת בני האדם, ולכל אדם $i\in I$ תהי A_i קבוצת החברים שלו. אז $\bigcup_{i\in I}A_i$ היא קבוצת כל האנשים שהם חבר של מישהו, ו־ $\bigcap_{i\in I}A_i$ היא קבוצת האנשים שהם חברים של כולם.



איור 2.3.4 ההפרש בין הקבוצה A לקבוצה B

ההפרש (difference) בין קבוצה A לקבוצה B הוא הקבוצה המכילה את כל אותם איברים של A שאינם איברים של B. ההפרש בין A ל־B מסומן $A\setminus B$.[8] בסימנים,

$$A\setminus B=\{x\,:\,x\in A\,,\,x\notin B\}$$

קבוצות סופיות ואינסופיות

מתבקש להבחין בין קבוצות סופיות לקבוצות אינסופיות. קבוצה A היא **סופית** אם היא קבוצה ריקה או שקיים מספר שלם וחיובי n ואיברים $x_1,x_2,x_3,\ldots,x_n$ כך ש־ $A=\{x_1,x_2,x_3,\ldots,x_n\}$. **קבוצה אינסופית** היא קבוצה שאינה סופית. נדון בכך באופן מעט יותר פורמלי בסעיף 3.5, לאחר שנגדיר את המספרים השלמים.

[8] יש ספרים בהם כותבים $A-B$ במקום $A\setminus B$.

זוגות סדורים וקבוצות מכפלה

לעתים מתעורר הצורך לייצג זוג של איברים a, b באופן שמבחין בין האיבר הראשון לאיבר השני. הקבוצה $\{a,b\}$ אינה מתאימה לכך כי היא אינה "זוכרת" מי מבין שני האיברים הוא הראשון ומי השני (הרי $\{a,b\} = \{b,a\}$). אנו זקוקים לאובייקט הדומה לקבוצה אך אשר "זוכר" את הסדר של האיברים בו. אובייקט זה נקרא **הזוג הסדור** (ordered pair) ומסומן על ידי (a,b). האיבר a נקרא **הרכיב הראשון** בזוג (a,b) ו־b נקרא **הרכיב השני**. שני זוגות סדורים $(a,b),(a',b')$ שווים אמ"מ מתקיימים השוויונות $a = a', b = b'$.[9] ייתכן ש־ $a = b$, ואז מתקבל הזוג הסדור (a,a) שבו הרכיב הראשון והשני זהים.

היחס בין זוג סדור לרכיבים שלו אינו כמו היחס בין קבוצה לאיבריה, ולעולם לא נרשום, למשל, $a \in (a,b)$. גם אין משמעות לדבר על איחודים או חיתוכים של זוגות סדורים. מצד שני, ייתכנו קבוצות המכילות זוגות סדורים כפי שייתכנו קבוצות בעלות כל סוג שהוא של איברים. בפרט, אם A, B קבוצות, אז **המכפלה הקרטזית**[10] (Cartesian product) של A ו־B, המסומנת $A \times B$, היא קבוצת הזוגות הסדורים בעלי רכיב ראשון מ־A ורכיב שני מ־B. בסימנים,

$$A \times B = \{(a,b) \,:\, a \in A\,,\, b \in B\}$$

למשל,

$$\{0,1,2\} \times \{a,b\} = \{(0,a),(0,b),(1,a),(1,b),(2,a),(2,b)\}$$

נדגיש שכפל קבוצות הוא פעולה שונה מכפל מספרים, וגם תכונותיו שונות. למשל, $A \times B$ אינו שווה בהכרח ל־ $B \times A$, כפי שאפשר לראות בדוגמה $A = \{0\}$, ו־ $B = \{1\}$. כאן $A \times B = \{(0,1)\}$ ואילו $B \times A = \{(1,0)\}$, והקבוצות שונות כיוון ש־ $(0,1) \neq (1,0)$.

כפי שהגדרנו זוג סדור ומכפלה של שתי קבוצות אפשר להגדיר שלשה סדורה, רביעייה סדורה, וכן הלאה. באופן כללי לכל מספר n אפשר להגדיר n־יה שהיא סדרה בת n איברים: הראשון, השני, השלישי וכן הלאה עד האיבר ה־n. כמו־כן אפשר להגדיר מכפלות של יותר משתי קבוצות. למשל, $A \times B \times C$ היא קבוצת השלישיות (a,b,c) עבור $a \in A\,,\, b \in B\,,\, c \in C$.

במתמטיקה מודרנית שמור לתורת הקבוצות מקום של כבוד כתורה העומדת בפני עצמה, ולא רק כשפת עזר בתחומי מתמטיקה אחרים. מי שמעוניין להרחיב את השכלתו בכיוון זה מופנה ל־[12],[13].

---

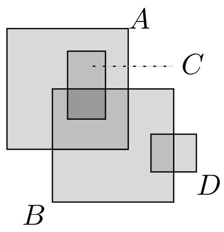
[9] הגדרנו את מושג הזוג הסדור כמושג בפני עצמו אך אפשר להגדיר אותו בעזרת מושגים בסיסיים יותר מתורת הקבוצות.

[10] על שם René Decartes, 1596־1650.

תרגילים

1. יהיו $A=\{1,2,3\}$ ו־ $B=\{3,4,5\}$. כתבו את הקבוצות $A\cup B, A\cap B, A\setminus B$ וקבעו אילו מהן מכילות את A או את B כתת־קבוצה.

2. באיור 2.3.5 כל גוון של אפור מגדיר קבוצה. תארו קבוצות אלה במונחים של הקבוצות A,B,C,D. למשל, הריבוע הכהה ביותר הוא $A\cap B\cap C$ (דרך אחרת לרשום אותו היא $B\cap C$).



איור 2.3.5

3. הוכיחו או הפריכו את היחסים הבאים:
   (א) $\emptyset=\{\emptyset\}$.
   (ב) $\emptyset=\{\}$.
   (ג) $\emptyset\in\{\emptyset\}$.
   (ד) $\emptyset\in\{\}$.
   (ה) $\emptyset\subseteq\{\emptyset\}$.
   (ו) $\emptyset\subseteq\{\}$.

4. נתונות שלוש קבוצות, A,B,C. ציירו דיאגרמות ון המתאימות לכל אחד מהתרחישים הבאים:
   (א) $A\cap B=\emptyset$, $C\subseteq B$.
   (ב) $A\cap B=\emptyset$, $C\subseteq A\cup B$.
   (ג) $A\subseteq B$, $C\subseteq B\setminus A$.
   (ד) $A\subseteq B$, $C\subseteq B$, $C\not\subseteq A$, $C\not\subseteq B\setminus A$.

5. הוכיחו את הטענות הבאות או הפריכו על ידי דוגמה נגדית (כדוגמה אפשר להסתפק בתרשים ון).
   (א) קיימות קבוצות A,B,C כך ש־ $A\cap B\cap C=\emptyset$ אבל $A\cap B\neq\emptyset$, $A\cap C\neq\emptyset$, $B\cap C\neq\emptyset$.
   (ב) קיימות קבוצות A,B,C כך ש־ $B\subseteq A$, $C\subseteq A$ אבל $(A\setminus B)\cap C=\emptyset$.
   (ג) קיימות קבוצות A,B,C,D כך ש־ $A\subseteq B$, $C\subseteq D$, $B\cap D=\emptyset$ אבל $A\cap C\neq\emptyset$.

6. הוכיחו שלכל A,B,C,D מתקיימים השוויונות הבאים:
   (א) $(A\cap B)\cap C=A\cap(B\cap C)$ (מכאן שניתן לרשום $A\cap B\cap C$ בלי דו־משמעות).
   (ב) $A\setminus(B\cup C)=(A\setminus B)\cap(A\setminus C)$.
   (ג) $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$ וגם $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$ (שוויונות אלה נקראים **כללי דה־מורגן**).[11]

[11] Augustus De Morgan, 1806־1871.

פרק 3

המספרים הממשיים

החשבון האינפיניטסימלי עוסק באובייקטים הקשורים למספרים כמו סדרות של מספרים, פונקציות שערכיהן מספרים, וכן הלאה. בכל אלה נדון בהמשך, אך תחילה עלינו להסביר מהו מספר ומהן התכונות של המספרים. נושאים אלו יעסיקו אותנו בשני הפרקים הקרובים, בהם נכיר את מערכת המספרים הממשיים.

בפרק הנוכחי נתרכז בתכונות הבסיסיות של המספרים: פעולות החיבור, החיסור, הכפל והחילוק, ויחס הסדר בין המספרים. נאפיין גם את המספרים השלמים והשברים, נדון בעיקרון האינדוקציה, ובייצוג העשרוני של המספרים השלמים.

3.1 ההצגה האקסיומטית של המספרים הממשיים

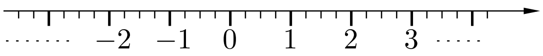
בבית־הספר לומדים על מספרים מנקודות מבט רבות. אחת מהן היא השיטה הסמלית, שבה מתארים מספרים באופנים שונים בעזרת סמלים. דוגמאות לשיטות רישום כאלה הן שיטת היצוג העברית א,ב,ג,...,ט,י,י"א,י"ב,... ושיטת היצוג הרומית I,II,III,IV,..., אבל השיטה המוכרת והיעילה ביותר היא השיטה העשרונית:

$$0,1,2,\ldots,9,10,11,\ldots,99,100,101,\ldots$$

משיש בידינו תיאור של המספרים השלמים אפשר לתאר שברים כמנות של מספרים שלמים. אפשר גם לתאר חלק מהשברים בכתיב עשרוני. למשל, את המספר שבע מאיות אפשר לרשום כ־$\frac{7}{100}$ או 0.07. לפעמים לומדים בבית־הספר גם על שברים עשרוניים אינסופיים כמו המספר $0.069999\ldots$ (זהו שוב המספר שבע מאיות!).

דרך נוספת לייצג את המספרים היא לזהות אותם עם נקודות על ישר. בשיטה זו קובעים ישר מאוזן אינסופי בשני הכיוונים המכונה **ציר המספרים** או **הישר הממשי** (number line), ובוחרים עליו נקודה מובחנת הנקראת 0 או **הראשית**. המספרים מזוהים עם הנקודות על ציר המספרים: אלה שמימין לאפס חיוביים ואלה שמשמאלו

שליליים. לאחר שקובעים יחידת אורך אפשר להגדיר את המספרים השלמים ואת השברים; למשל, המספר 1 הוא הנקודה מימין לראשית שלמרחקו מהראשית הוא יחידת אורך אחת. איור 3.1.1 למטה מייצג את ציר המספרים. הכיוון של החץ מרמז על הכיוון החיובי.



איור 3.1.1 ציר המספרים

שיטות אלה טובות ויעילות אך לא נבחר בהן כבסיס להגדרת המספרים. אחת הסיבות לכך היא שבבית הספר לרוב אין מגדירים אותן באופן מדויק, ולא מוכיחים את התכונות השונות של המספרים. למשל, לא מוכיחים את הכלל הידוע שאם x, y מספרים אז $x+y=y+x$. במקום הוכחה סומכים על כך שלאחר שראו מספר רב של דוגמאות התלמידים ישתכנעו שהכלל מתקיים תמיד. זו כמובן אינה הוכחה.

התבוננות ביקורתית מעלה שאלות נוספות, ובראשן השאלה האם כל השיטות מתארות את אותם מספרים. התשובה שלילית: המספר $\frac{1}{3}$, למשל, אינו ניתן לייצוג כשבר עשרוני סופי. מאידך יש מספרים הניתנים לייצוג כשבר עשרוני אינסופי אך לא כמנה של מספרים שלמים, כמו $\sqrt{2}$ ו־π (אלה עובדות שלעתים אינן מוזכרות בבית הספר). ועוד: האם כל נקודה על ציר המספרים מתאימה למנה של מספרים שלמים? או לשבר עשרוני סופי? או לשבר עשרוני אינסופי? ולהפך, האם כל שבר עשרוני אינסופי מתאים לנקודה על הישר? לאור שאלות אלה עולה השאלה, האם יש בין השיטות האלה אחת שהיא כללית מספיק כדי לתאר את כל המספרים, או שמא כולן מתארות רק חלק מהמספרים?

לו רצינו להתבסס על הידע שלנו מבית הספר היינו נדרשים לענות על השאלות לעיל. הדבר אפשרי, אך הוא היה גוזל זמן רב. במקום זאת ניתן הגדרה עצמאית של המספרים, שנועדה להביע את מושג המספר המופשט המסתתר מאחורי כל השיטות שלומדים בבית הספר, ואת התכונות המשותפות שלהן.

ההגדרה לא תשען על תיאור קונקרטי מסוים של המספרים אלא תיעשה בשיטה האקסיומטית, כלומר, לא נגדיר מהם המספרים אלא נתאר את התכונות שלהם. המילה אקסיומה, שמקורה בשפה היוונית, שמשה בידי המתמטיקאים ביוון העתיקה כדי לתאר הנחה שאיה ניתנת להצדקה או שמובנת מאליה. כיום המילה מציינת הנחת יסוד של תורה מתמטית, שתפקידה להגדיר את המערכת בה עוסקים. אי אפשר להיפטר לגמרי מהנחות כאלה, והן קיימות בכל תורה מתמטית.

על אף שפרטי ההגדרה מפוזרים על פני שני הפרקים הקרובים, למען הסדר הטוב נרכז בקיצור את החלקים כאן.

הגדרה 3.1.1 מערכת המספרים הממשיים (real number system) מורכבת מקבוצה $\mathbb{R}$ שאיבריה נקראים **מספרים ממשיים** (real numbers), מפעולות חיבור וכפל, ומיחס סדר, באופן שמתקיימים אקסיומות השדה (סעיף 3.2), אקסיומות הסדר (סעיף 3.3), ואקסיומת השלמות (סעיף 4.2 בפרק הבא).

הערות

1. ההגדרה הזו היא רק שלד. התוכן האמתי מצוי באקסיומות.
2. לפי המינוח שבחרנו, הסימון $x \in \mathbb{R}$ והאמירה "x הוא מספר (ממשי)" מבטאים בדיוק את אותו הדבר.
3. ההגדרה אינה עונה על השאלה מהו מספר: האם יש לו צורה, משקל, וכן הלאה. גם לא נענה על שאלה זו בהמשך. אי הוודאות הזו לא תיצור שום קושי, כי השאלה לעולם לא תתעורר.
4. בסך הכול נזדקק ל־11 אקסיומות כדי להגדיר את $\mathbb{R}$. זו רשימה קצרה מאד של תכונות, אך כפי שנראה כל התכונות האחרות של המספרים נובעות מהן ומהן בלבד.
5. שתי שאלות עולות מהגדרה 3.1.1. ראשית, האם בכלל קיימת מערכת המקיימת את האקסיומות של הממשיים? התשובה חיובית. שאלה שנייה היא, האם יש רק מערכת אחת כזו או שמא יש מערכות שונות רבות. התשובה היא שיש רק אחת, או ליתר דיוק שכל המערכות המקיימות את האקסיומות של הממשיים ניתנות לזיהוי אחת עם השנייה באופן שאין טעם להבדיל ביניהן (הדבר דומה לכך שכשדנים במשחק השח־מט אין זה חשוב באיזה ערכת משחק מדובר. אם נחליף את הלוח ואת הכלים בלוח אחר ובכלים אחרים, לא ישתנה שום דבר מהותי). את שתי הטענות האלה לא נוכיח כאן.[1] אפשר לקרוא עליהן למשל בספר [13].

ההצגה האקסיומטית של המספרים והדיון הפורמלי בהם עלולים לעורר את הרושם שהמספרים הממשיים שונים מהמספרים שעליהם לומדים בבית הספר. לא כך הדבר: מהר מאד יתברר שכל ההרגלים והשיטות שלומדים בבית הספר נכונים במערכת $\mathbb{R}$. כך לגבי כללי החשבון, שנובעים מיד או עם מאמץ קל מהאקסיומות. עם מעט יותר תחכום נצליח לזהות בתוך $\mathbb{R}$ מחלקות של מספרים שלמים ושל שברים, ונראה שהייצוג העשרוני עובד כמצופה. מאוחר יותר (בפרק 6) נראה שכל מספר ניתן לתיאור כשבר עשרוני אינסופי. נראה גם שיטות ייצוג אחרות.

לא נעסוק באופן פורמלי במודל הגאומטרי של המספרים. היה אפשר לעשות זאת, אך לשם כך יש לפתח באופן פורמלי את תורת הגאומטריה. לא נרצה להתעכב על כך. למרות זאת, הזיהוי של מספרים עם נקודות על ישר, או עם אורכים של קטעים, הוא דרך מצוינת לדמיין מספרים, וכדאי מאד להיעזר בו. המודל הגאומטרי ייתן בהמשך מוטיבציה ואינטואיציה למהלכים הפורמליים שלנו, ולעתים קרובות אף נקרא למספרים בשם נקודות כדי לרמוז על הפירוש הגאומטרי שלהם.

במהלך הקריאה של פרק זה יש לשמור על פיצול אישיות מסוים. מצד אחד, מכאן ואילך **התכונות היחידות של המספרים שמותר להסתמך עליהן הן אלה שנובעות מהאקסיומות**. לכן כשאתם קוראים את הפרק ופותרים תרגילים בדקו תמיד

[1] כדי להוכיח את קיום הממשיים משתמשים רק בתורת הקבוצות, ואין צורך בהנחות נוספות.

שמא אתם מסתמכים על הרגלים ישנים ולא על תכונות שהוכחו מהאקסיומות. מצד שני, המערכת $\mathbb{R}$ נועדה לתאר את אותם המספרים שמוכרים לכם זה מכבר. האינטואיציה שפיתחתם בבית הספר לעתים קרובות תהיה נכונה, וכדאי להיעזר בה. בסיום הפרק המצב ישוב פחות או יותר לקדמותו, שכן כל התכונות המוכרות של המספרים יוכחו במהלך הפרק, ותוכלו אז לשוב ולנהוג במספרים לפי ההרגלים הישנים.

3.2 הפעולות החשבוניות

סעיף זה מביא קבוצת אקסיומות המכונה **אקסיומות השדה**[2] העוסקת בפעולות החשבוניות חיבור כפל.

לכל שני מספרים ממשיים x, y מתאים מספר $x+y$ הנקרא **הסכום** של x ו־y, ומספר $x \cdot y$ הנקרא **המכפלה** של x ו־y. לרוב נשמיט את סימן הכפל $\cdot$ ונרשום xy במקום $x \cdot y$.

תכונה ראשונה של המספרים היא שסדר האיברים וסדר ביצוע הפעולות אינו משנה את התוצאה:

אקסיומה I חילוף (commutativity): לכל $x, y \in \mathbb{R}$ מתקיים $x + y = y + x$ וגם $x \cdot y = y \cdot x$.

אקסיומה II קיבוץ (associativity): לכל $x, y, z \in \mathbb{R}$ מתקיים $(x+y)+z = x+(y+z)$ וגם $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.

פעולות החיבור והכפל הוגדרו כפעולות בין זוגות של מספרים, ולכן הביטוי $x+y+z$, למשל, כלל אינו מוגדר. ניתן להפוך אותו לביטוי חוקי על יד הכנסת סוגריים. יש כמה דרכים לעשות זאת, למשל: $(x + y) + z$, או $x + (y + z)$. אם נרשה לשנות את סדר המחוברים יש עוד אפשרויות, כגון $(x + z) + y$. בעזרת חוקי החילוף והקיבוץ אפשר להראות שכל דרכי הכתיבה השונות נותנות אותה תוצאה (כדי להוכיח זאת יש למנות את כל הדרכים האלה ולבדוק שאפשר לעבור בין כל שתים מהן על ידי שרשרת של שוויונות, שכל אחת מהן נובעת מאחת האקסיומות). מסיבה זו לא נטרח להבחין בין צורות הכתיבה השונות, ונרשה לעצמנו לכתוב $x+y+z$ כדי לציין את המספר המתקבל מכל אחת מדרכי חישוב אלה. הערה דומה חלה על סכומים של יותר משלושה מספרים, וגם על מכפלות.

לגבי ביטויים המכילים פעולות חיבור וכפל יחד נאמץ את המוסכמה הקובעת שפעולת הכפל קודמת לפעולת החיבור. כך הביטוי $x \cdot y + z$ מפורש תמיד בתור $(x \cdot y) + z$ ולעולם אינו מפורש כ־ $x \cdot (y + z)$.

[2] מערכת המקיימת את אקסיומות השדה מכונה **שדה** (field). בנוסף למספרים הממשיים יש גם שדות אחרים.

אקסיומה III פילוג (distributivity): לכל $x,y,z\in\mathbb{R}$ מתקיים $x\cdot(y+z)=x\cdot y+x\cdot z$.

ישנה גרסה סימטרית לחוק הפילוג: לכל $x,y,z\in\mathbb{R}$ מתקיים $(x+y)\cdot z=x\cdot z+y\cdot z$. הנה הוכחה מפורטת של השוויון, המסתמכת על האקסיומות הנתונות. בהינתן מספרים $x,y,z\in\mathbb{R}$, נתבונן בשרשרת השוויונות הבא:

$$(x+y)\cdot z=z\cdot(x+y)=z\cdot x+z\cdot y=x\cdot z+y\cdot z$$

ההוכחה תושלם אם נצדיק כל שוויון, שכן אז הביטוי הימני ביותר והשמאלי ביותר שווים, וזה מה שרצינו להוכיח. ואמנם, השוויון האמצעי נובע מאקסיומת הפילוג ושני השוויונות האחרים נובעים מאקסיומת החילוף.

אקסיומה IV איברים ניטרליים (unity): קיימים שני מספרים שונים $0,1\in\mathbb{R}$ כך שלכל $x\in\mathbb{R}$ מתקיים $x+0=x\cdot 1=x$. האיבר 0 נקרא **אפס**, ואילו 1 נקרא **אחד**.[3]

מאקסיומות החילוף נובע שמתקיים שלכל מספר x מתקיים $0+x=1\cdot x=x$.

אקסיומה V איברים נגדיים (additive inverses): לכל $x\in\mathbb{R}$ קיים $y\in\mathbb{R}$ יחיד כך ש־ $x+y=0$. מספר זה מסומן $-x$ ונקרא **האיבר הנגדי ל־**x.

אקסיומה VI איברים הפכיים (multiplicative inverses): לכל $0\neq x\in\mathbb{R}$ קיים $y\in\mathbb{R}$ יחיד כך ש־ $x\cdot y=1$. מספר זה מסומן x^{-1} ונקרא **האיבר ההפכי ל־**x.

בעזרת איברים נגדיים והפכיים ניתן להגדיר את פעולות החיסור והחילוק. אם $x,y\in\mathbb{R}$ אז **ההפרש** בין x ו־y הוא המספר $x+(-y)$ ומסומן על ידי $x-y$. אם $y\neq 0$ אז **המנה** של מספרים x ו־y הוא המספר $x\cdot y^{-1}$, ומסומן x/y או $\frac{x}{y}$. שימו לב שלפי סימון זה, לכל $x\neq 0$ מתקיים $\frac{1}{x}=1\cdot x^{-1}=x^{-1}$.

מפליא לגלות שקומץ האקסיומות שמנינו עד כה מאפשר להוכיח תכונות מוכרות רבות של המספרים. אנו מרכזים תכונות אלה במשפט הבא:

משפט 3.2.1 (התכונות החשבוניות של המספרים הממשיים):

1. לכל $x,y,a\in\mathbb{R}$, אם $x+a=y+a$ אז $x=y$. כמו כן אם $ax=ay$ ובנוסף $a\neq 0$ אז $x=y$.
2. לכל $x\in\mathbb{R}$ מתקיים $0\cdot x=0$.
3. אם $a,b\in\mathbb{R}$ ו־ $a\neq 0$ אז קיים מספר יחיד x המקיים את השוויון $ax+b=0$.
4. $-0=0$ וכן $1^{-1}=1$.
5. לכל $x,y\in\mathbb{R}$ מתקיים $-(x+y)=(-x)+(-y)$.

[3] אפס נקרא לפעמים גם **האיבר הניטרלי לחיבור** או **היחידה החיבורית**, ואחד נקרא לפעמים **האיבר הניטרלי לכפל** או **היחידה הכפלית**.

6. לכל $x \in \mathbb{R}$ מתקיים $-(-x) = x$.
7. לכל $x, y \in \mathbb{R}$, אם $x, y \neq 0$ אז $x \cdot y \neq 0$ ומתקיים $(x \cdot y)^{-1} = x^{-1} \cdot y^{-1}$.
8. לכל $x \in \mathbb{R}$, אם $x \neq 0$ אז $x^{-1} \neq 0$ ומתקיים $(x^{-1})^{-1} = x$.
9. לכל $x \in \mathbb{R}$ מתקיים $(-1) \cdot x = -x$.
10. לכל $x, y \in \mathbb{R}$ מתקיים $-(x \cdot y) = (-x) \cdot y = x \cdot (-y)$.
11. לכל $x \in \mathbb{R}$ אם $x \neq 0$ אז $-x \neq 0$ ומתקיים $-(x^{-1}) = (-x)^{-1}$.

הוכחה סעיף (1) לפעמים נקרא לפעמים **עקרון הצמצום**. יהיו x, y מספרים שמקיימים $x + a = y + a$ למספר a כלשהו. נחבר $-a$ לשני האגפים של השוויון $x + a = y + a$ נקבל את השוויון

$$(x + a) + (-a) = (y + a) + (-a)$$

מקיבוץ מקבלים שוויון שקול

$$x + (a + (-a)) = y + (a + (-a))$$

מההגדרה של המספר $-a$ נובע $a + (-a) = 0$, ולכן השוויון הקודם שקול לשוויון

$$x + 0 = y + 0$$

ומהגדרת האפס מקבלים $x = y$. המקרה הכפלי נובע בצורה דומה.

נוכיח את (2). יהי $x \in \mathbb{R}$. נשים לב שהשוויון $0 = 0 + 0$ גורר $0 \cdot x = (0 + 0) \cdot x$. חוק הפילוג נותן שוויון שקול:

$$0 \cdot x + 0 \cdot x = 0 \cdot x$$

ומכיוון ש־ $0 \cdot x = 0 + 0 \cdot x$ לפי תכונה האפס, מקבלים

$$0 \cdot x + 0 \cdot x = 0 + 0 \cdot x$$

בעזרת הסעיף הראשון במשפט נצמצם מחובר $0 \cdot x$ מכל אגף ונקבל ש־ $0 \cdot x = 0$.

נוכיח את (3). יהיו $a, b \in \mathbb{R}$ עם $a \neq 0$. ראשית נראה שהמספר $x = a^{-1} \cdot (-b)$ מקיים את השוויון $ax + b = 0$, ואז נראה שזהו המספר היחיד שמקיים זאת. ואמנם,

$$\begin{aligned}
ax + b &= a(a^{-1} \cdot (-b)) + b \\
&= (aa^{-1})(-b) + b \\
&= 1 \cdot (-b) + b \\
&= (-b) + b \\
&= 0
\end{aligned}$$

השוויונות לעיל נובעים מכלל הקיבוץ, מתכונות ההפכי והנגדי ומהתכונות של $0,1$. לגבי היחידות, נניח ש־y הוא מספר המקיים $ay+b=0$. אז מתקיים $ax+b=ay+b$, והפעלת עקרון הצמצום פעמיים גורר ש־ $x=y$, כפי שרצינו.

נוכיח את (4). כדי לראות ש־ $-0=0$, נשים לב ששני המספרים $0,-0$ פותרים את המשוואה $x+0=0$. העובדה ש־0 הוא פתרון נובעת מתכונת הניטרליות של 0, והעובדה ש־-0 הוא פתרון נובע מהגדרת הנגדי. כעת, לפי הסעיף הקודם למשוואה הזו יש פתרון יחיד, כלומר שני הפתרונות שמצאנו שווים זה לזה: $0=-0$ (המשוואה $x+0=0$ היא מהסוג $ax+b=0$ עבור $a=1$ ו־$b=0$). ההוכחה ש־ $1^{-1}=1$ דומה: מראים ששניהם פותרים את המשוואה $1\cdot x=1$.

הוכחת יתר הסעיפים מושארת כתרגיל למעוניינים. בכולם הרעיון דומה: על מנת להוכיח ששני מספרים שווים, מראים ששניהם פותרים המשוואה מתאימה שצורתה $ax+b=0$. הוכחות אפשר גם למצוא בספרים [16][6],[10]. ■

מהמשפט אפשר להסיק את כללי החשבון המוכרים של שברים. נוכיח למשל שלכל $x,y,a\in\mathbb{R}$, אם $y,a\neq 0$ אז $\frac{x}{y}=\frac{xa}{ya}$. ואמנם,

$$\frac{xa}{ya}=(xa)(ya)^{-1}=(xa)(y^{-1}a^{-1})=x(aa^{-1})y^{-1}=x\cdot 1\cdot y^{-1}=x\cdot y^{-1}=\frac{x}{y}$$

כאן השוויון הראשון משמאל נובע מהגדרת המנה, השני מסעיף (7) של המשפט, השלישי מאקסיומת החילוף והקיבוץ (ראו דיון אחרי אקסיומה זו על סכומים ומכפלות של יותר משני איברים), הרביעי מהגדרת ההפכי, החמישי מהגדרת המספר 1 והשישי שוב לפי הגדרת המנה. יתר כללי החשבון המוכרים של שברים מתקיימים גם־כן. ראו תרגיל (5) בסוף הסעיף.

התוצאות שראינו עד כה עלולות לעורר רושם שאקסיומות השדה כבר גוררות את כל התכונות המוכרות של המספרים. אין הדבר כך: לא נוכל בשלב זה להוכיח אפילו ש־ $1+1\neq 0$. על מנת להוכיח שאי אפשר להוכיח זאת לא די לחפש הוכחה ולהיכשל, אלא יש למצוא דוגמה למערכת המקיימת את כל אקסיומות השדה ובכל זאת מתקיים בה $1+1=0$. ואמנם יש מערכות כאלה, כמו המערכת הבאה המסומנת $\mathbb{F}_2$ ומכילה בדיוק שני איברים, המסומנים $0,1$. הפעולות מוגדרות בעזרת הטבלאות הבאות (שהן "לוח החיבור" ו"לוח הכפל" ב־$\mathbb{F}_2$):

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

שימו לב שאכן מתקיים $1+1=0$. הבדיקה שאקסיומות השדה מתקיימות קלה בדוגמה זו כי כל שיש לעשות הוא לבדוק מספר סופי של תנאים. למשל כדי לוודא שהחיבור מקיים את אקסיומת החילוף די לבדוק שלכל שני איברים $x,y\in\mathbb{F}_2$

מתקיים $x+y=y+x$. אבל יש רק ארבעה זוגות של איברים, ואפשר לבדוק זאת מהטבלה.

תורת השדות היא תורה מעניינת ועשירה עם שימושים מגוונים במתמטיקה. אפשר לקרוא עליה בספר [16], או בספר המתקדם יותר [17].

תרגילים

1. הוכיחו בפירוט שלכל $a,b,c,d \in \mathbb{R}$ מתקיים $(a+b)(c+d)=ac+ad+bc+bd$. הסיקו את הנוסחה המוכרת $(a+b)^2=a^2+2ab+b^2$ (כאן $x^2=x\cdot x$ ו־$2=1+1$)

2. השלימו את הוכחת משפט 3.2.1. בכל שלב אפשר להיעזר בסעיפים הקודמים!

3. יהי $x \in \mathbb{R}$. הוכיחו שאם a מספר המקיים $x+a=x$ אז $a=0$. באותו אופן, הראו שאם $x \neq 0$ ואם a מקיים $x\cdot a=x$ אז $a=1$. (מכאן אנו למדים שהתכונות המגדירות את $0,1$ מאפיינות אותם לגמרי, כלומר, אין עוד מספרים עם תכונות אלה).

4. הוכיחו את הגרסאות הבאות של חוק הפילוג עבור פעולות החיסור והחילוק (היעזרו במשפט 3.2.1):

   (א) לכל $x,y,z \in \mathbb{R}$ עם $z \neq 0$ מתקיים $\frac{x+y}{z}=\frac{x}{z}+\frac{y}{z}$.

   (ב) לכל $x,y,z \in \mathbb{R}$ מתקיים $z(x-y)=zx-zy$.

5. הוכיחו שאם x,y,u,v מספרים ו־$y,v \neq 0$ אז מתקיימים כללי המנה הרגילים:

$$\frac{x}{y}\cdot\frac{u}{v} = \frac{x\cdot u}{y\cdot v}$$
$$\frac{x}{y}+\frac{u}{v} = \frac{xv+yu}{yv}$$

ואם גם $u \neq 0$ אז

$$\frac{x/y}{u/v}=\frac{xv}{yu}$$

היעזרו במשפט 3.2.1 ובעובדה ש־$\frac{xa}{ya}=\frac{x}{y}$ אשר הוכחה אחרי המשפט.

6. הוכיחו את השוויון $(x+y)(x-y)=x^2-y^2$ (כאן סימנו $x^2=x\cdot x$ ו־$y^2=y\cdot y$) והסיקו שיש ב־$\mathbb{R}$ לכל היותר שני מספרים המקיימים $x^2=1$. (*) האם תוכלו להוכיח שיש בדיוק שני מספרים כאלה?

7. (*) תהי A קבוצה ותהי $\mathcal{P}$ קבוצת התת־קבוצות של A, כלומר

$$\mathcal{P}=\{B : B \subseteq A\}$$

במערכת זו 0 יסמן את הקבוצה הריקה ו־1 את A. נגדיר פעולות בין איברי $\mathcal{P}$ באופן הבא: לקבוצות $B, C \subseteq A$

$$\begin{aligned} B + C &= (B \setminus C) \cup (C \setminus B) \\ B \cdot C &= B \cap C \end{aligned}$$

(ציירו דיאגרמות ון להמחשת הפעולות!)

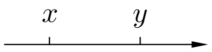
מההגדרה נובע בקלות ש־ $B + 0 = B$ ו־ $B \cdot 1 = B$ לכל $B \in \mathcal{P}$ (וודאו זאת!). בדקו שהמערכת שהגדרנו עתה מקיימת גם את האקסיומות I־IV. מתי היא מקיימת את אקסיומות הנגדי וההפכי?

8. הוכיחו בפירוט שהמערכת $\mathbb{F}_2$ היא שדה.

9. נניח ש־$\mathbb{F}$ שדה ו־ $a, b \in \mathbb{F}$ מקיימים $a - b = b - a$. האם $a = b$?

3.3 תכונות הסדר

סעיף זה מביא קבוצה נוספת של אקסיומות שיחד עם האקסיומות מהסעיף הקודם מכונה **אקסיומות השדה הסדור**.[4]



איור 3.3.1 x קטן מ־y

במערכת $\mathbb{R}$ קיים **יחס סדר** (order) בין מספרים. היחס "x גדול מ־y" (או באופן שקול "y קטן מ־x") מסומן $x > y$ (או באופן שקול $y < x$). יחסים אלה נקראים **אי־שוויונות**. מבחינה גאומטרית, $x < y$ אמ"מ הנקודה המתאימה ל־x על ציר המספרים נמצאת משמאל לנקודה המתאימה ל־y.

בעזרת היחס "גדול מ..." נגדיר יחס נוסף, "גדול או שווה ל...". היחס $x \leq y$ (המסומן גם $y \geq x$) פירושו ש־y גדול מ־x או שווה לו, כלומר

$$x \leq y \iff (x = y) \vee (x < y)$$

היחס $\leq$ מכונה **אי־שוויון חלש** weak inequality, לעומת היחס $<$ שנקרא **אי־שוויון חזק** strict inequality. נשתמש במונח אי־שוויון כדי להתייחס לאי־שוויונות משני הסוגים. שימו לב שאי־שוויון חלש בין מספרים אינו פוסל אי־שוויון חזק. אם $x < y$ אז מתקיים גם $x \leq y$. באי־שוויון החלש יש פחות מידע אך הוא עדיין נכון.

מספר x המקיים $x > 0$ נקרא **חיובי** (positive), ומספר המקיים $x < 0$ נקרא **שלילי** (negative). קבוצות המספרים החיוביים והשליליים מסומנות על ידי

$$\begin{aligned} \mathbb{R}^{+} &= \{x \in \mathbb{R} : x > 0\} \\ \mathbb{R}^{-} &= \{x \in \mathbb{R} : x < 0\} \end{aligned}$$

[4] מערכת המקיימת את אקסיומות השדה ואת אקסיומות השדה הסדור נקראת, כמובן, **שדה סדור**.

מספר x נקרא **מספר אי־שלילי** אם $x \geq 0$, ונקרא **אי־חיובי** אם $x \leq 0$. נאמר ששני מספרים x, y הם **שווי סימן** אם שניהם אי־שליליים או שניהם אי־חיוביים.

כשקוראים אי־שוויונות חשוב לזכור שקוראים ביטויים מתמטיים משמאל לימין. כשכתוב "יהי $0 < y$" הכוונה היא "יהי y גדול מ־0" ומתכוונים "יהי y מספר גדול מאפס". קריאת הטענה כאילו כתוב "יהי 0 קטן מ־y" היא מבלבלת.

התכונות הבסיסיות של יחס הסדר הם:

אקסיומה VII השוואה (trichotomy): לכל $x, y \in \mathbb{R}$ מתקיימת אחת ובדיוק אחת מהאפשרויות $x = y$, $x < y$ או $y < x$.

אקסיומה VIII תורשתיות (transitivity): לכל $x, y, z \in \mathbb{R}$, אם $x < y$ וגם $y < z$ אז $x < z$.

בהינתן זוג אי־שוויונות $y < z$, $x < y$ נקצר ונרשום $x < y < z$. שימו לב ששרשרת כזאת גוררת, לפי אקסיומת התורשתיות, את האי־שוויון $x < z$

הקשר בין יחס הסדר לפעולות החיבור והכפל מוסדר באקסיומות הבאות:

אקסיומה IX אי־רגישות הסדר לחיבור (invariance to addition): לכל $x, y, a \in \mathbb{R}$ מתקיים $x < y$ אמ"מ $x + a < y + a$.

אקסיומה X אי־רגישות הסדר לכפל במספר חיובי (invariance to multiplication): לכל $x, y, a \in \mathbb{R}$ אם $a > 0$ אז $x < y$ אמ"מ $a \cdot x < a \cdot y$.

כמו בסעיף הקודם, מעט האקסיומות שהצגנו גוררות את רוב התכונות המוכרות של יחס הסדר:

משפט 3.3.1 (תכונות הסדר של הממשיים):

1. יהיו $x, y, u, v \in \mathbb{R}$ כך ש־ $x < y$ ו־ $u < v$. אז $x + u < y + v$, ואם בנוסף x, u חיוביים אז $x \cdot u < y \cdot v$.
2. לכל $x, y \in \mathbb{R}$ מתקיים $x < y$ אמ"מ $-y < -x$.
3. אם $0 \neq x \in \mathbb{R}$, אז מתקיימת בדיוק אחת מהאפשרויות $x > 0$ או $-x > 0$.
4. אם $x, y, a \in \mathbb{R}$, ואם $a < 0$ אז $x < y$ אמ"מ $a \cdot x > a \cdot y$.
5. אם $x, y \in \mathbb{R}$ מקיימים $x < 0$ וגם $y < 0$ אז $x \cdot y > 0$.
6. אם $0 \neq x \in \mathbb{R}$ אז $x \cdot x > 0$, ובפרט מתקיים $-1 < 0 < 1$.
7. לכל $x \in \mathbb{R}$, אם $x > 0$ אז $x^{-1} > 0$ ואילו אם $x < 0$ אז $x^{-1} < 0$.
8. אם $x < y$ מספרים שווי סימן ושניהם שונים מאפס אז $\frac{1}{y} < \frac{1}{x}$.

הוכחה נוכיח את (1): מאי־רגישות הסדר לחיבור מקבלים $x+u<y+u$, ושוב מאותה סיבה מקבלים $y+u<y+v$. שילוב אי השוויונות נותן, הודות לתורשתיות הסדר, $x+u<y+v$. את המקרה הכפלי מוכיחים באופן דומה.

נוכיח את (2): אם $x<y$ נוכל לחבר $-x-y$ לשני האגפים ומאי־רגישות הסדר לחיבור מקבלים

$$-y=x+(-x-y)<y+(-x-y)=-x$$

לכן אם $x<y$ אז $-y<-x$.

כדי לסיים את ההוכחה יש להראות שאם $-y<-x$ אז $x<y$. ניתן לחזור על ההוכחה הקודמת אך אפשר גם להסתמך עליה. אנו כבר יודעים שלכל $a,b\in\mathbb{R}$ מתקיים

$$a<b \quad\Longrightarrow\quad -b<-a$$

אם נציב $a=-y$ ו־ $b=-x$ נקבל את הגרירה המיוחלת. ■

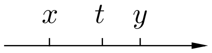
הוכחת יתר הסעיפים מושארת כתרגיל למעוניינים. תוכלו גם למצוא אותה בספרים [16][6],[10].

המשפט מאפשר לטפל באי־שוויונות לפי הכללים הרגילים המוכרים מבית הספר. למשל, נחפש תיאור מפורש של המספרים x המקיימים $\frac{x+1}{x-1}>1$. כדי להיפטר מהמכנה נרצה לכפול את הביטוי ב־$x-1$, אך אם $x-1$ שלילי הדבר ישנה את כיוון האי־שוויון (זהו סעיף (4) של המשפט). לכן נפריד למקרים. הרי $x-1>0$ אמ"מ $x>1$ (הוספנו 1 לשני האגפים והסתמכנו על אקסיומה IX) ולכן אם $x>1$ אז האי־שוויון המקורי שקול לאי־שוויון $x+1>x-1$. חיסור של x משני האגפים נותן א־שוויון שקול $1>-1$ אשר מתקיים תמיד (סעיפים (6) ו־(2) של המשפט). לכן האי־שוויון נכון לכל $x>1$. מאידך אם $x<1$ אז כפל האי־שוויון $\frac{x+1}{x-1}>1$ ב־ $x-1$ נותן אי־שוויון שקול $x+1<x-1$ וזה שקול לאי־שוויון $1<-1$, שאינו נכון. מכך אנו מסיקים ש־ $\frac{x+1}{x-1}>1$ אמ"מ $x>1$.

העובדה ש־ $1>0$ (סעיף (6) של המשפט האחרון) גוררת שיש אינסוף מספרים שונים זה מזה. ואמנם, אם נוסיף לאי־שוויון $0<1$ את המספר 1 נקבל $1<1+1$. שוב נוסיף 1 לשני האגפים ונקבל $1+1<1+1+1$, וכן הלאה. אם נמשיך להוסיף 1 נקבל שרשרת אי־שוויונות

$$1\ <\ 1+1\ <\ 1+1+1\ <\ 1+1+1+1\ <\ldots$$

אף שני מספרים בשרשרת אינם שווים, כי מבין כל שניים הימני גדול מהשמאלי (זו מסקנה של אקסיומת התורשתיות).



איור 3.3.2 t בין x ל־y

ניתן להסיק שיש אינסוף מספרים שונים גם מהטענה הבאה. לפני כן נבהיר שכאשר אומרים שמספר t נמצא **בין** שני מספרים x,y, הכוונה היא ש־t גדול מהמספר הקטן מבין x,y וקטן מהמספר הגדול מבין x,y, דהיינו ש־ $x<t<y$ אם $x<y$, וש־ $y<t<x$ אם $y<x$.

טענה 3.3.2 (צפיפות הסדר) בין כל שני מספרים קיים מספר נוסף.

הוכחה יהיו $a<b$ מספרים. עלינו להראות שיש מספר x שמקיים $a<x<b$. נבחר את x להיות נקודת האמצע בין a ל־b, כלומר, $x=\frac{a+b}{2}$ (המספר 2 לא הוגדר, אך הוא כמובן $1+1$. שימו לב ש־ $2\neq 0$ לפי הדיון אחרי משפט 3.3.1). מכיוון ש־ $a<b$ מתקיים

$$x=\frac{a+b}{2}=\frac{a}{2}+\frac{b}{2}<\frac{b}{2}+\frac{b}{2}=b(\frac{1}{2}+\frac{1}{2})=b$$

רוב השלבים ברורים למעט האי־שוויון, אותו נוכיח בפירוט למען הסדר הטוב. ראשית, $1>0$ ולכן $2=1+1>1>0$. לפי סעיפים (7),(8) של משפט 3.3.1 נובע ש־$1=\frac{1}{1}<\frac{1}{2}<0$ ומכיוון ש־ $a<b$ מאקסיומה X נובע $\frac{a}{2}<\frac{b}{2}$. אז האי־שוויון $\frac{a}{2}+\frac{b}{2}<\frac{b}{2}+\frac{b}{2}$ נובע מאקסיומה IX. באותו אופן מקבלים את האי־שוויון

$$x=\frac{a+b}{2}=\frac{a}{2}+\frac{b}{2}>\frac{a}{2}+\frac{a}{2}=a(\frac{1}{2}+\frac{1}{2})=a$$

ובסיכום $a<x<b$. ■

יתר הסעיף מוקדש לכמה מושגים שהגדרתם דורשת את יחס הסדר. נתחיל עם הערך המוחלט של מספר:

הגדרה 3.3.3 לכל $x\in\mathbb{R}$ **הערך המוחלט** של x (absolute value או modulus) מסומן על־ידי $|x|$ ומוגדר על־ידי

$$|x|=\begin{cases} x & x\geq 0\\ -x & x<0\end{cases}$$

על ידי הפרדה למקרים לפי הסימן של x רואים ש־ $|x|\geq 0$ לכל x. יתרה מזאת, לכל x מתקיימים שני האי־שוויונות $x\leq|x|$ וגם $-x\leq|x|$. אחד מהם הוא שוויון, ומתקיים שוויון בשניהם אמ"מ $x=0$ (הוכיחו טענות אלה!).

אפשר לחשוב על הערך המוחלט מבחינה גאומטרית: אם x נקודה על ציר המספרים אז הערך המוחלט של x הוא המרחק של x מהנקודה 0 (זיהינו כאן את המרחקים עם המספרים הממשיים האי־שליליים). באופן דומה, אם x,y מספרים אז אפשר לפרש את $|x-y|$ בתור אורך הקטע שקצותיו הנקודות המתאימות ל־x,y על ציר המספרים (הבהירו לעצמכם שאתם מבינים מדוע!).

המשפט הבא קל להוכחה אך חשוב ביותר.

משפט 3.3.4 (אי־שוויון המשולש (triangle inequality)) לכל $x,y\in\mathbb{R}$,

$$|x+y|\leq|x|+|y|$$

ויש שוויון אמ"מ x,y שווי סימן.

הוכחה אם $x,y \ge 0$ אז $|x| = x$ ו־ $|y| = y$. כמו־כן $x + y \ge 0$ ולכן

$$|x+y| = x+y = |x|+|y|$$

באותו אופן מראים שם $x,y \le 0$ אז מתקיים השוויון המבוקש.

אחרת x,y אינם שווי סימן. נניח בלי הגבלת הכלליות ש־ $x \ge 0 > y$. אם $x+y \ge 0$ אז

$$|x+y| = x+y = |x|-|y| < |x|+|y|$$

כאשר האי־שוויון נובע מכך ש־ $-|y| < 0 < |y|$. במקרה $x+y<0$ מטפלים בצורה דומה. ■

מסקנה 3.3.5 (אי־שוויון המשולש ההפוך) לכל $x,y \in \mathbb{R}$ מתקיים $|x-y| \ge |x|-|y|$ וגם $|x-y| \ge ||x|-|y||$.

הוכחת המסקנה מושארת כתרגיל. נעבור לנושא אחר:

הגדרה 3.3.6 תהי $A \subseteq \mathbb{R}$ קבוצה של מספרים ו־$s \in \mathbb{R}$. אז s נקרא **המקסימום** של A, ומסומן $\max A$, אם $s \in A$ ולכל $x \in A$ מתקיים $x \le s$. באופן דומה, s נקרא **המינימום** של A, ומסומן $\min A$, אם $s \in A$ ולכל $x \in A$ מתקיים $s \le x$.

יש לוודא שלקבוצה יש לכל היותר מקסימום ומינימום אחד, אחרת אין הצדקה לדבר על המקסימום והמינימום של קבוצה בהא הידיעה. לא קשה להוכיח זאת: נניח ש־s,s' שניהם מקסימום של A. אז $s,s' \in A$, מכיון ש־s מקסימום של A מתקיים $s' \le s$, ומכיוון ש־s' מקסימום מתקיים $s \le s'$. לכן $s = s'$, כפי שרצינו. ההוכחה לגבי מינימום דומה.

נציין שלא לכל קבוצת מספרים יש מינימום. למשל, לקבוצה $\mathbb{R}$ אין מינימום, כי אם $x \in \mathbb{R}$ אז $x-1 \in \mathbb{R}$, וכיוון ש־ $x-1 < x$ הרי ש־x אינו מינימום של $\mathbb{R}$. נשוב ונדון במושג זה שוב בסעיף 3.5.

נעבור להגדרה של משפחה חשובה של תת־קבוצות של הישר: הקטעים. כפי שהשם מרמז, הפירוש הגאומטרי של קבוצות אלה הם קטעים בישר הממשי, עם או בלי נקודות הקצה.

איור 3.3.3 קטע שקצותיו a ו־b

הגדרה 3.3.7 יהיו $a \le b$ מספרים ממשיים. **קטע** (interval) הוא קבוצה מאחת הצורות הבאות:

$$\begin{aligned} [a,b] &= \{x \in \mathbb{R} : a \le x \le b\} \\ (a,b) &= \{x \in \mathbb{R} : a < x < b\} \\ [a,b) &= \{x \in \mathbb{R} : a \le x < b\} \\ (a,b] &= \{x \in \mathbb{R} : a < x \le b\} \end{aligned}$$

בכל המקרים a נקרא **הקצה שמאלי** של הקטע ו־b נקרא **הקצה הימני**. הקטע $[a,b]$ נקרא קטע **סגור** (closed), הקטע (a,b) נקרא קטע **פתוח** (open),[5] והקטעים $[a,b)$ ו־$(a,b]$ נקראים קטעים **חצי פתוחים** (half-open) מימין ומשמאל בהתאמה.[6] קטע נקרא **מנוון** אם נקודות הקצה שלו שוות זו לזו.

כמו כן, **קרן** (ray) היא קבוצה מאחת הצורות הבאות:

$$\begin{aligned}
(-\infty,a] &= \{x\in\mathbb{R} : x\leq a\}\\
[a,\infty) &= \{x\in\mathbb{R} : x\geq a\}\\
(-\infty,a) &= \{x\in\mathbb{R} : x< a\}\\
(a,\infty) &= \{x\in\mathbb{R} : x> a\}\\
(-\infty,\infty) &= \mathbb{R}
\end{aligned}$$

הנקודה a נקראת **הקצה** של הקרן. שתי הקרניים הראשונות הן סגורות ב־a, והשתיים שאחריהן פתוחות ב־a.

לפעמים קטע נקרא **קטע חסום** וקרן נקראת **קטע לא חסום**.[7]

הערות

1. אם I קטע אז קל לראות מההגדרה שלכל $x,y\in I$, הקטע I מכיל גם את כל הנקודות בין x ל־y.
2. קטע מנוון מכיל נקודה אחת (אם מדובר בקטע מהצורה $[a,a]$) או שהוא קבוצה ריקה (זה המקרה לקטעים מהצורה (a,a), $[a,a)$ או $(a,a]$).
3. הסימן "∞" נקרא **אינסוף** והסימן "$-\infty$" נקרא **מינוס אינסוף**. אלו סימנים שהומצאו כדי לסמן קרניים בצורה דומה לאופן בו מסמנים קטעים, ויהיו להם גם שימושים נוספים בהמשך. חשוב לזכור: **אינסוף ומינוס אינסוף אינם מספרים**. לא ניתן לבצע בהם פעולות חשבוניות או פעולות השוואה, גם אם נדמה שיש הצדקה לרשום, למשל, "$\infty+1=\infty$" או "$0<\infty$" .

הפנים (interior) של קטע היא קבוצת הנקודות בקטע שאינן נקודות קצה. **האורך** (length) של קטע חסום I שקצותיו $a,b\in\mathbb{R}$ הוא $|b-a|$ ומסומן על ידי $|I|$.[8]

קל לבדוק שהפנים של קטע או קרן הוא קטע פתוח או קרן פתוחה בהתאמה, והפנים של קטע הוא ריק רק אם הקטע המקורי הוא מנוון.

---

[5] הסימון (a,b) זהה לסימון של זוג סדור אך תמיד יהיה אפשר להבדיל ביניהם לפי ההקשר.

[6] ישנם ספרים בהם קטע פתוח מסומן $]a,b[$, קטע חצי פתוח מימין מסומן $[a,b[$ וקטע חצי פתוח משמאל מסומן $]a,b]$.

[7] לפעמים קטע חסום נקרא **קטע סופי**, וקטע לא חסום נקרא **קטע אינסופי**. הסיבה לכך היא שלקטע סופי יש אורך סופי, ואילו לקטע אינסופי יש "אורך אינסופי".

[8] הסימון $|\cdot|$ מציין כעת הן ערך מוחלט של מספר והן אורך קטע, ובהמשך תהיה לו גם משמעות נוספת. הכוונה תהיה תמיד ברורה מההקשר ואין סיבה שתיווצר דו־משמעות.

למשל, האורך של הקטעים $[0,1],[0,1),(0,1],(0,1)$ הוא 1 והפנים של כל הקטעים האלה הוא $(0,1)$. האורך של קטע הוא חיובי אמ"מ הקטע אינו מנוון. אם I קטע ריק או בעל נקודה אחת האורך שלו אפס. בכל מקרה אחר האורך שלו חיובי ממש. הוכיחו טענות אלה!

תרגילים

1. ציירו את הקבוצות הבאות על ציר המספרים וכתבו אותם כאיחודים של קטעים:
   (א) $A = \{x \in \mathbb{R} : x \cdot x < 1\}$.
   (ב) $B = \{x \in \mathbb{R} : (x-1)(x+2) < 0\}$.
   (ג) $C = \{x \in \mathbb{R} : (x \neq 0) \wedge \frac{x^2-4}{x} > 0\}$.
   (ד) $D = \{x \in \mathbb{R} : x \neq 2 \wedge \frac{x(x+1)}{x-2} < 0\}$.
   (ה) $E = \{x \in \mathbb{R} : x \neq -1 \wedge x \neq 2 \wedge ((x+1)(x-2))^{-1} < 0\}$.
   (ו) $F = \{x \in \mathbb{R} : x \neq 0 \wedge x \neq -1 \wedge \frac{1}{x} - \frac{1}{x+1} > 0\}$.
2. הוכיחו את התכונות הבאות של האי־שוויון החלש (יש להסתמך רק על תכונות הסדר והגדרת היחס $\leq$):
   (א) (השוואה) לכל $x, y \in \mathbb{R}$ מתקיים לפחות אחד האי־שוויונות $x \leq y$ או $y \leq x$, ומתקיים $x = y$ אמ"מ שני האי־שוויונות מתקיימים.
   (ב) (תורשתיות) אם $x \leq y$ ו־ $y \leq z$ אז $x \leq z$. יתרה מזאת יתקיים $x < z$ אמ"מ לפחות אחד מהאי־שוויונות המקוריים הוא אי־שוויון חזק.
   (ג) (אי־רגישות לחיבור) אם $x \leq y$ ו־ $a \in \mathbb{R}$ אז $x + a \leq y + a$.
   (ד) (אי־רגישות לכפל בחיובי) אם $x \leq y$ ו־ $0 \leq a$ אז $ax \leq ay$.
3. השלימו את הוכחת משפט 3.3.1.
4. יהי $x < 0$. הוכיחו כי $\frac{x}{2} > x$.
5. בשאלה זו נסמן $z^2 = z \cdot z$. הוכיחו את הטענות הבאות:
   (א) לכל $x > 0$ מתקיים $x + \frac{1}{x} \geq 2$, ויש שוויון אמ"מ $x = 1$.
   (ב) לכל $x, y \in \mathbb{R}$ שאינם שניהם אפס מתקיים $x^2 + xy + y^2 > 0$.
   (ג) לכל $x, y \in \mathbb{R}$ מתקיים $xy \leq \frac{1}{2}(x^2 + y^2)$, ויש שוויון אמ"מ $x = y$.
   (ד) יהיו a, b, u, v מספרים המקיימים $a^2 + b^2 = u^2 + v^2 = 1$. נניח גם ש־ $au + bv = 1$. הוכיחו כי $a = u\,,\; b = v$.
6. כתבו את הביטויים הבאים בלי סימן הערך המוחלט, על ידי הפרדה למקרים. למשל,

$$b + |a| = \begin{cases} b + a & a \geq 0 \\ b - a & a < 0 \end{cases}$$

(א) $|a-1|+|b-2|$.

(ב) $|a+b|+|a|$.

(ג) $|a+2b|\cdot|a+b|$.

7. הוכיחו את התכונות הבאות של הערך המוחלט:

(א) $|x|=|-x|$

(ב) $|x\cdot y|=|x|\cdot|y|$

(ג) $|\frac{1}{x}|=\frac{1}{|x|}$

8. הוכיחו שלכל $a,b\in\mathbb{R}$ מתקיימים השוויונות $\max\{a,b\}=\frac{1}{2}(a+b+|a-b|)$ ו־ $\min\{a,b\}=\frac{1}{2}(a+b-|a-b|)$.

9. הוכיחו שהאי־שוויונות הבאים מתקיימים לכל $x,y\in\mathbb{R}$ (זו מסקנה 3.3.5). מתי יש שוויון?

(א) $|x-y|\geq|x|-|y|$.

(ב) $|x-y|\geq|\,(\,|x|-|y|\,)\,|$.

10. ציירו את הקבוצות הבאות על ציר המספרים וכתבו אותם כאיחודים של קטעים:

(א) $A=\{x\in R\,:\,|x-2|<1\}$.

(ב) $B=\{x\in\mathbb{R}\,:\,|x+3|+x<1\}$

(ג) $C=\{x\in\mathbb{R}\,:\,|x-1|+|x-3|\leq1\}$

(ד) $D=\{x\in\mathbb{R}\,:\,|x-1|/|x-2|>1/2\}$

11. נניח ש־ $0\geq x$ ולכל מספר חיובי ε מתקיים $x\leq\varepsilon$. הוכיחו ש־ $x=0$. כמו־כן, אם $y\in\mathbb{R}$ ולכל מספר חיובי ε מתקיים $|y|\leq\varepsilon$ אז $y=0$.

12. הראו שבהנחה של־ $A,B\subseteq\mathbb{R}$ יש מקסימום מתקיים

$$\max A\cup B=\max\{\max A,\max B\}$$

נסחו והוכיחו טענה דומה לגבי מינימום. נניח של־A יש מקסימום ול־B יש מינימום. האם אפשר לומר משהו על קיומם של מקסימום או מינימום לקבוצה $A\cup B$ או לקבוצה $A\cap B$?

13. תהי A קבוצה ויהי s המקסימום של A. נגדיר $B=\{-a\,:\,a\in A\}$. הוכיחו ש־$-s$ הוא מינימום של B.

14. הוכיחו שהחיתוך של קטעים הוא קטע.

15. הוכיחו שלקטע יש איבר מקסימלי אמ"מ זהו קטע חסום וסגור מימין. נסחו והוכיחו טענה דומה לגבי קיום מינימום של קטע.

16. הוכיחו שהפנים של קטע לא מנוון הוא קטע פתוח לא מנוון.

17. יהיו $x<y$ מספרים ממשיים.

(א) יהי $0 \le \alpha \le 1$. הוכיחו שהמספר $z = \alpha x + (1-\alpha)y$ מקיים $x \le z \le y$. מתי מתקיימים אי־שוויונות חזקים?

(ב) הראו שכל מספר $z \in [x, y]$ ניתן לכתיבה כמו בסעיף (א), ושהמספר α בכתיבה זו הוא יחיד.

18. קל לבדוק מההגדרות וממשפט 3.3.1 שלכל $x \neq 0$ מתקיימת בדיוק אחת האפשרויות, $x \in \mathbb{R}^+$ או $-x \in \mathbb{R}^+$, ושאם $a, b \in \mathbb{R}^+$ אז $a + b, a \cdot b \in \mathbb{R}^+$. בשאלה זו נראה שבעזרת קבוצת מספרים בעלת התכונות האלה אפשר להגדיר יחס סדר המקיים את אקסיומות הסדר. מכאן שקיומו של יחס סדר שקול לקיום קבוצה עם תכונות אלה.

נשכח לעת עתה את יחס הסדר ונחשוב על $\mathbb{R}$ כפי שהייתה בסוף הסעיף הקודם, דהיינו יש בה חיבור וכפל המקיימים את אקסיומות I־VI. נניח שנתונה קבוצה $P \subseteq \mathbb{R}$ המקיימת (א) כל $x \in \mathbb{R}$ מקיים בדיוק אחת מהאפשרויות $x = 0$, $x \in P$ או $-x \in P$, (ב) אם $x, y \in P$ אז גם $x + y, x \cdot y \in P$. כעת נגדיר את היחס $x < y$ בין מספרים ממשיים על ידי

$$x < y \quad \Longleftrightarrow \quad y - x \in P$$

הוכיחו שהיחס הזה מקיים את אקסיומות הסדר.

3.4 המספרים הטבעיים, השלמים והרציונליים

בסעיף זה נגדיר כמה תת־מערכות חשובות של $\mathbb{R}$. הפשוטה ביותר היא מערכת המספרים הטבעיים, שאיבריה הם

$$0\,,\;1\,,\;1+1\,,\;1+1+1\,,\;\ldots$$

וכן הלאה, כלומר כל אותם המספרים המתקבלים מ־0 על ידי הוספה חוזרת של 1. הגדרה זו ברורה מבחינה אינטואיטיבית אך אינה הגדרה פורמלית. במקומה נגדיר את קבוצת המספרים הטבעיים להיות הקבוצה הקטנה ביותר המכילה את 0 ובעלת התכונה שאם n בקבוצה אז $n + 1$ בקבוצה. ניעזר בהגדרה הבאה:

הגדרה 3.4.1 קבוצה $I \subseteq \mathbb{R}$ של מספרים נקראת ***אינדוקטיבית*** (inductive) אם $0 \in I$, ובנוסף ל־I יש התכונה שאם $x \in I$ אז $x + 1 \in I$.

ברור מבחינה אינטואיטיבית שהקבוצה $\{0, 1, 1+1, \ldots\}$ שאותה אנו רוצים להגדיר היא אינדוקטיבית, אבל יש קבוצות אינדוקטיביות אחרות. ישנה למשל הקבוצה $\mathbb{R}$ עצמה, וקבוצות כגון $\{\ldots, -1, 0, 1, 2, \ldots\}$ ו־ $\{0, \frac{1}{2}, 1, 1\frac{1}{2}, 2, 2\frac{1}{2}, 3, \ldots\}$. מה שמיוחד בקבוצה $\{0, 1, 1+1, \ldots\}$ הוא שכל קבוצה אינדוקטיבית אחרת מכילה

אותה, כי אם I אינדוקטיבית אז $0 \in I$, ולכן $1 = 1 + 0 \in I$, ולכן $1 + 1 \in I$, ולכן $1 + 1 + 1 \in I$, וכן הלאה. אנו נבחר תכונה זו כבסיס להגדרה של המספרים הטבעיים:

הגדרה 3.4.2 קבוצת **המספרים הטבעיים** (natural numbers) מסומנת ב־$\mathbb{N}$ והיא מורכבת מהמספרים אשר שייכים לכל קבוצה אינדוקטיבית.[9]

הערות

1. שימו לב שמההגדרה נובע ש־ $I \supseteq \mathbb{N}$ לכל קבוצה אינדוקטיבית I.
2. נשסמן בדרך הרגילה את עשרת המספרים הטבעיים הראשונים: הספרה 2 מייצגת את המספר הטבעי $1 + 1$, הספרה 3 את $1 + 1 + 1$, וכן הלאה עד הספרה 9. בסעיף 3.7 להלן נראה שהשיטה העשרונית מאפשרת לייצג באופן יחיד כל מספר טבעי.
3. דרך שקולה לתאר את $\mathbb{N}$ הוא בתור חיתוך כל הקבוצות האינדוקטיביות, דהיינו $\mathbb{N} = \cap_{I \in \mathcal{U}} I$, כאשר $\mathcal{U} = \{I \subseteq \mathbb{R} : \text{ אינדוקטיבית } I\}$ היא קבוצת כל התת־קבוצות האינדוקטיביות של $\mathbb{R}$.

מתקבל על הדעת ש־$\mathbb{N}$ היא בעצמה קבוצה אינדוקטיבית. הנה הוכחה לכך:

למה 3.4.3 $\mathbb{N}$ היא קבוצה אינדוקטיבית.

הוכחה יש לבדוק ש־$\mathbb{N}$ מקיימת את התנאים שבהגדרה של קבוצה אינדוקטיבית.

נראה ש־ $0 \in \mathbb{N}$. ואמנם אם I קבוצה אינדוקטיבית אז $0 \in I$ (לפי ההגדרה של קבוצה אינדוקטיבית), כלומר 0 שייך לכל קבוצה אינדוקטיבית. לכן לפי ההגדרה $0 \in \mathbb{N}$.

נבדוק את הסגירות להוספת 1: נניח $x \in \mathbb{N}$. יש להראות $x + 1 \in \mathbb{N}$. לשם כך יש להראות שלכל קבוצה אינדוקטיבית I מתקיים $x + 1 \in I$. תהי אם כן I קבוצה אינדוקטיבית. מכיוון $x \in \mathbb{N}$ אנו מסיקים ש־ $x \in I$. מכיוון ש־I אינדוקטיבית, $x + 1 \in I$, וזה מה שרצינו להראות. ∎

אנו מסיקים מהלמה ש־$\mathbb{N}$ היא באמת הקבוצה האינדוקטיבית הקטנה ביותר, כי $\mathbb{N}$ היא קבוצה אינדוקטיבית ולפי הגדרתה היא מוכלת בכל קבוצה אינדוקטיבית אחרת.

אחת התכונות החשובות ביותר של המספרים הטבעיים היא עקרון האינדוקציה. בספרים רבים מקבלים תכונה זו כאקסיומה, אך אנו נוכל להוכיח מההגדרה שנתנו:

[9] יש ספרים בהם המספרים הטבעיים מוגדרים להיות הקבוצה $\{1, 2, 3 \ldots\}$ (כלומר, 0 אינו נחשב טבעי).

משפט 3.4.4 (עקרון האינדוקציה) תהי $I \subseteq \mathbb{N}$ קבוצה המכילה את 0 וכך שאם $n \in I$ אז $n+1 \in I$. אז $I = \mathbb{N}$.

הוכחה נתון שהקבוצה I אינדוקטיבית, ולכן $\mathbb{N} \subseteq I$. כמו־כן נתון ש־ $I \subseteq \mathbb{N}$. מצירוף שתי ההכלות האלה נובע $I = \mathbb{N}$, כנדרש. ■

הגדרנו מיהם המספרים הטבעיים, אך לא בררנו עדיין מה התכונות שהם. המשפט הבא מרכז את תכונותיהם העיקריות.

משפט 3.4.5 סכום ומכפלה של מספרים טבעיים הוא מספר טבעי. כל מספר טבעי הוא אי־שלילי, ואם $m \leq n$ מספרים טבעיים אז $n - m$ טבעי. אם n טבעי אז אין אף מספר טבעי בין n ל־$n+1$.

הוכחה נוכיח לדוגמה את הטענה שסכום של טבעיים הוא טבעי. את יתר הסעיפים אנו משאירים כתרגיל. די שנראה שלכל n טבעי הקבוצה

$$I = \{m \in \mathbb{N} \,:\, n + m \in \mathbb{N}\}$$

שווה ל־$\mathbb{N}$ (וודאו שאתם מבינים מדוע זה מספיק!). נקבע אם כן $n \in \mathbb{N}$ ותהי I הקבוצה האמורה. ברור מההגדרה ש־ $I \subseteq \mathbb{N}$ ולכן לפי עקרון האינדוקציה די להראות ש־I אינדוקטיבית. ואמנם, מתקיים $n + 0 = n \in \mathbb{N}$ ולכן $0 \in I$. נניח כעת ש־ $m \in I$, ונראה ש־ $m + 1 \in I$. מההנחה $m \in I$ נובע ש־ $n + m \in \mathbb{N}$. לכן

$$n + (m + 1) = (n + m) + 1 \in \mathbb{N}$$

כי $n + m$ טבעי ולכן $(n + m) + 1$ טבעי (הרי $\mathbb{N}$ קבוצה אינדוקטיבית). בכך סיימנו את ההוכחה. ■

מהמשפט נובע, למשל, ש־-1 אינו מספר טבעי, כי $-1 < 0$ בעוד ש־ $0 \geq n$ לכל n טבעי. באופן דומה נסיק ש־$\frac{1}{2}$ אינו מספר טבעי, כי מאחר ו־ $1 < 2$ מתקיים $0 < \frac{1}{2} < \frac{1}{1} = 1$, בסתירה לחלק האחרון של המשפט.

כפי שראינו, הנגדי של מספר טבעי אינו בהכרח מספר טבעי. אם מוסיפים לטבעיים את הנגדיים שלהם מקבלים את מערכת המספרים השלמים:

הגדרה 3.4.6 מספר טבעי או הנגדי של מספר טבעי נקרא **מספר שלם** (integer). קבוצת המספרים השלמים מסומנת ב־$\mathbb{Z}$,[10] דהיינו

$$\mathbb{Z} = \{\pm n \,:\, n \in \mathbb{N}\} = \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}$$

[10] האות $\mathbb{Z}$ באה מהמילה הגרמנית Zahlen, שפירושה "מספרים".

כל מספר שלם הוא הפרש של מספרים טבעיים, כי אם $n \in \mathbb{N}$ אז $n = n - 0$ ואילו $-n = 0 - n$, ובכך כיסינו את כל השלמים. גם ההפך הוא נכון, כלומר, הפרש של מספרים טבעיים הוא שלם. ואמנם, יהיו m, n טבעיים ויהי $k = m - n$. כדי לראות ש־k שלם נפריד לשני מקרים. אם $m \geq n$ אז לפי משפט 3.4.5 המספר $m - n$ הוא טבעי, כלומר k הוא טבעי ובפרט הוא שלם. מאידך אם $m < n$ אז $n - m$ טבעי, ולכן $k = -(n - m)$ הוא שלם. ראינו אם כן ש־

$$\mathbb{Z} = \{m - n \,:\, m, n \in \mathbb{N}\}$$

משפט 3.4.7 סכום, מכפלה והפרש של מספרים שלמים הוא מספר שלם, ואם n שלם אז אין מספרים שלמים בין n ל־$n + 1$.

הוכחה נוכיח למשל שסכום של מספרים שלמים הוא שלם. יהיו $m, n \in \mathbb{Z}$. אז יש מספרים טבעיים i, j, i', j' כך ש־ $m = i - j$ ו־ $n = i' - j'$. לכן

$$m + n = i - j + (i' - j') = (i + i') - (j + j')$$

מכיוון ש־ $i + i', j + j' \in \mathbb{N}$, הצלחנו להציג את $m + n$ כהפרש של מספרים טבעיים, ולכן $m + n \in \mathbb{Z}$. הוכחת התכונות האחרות מושארת כתרגיל. ■

מסקנה 3.4.8 הסכום של מספר שלם n ומספר ממשי x הוא מספר שלם אמ"מ x שלם, והוא מספר טבעי אמ"מ x שלם ומקיים $x \geq -n$.

הוכחה יהי $n \in \mathbb{Z}$ ו־ $x \in \mathbb{R}$. אם x שלם אז לפי המשפט הקודם, $n + x$ שלם. להפך, אם $n + x$ שלם אז $x = (n + x) - n$ הוא הפרש של שלמים ולכן שלם. את הוכחת התכונות הנותרות אנו משאירים כתרגיל. ■

המערכת $\mathbb{Z}$ עדיין אינה סגורה לאיברים הפכיים: למשל, $2 \in \mathbb{Z}$ אבל $\frac{1}{2} \notin \mathbb{Z}$, כי $\frac{1}{2}$ אינו טבעי וגם $-\frac{1}{2}$ אינו טבעי. המערכת האחרונה שנגדיר היא המערכת הקטנה ביותר שסגורה לחבור וכפל ומכילה את השלמים וההפכיים שלהם:

הגדרה 3.4.9 מספר נקרא **רציונלי** (rational) אם הוא מנה של מספרים שלמים. קבוצת המספרים הרציונליים מסומנת ב־$\mathbb{Q}$,[11] דהיינו,

$$\mathbb{Q} = \{\frac{m}{n} \,:\, n, m \in \mathbb{Z}\,,\, n \neq 0\}$$

[11] האות $\mathbb{Q}$ באה מהמילה quitient, שפירושה מנה. המילה "רציונלי" באה מהמילה ratio, שפירושה יחס.

אם m,n שלמים אז המנה $r=\frac{m}{n}$ נקראת **שבר פשוט**, המספר m נקרא **המונה** ו־n **המכנה**. יש דרכים רבות לרשום את r כשבר פשוט: אם $r=\frac{m}{n}$ היא הצגה אחת כזו אז ניתן לרשום את r גם בתור $r=\frac{2m}{2n}$, ובאופן כללי אם $a\neq 0$ שלם אז $r=\frac{am}{an}$. כיוון ש־a יכול להיות חיובי או שלילי אנו מסיקים שלכל מספר רציונלי r יש הצגות כשבר פשוט עם מכנה חיובי, וגם הצגות כשבר פשוט עם מכנה שלילי.

משפט 3.4.10 אם $r,s\in\mathbb{Q}$ אז $r+s, r\cdot s, -r\in\mathbb{Q}$ ואם $r\neq 0$ אז $r^{-1}\in\mathbb{Q}$

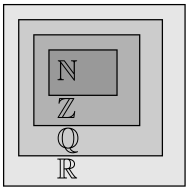
הוכחה נוכיח למשל שאם $0\neq r\in\mathbb{Q}$ אז $r^{-1}\in\mathbb{Q}$. אם $r=\frac{i}{j}\neq 0$ $(i,j\in\mathbb{Z})$ אז בהכרח $i\neq 0$ ולכן $\frac{j}{i}$ מוגדר. אז מתקיים

$$r\cdot\frac{j}{i}=\frac{i}{j}\cdot\frac{j}{i}=\frac{ij}{ji}=1$$

ולכן $r^{-1}=\frac{j}{i}\in\mathbb{Q}$, כנדרש (השתמשנו כאן בתכונות ממשפט 3.2.1). יתר התכונות מושארות כתרגיל. ■

מערכות המספרים שפגשנו עד כה מקיימות את יחסי ההכלה

$$\mathbb{N}\subseteq\mathbb{Z}\subseteq\mathbb{Q}\subseteq\mathbb{R}$$



איור 3.4.1 יחסי ההכלה בין מערכות המספרים

באיור 3.4.1 מופיעה דיאגרמת ון המתארת הכלות אלה. ציירנו את התרשים כך שמתקבל הרושם שכל מערכת בשרשרת $\mathbb{N}\subseteq\mathbb{Z}\subseteq\mathbb{Q}\subseteq\mathbb{R}$ מכילה ממש את קודמתה, כלומר, שאף אחת מההכלות בין הקבוצות אינה שוויון. ואמנם, מתקיים $-1\in\mathbb{Z}$ ולכן $\mathbb{Z}\neq\mathbb{N}$, וכן $\frac{1}{2}\in\mathbb{Q}$ ולכן $\mathbb{Q}\neq\mathbb{Z}$. אך האם $\mathbb{Q}\neq\mathbb{R}$? הטענה הבאה שופכת אור על השאלה הזו:

טענה 3.4.11 המערכת $\mathbb{Q}$ מקיימת את אקסיומות השדה הסדור.

הוכחה המשפט האחרון מבטיח ש־$\mathbb{Q}$ היא תת־מערכת של $\mathbb{R}$ במובן שפעולות החיבור והכפל אינן מוציאות אותנו אל מחוץ ל־$\mathbb{Q}$. אקסיומות החילוף, הקיבוץ והפילוג חלות ב־$\mathbb{Q}$ באופן אוטומטי, כי המספרים הרציונליים הם בפרט ממשיים. המספרים $0,1$ שייכים ל־$\mathbb{Q}$, ואם $r\in\mathbb{Q}$ אז גם $-r\in\mathbb{Q}$ וגם $r^{-1}\in\mathbb{Q}$ בתנאי ש־ $r\neq 0$, כלומר, אקסיומות הנגדי וההפכי מתקיימות ב־$\mathbb{Q}$. אקסיומות הסדר מתקיימות ב־$\mathbb{Q}$ כי הן מתקיימות ב־$\mathbb{R}$. ■

מהטענה נובע שעל סמך האקסיומות הנתונות בשלב זה לא ניתן להוכיח שקיימים מספרים ממשיים שאינם רציונליים. ואמנם, אילו על סמך אקסיומות השדה הסדור יכולנו להוכיח את קיומם של מספרים שאינם שייכים ל־$\mathbb{Q}$, הרי שההוכחה, והמסקנה שלה, היו תקפים בכל מערכת המקיימת אקסיומות אלה, ובפרט ב־$\mathbb{Q}$ עצמה. כלומר, היה נובע שב־$\mathbb{Q}$ יש איבר שאינו איבר של $\mathbb{Q}$, וזה כמובן לא ייתכן, כי $\mathbb{Q}\setminus\mathbb{Q}$ היא הקבוצה הריקה.

אם כן, בשלב זה אין אנו יכולים להוכיח ש־ $\mathbb{Q}\neq\mathbb{R}$. נוכל להוכיח זאת בפרק הבא, לאחר שנוסיף עוד אקסיומה לאקסיומות של $\mathbb{R}$.

תרגילים

1. תהי $I \subseteq \mathbb{N}$ קבוצה בעלת התכונה שאם $n \in I$ אז $n+1 \in I$. האם $I = \mathbb{N}$?
2. השלימו את החלקים החסרים בהוכחות של משפטים 3.4.5, 3.4.7 ו־3.4.10.
3. נאמר שקבוצה $I \subseteq \mathbb{R}$ היא סגורה לפעולת הנגדי אם לכל $n \in I$ מתקיים $-n \in I$. הוכיחו ש־$\mathbb{Z}$ היא הקבוצה הקטנה ביותר שהיא גם אינדוקטיבית וגם סגורה לנגדי, דהיינו הוכיחו ש־ $\mathbb{Z} = \cap_{I \in \mathcal{V}} I$, כאשר

$$\mathcal{V} = \{I \subseteq \mathbb{R} : \text{ אינדוקטיבית וסגורה לנגדי } I\}$$

(רמז: ההוכחה קלה אם מתבססים על הטענות שהוכחו בסעיף זה. הקושי העיקרי כאן הוא להבין מה בדיוק הטענה שאתם מנסים להוכיח. לפני שתתחילו, נסחו במדויק מהן שתי הקבוצות שאתם מעוניינים להוכיח שוויון ביניהן).
4. יהי $x \in \mathbb{R}$. הראו שקיימת קבוצה אינדוקטיבית קטנה ביותר המכילה את x, ותארו אותה.
5. נסחו במדויק והוכיחו את הטענה הבאה: $\mathbb{Q}$ היא המערכת האינדוקטיבית הקטנה ביותר שהיא סגורה לנגדי וסגורה להפכי.

3.5 עקרון האינדוקציה ושימושיו

עיקרון האינדוקציה, אותו פגשנו בסעיף הקודם, לובש צורות רבות.

עקרון ההוכחה באינדוקציה

טענה 3.5.1 (עקרון ההוכחה באינדוקציה) תהי $P(n)$ טענה התלויה ב־n. על מנת ש־$P(n)$ תתקיים לכל n טבעי מספיק שיתקיימו שני התנאים הבאים:

1. **בסיס האינדוקציה:** $P(0)$ טענה נכונה.
2. **שלב האינדוקציה:** אם $P(n)$ טענה נכונה אז $P(n+1)$ טענה נכונה.

הוכחה (של עיקרון ההוכחה באינדוקציה) תהי I קבוצת המספרים הטבעיים n שעבורם $P(n)$ מתקיימת. אז $I \subseteq \mathbb{N}$. כמו־כן על סמך בסיס האינדוקציה $0 \in I$, ועל סמך שלב האינדוקציה, אם $n \in I$ אז $n+1 \in I$. לכן לפי עיקרון האינדוקציה, $I = \mathbb{N}$, וזה אומר בדיוק ש־$P(n)$ נכונה לכל מספר טבעי n. ∎

האינטואיציה העומדת מאחורי עקרון ההוכחה באינדוקציה הוא כדלקמן. מבסיס האינדוקציה ידוע ש־$P(0)$ נכונה. נפעיל את שלב האינדוקציה עם $n=0$, ונסיק מ־$P(0)$ ש־$P(1)$ טענה נכונה. נפעיל שוב את שלב האינדוקציה עם $n=1$ ונקבל

מ־$P(1)$ את ש־$P(2)$ נכונה. נפעיל את שלב האינדוקציה שוב ונקבל ש־$P(3)$ נכונה, וכן הלאה. אינטואיטיבית ברור שלכל n טבעי שרשרת המסקנות תגיע בשלב כלשהו ל־n ונקבל ש־$P(n)$ נכונה.

לעתים קרובות מתעורר צורך להוכיח שטענה $P(n)$ מתקיימת לכל מספר טבעי החל ממספר טבעי N נתון. לשם כך מספיק להראות ש־$P(N)$ מתקיימת ושלכל $n \geq N$, אם $P(n)$ מתקיימת אז $P(n+1)$ מתקיימת. הסיבה היא שתנאים אלה מבטיחים, על פי עקרון ההוכחה באינדוקציה, שהטענה $Q(n) = P(n+N)$ מתקיימת לכל $n \in \mathbb{N}$, ומכאן ש־ $P(n)$ מתקיימת לכל $n \geq N$ טבעי.

בשימושים של עקרון ההוכחה באינדוקציה בדרך כלל מראים ששלב האינדוקציה מתקיים בעזרת נימוק בעל הצורה "יהי $n \in \mathbb{N}$. נניח ש־$P(n)$ נכונה. אז ... [פרטי ההוכחה] ... ולכן $P(n+1)$ נכונה". בתוך פרטי ההוכחה בדרך כלל משתמשים בהנחה ש־$P(n)$ נכונה, הנקראת **הנחת האינדוקציה**.

דוגמאות

1. לכל מספר n, נאמר ש־n **מספר זוגי** אם קיים $n' \in \mathbb{N}$ כך ש־ $n = 2n'$, ונאמר שהוא **אי זוגי** אם קיים $n' \in \mathbb{N}$ כך ש־ $n = 2n' + 1$. נוכיח כעת שכל מספר $n \in \mathbb{N}$ מקיים אחת ובדיוק אחת משתי האפשרויות האלה.

   ראשית, נוכיח שכל מספר טבעי הוא או זוגי או אי־זוגי. לשם כך נוכיח באינדוקציה את שהטענה

   $$P(n) = \text{" או ש־ } n \text{ זוגי, או ש־ } n \text{ אי־זוגי "}$$

   מתקיימת לכל n טבעי.

   ואמנם, בסיס האינדוקציה $P(0)$ מתקיים כי $0 = 2 \cdot 0$, ולכן 0 זוגי.

   עלינו להראות את שלב האינדוקציה, כלומר את הטענה $P(n) \implies P(n+1)$. ביתר פירוט, עלינו להראות שאם n הוא זוגי או אי־זוגי, אז $n+1$ הוא זוגי או אי־זוגי. נפריד למקרים:

   - אם n זוגי אז קיים $k \in \mathbb{N}$ כך ש־ $n = 2k$. אז $n+1 = 2k+1$, ולכן $n+1$ אי־זוגי.
   - אם n אי־זוגי אז קיים $k \in \mathbb{N}$ כך ש־ $n = 2k+1$. אז מתקיים $n+1 = (2k+1)+1 = 2(k+1)$, ומאחר ו־ $k+1 \in \mathbb{N}$ קיבלנו ש־ $n+1$ זוגי.

   בכל מקרה קיבלנו ש־ $n+1$ הוא או זוגי או אי־זוגי, כפי שרצינו.

   נותר להראות שאף מספר אינו גם זוגי וגם אי־זוגי. נניח בשלילה שקיים מספר טבעי n שהוא גם זוגי וגם אי־זוגי. אז קיימים מספרים $k, m \in \mathbb{N}$ כך

ש־ $n = 2k = 2m + 1$. מכאן נובע ש־ $2k > 2m$, ולכן $k > m$. נעביר אגפים בשוויון למעלה

$$2(k - m) = 1$$

ומכיוון ש־ $k > m$ קיבלנו ש־ $k - m$ מספר טבעי. אבל חלוקה ב־2 של השוויון $2(k - m) = 1$ נותן ש־ $k - m = \frac{1}{2}$, בסתירה לכך ש־ $k - m$ טבעי. מכאן שאין מספר טבעי שהוא גם זוגי וגם אי־זוגי.

2. לכל $n \geq 1$ טבעי יהי S_n סכום n המספרים הטבעיים החיוביים הראשונים, כלומר

$$S_n = 1 + 2 + \ldots + n$$

(למשל $S_1 = 1\,,\ S_2 = 1 + 2\,,\ S_3 = 1 + 2 + 3$ וכן הלאה). נראה שבאופן כללי מתקיים $S_n = \frac{1}{2}n(n + 1)$. נוכיח זאת באינדוקציה על n. ליתר דיוק, נראה שהנוסחה מתקיימת עבור $n = 1$ ושאם היא מתקיימת ל־n אז היא מתקיימת ל־$n + 1$, ונסיק שהיא מתקיימת לעל $n \geq 1$ טבעי.

עבור $n = 1$ הטענה נכונה, כי $S_1 = 1 = \frac{1}{2}1(1 + 1)$.

נניח שהטענה נכונה ל־n, כלומר נניח ש־ $S_n = \frac{1}{2}n(n+1)$, ונראה שנובע מכך ש־ $S_{n+1} = \frac{1}{2}(n + 1)(n + 2)$. ואמנם,

$$\begin{aligned} S_{n+1} &= S_n + (n + 1) \\ &= \frac{n(n + 1)}{2} + (n + 1) \\ &= \frac{(n + 1)(n + 2)}{2} \end{aligned}$$

כנדרש.

עקרון המינימום

לא לכל קבוצת מספרים טבעיים יש מקסימום (הגדרה 3.3.6). למשל, ל־$\mathbb{N}$ עצמה אין איבר מקסימלי. אולם מסתבר שאחת הצורות השקולות של עקרון האינדוקציה היא הטענה הבאה:

משפט 3.5.2 (עקרון המינימום)[12] לכל תת־קבוצה לא ריקה של $\mathbb{N}$ יש איבר מינימלי.

הוכחה יהי $A \subseteq \mathbb{N}$. נניח בשלילה של־A אין מינימום. נגדיר את B להיות קבוצת כל המספרים הטבעיים הקטנים מכל מספר ב־A, כלומר

$$B = \{n \in \mathbb{N} : a > n \text{ מתקיים } a \in A \text{ לכל } \}$$

[12] לעתים עקרון המינימום מכונה **עקרון הסדר הטוב** (well-ordering principle).

שימו לב שאין ל־A ול־B איברים משותפים, ולכן אם נוכיח ש־ $B = \mathbb{N}$ אז ינבע $A = \emptyset$, בסתירה לנתון. מכך נסיק של־A יש מינימום.

כדי להראות ש־ $B = \mathbb{N}$ נראה ש־B קבוצה אינדוקטיבית.

ראשית, נשים לב שאילו היה מתקיים $0 \in A$ אז 0 היה בוודאי המינימום של A, כי A מכיל רק מספרים טבעיים. מכיוון שאנו מניחים של־A אין מינימום, הרי $0 \notin A$, ולכן לכל $a \in A$ מתקיים $a > 0$. לכן $0 \in B$.

שנית, נראה שאם $b \in B$ אז $b + 1 \in B$. לפי הגדרת הקבוצה B, לכל $a \in A$ מתקיים $a > b$, ולכן $a \geq b+1$ (זו תכונה של המספרים הטבעיים שהוזכרה במשפט 3.4.5). אילו היה מתקיים $b+1 \in A$ היה נובע ש־ $b+1$ הוא מינימום של A, ומכיוון שאנו מניחים של־A אין מינימום, הרי $b+1 \notin A$. קיבלנו אפוא שכל $a \in A$ מקיים $a \geq b+1$ וגם $a \neq b+1$, כלומר $a > b+1$. לכן $b+1 \in B$.

הראינו ש־ $0 \in B$ ושאם $b \in B$ אז $b+1 \in B$, ולכן לפי עקרון האינדוקציה $B = \mathbb{N}$, כפי שרצינו. ■

הנה שימוש של עקרון המינימום. תהי $r = \frac{m}{n}$ הצגה של מספר רציונלי r כמנה של מספרים שלמים עם $n > 0$. ההצגה נקראת הצגה **מצומצמת** (reduced או in least terms) אם לא ניתן לרשום $r = \frac{m'}{n'}$ לאף מספר טבעי n' קטן יותר מ־n. אז אומרים גם שהמנה $\frac{m}{n}$ היא **שבר מצומצם**.

אם $\frac{m}{n}$ הוא שבר מצומצם אז אין מספר טבעי $k > 1$ המחלק את m, n. ואמנם אילו היה כזה אז היה אפשר לרשום $m = km'$ ו־ $n = kn'$ למספרים שלמים m', n' כלשהם, והיה מתקיים $\frac{m}{n} = \frac{km'}{kn'} = \frac{m'}{n'}$ (בשוויון האחרון "צמצמנו" את k מהמונה והמכנה). אבל זו סתירה לכך ש־$\frac{m}{n}$ שבר מצומצם, כי השוויון $n = kn'$ גורר $0 < n' < n$.

את העובדה שלכל מספר רציונלי יש הצגה מצומצמת אפשר להוכיח בעזרת עקרון המינימום. ואמנם, יהי r מספר רציונלי. אם $r = 0$ אז $r = \frac{0}{1}$ היא בבירור ההצגה המצומצמת היחידה. אחרת, נגדיר

$$A = \{n \in \mathbb{N} :\ r = \frac{m}{n} \text{ למספר שלם } m \text{ כלשהו}\}$$

הקבוצה A היא תת־קבוצה של $\mathbb{N}$, והיא לא ריקה כי לכל מספר רציונלי יש הצגות שבהם המכנה אי־שלילי. לכן לפי עיקרון המינימום יש ל־A איבר מינימלי. אם $n = \min A$ אז יש m יחיד כך ש־ $r = \frac{m}{n}$ (למה?). לפי ההגדרה, $\frac{m}{n}$ היא הצגה מצומצמת של r, ומכיוון של־A יש מינימום יחיד, זו גם ההצגה המצומצמת היחידה של r.

עקרון ההוכחה באינדוקציה מלאה

בשלב האינדוקציה של הוכחות באינדוקציה מתעורר לפעמים צורך להסתמך לא רק על השלב הקודם, אלא על כמה מהשלבים הקודמים או אפילו על כולם. המשפט הבא מאפשר לעשות זאת:

טענה 3.5.3 (עקרון ההוכחה באינדוקציה מלאה) תהי $P(n)$ טענה כך שלכל n טבעי מתקיים

$$P(n) \quad \Longrightarrow \quad (P(k) \text{ מתקיים לכל מספר טבעי } k < n)$$

אזי $P(n)$ נכונה לכל n טבעי.

הערה במבט ראשון טענה זו נראית מוזרה כי אין בה זכר לבסיס האינדוקציה. אבל קיום התנאי במשפט עבור $n = 0$ אומר שמתקיימת הטענה "אם $P(k)$ נכונה לכל $k < 0$ טבעי, אז $P(0)$ נכונה". שימו לב שאין מספרים טבעיים שליליים, ולכן הטענה "$P(k)$ נכונה לכל $k < 0$ טבעי" מתקיימת באופן ריק. מכאן שהתנאי במשפט גורר ש־$P(0)$ מתקיימת. מכאן שבסיס האינדוקציה אינו חסר מן הניסוח, הוא רק מסתתר.

הוכחה תהי $A \subseteq \mathbb{N}$ קבוצת המספרים n שעבורם $P(n)$ אינה מתקיימת. אם נראה ש־A ריקה אז סיימנו, כי פירוש הדבר ש־$P(n)$ נכונה לכל $n \in \mathbb{N}$.

נניח בשלילה ש־A אינה ריקה. אז לפי טענה 3.5.2 יש ל־A איבר קטן ביותר שנקרא לו n. משמע, לכל מספר k כך ש־ $k < n$ מתקיימת הטענה $P(k)$. אבל אז לפי ההנחה גם $P(n)$ מתקיימת, בסתירה לכך ש־ $n \in A$. ההנחה ש־A אינה ריקה הובילה לסתירה, ולכן נסיק שההנחה $A \neq \emptyset$ אינה נכונה, כלומר $A = \emptyset$, כנדרש. ■

הגדרה ברקורסיה

לעקרון ההוכחה באינדוקציה מתלווה **עקרון ההגדרה ברקורסיה**. עקרון זה מבטיח שאפשר להגדיר בצורה יחידה איבר $a(n)$ התלוי במספר טבעי n על ידי כך שניתן כלל הקובע את $a(n)$ על סמך המספר n והאיברים $a(i)$ עבור $i < n$. כלל כזה נקרא **נוסחת נסיגה**.[13]

למשל, נגדיר את $a(n)$ לפי הכלל שאם $n = 1, 2$ אז $a(n) = 1$, ואם $n \geq 2$ אז $a(n)$ מוגדרת על ידי נוסחת הנסיגה $a(n) = a(n-1) + a(n-2)$. בפרט $a(3) = a(1) + a(2) = 1 + 1 = 2$, והאיברים הראשונים בסדרה הם $1, 1, 2, 3, 5, 8, \ldots$, וכן הלאה (חשבו את חמשת האיברים הבאים!).

קבוצות סופיות

תהי A קבוצה. אם יש מספר טבעי $n \geq 1$ ואיברים $x_1, x_2, \ldots, x_n$ שכולם שונים זה מזה וכך ש־ $A = \{x_1, x_2, \ldots, x_n\}$, אז n נקרא **מספר האיברים**[14] **ב־**A ומסומן

[13] קל להראות בעזרת עקרון האינדוקציה שבהינתן נוסחת הנסיגה יש לכל היותר דרך אחת להגדיר את סדרת האיברים $a(0), a(1), a(2), \ldots$. עקרון האינדוקציה מאפשר גם להוכיח שיש סדרה המקיימת את נוסחת הנסיגה כזו. הדבר ברור למדי מבחינה אינטואיטיבית ולא ניכנס לפרטים.

[14] מספר האיברים בקבוצה A נקרא לפעמים **העוצמה** (cardinality) של A.

ב־$|A|$. אם A ריקה אומרים שמספר האיברים הוא 0. בשני המקרים האלה A נקראת **סופית** (finite), אחרת היא נקראת **אינסופית** (infinite).

קבוצות סופיות מקיימות את התכונות הבאות:

טענה 3.5.4 יהיו A, B קבוצות.

1. אם A סופית ואם $|A| = m$ וגם $|A| = n$ אז $m = n$ (כלומר מספר האיברים בקבוצה נקבע באופן יחיד).
2. אם $A \subseteq B$ ו־B סופית אז A סופית ומתקיים $|A| \leq |B|$.
3. אם A, B סופיות אז $A \cap B$ סופית ומתקיים $|A \cap B| \leq \min\{|A|, |B|\}$ כאשר יש שוויון אמ"מ $A \subseteq B$ או $B \subseteq A$.
4. אם A, B סופיות אז $A \cup B$ סופיות, מתקיים $|A \cup B| \leq |A| + |B|$, ויש שוויון אמ"מ $A \cap B = \emptyset$. בפרט אם $A \cup B$ אינסופית אז לפחות אחת הקבוצות A, B אינסופית.

מבחינה אינטואיטיבית התכונות ברורות, ואנו משאירים את ההוכחה הפורמלית כתרגיל. במקום זאת נוכיח את הטענה החשובה הבאה:

טענה 3.5.5 לכל קבוצה סופית ולא ריקה של מספרים (לאו דווקא מספרים טבעיים) יש מקסימום ומינימום.

הוכחה נוכיח את הטענה לגבי מינימום. ההוכחה שיש מקסימום דומה.

כיוון שלקבוצה לא ריקה יש גודל לפחות 1, די להוכיח לכל $n \geq 1$ טבעי את הטענה שאם $A \subseteq \mathbb{R}$ סופית ואם $|A| \geq 1$ אז ל־A יש מינימום. ההוכחה לכך באינדוקציה על n.

עבור $n = 1$ העובדה ש־ $|A| = 1$ פירושה ש־A מכיל איבר יחיד ולכן אותו איבר הוא כמובן המינימום של A.

נניח שלכל קבוצת מספרים בגודל n יש מינימום ונוכיח שלכל קבוצת מספרים בגודל $n+1$ יש מינימום גם־כן. תהי $A = \{a_1, a_2, \ldots, a_n, a_{n+1}\}$ קבוצה בגודל $n+1$, ונגדיר $B = \{a_1, a_2, \ldots, a_n\}$ כך ש־ $|B| = n$. מהנחת האינדוקציה נובע של־B יש מינימום, כלומר יש $1 \leq i \leq n$ כך ש־ $a_i \leq a_j$ לכל $1 \leq j \leq n$. אם $a_i \leq a_{n+1}$ אז a_i הוא גם המינימום של A. אחרת, $a_{n+1} < a_i$ וממילא לכל $1 \leq j \leq n$ מתקיים $a_{n+1} < a_i \leq a_j$, וקיבלנו ש־ a_{n+1} הוא המינימום של A. בכל מקרה, הראינו של־A יש מינימום, כנדרש. ■

נסיים את הסעיף באזהרה: יש להשתמש בעיקרון האינדוקציה בזהירות, אחרת ניתן בקלות להוכיח טענות שגויות. למשל,

"טענה" אם A קבוצה סופית של אנשים אז או ש־A מכילה רק גברים, או ש־A מכילה רק נשים.

"הוכחה" נוכיח באינדוקציה את הטענה ש־

$$P(n) = \text{" אם } A \text{ קבוצה בת } n \text{ אנשים אז או שכולם גברים או שכולן נשים "}$$

בסיס האינדוקציה: עבור $n = 1$ הטענה אומרת שקבוצה המכילה אדם יחיד היא קבוצה המכילה רק גברים או רק נשים, וזו טענה נכונה.

שלב האינדוקציה: נניח שהטענה $P(n)$ נכונה. תהי כעת $A = \{a_1, a_2, \ldots, a_n, a_{n+1}\}$ קבוצה של $n+1$ אנשים. נגדיר

$$A' = \{a_1, a_2, \ldots, a_n\}\ ,\ A'' = \{a_2, a_3, \ldots, a_{n+1}\}$$

יש שתי אפשרויות: או ש־a_2 גבר או ש־a_2 אישה. נניח למשל ש־ a_2 הוא גבר. מכיוון ש־A', A'' הן קבוצות בגודל n הרי שלפי הטענה $P(n)$ כל אחת מכילה רק אנשים ממין אחד, ומכיוון ש־a_2 שייך לשתיהן הרי ששתיהן מכילות רק גברים. אבל $A = A' \cup A''$ ולכן A מכילה רק גברים. באותו אופן אם a_2 היא אישה אז A מכילה רק נשים. הראינו שבכל מקרה A מכילה רק אנשים ממין אחד, כלומר, ש־$P(n+1)$ נובע מ־$P(n)$, כנדרש.

היכן השגיאה בהוכחה?

תרגילים

1. יהיו $k, n \in \mathbb{N}$ טבעיים. בשאלה זו נדון **בחלוקה עם שארית** (division with remainder) של n ב־k: הוכיחו שקיימים מספרים טבעיים יחידים $d, r \in \mathbb{N}$ כך ש־ $0 \le r < k$ ו־ $n = dk + r$. המספר r נקרא **השארית** בחלוקת n ב־k. שימו לב ששאלה זו היא הכללה של הטענה שכל מספר טבעי הוא זוגי או אי זוגי.

2. הוכיחו שהסכום S_n של n הריבועים הראשונים, $s_n = 1^2 + 2^2 + \ldots + n^2$, שווה ל־$\frac{1}{6} n \cdot (n+1) \cdot (2n+1)$. כאן סימנו $k^2 = k \cdot k$.

הערה תחת ההנחה ש־S_n מקיים נוסחה מהצורה $S_n = an^3 + bn^2 + cn + d$, אפשר לחשב את a, b, c, d כדלקמן. ראשית, מחשבים את S_1, S_2, S_3, S_4, למשל: $S_2 = 1^2 + 2^2 = 5$. שנית מציבים בנוסחה $S_n = an^3 + bn^2 + cn + d$ את הערכים $n = 1, 2, 3, 4$, ומקבלים עבור S_1, S_2, S_3, S_4 ביטויים המכילים את a, b, c, d, למשל: $S_2 = a \cdot 2^3 + b \cdot 2^2 + c \cdot 2 + d$. השוואת הביטויים שקיבלנו עבור S_2 נותן את המשוואה $5 = 8a + 4b + 2c + d$, ומ־S_1, S_3, S_4 מקבלים שלוש משוואות נוספות, או בסך הכל ארבע משוואות בארבעת הנעלמים a, b, c, d. פתרון מערכת המשוואות נותן את a, b, c, d. ראו גם תרגיל (20) בסעיף 3.6.

3. הראו שעיקרון האינדוקציה, עיקרון ההוכחה באינדוקציה, עיקרון המינימום ועיקרון ההוכחה באינדוקציה מלאה כולם שקולים (חלק מהגרירות הוכחו בסעיף זה. עליכם להשלים את הגרירות החסרות).
4. הוכיחו שאם $A \subseteq \mathbb{Z}$ קבוצה, ואם קיים $n \in \mathbb{Z}$ כך ש־ $n \leq a$ לכל $a \in A$, אז ל־A יש מינימום. הוכיחו שאם יש $m \in \mathbb{Z}$ כך ש־ $a \leq m$ לכל $a \in A$ אז ל־A יש מקסימום.
5. תהי $A \subseteq \mathbb{N}$ קבוצה כך ש־ $A \neq \emptyset$ ו־ $A \neq \mathbb{N}$. הוכיחו שאו שיש $n \in A$ כך ש־ $n+1 \notin A$, או שיש $n \in \mathbb{N} \setminus A$ כך ש־ $n+1 \in A$. הראו שאם שתי הקבוצות $A, \mathbb{N} \setminus A$ אינסופיות אז שתי האפשרויות תקפות.
6. הוכיחו את התכונות של קבוצות סופיות מטענה 3.5.4.
7. הוכיחו ש־$\mathbb{N}$ אינסופית (רמז: טענה 3.5.5).

3.6 חזקות, סכומים ומכפלות

בסעיף זה נגדיר כמה פעולות המתקבלות מהפעלה חוזרת של פעולות חיבור וכפל.

חזקות עם מעריך טבעי

פעולת החזקה מוגדרת ככפל חוזר של מספר עם עצמו. באופן פורמלי,

הגדרה 3.6.1 לכל $a \in \mathbb{R}$ ולכל $n \in \mathbb{N}$ נגדיר ברקורסיה מספר a^n באופן הבא:

$$\begin{aligned} a^0 &= 1 \\ a^{n+1} &= a^n \cdot a \end{aligned}$$

a^n נקרא a **בחזקת** n (a to the power n) או לפעמים **החזקה ה־n־ית של** a. המספר n נקרא **המעריך** (exponent), והמספר a נקרא **הבסיס** (base).

נדון בפירוט בתכונות החזקה בפרקים הבאים, ונכליל אותן במידה ניכרת. בשלב זה נציין את התכונות הבאות:

טענה 3.6.2 לכל $m, n \in \mathbb{N}$ ולכל $a, b \in \mathbb{R}$ מתקיים $a^{m+n} = a^m \cdot a^n$, $(a^m)^n = a^{mn}$ ו־ $(ab)^n = a^n b^n$.

ההוכחה הפורמלית של הטענה כוללת סדרה של הוכחות באינדוקציה, ומושארת כתרגיל (ראו תרגיל (7) בסוף הסעיף). אפשר גם לתת לה הסבר פחות פורמלי אך משכנע. למשל, כיוון ש־ a^{m+n} הוא המכפלה של a עם עצמו $m+n$ פעמים, מתקיים

$$a^{m+n} = \underbrace{a \cdot a \cdot \ldots \cdot a}_{m+n \text{ פעמים}} = \underbrace{a \cdot a \cdot \ldots \cdot a}_{m \text{ פעמים}} \cdot \underbrace{a \cdot a \cdot \ldots \cdot a}_{n \text{ פעמים}} = a^m \cdot a^n$$

ובאופן דומה אפשר להצדיק את שני השוויונות האחרים.

סכומים סופיים

לעתים קרובות נטפל בסכומים ומכפלות של יותר משני מספרים. כפי שכבר ציינו, עבור מספרים $a_1, a_2, \dots, a_n$ ניתן להוכיח (באינדוקציה, כמובן!) שאם נחבר את המספרים האלה בכל אחד מהסידורים האפשריים נקבל את אותה התוצאה (זו מסקנה מאקסיומת הקיבוץ). לכן אפשר לדבר על הסכום של המספרים האלה מבלי לציין את סדר הסכימה. כדי לסמן את הסכום הזה ניתן להיעזר ב"שלוש נקודות", כך:

$$a_1 + a_2 + \ldots + a_n$$

סימון זה קל לקריאה אך יש לו מספר חסרונות. ראשית אם $n = 1$ סימון זה מטעה למדי, כי מופיעים בו לכאורה יותר מדי מחוברים. שנית, הוא עלול להסתיר מידע: האם $0 + 1 + \ldots + n^2$ הוא סכום $n^2 + 1$ המספרים העוקבים $0 + 1 + 2 + \ldots + n^2$, או שמא הוא סכום $n + 1$ הריבועים הראשונים, $0 + 1 + 4 + 9 + \ldots + n^2$? כדי למנוע אי־בהירות כזאת קיים סימון מקוצר ומדויק לסכום:

סימון 3.6.3 אם $n \in \mathbb{N}$ ו־$a_1, a_2, \dots, a_n$ הם מספרים אז הסכום שלהם מסומן על ידי $\sum_{i=1}^{n} a_i$. באופן כללי יותר, הסכום של מספרים $a_k, a_{k+1}, \dots, a_n$ מסומן $\sum_{i=k}^{n} a_i$. ביטוי מהצורה $\sum_{i=k}^{n} a_i$ כאשר $n < k$ מוגדר להיות 0 (כי זהו סכום של אפס מחוברים).

הערות

1. את הביטוי $\sum_{i=k}^{n} a_i$ קוראים כך: "הסכום, כאשר i רץ מ־k עד n, של a_i". סימן הסכימה $\sum$ היא האות היוונית סיגמא גדולה. המשתנה i בסכום $\sum_{i=k}^{n} a_i$ נקרא **אינדקס הסכימה** של הסכימה, או **משתנה הסכימה**. המספר k בביטוי $i = k$ והמספר n מעל סימן הסכימה נקראים **גבולות הסכימה**. כדי לדעת אילו מהמספרים a_i משתתפים בסכימה $\sum_{i=k}^{n} a_i$, יש להציב במקום האינדקס הרץ i את כל המספרים הטבעיים בטווח בין הגבול התחתון לעליון. למשל,

$$\sum_{i=2}^{5} a_i = a_2 + a_3 + a_4 + a_5$$

2. שימו לב שכאשר $k \leq n$, בסכום $\sum_{i=k}^{n} a_i$ יש $n - k + 1$ מחוברים, ולא $n - k$ מחוברים.

3. בחירת הסמל המציין את אינדקס הסכימה אינה חשובה כל עוד אין לו תפקיד נוסף שיכול להוביל לדו־משמעות. הסכום של המספרים $a_1, a_2, \ldots, a_n$ ניתן לכתיבה גם בכל אחת מהדרכים $\sum_{j=1}^{n} a_j$ או $\sum_{m=1}^{n} a_m$, וכן הלאה. אולם הכתיבה $\sum_{n=1}^{n} a_n$ אינה חוקית (כאן n מופיע גם בתפקיד האינדקס וגם בתפקיד הגבול העליון של הסכימה). כמו־כן ביטוי כמו $i + \sum_{i=1}^{n} a_i$ אינו חוקי, כי לסימן i יש כבר משמעות לפני שהוא מופיע כאינדקס בסכום.

   אגב, ביטוי כגון $\sum_{i=1}^{n} a_j$ הוא חוקי, אך כיוון שהאינדקס j איננו משתנה הסכימה, הסכום שווה לסכום של המספר a_j עם עצמו n פעמים, כלומר ל־$n \cdot a_j$, ולא לסכום $a_1 + \ldots + a_n$.

4. אם I קבוצה סופית ולכל $i \in I$ מתאים מספר a_i אז הסכום $\sum_{i \in I} a_i$ מוגדר להיות סכום כל ה־a_i־ים עבור $i \in I$. נציין שבניגוד לסימון $\sum_{j=1}^{n} a_j$, שבו בסימון מצוין גם סדר הסכימה, בסימון $\sum_{i \in I} a_i$ לא קבענו סדר, אך סדר הסכימה לא משנה, כמובן.

5. אפשר לתת הגדרה פורמלית של סימן הסכום ברקורסיה על מספר האיברים המשתתפים בסכימה. אנו משאירים את הפרטים כתרגיל למעוניינים.

דוגמאות

1. $\sum_{k=1}^{n} 1 = n$ כי זה המספר שהתקבל מחיבור המספר 1 עם עצמו n פעמים.

2. $\sum_{k=1}^{n} k = 1 + 2 + \ldots + n$. מצאנו נוסחה לסכום הזה בדוגמה (2) בעמוד 44. סכום כזה נקרא **סכום חשבוני** (arithnmetic sum). כך גם נקרא סכום מהצורה $\sum_{k=1}^{n}(a + kb)$ (מצאו נוסחה לסכום כזה!).

3. $\sum_{k=0}^{n} a^k = 1 + a + a^2 + \ldots + a^n$. סכום כזה נקרא **סכום הנדסי** או **גאומטרי** (geometric sum). נראה שלכל מספר $a \neq 1$ ו־n טבעי, מתקיים $\sum_{k=0}^{n} a^k = \frac{a^{n+1}-1}{a-1}$. אפשר להוכיח את השוויון באינדוקציה, אך הנה דרך אחרת. יהי $c = \sum_{i=0}^{n} a^i$ הגודל שאותו אנו רוצים לחשב. נתבונן בסכום $\sum_{i=0}^{n+1} a^i$. ניתן לרשום אותו בשתי דרכים. דרך אחת היא

$$1 + a + a^2 + \ldots + a^n + a^{n+1} = c + a^{n+1}$$

   ואפשר לרשום אותה גם בצורה

$$1 + a + a^2 + \ldots + a^n + a^{n+1} = 1 + a(1 + a + a^2 + \ldots + a^n) = 1 + ac$$

   (הצדיקו כל שוויון!). השוואת הביטויים נותנת

$$c + a^{n+1} = 1 + ac$$

   והטענה נובעת על ידי העברת אגפים.

כאשר משתמשים בסימן הסכום נעזרים לעתים בכמה כללים בסיסיים המאפשרים לעבור מדרך רישום אחת לאחרת. הנה לדוגמה כמה כללים כאלה, שאנו מביאים ללא הוכחה. נתקל גם באחרים במהלך הספר.

פיצול סכום לשני סכומים: יהי $k \le m < n$. סכום המספרים $a_k, \ldots, a_n$ מתפרק לסכומים כדלקמן:

$$\begin{aligned} \sum_{i=k}^{n} a_i &= a_k + a_{k+1} + \ldots + a_n \\ &= (a_k + a_{k+1} + \ldots + a_m) + (a_{m+1} + \ldots + a_n) \\ &= \sum_{i=k}^{m} a_i + \sum_{i=m+1}^{n} a_i \end{aligned}$$

שוויון זה הוא ביטוי לאקסיומת הקיבוץ.

שינוי גבולות סכימה: כשסוכמים את המספרים $a_k, \ldots, a_n$ אפשר לרשום את המחוברים גם בתור $a_{0+k}, a_{1+k}, \ldots, a_{(n-k)+k}$, ולכן

$$\sum_{i=k}^{n} a_i = \sum_{j=0}^{n-k} a_{j+k}$$

ניתן לראות את השוויון הזה כאילו "החלפנו משתנה" והצבנו $i = j + k$. באותו אופן, לכל m מתקיים

$$\sum_{i=k}^{n} a_i = \sum_{j=m}^{m+(n-k)} a_{j+(k-m)}$$

שינוי סדר סכימה בסכום כפול: יהיו $n, m \in \mathbb{N}$ ולכל $0 \le i < m\,,\, 0 \le j < n$ יהי נתון מספר $a_{i,j}$. נרשום את המספרים בטבלה, כך:

$$\begin{array}{cccc} a_{0,0} & a_{0,1} & \cdots & a_{0,n-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,n-1} \\ \vdots & \vdots & & \vdots \\ a_{m-1,0} & a_{m-1,1} & \cdots & a_{m-1,n-1} \end{array}$$

כל אחד מהמספרים $a_{i,j}$ עבור $0 \le i < n\,,\, 0 \le j < m$ מופיע בריבוע פעם אחת.

אפשר לתאר את סכום כל המספרים $a_{i,j}$ הללו בכמה דרכים. דרך אחת היא לסמן

$$I = \{(i,j) \,:\, 0 \le i < m\,,\, 1 \le j < n\}$$

ואז מדובר בסכום $\sum_{(i,j)\in I} a_{i,j}$. דרך אחרת היא לסכם קודם את השורות בטבלה. לכל שורה מתוך m השורות נקבל מספר: את סכום n המספרים בשורה ה־i נסמן ב־ $b_i = \sum_{j=0}^{n-1} a_{i,j}$. הסכום של כל המספרים בטבלה הוא $\sum_{i=0}^{m-1} b_i$, כלומר

$$\sum_{(i,j)\in I} a_{i,j} = \sum_{i=0}^{m-1} \left(\sum_{j=0}^{n-1} a_{i,j} \right)$$

לרוב משמיטים את הסוגריים ובאגף ימין רושמים פשוט $\sum\sum$. באופן דומה, אם נסכם את m המספרים בעמודה ה־j נקבל מספר $c_j = \sum_{i=0}^{m-1} a_{i,j}$, וסכום כל המספרים בטבלה הוא סכום כל ה־c_j־ים, כלומר $\sum_{j=0}^{n-1} c_j$. קיבלנו בסיכום ש־

$$\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} a_{i,j} = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} a_{i,j} = \sum_{(i,j)\in I} a_{i,j}$$

מכפלות סופיות

בנוסף לסימן הסכום קיים סימן מקוצר למכפלה. במקום לרשום מכפלה של n מספרים $a_1, a_2, \ldots, a_n$ בעזרת שלוש נקודות, כמו ב־ $a_1 \cdot a_2 \cdot \ldots \cdot a_n$, נגדיר:

הגדרה 3.6.4 המכפלה סופית של n מספרים $a_1, \ldots, a_n$ מסומנת על ידי $\prod_{i=1}^{n} a_i$. באופן כללי יותר, המכפלה של מספרים $a_k, a_{k+1}, \ldots, a_n$ מסומנת ב־$\prod_{i=k}^{n} a_i$. ביטוי מהצורה $\prod_{i=k}^{n} a_i$ כאשר $n < k$ מוגדר להיות 1.

האות Π היא האות היוונית פאי גדולה. ההערות אחרי הגדרת סימן הסכום חלות על מכפלות עם השינויים המתבקשים. בתרגילים בסוף הסעיף תתבקשו גם לנסח עבור מכפלות כללים המקבילים לכללים שתיארנו לעיל עבור סכומים.

למכפלה של n הטבעיים החיוביים הראשונים יש סימון מיוחד:

הגדרה 3.6.5 עבור $n > 0$ טבעי, המכפלה $\prod_{i=1}^{n} i = 1 \cdot 2 \cdot \ldots \cdot n$ מסומנת ב־$n!$, ונקראת n **עצרת** (factorial). המספר $0!$ מוגדר להיות 1.

גדלים חשובים נוספים המוגדרים בעזרת עצרות הם:

הגדרה 3.6.6 יהי $n > 0$ טבעי ו־ $0 \le k \le n$. **המקדם הבינומי** n **מעל** (the binomial coefficient n over k), המסומן $\binom{n}{k}$, מוגדר על ידי

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

המקדמים הבינומיים הם מספרים טבעיים, על אף שבמבט ראשון הם כלל לא נראים כאלה (תוכיחו זאת בתרגילים שבסוף הסעיף).

אחד השימושים של המקדמים הבינומיים הוא במשפט הבא, שמתאר כיצד לכתוב חזקה של סכום כסכום של חזקות:

משפט 3.6.7 (נוסחת הבינום (binomial theorem)) יהיו a, b מספרים ו־ $n \in \mathbb{N}$. אז

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

למשל, הנוסחה $(x+y)^2 = x^2+2xy+y^2$ היא מקרה פרטי של המשפט, וגם הנוסחה $(x+y)^3 = x^3+3x^2y+3xy^2+y^3$ (הבהירו לעצמכם מדוע!). ההוכחה של המקרה הכללי מושארת כתרגיל.

תרגילים

1. כתבו בקצרה את הביטויים הבאים בעזרת הסימנים $\sum, \prod$. לדוגמה,

$$x_1^2 + x_2^2 + \ldots + x_n^2 = \sum_{k=1}^{n} x_k^2$$

(א) $1 + a^2 + a^4 + \ldots + a^{2n}$.

(ב) $x_1 + x_1x_2 + x_1x_2x_3 + \ldots + (x_1x_2 \ldots x_{100})$.

(ג) $a^n + a^{n-1}b + a^{n-2}b^2 + \ldots + b^n$.

2. חשבו את הערך של הביטויים הבאים (כלומר מצאו דרך לכתוב אותם ללא הסימנים $\sum, \prod$):

(א) $\sum_{k=1}^{5} k \cdot (k-1)$.

(ב) $\prod_{n=1}^{1000} 2$.

(ג) $\prod_{r=1}^{3} \sum_{k=-r}^{r} k$.

(ד) $\sum_{i=1i}^{n} \frac{1}{i(i+1)}$ (רמז: $\frac{1}{i(i+1)} = \frac{1}{i} - \frac{1}{i+1}$).

(ה) $\sum_{i=0}^{100} \frac{i \cdot i}{i+1} - \sum_{i=0}^{100} \frac{(i+1)(i+1)}{i+2}$.

(ו) $\sum_{s=1}^{100} (7^i - 7^{i-1})$.

(ז) $\sum_{i=1}^{10} \sum_{j=1}^{10} (j \cdot i^2 - j \cdot (i-1)^2)$.

3. יהיו $x_1, x_2, \ldots, x_n$ מספרים ממשיים ויהיו $\lambda_1, \ldots, \lambda_n$ מספרים אי־שליליים כך ש־ $\sum_{k=0}^{n} \lambda_k = 1$. הוכיחו ש־

$$\min\{x_1, \ldots, x_n\} \leq \sum_{k=1}^{n} \lambda_k x_k \leq \max\{x_1, \ldots, x_n\}$$

4. התבוננו בסכום $\sum_{i=1}^{n} \sum_{j=i}^{n} a_{i,j}$. כתבו את הסכום במפורש עבור $n = 3$, ומצאו מספרים k, m (התלויים אולי ב־j, n) כך שלכל אוסף מחוברים $a_{i,j}$ מתקיים השוויון $\sum_{i=1}^{n} \sum_{j=i}^{n} a_{i,j} = \sum_{j=1}^{n} \sum_{i=k}^{m} a_{i,j}$ (רמז: שימו לב שהסכום $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j}$ מתאר סכימה על הזוגות (i, j) במלבן n על n. איזה חלק מהמלבן מתאר הסכום בשאלה זו?).

5. יהי n מספר שלם ונתבונן בסכום $\sum_{k=1}^{n} k = 1 + 2 + 3 + \ldots + n$. נסמן את הסכום ב־$S$. השיטה הבאה לחישוב S מיוחסת למתמטיקאי קארל פרידריך גאוס.[15] נרשום את המחוברים פעמיים, פעם בסדר הרגיל ופעם בסדר ההפוך:

$$\begin{array}{ccccccccc} 1 & , & 2 & , & \ldots & , & n-1 & , & n \\ n & , & n-1 & , & \ldots & , & 2 & , & 1 \end{array}$$

הסכום של כל שורה הוא S ולכן סכום כל המספרים בריבוע הוא $2S$. מאידך, הסכום של כל עמודה הוא $n + 1$. כיוון שיש n עמודות, סכום כל המספרים הוא $n \cdot (n+1)$, וקיבלנו $2S = n(n+1)$, כלומר $S = \frac{1}{2}n(n+1)$. כתבו הוכחה זו בעזרת סימן הסכום.

6. הוכיחו באינדוקציה ש־ $(-1)^n = 1$ כאשר n זוגי ו־ $(-1)^n = -1$ כאשר n אי־זוגי.

7. הוכיחו את התכונות הבאות של פעולת החזקה.
   (א) לכל $a \in \mathbb{R}$ ולכל $n, m \in \mathbb{N}$ מתקיים $a^{n+m} = a^n \cdot a^m$ (רמז: קבעו את a, n ובצעו אינדוקציה על m)
   (ב) לכל $a \in \mathbb{R}$ ולכל $n, m \in \mathbb{N}$ מתקיים $a^{mn} = (a^m)^n$
   (ג) אם $a, b \in \mathbb{R}$ ו־ $n \in \mathbb{N}$ אז $(ab)^n = a^n b^n$
   (ד) אם $0 < a < b$ ו־$0 < n \in \mathbb{N}$ אז $a^n < b^n$
   (ה) אם $m < n$ טבעיים, אז אם $a > 1$ מתקיים $a^m < a^n$, ואם $0 < a < 1$ אז $a^m > a^n$.

8. תהי $P(n)$ טענה התלויה ב־n ונניח:
   (א) $P(1)$ נכונה
   (ב) אם $P(n)$ נכונה אז $P(2n)$ נכונה

   הוכיחו ש־$P(2^k)$ נכונה לכל $k \in \mathbb{N}$.

[15] Carl Friedrich Gauss, 1777-1855.

9. הוכיחו ש־ $2^k > k$ לכל $k \in \mathbb{N}$.

10. הוכיחו את **אי־שוויון ברנולי**:[16] אם $x > -1$ אז $(1+x)^n \geq 1 + nx$.

11. הראו שלכל $x, y > 0$ ולכל k טבעי מתקיים $(x+y)^k \leq x^k + y^k$.

12. הראו שלכל שני מספרים x, y מתקיים $(x^n - y^n) = (x-y)\sum_{i=1}^{n} x^{n-i}y^{i-1}$. שימו לב שזו הכללה של דוגמה (3) בעמוד 51 (למה?).

13. נסחו עבור מכפלות את הכללים המקבילים לכללים על פיצול סכום, שינוי גבולות סכימה ושינוי סדר סכימה בגבול כפול.

14. לכל מספר c ומספרים $a_1, \ldots, a_n$ מתקיים $\sum_{i=1}^{n}(c \cdot a_i) = c \cdot \sum_{i=1}^{n} a_i$ (זו מסקנה מחוק הפילוג). כאשר c מספר טבעי, מה הטענה המקבילה לגבי מכפלות סופיות?

15. הוכיחו שלכל n ולכל $0 \leq i \leq n$ טבעיים מתקיים $\binom{n}{i} = \binom{n}{n-i}$. הוכיחו שאם $i < j \leq \frac{n}{2}$ אז $\binom{n}{i} < \binom{n}{j}$.

16. הוכיחו את נוסחת הנסיגה $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ (שימו לב שמשמאל מופיע n, ומימין $n-1$). הסיקו שהמקדמים הבינומיים הם מספרים שלמים.

נוסחת הנסיגה מאפשרת להציג את המקדמים הבינומיים כאיברים של מערך המשולשי הנקרא **משולש פסקל**,[17] ראו איור 3.6.1. המקדם $\binom{n}{k}$ הוא האיבר ה־k בשורה ה־n של משולש פסקל. נוסחת הנסיגה פירושה שכל איבר פנימי במערך הוא סכום שני האיברים הקרובים בשורה שמעליו.

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1
\end{array}
$$

איור 3.6.1 משולש פסקל

17. הוכיחו את משפט הבינום (משפט 3.6.7).

18. (*) הוכיחו ש־$\binom{n}{k}$ שווה למספר התת־קבוצות בגודל k של קבוצה בגודל n (רמז: הוכיחו זאת באינדוקציה על n. שימו לב שאם A קבוצה בגודל n ו־$a \in A$ אז אפשר לחלק את התת־קבוצות של A לשני סוגים: כאלה המכילות את a, שהן תת־קבוצות מהצורה $B \cup \{a\}$ כאשר $B \subseteq A \setminus \{a\}$ מגודל $k-1$, וכאלה שאינן מכילות את a, שהן תת־קבוצות מגודל k של $A \setminus \{a\}$).

מה הקשר בין שאלה זו לנוסחת הבינום? (רמז: כשפותחים את הסוגריים בביטוי $(a+b)(a+b)(a+b)$, המחובר ab^2, למשל, מקבל מקדם השווה למספר הדרכים לבחור a אחד ושני b־ים משלושת הגורמים במכפלה המקורית).

19. הוכיחו בעזרת נוסחת הבינום את השוויונות:

(א) $\sum_{k=0}^{n} \binom{n}{k} = 2^n$

(ב) $\sum_{k=0}^{n} (-1)^n \binom{n}{k} = 0$

20. הוכיחו שלכל $k \in \mathbb{N}$ קיימים מספרים $a_{k+1}, a_k, \ldots, a_0 \in \mathbb{Q}$ כך שלכל $n \in \mathbb{N}$ מתקיים

$$1^k + 2^k + \ldots + n^k \quad = \quad a_{k+1}n^{k+1} + a_k n^k + \ldots + a_1 n + a_0$$

---

[16] Johann Bernoulli, 1667־1748.

[17] Blaise Pascal, 1623־1662

(שימו לב שבצד שמאל יש n מחוברים, בעוד שבצד ימין יש $k+2$ מחוברים, כלומר, ניתן לחשב את הסכום של n חזקות ה־k הראשונות במספר פעולות חיבור וכפל קבוע. כאשר n גדול מ־k, זהו חיסכון משמעותי). (רמז: אפשר להוכיח זאת באינדוקציה. מתקיים

$$(n+1)^k - n^k = kn^{k-1} + \dots$$

כאשר במקום שלוש נקודות מופיעות חזקות n^j עבור $0 \leq j < k-1$ עם מקדמים קבועים שאינם תלויים ב־n).

הערה שאלה זו מכלילה את תרגיל (2) מעמוד 48, שם חישבנו את המקדמים a_i במקרה $k=2$. מעניין לדעת שלא ידועה נוסחה מפורשת למקדמים עבור k כללי.

21. נסמן $I = \{1, \dots, n\}$, ועבור $i \in I$ יהיו a_i, b_i מספרים ממשיים. הוכיחו ש־

$$\prod_{i=1}^{n}(a_i+b_i) = \sum_{I_0 \subseteq I}\left((\prod_{i \in I_0} a_i)(\prod_{j \in I \setminus I_0} b_j) \right)$$

(כדאי קודם לחשב את $(a_1+b_1)(a_2+b_2)$ ולראות איך התוצאה קשורה לנוסחה למעלה).

22. מספר טבעי $p > 1$ נקרא **מספר ראשוני** (prime number) אם אינו מתחלק באף מספר טבעי חוץ מ־1 ו־p, דהיינו אם $p = mn$ ו־m, n טבעיים אז אחד מהם הוא p והאחר 1. למשל, המספרים $2, 3, 5, 7, 11, 13, 17, 19, 23, 29$ הם עשרת המספרים הראשוניים הראשונים. הוכיחו שלכל מספר טבעי $n > 1$ יש מספרים ראשוניים $p_1, p_2, \dots, p_k$, לאו דווקא שונים זה מזה, כך שמתקיים $n = \prod_{i=1}^{k} p_i$.

למעשה נכונה טענה חזקה יותר: כל מספר טבעי ניתן לכתיבה יחידה כמכפלה של מספרים ראשוניים (עד כדי שינוי סדר הגורמים), אך לא נוכיח זאת. לעומת זאת בתרגיל (5) בעמוד 206 נוכיח שקיימים אינסוף מספרים ראשוניים.

3.7 ההצגה העשרונית של המספרים השלמים

נסיים באחד הנושאים בהם פתחנו את הפרק: שאלת הייצוג של המספרים. בסעיף זה נפתח את ההצגה העשרונית שתאפשר לייצג מספרים טבעיים ושלמים ובאופן עקיף כל מספר רציונלי.

הצגה עשרונית היא למעשה הצגה של מספר שלם כסכום: למשל, 127 הוא המספר

$$100 + 20 + 7 = 1 \cdot 100 + 2 \cdot 10 + 7 \cdot 1 = 1 \cdot 10^2 + 2 \cdot 10^1 + 7 \cdot 10^0$$

(זכרו ש־$10^0 = 1$). באופן כללי,

הגדרה 3.7.1 המספרים הטבעיים $0,1,2,3,4,5,6,7,8,9$ נקראים **ספרות** עשרוניות (decimal digits). בהינתן ספרות $a_0,a_2,\dots,a_n$, המספר $x=\sum_{k=0}^n a_k\cdot 10^k$ מסומן ב־ $a_na_{n-1}\dots a_1a_0$, ואומרים שזו **הצגה עשרונית** (decimal representation) של x.

משפט 3.7.2 לכל מספר טבעי יש הצגה עשרונית.

הוכחה נראה שלכל n טבעי יש הצגה עשרונית. ההוכחה באינדוקציה על n. עבור $n=0$ קיימת ההצגה העשרונית 0, ולכן הטענה נכונה במקרה זה.

נניח ש־ $n=\sum_{k=0}^m a_k10^k$ היא הצגה עשרונית של n וננסה לקבל הצגה עשרונית של $n+1$. ההוכחה אינה אלא תיאור פורמלי של התהליך של הוספת 1 לפי כללי החיבור הנלמדים בבית הספר. הנה הפרטים.

אם $a_0\neq 9$ אז a_0+1 היא ספרה, ולכן

$$n+1=\sum_{k=0}^m a_k10^k+1=\sum_{k=1}^m a_k10^k+(a_0+1)10^0$$

ואגף ימין הוא הצגה עשרונית של n.

אחרת, הייצוג העשרוני $a_ma_{m-1}\dots a_0$ של n מסתיים ב־9, ואולי ב־99 או 999. בכל מקרה, קיים אינדקס $m_0\in\{1,2,\dots,m\}$ מקסימלי כך שהספרות $a_0,a_1,\dots,a_{m_0}$ כולם שווים ל־9.

אם $m_0=m$ אז

$$n=\sum_{k=0}^m 9\cdot 10^k=9\cdot\sum_{k=0}^m 10^k=9\cdot\frac{10^{m+1}-1}{10-1}=10^{m+1}-1$$

(השתמשנו בנוסחה לסכום גאומטרי מדוגמה (3) בעמוד 51). לכן $n+1=10^{m+1}$, ואפשר לרשום את $n+1$ בכתיב עשרוני בתור $n+1=\sum_{k=0}^{m+1} b_k10^k$ כאשר $b_{m+1}=1$ ו־ $b_k=0$ לכל $0\le k\le m$.

אחרת $m_0<m$, ואז מהגדרת m_0 מתקיים $a_{m_0+1}\neq 9$. מאותו שיקול כמו במקרה $m_0=m$ מתקיים $\sum_{k=0}^{m_0} a_k10^k=10^{m_0+1}-1$ ולכן

$$\begin{aligned}n+1=\sum_{k=0}^m a_k10^k+1&=\sum_{k=m_0+1}^m a_k10^k+\sum_{k=0}^{m_0} a_k10^k+1\\&=\sum_{k=m_0+1}^m a_k10^k+10^{m_0+1}\end{aligned}$$

אבל $a_{m_0+1}<9$ ולכן $a_{m_0+1}+1$ היא ספרה, וקיבלנו ש־

$$n+1=\sum_{k=m_0+2}^m a_k10^k+(a_{m_0+1}+1)\cdot 10^{m_0+1}$$

זו הצגה עשרונית של $n+1$. ■

ההצגה העשרונית של מספר אינה יחידה: למשל, $7\,,07\,,007$ מתארים כולם את אותו מספר. אך פרט לאפשרות להוסיף אפסים מצד שמאל ההצגה היא יחידה, כפי שנובע מהטענה הבאה:

טענה 3.7.3 אם $\sum_{i=0}^{n} a_i \cdot 10^i = \sum_{i=0}^{n} b_i \cdot 10^i$ ו־a_i, b_i הם מספרים טבעיים המקיימים $0 \le a_i, b_i \le 9$ אז $a_i = b_i$.

ההוכחה מושארת כתרגיל.

ראינו כיצד לייצג כל מספר טבעי. על ידי הוספת הסימן "$-$" משמאל לסדרת הספרות כדי לציין נגדי אפשר גם לייצג מספרים שלמים שליליים, ובסיכום מצאנו דרך לייצג כל מספר שלם. אפשר לייצג כל מספר רציונלי כמנה של מספרים שלמים. מחלקה קטנה יותר של מספרים רציונליים אפשר להציג בדרך הבאה.

הגדרה 3.7.4 יהיו $a_1, \ldots, a_n$ ספרות. הסימן $0.a_1a_2\ldots a_n$ נקרא **ההצגה העשרונית** (decimal representation) של השבר $\frac{1}{10^n} \cdot b$, כאשר b השלם בעל ההצגה העשרונים $a_1a_2\ldots a_n$; כלומר

$$0.a_1a_2\ldots a_n = \sum_{k=1}^{n} \frac{a_k}{10^k}$$

כמו־כן בהינתן ספרות עשרוניות $b_0, b_1, \ldots, b_m$ נסמן ב־$b_m b_{m-1}\ldots b_0.a_1a_2\ldots a_n$ את המספר $b_m\ldots b_0 + 0.a_1\ldots a_n$.

אפשר למצוא דיון נוסף בנושא הייצוג העשרוני בתרגילים למטה.

הבחירה של הספרות $0, 1, \ldots, 9$ הייתה שרירותית, ואפשר היה לבחור בתור ספרות את המספרים $0, 1, \ldots, b-1$ לכל $b \ge 2$ טבעי. שיטת ההצגה המתקבלת, שבה $a_n a_{n-1}\ldots a_0$ מייצג את $\sum_{k=0}^{n} a_k b^k$ ו־ $0.a_1a_2\ldots a_n$ מייצג את $\sum_{k=1}^{n} a_k b^{-k}$, נקראת שיטת ההצגה **בבסיס** b. בפרט כאשר $b=2$ השיטה נקראת הצגה **בינרית** (binary), ובה המספרים מיוצגים רק בעזרת הספרות $0,1$. זו שיטה מקובלת לאכסון מידע במחשבים. כאשר $b=3$ השיטה המתקבלת נקראת שיטת הצגה **שלישונית** (ternary).

הייצוג בעזרת שברים עשרוניים מאפשר לייצג רק חלק מהשברים. למשל, המספר $\frac{1}{3}$ אינו ניתן לייצוג כזה (ראו תרגיל (3) בסוף הסעיף). מאידך האפשרות לייצג בכתיב עשרוני את השלמים מאפשר לייצג כל מספר רציונלי כמנה של מספרים כאלה. בפרק הבא נראה שיש מספרים שאינם רציונליים, ושאלה חשובה היא כיצד לייצג אותם. נענה על שאלה זו בסעיף 6.7, שם נגדיר שברים עשרוניים אינסופיים ונראה שכל מספר ניתן לייצוג באופן זה.

תרגילים

1. הוכיחו את טענה 3.7.3.

2. יהי n מספר טבעי. נתאר להלן תהליך המייצר את הספרות של הפיתוח העשרוני שלו. נגדיר $r_0 = n$, ובהינתן r_k נגדיר את a_k להיות השארית של חלוקת r_k ב־10 ואת r_{k+1} להיות המנה השלמה של r_k ב־10, דהיינו a_k, r_{k+1} הם המספרים טבעיים היחידים כך ש־ $0 \leq a_k < 10$ ו־ $r_k = 10r_{k+1}+a_k$ (ראו תרגיל (1) בסעיף 3.5). הוכיחו שלאחר לכל היותר n צעדים מתקיים $r_k = 0$. הראו שאם k הוא האינדקס המקסימלי כך ש־ $a_k \neq 0$ אז $n = \sum_{i=0}^{k} a_i \cdot 10^i$. תנו הוכחה תחת ההנחה של־n יש ייצוג עשרוני (זו הוכחה קלה). אחר־כך הוכיחו זאת בלי הנחה זו (ותקבלו הוכחה חדשה למשפט 3.7.2).

3. הוכיחו שלמספר $\frac{1}{3}$ אין הצגה כשבר עשרוני.

4. כתבו את המספר 107 בייצוג בינרי ושלישוני.

5. יהי n מספר טבעי בעל ייצוג עשרוני $a_k a_{k-1} \ldots a_0$.
   (א) הוכיחו שאם $\sum_{i=0}^{k} a_i$ מתחלק ב־9 אז n מתחלק ב־9.
   (ב) הוכיחו שגם ההפך נכון, כלומר: אם n מתחלק ב־9 אז גם $\sum_{i=0}^{k} a_i$ מתחלק ב־9 (זה קשה יותר).
   (ג) מה הטענה המקבילה למספרים המיוצגים בבסיס 8?

6. נאמר שלמספר טבעי n יש ייצוג מסוג א' אם קיימות ספרות עשרוניות $a_0, \ldots, a_k$ כך ש־ $n = \sum_{i=0}^{k} a_k i^2$, ונאמר שיש לו ייצוג מסוג ב' אם קיימות ספרות עשרוניות $b_0, \ldots, b_m$ כך ש־ $n = \sum_{j=0}^{m} b_j 2^{j^2}$. לכל אחת מהשיטות הללו ענו על השאלות הבאות:
   (א) האם לכל מספר טבעי יש ייצוג כזה?
   (ב) האם הייצוג של מספר בשיטה הוא יחיד?

7. ראינו שהמספר $\frac{1}{3}$ אינו ניתן לייצוג בבסיס עשר (שאלה (3) למעלה) אבל אפשר לייצג אותו בבסיס 3 בצורה 0.1. האם יש מספר רציונלי שאינו ניתן לייצוג באף בסיס?

8. (*) הוכיחו שהשיטות שלומדים בבית הספר לחיבור, כפל וחילוק של מספרים בכתיב עשרוני נותנות תוצאות נכונות (לגבי חילוק, הראו שהתוצאה נכונה בתנאי שהחישוב מסתיים).

פרק 4

תכונת השלמות של המספרים הממשיים

כפי שהדברים עומדים עתה, המערכת $\mathbb{R}$ מקיימת את רוב התכונות שאנו מצפים שיקיימו המספרים. היא גם מכילה את המערכות המוכרות $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, ואלה מתאימות לרוב המטרות המעשיות שלשמן אנו צריכים מספרים. אולם מסתבר שהמערכת $\mathbb{R}$ אינה משיגה עדיין את אחת המטרות שהצבנו בתחילת הפרק הקודם: אי אפשר להשתמש בה כדי לתאר אורכים או כדי לתאר את כל נקודות על ישר. אנו נראה גם שיש משוואות אלגבריות פשוטות שאינן ניתנות כרגע לפתרון.

לאחר שנדון בבעיות האלה מעט יותר בפירוט, נציג את האקסיומה האחרונה של הממשיים. לצורך כך נכיר את המושג החשוב חסם עליון של קבוצה. נדון גם בקשר בין המספרים הרציונליים והמערכת $\mathbb{R}$. לבסוף נוכיח את קיומם של שורשים ונתחיל בהגדרה של פעולת החזקה עם מעריך ממשי כללי (ההגדרה תושלם רק בפרק הבא).

4.1 הפרדוקס של פיתגורס

המספרים הרציונליים אינם מתאימים לתיאור של גדלים חשובים רבים. הדוגמה הפשוטה ביותר היא:

משפט 4.1.1 (אסכולת פיתגורס, המאה ה־5 לפני הספירה)[1] לא קיים מספר רציונלי r שמקיים $r^2 = 2$.

[1] העובדה שאין לשתיים שורש ריבועי הייתה ידועה כבר במאה החמישית לפני הספירה והתגלתה כנראה על ידי אחד מבני הכת הפיתגוראית, שנוסדה על ידי המדען והפילוסוף היווני פיתגורס (Pythagoras, 587־502 לפנה"ס). אחד ממאפייני הכת הייתה תורה נומרולוגית, ומכאן החשיבות שבני הכת יחסו למספרים. העובדה שאין מנה של שלמים שריבועו שתיים הייתה עבורם בלתי מתקבלת על הדעת. לפי האגדה, גילה זאת הפילוסוף היפסוס במהלך שיט בים. כאשר סיפר לחבריו על כך הם השליכו אותו מהאנייה, והשאירו אותו לטבוע.

הוכחה נניח בשלילה שקיים שבר r כזה, כלומר, נניח שקיימים מספרים שלמים m, n כך ש־ $(\frac{m}{n})^2 = 2$. ניתן להניח שבחרנו הצגה מצומצמת של r, כלומר ש־ $n > 0$ ואין זוג $m', n' \in \mathbb{Z}$ כך ש־ $r = \frac{m'}{n'}$ וגם $|n'| < n$.

מההנחה $(\frac{m}{n})^2 = 2$ נקבל לאחר פתיחת סוגריים והעברת אגפים

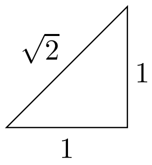
$$m^2 = 2 \cdot n^2$$

מכיוון שאגף ימין זוגי, גם אגף שמאל זוגי, ולכן m זוגי (כי אילו m היה אי־זוגי, גם m^2 היה אי־זוגי. הוכיחו זאת!). קיים לכן מספר שלם m' כך ש־ $m = 2 \cdot m'$. נציב זאת בשוויון הקודם ונקבל $4 \cdot (m')^2 = 2 \cdot n^2$, ולאחר צמצום גורם 2 נקבל

$$2 \cdot (m')^2 = n^2$$

נפעיל את אותו שיקול פעם נוספת: מכיוון שאגף שמאל זוגי, נסיק שגם אגף ימין זוגי, ולכן n זוגי. מכאן שקיים מספר שלם n' כך ש־ $n = 2 \cdot n'$.

לסיכום, קיבלנו ש־ $m = 2m'$ וגם $n = 2n'$, ולכן $r = \frac{m}{n} = \frac{m'}{n'}$. אבל $0 < n' < n$, בסתירה להנחה ש־$\frac{m}{n}$ היא הצגה מצומצמת של r. לכן ההנחה שיש שבר $\frac{m}{n}$ כך ש־ $(\frac{m}{n})^2 = 2$ הובילה לסתירה, ולכן אינה נכונה, כלומר: אין שבר כזה. ■



איור 4.1.1

דוגמה זו היא בעלת אופי אלגברי, אך נובעת ממנה מסקנה גיאומטרית: המספרים הרציונליים אינם מתאימים לצורך ייצוג של אורכים. אם נבנה משולש ישר זווית שניצביו באורך יחידה אחת, כמו באיור 4.1.1, אז משפט פיתגורס מתורת הגאומטריה של המישור קובע שהאורך r של היתר במשולש מקיים

$$r^2 = 1^2 + 1^2 = 2$$

ולכן אינו ניתן לתיאור על ידי מספר רציונלי.

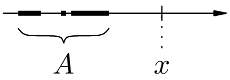
נשאלת השאלה, האם קיים מספר ממשי (שאינו רציונלי) שריבועו שתיים? אם מסתמכים רק על האקסיומות העומדות בשלב זה לרשותנו אי אפשר להוכיח את הקיום של מספר כזה. הרי אילו יכולנו להוכיח זאת היה נובע שיש מספרים ממשיים שאינם רציונליים, וכבר הערנו שלא ניתן להוכיח זאת על סמך האקסיומות הנתונות עד כה (ראו טענה 3.4.11 והדיון שאחריה).

תרגילים

1. הוכיחו שאם m^2 זוגי אז m זוגי ושאם m^2 אי־זוגי אז m אי־זוגי (השתמשנו בכך בהוכחה של משפט 4.1.1).
2. הוכיחו שלא קיים $r \in \mathbb{Q}$ המקיים $r^2 = 3$ ושלא קיים $r \in \mathbb{Q}$ המקיים $r^3 = 2$.
3. הכלילו את השאלה הקודמת והראו שאם p מספר ראשוני ו־ $n > 1$ מספר טבעי אז אין $r \in \mathbb{Q}$ המקיים $r^n = p$ (מספרים ראשוניים הוגדרו בשאלה (22) בעמוד 57).

4.2 אקסיומת החסם העליון

בסעיף זה ננסח את האקסיומה האחרונה של $\mathbb{R}$. אחד המניעים להגדרתה הוא האינטואיציה הגיאומטרית והפיזיקלית שלנו, ובעמודים הבאים ניעזר תכופות בזיהוי של המספרים עם נקודות על ציר המספרים. נחזור ונדגיש שאנו משתמשים בו להמחשה בלבד, ולא נתבסס על נימוקים גאומטריים בהגדרות או בהוכחות.



איור 4.2.1 x הוא חסם מלעיל של A

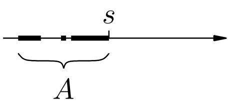
הגדרה 4.2.1 תהי $A \subseteq \mathbb{R}$ קבוצת מספרים. מספר x נקרא **חסם מלעיל** (upper bound) של A אם לכל $a \in A$ מתקיים $a \leq x$. אומרים ש־A **קבוצה חסומה מלעיל** (bounded from above) אם קיים לה חסם מלעיל.

קבוצה A היא חסומה מלעיל אם אינה מכילה מספרים גדולים "בלי סוף". מבחינה גאומטרית פירוש הדבר שאם חושבים על A כקבוצת נקודת בישר אז יש נקודות בישר הנמצאות מימין לכל נקודה ב־A. ראו איור 4.2.1.

דוגמאות

1. 0 הוא חסם מלעיל של הקבוצה $A = \{x \in \mathbb{R} \,:\, x < 0\}$. כמובן גם כל $x > 0$ הוא חסם מלעיל שלה.
2. הקבוצה $\mathbb{R}$ אינה חסומה מלעיל. ואמנם, יהי x מספר ונראה ש־x אינו חסם מלעיל של $\mathbb{R}$. כל שיש לעשות הוא להראות שיש מספרים ב־$\mathbb{R}$ אשר גדולים מ־x, וזה בוודאי נכון, כי למשל $x + 1 > x$ ו־ $x + 1 \in \mathbb{R}$.
3. אם A קבוצה חסומה ו־x חסם מלעיל שלה אז כל מספר $y > x$ הוא גם חסם מלעיל של A, כי אם $a \in A$ אז $a \leq x$ ובוודאי $a \leq y$. מכאן שאם לקבוצה יש חסם מלעיל אחד אז יש לה אינסוף חסמים.
4. תהי A קבוצה ויהי x המקסימום של A. אז מהגדרת המקסימום ברור ש־x חסם מלעיל של A.

כפי שהערנו, לקבוצה חסומה מלעיל יש חסמים מלעיל רבים. ההגדרה הבאה מבחינה בחסם אחד מביניהם שהוא בעל תכונה מיוחדת.



איור 4.2.2 s הוא חסם עליון של A

הגדרה 4.2.2 תהי A קבוצה לא ריקה וחסומה מלעיל של מספרים. מספר s נקרא **חסם עליון** (supremum) של A אם הוא החסם מלעיל הקטן ביותר של A, כלומר: s הוא חסם מלעיל של A, וכל חסם מלעיל t של A מקיים $s \leq t$. החסם העליון של A, אם הוא קיים, מסומן $\sup A$.[2]

[2] לחסם עליון קוראים גם least upper bound ומסמנים אותו לפעמים ב־l.u.b..

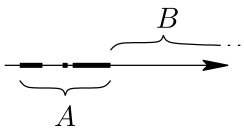
הערות

1. שימו לב שלפי ההגדרה, s הוא חסם עליון של A אמ"מ s הוא המינימום של קבוצת החסמים מלעיל של A.
2. לפעמים כאשר A קבוצה שאינה חסומה מלעיל נרשום $\sup A = \infty$. בסימון זה איננו מתכוונים למושג החסם העליון כפי שהוגדר למעלה, אלא זהו פשוט סימון שנועד לרמוז על כך ש־A אינה חסומה.

הנה תיאור אינטואיטיבי יותר של מושג החסם העליון. תהי $A \subseteq \mathbb{R}$ קבוצה חסומה מלעיל ולא ריקה, ונחשוב עליה כקבוצת נקודות על הישר. יהי x חסם מלעיל של A ונדמיין חלקיק הנמצא ב־x ונע לכיוון שמאל. אם נניח שהחלקיק אינו יכול לעבור דרך הנקודות ב־A אז הוא ייעצר בנקודה כלשהי על הישר. נקרא לנקודה זו s. מכיוון שהחלקיק לא עבר דרך אף נקודה ב־A והוא התחיל מימין ל־A, הרי שכל הנקודות ב־A נמצאות משמאל לנקודת העצירה של החלקיק, ולכן s חסם מלעיל של A. יתר על כן s הוא החסם מלעיל המינימלי, כי אילו היה חסם מלעיל t קטן יותר החלקיק יכול היה להמשיך לנוע לפחות עד שיגיע ל־t, ולא היה עוצר ב־s.

ההסבר הפיזיקלי למעלה נותן את הרושם שהחסם העליון תמיד קיים, ושהוא מוגדר בצורה יחידה, אך טענות אלה אינן מובנות מאליהן. נתייחס תחילה לשאלת היחידות. עלינו לפסול את האפשרות שלקבוצה יהיו שני חסמים עליונים שונים.

טענה 4.2.3 לקבוצה יש לכל היותר חסם עליון אחד.



איור 4.2.3 קבוצת החסמים מלעיל של A

הוכחה תהי $A \subseteq \mathbb{R}$ קבוצה ונגדיר ש־

$$B = \{x \in \mathbb{R} : \text{חסם מלעיל של } A \ x\}$$

חסם עליון של A הוא מינימום של B. באופן כללי ראינו שאם לקבוצה יש מינימום אז הוא יחיד (עמוד 33), ובפרט ל־B יש לכל היותר מינימום אחד. לכן ל־A לכל היותר חסם עליון אחד. ■

יחידות החסם העליון מצדיקה את השימוש בהא הידיעה בביטוי "החסם העליון".

נעבור לשאלה, לאילו קבוצות של מספרים יש חסם עליון. לקבוצה שאינה חסומה מלעיל לא יכול להיות חסם עליון כי חסם עליון הוא בפרט חסם מלעיל. כמו־כן לקבוצה הריקה לא יכול להיות חסם עליון. למעשה, לכל מספר x מתקיים שאם $a \in \emptyset$ אז $a \leq x$ (זה מתקיים באופן ריק) ולכן כל מספר x הוא חסם מלעיל של הקבוצה הריקה, ומכאן של־$\emptyset$ אין חסם מלעיל קטן ביותר, ואין לו חסם עליון.[3] עם זאת, ישנם מקרים שבהם נוכל להראות שקיים חסם עליון.

[3] כיוון שכל מספר הוא חסם מלעיל של $\emptyset$, נהוג לפעמים לרשום $\sup \emptyset = -\infty$.

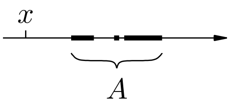
דוגמאות

1. תהי $J=\{x\in\mathbb{R}\,:\,x<0\}$. אז 0 הוא חסם מלעיל של J מעצם ההגדרה. כמו־כן אף מספר $x<0$ אינו חסם מלעיל של J כי אם $x<0$ אז גם $\frac{x}{2}<0$ ולכן $\frac{x}{2}\in J$, אבל $\frac{x}{2}>x$ ולכן x אינו חסם מלעיל של J. מכאן ש־0 הוא החסם מלעיל הקטן ביותר של J ולכן הוא החסם העליון.
2. תהי A קבוצת מספרים ויהי s מקסימום של A. אז s הוא החסם העליון של A, שכן (א) לכל $a\in A$ מתקיים $a\leq s$ ולכן s חסם מלעיל של A, ו־(ב) אם $t<s$ אז t אינו חסם מלעיל של A, כי $s\in A$.
3. לכל קבוצת מספרים סופית יש חסם עליון, כי לכל קבוצה סופית יש מקסימום.

כפי שראינו, משיקולים גאומטריים ופיזיקליים אנו מצפים שלכל קבוצה חסומה מלעיל ולא ריקה יהיה קיים חסם עליון. ואולם, תכונה זו אינה נובעת מהאקסיומות הנתונות, וכדי להבטיחה דרושה אקסיומה נוספת.

אקסיומה XI **אקסיומת החסם העליון**[4] לכל תת־קבוצה לא ריקה וחסומה מלעיל של $\mathbb{R}$ יש חסם עליון.

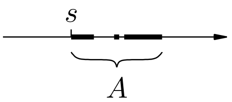
ההעדפה שהפגנו לכיוון החיובי בהגדרות של המושגים של חסם מלעיל וחסם עליון הייתה שרירותית. הנה ההגדרות המקבילות בכיוון ההפוך:



איור 4.2.4 x הוא חסם מלרע של A

הגדרה 4.2.4 תהי $A\subseteq\mathbb{R}$. מספר x נקרא **חסם מלרע** (lower bound) של A אם לכל $a\in A$ מתקיים $x\geq a$. הקבוצה A נקראת **חסומה מלרע** (bounded from below) אם קיים לה חסם מלרע.

קבוצה $A\subseteq\mathbb{R}$ שהיא חסומה גם מלעיל וגם מלרע נקראת **חסומה**.



איור 4.2.5 s הוא חסם תחתון של A

הגדרה 4.2.5 תהי A קבוצה לא ריקה וחסומה מלרע של מספרים. מספר s נקרא **חסם תחתון** (infimum) של A אם הוא החסם מלרע הגדול ביותר של A. כלומר, אם הוא חסם מלרע של A, ולכל חסם מלרע t של A, מתקיים $s\leq t$. החסם התחתון של A, אם הוא קיים, מסומן $\inf A$.[5]

כל ההערות והתכונות שהזכרנו עבור חסמים עליונים וחסמים מלעיל תקפים עבור חסמים תחתונים וחסמים מלרע לאחר השינויים המתבקשים בניסוח.

מסתבר שאין צורך לנסח אקסיומה חדשה שתבטיח קיום של חסמים תחתונים, אלא קיומם נובע מאקסיומת החסם העליון.

למה 4.2.6 תהי $A\subseteq\mathbb{R}$ ויהי x מספר. אז A חסומה מלעיל על ידי x אמ"מ הקבוצה $B=\{-a\,:\,a\in A\}$ חסומה מלרע על ידי $-x$.

---

[4] לאקסיומת החסם העליון קוראים לפעמים **אקסיומת השלמות** (axiom of completeness).
[5] לחסם תחתון קוראים גם greatest lower bound ומסמנים אותו לפעמים ב־g.u.b..

הוכחה כל מה שהלמה אומרת הוא שאם x נמצא מימין לכל נקודה ב־A בציר המספרים אז לאחר שיקוף המחליף ימין ושמאל היחס בין הנקודה לקבוצה מתהפך, כלומר $-x$ נמצא משמאל לכל נקודה בשיקוף של A.

הנה הוכחה מפורטת. נניח ש־A חסומה מלעיל על ידי x, ונראה ש־B חסומה מלרע ע"י $-x$. יהי $b \in B$, עדינו להראות ש־ $-b > -x$. ואמנם, לפי ההגדרה של B יש $a \in A$ כך ש־ $b = -a$. מאחר ש־ $a \leq x$ (כי x חסם מלעיל של A) הרי $b = -a \geq -x$, כפי שרצינו. הכיוון השני של הטענה, דהיינו שאם $-x$ חסם מלרע של $-A$ אז x חסם מלעיל של A, מוכח באופן דומה (השלימו בעצמכם את הפרטים!). ■

מסקנה 4.2.7 יהיו A, B כמו בלמה הקודמת. אז x חסם עליון של A אמ"מ $-x$ חסם תחתון של B.

הוכחה נניח ש־s חסם עליון של A. מהלמה הקודמת נובע ש־$-s$ חסם מלרע של B. יהי x חסם מלרע של B. עלינו להראות ש־ $x \leq -s$. מהלמה הקודמת נובע ש־ $-x$ חסם מלעיל של A, ולכן $s \leq -x$, כלומר $x \leq -s$, כפי שרצינו. הכיוון השני (כלומר, שאם $-s$ חסם תחתון של $-A$ אז s חסם עליון של A) מוכח בצורה דומה. ■

אנו מקבלים מכאן:

משפט 4.2.8 (תכונת החסם התחתון) לכל קבוצה לא ריקה וחסומה מלרע של מספרים ממשיים יש חסם תחתון.

הוכחה תהי B הקבוצה הנתונה. הרעיון הוא לשקף את B ולקבל קבוצה A, למצוא לה חסם עליון, ולשקף חזרה. אחרי שני שיקופים הקבוצה B חוזרת לעצמה, והחסם העליון של A עובר לחסם מלרע של B.

נסמן $A = \{-b \,:\, b \in B\}$. מכיוון ש־B חסומה מלרע לפי למה 4.2.6 A חסומה מלעיל. A אינה ריקה כי B אינה ריקה, ולכן מאקסיומת החסם העליון אנו מסיקים שיש לה חסם עליון s. מהלמה הקודמת נובע ש־$-s$ הוא החסם התחתון של הקבוצה $C = \{-a \,:\, a \in A\}$. אבל קל לוודא ש־ $C = B$ (הוכיחו זאת!) וקיבלנו ש־$-s$ החסם התחתון של B, ובפרט יש לה חסם תחתון. ■

תרגילים

1. הוכיחו שאיחוד, חיתוך והפרש של שתי קבוצות חסומות מלעיל היא קבוצה חסומה מלעיל, וכנ"ל לקבוצות חסומות מלרע.
2. הוכיחו או הפריכו:
   (א) אם $A \subseteq B \subseteq \mathbb{R}$ ואם A חסומה מלעיל אז B חסומה מלעיל.
   (ב) אם $A \subseteq B \subseteq \mathbb{R}$ ואם B חסומה מלרע אז A חסומה מלרע.

(ג) אם A,B קבוצות לא ריקות של מספרים וכל $b \in B$ הוא חסם מלעיל של A, אז כל $a \in A$ הוא חסם מלרע של B.
(ד) A חסומה מלעיל אמ"מ $A \cap \mathbb{Z}$ חסומה מלעיל.

3. תהי A קבוצה ו־s מספר. לכל אחד מהתנאים הבאים קבעו האם הוא גורר, נובע, שקול או סותר את השוויון $s = \sup A$, או שאף אחד מהקשרים אינו חל.
   (א) לכל $a \in A$ מתקיים $s \geq a$.
   (ב) לכל $a \in A$ קיים $t < s$ עם $a < t$.
   (ג) לכל $x > s$ קיים $a \in A$ עם $s < a < x$.
   (ד) לכל תת־קבוצה סופית $B \subseteq A$ מתקיים $s \geq \max B$.
   (ה) s הוא החסם העליון של $A \cup \{s\}$.

4. הוכיחו שאם $A \subseteq \mathbb{R}$ לא ריקה וחסומה מלעיל ומלרע אז $\inf A \leq \sup A$, ויש שוויון אמ"מ A מכילה בדיוק מספר אחד.

5. הוכיחו שאם $A \subseteq B \subseteq \mathbb{R}$ קבוצות לא ריקות וחסומות אז

$$\inf B \leq \inf A \leq \sup A \leq \sup B$$

6. יהיו $A,B \subseteq \mathbb{R}$ קבוצות לא ריקות וחסומות מלעיל, ונניח שלכל $a \in A$ יש $b \in B$ כך ש־ $a \geq b$. הוכיחו כי $\sup B \geq \sup A$.

7. הוכיחו את הטענה הבאה, שהשתמשנו בה בהוכחת משפט 4.2.8. תהי $A \subseteq \mathbb{R}$ לא ריקה, ונסמן $B = \{-a \,:\, a \in A\}$ ו־ $C = \{-b \,:\, b \in B\}$. הראו ש־ $C = A$.

8. בתרגיל זה נראה שאקסיומת החסם העליון שקולה לתכונה שלכל קטע יש נקודות קצה. ליתר דיוק,

הגדרה 4.2.9 קבוצה $I \subseteq \mathbb{R}$ נקראת **מרווח** אם יש לה התכונה שלכל שתי נקודות $x,y \in I$, כל מספר t בין x ל־y שייך ל־I.

   (א) הראו שכל קטע הוא מרווח.
   (ב) יהי $I \subseteq \mathbb{R}$ מרווח חסום (גם מלעיל וגם מלרע). הוכיחו ש־I קטע סופי.
   (ג) נניח שידוע לנו שכל מרווח חסום הוא קטע. הוכיחו שאקסיומת החסם העליון נובעת מהנחה זו.

9. בתרגיל זה נראה שאקסיומת החסם העליון שקולה לטענה שלכל שתי קבוצות $A,B \subseteq \mathbb{R}$, אם A נמצאת כולה "משמאל" ל־B אז קיים מספר המפריד בין A ל־B. באופן מדויק, עבור $A,B \subseteq \mathbb{R}$ נסמן $A \prec B$ אם A,B לא ריקות ולכל $a \in A$, $b \in B$ מתקיים $a \leq b$.

(א) בעזרת אקסיומת החסם העליון, הוכיחו שלכל A, B כך ש־ $A \prec B$ קיים $x_0 \in \mathbb{R}$ כך ש־ $A \prec \{x_0\} \prec B$
(ב) נניח שלכל A, B כך ש־ $A \prec B$ קיים $x_0 \in \mathbb{R}$ כך ש־ $A \prec \{x_0\} \prec B$. הסיקו מכך את אקסיומת החסם העליון (ניתן להסתמך כמובן על כל אקסיומות השדה והסדר אך לא על אקסיומת החסם העליון עצמה!).

4.3 תכונת הארכימדיות וצפיפות המספרים רציונליים

סעיף זה עוסק בקשר בין המערכת $\mathbb{R}$ לתת־המערכות $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$. למשל, אם מתבוננים בקבוצת הנקודות על ציר המספרים המתאימים למספרים השלמים, נדמה שזו אינה קבוצה חסומה. ההוכחה הפורמלית של טענה זו מצריכה את אקסיומת החסם העליון:[6]

משפט 4.3.1 (תכונת הארכימדיות[7]) הקבוצה $\mathbb{N}$ אינה חסומה מלעיל ב־$\mathbb{R}$.

הוכחה נניח בשלילה ש־$\mathbb{N}$ חסומה מלעיל. היא גם אינה ריקה, ולכן יש לה חסם עליון. נסמן אותו ב־s. אז קיים $n \in \mathbb{N}$ כך ש־ $n > s - 1$, כי אחרת $s - 1$ היה חסם מלעיל של $\mathbb{N}$, בסתירה לכך ש־s החסם מלעיל המינימלי. אבל נובע כעת ש־ $s < n + 1$, וכמובן $n + 1 \in \mathbb{N}$. פירוש הדבר ש־s אינו חסם מלעיל של $\mathbb{N}$, בסתירה לאופן שבחרנו את s. מכאן ש־$\mathbb{N}$ אינה חסומה מלעיל. ■

מהמשפט נובע ש־$\mathbb{Z}, \mathbb{Q}$ אינן חסומות מלעיל או מלרע. ההוכחה לכך מושארת כתרגיל.

מסקנה 4.3.2 לכל מספר ממשי $\varepsilon > 0$ קיים מספר טבעי n כך ש־ $0 < \frac{1}{n} < \varepsilon$.

הוכחה יהי $\varepsilon > 0$. אם לכל מספר טבעי $n \in \mathbb{N}$ היה מתקיים $\varepsilon \leq \frac{1}{n}$, אז לכל $n \in \mathbb{N}$ היינו מקבלים $n \leq \frac{1}{\varepsilon}$, כלומר, $\frac{1}{\varepsilon}$ חסם מלעיל של $\mathbb{N}$, בסתירה לתכונת הארכימדיות של הממשיים. ■

מסקנה 4.3.3 אם (a, b) הוא קטע מאורך גדול מ־1 אז הוא מכיל מספר שלם.

הוכחה נתבונן בקבוצת המספרים השלמים שאינם גדולים מ־a:

$$L = \{n \in \mathbb{Z} \,:\, n \leq a\}$$

L אינה ריקה (כי אחרת a חסם מלרע של $\mathbb{Z}$ בסתירה לתכונת הארכימדיות), והיא חסומה מלעיל על ידי a. מארכימדיות נובע שיש מספר טבעי p גדול מ־a, ובפרט

[6] אפשר להוכיח שבלי אקסיומת החסם העליון לא ניתן להוכיח את הטענה.
[7] Archimedes, 287־212 לפנה"ס.

p חסם מלעיל של L. לכן ל־L יש מקסימום (ראו תרגיל (4) בעמוד 49). נסמן $n = \max L$, ונראה ש־ $n+1 \in (a,b)$.

ואמנם, $a < n+1$ כי $n+1 \notin L$ (אילו $n+1 \in L$ היינו מקבלים סתירה לכך ש־$n = \max L$). מצד שני $n \geq a$ ולכן $n+1 \leq a+1$. לבסוף, $b > a+1$ כי (a,b) מאורך גדול מ־1. קיבלנו בסיכום $a < n+1 < b$, כנדרש. ■

המסקנה האחרונה פירושה שהמספרים הרציונליים נמצאים בכל קטע ארוך מספיק, אך האמת היא שהם נמצאים "בכל מקום". באופן מדויק,

הגדרה 4.3.4 קבוצה $A \subseteq \mathbb{R}$ נקראת **צפופה** (dense) ב־$\mathbb{R}$ אם כל קטע לא מנוון מכיל נקודה מ־A (כלומר, אם לכל $x < y$ קיים $a \in A$ בין x ל־y).

שימו לב שהגדרה זו מתארת תכונה שאפשר ליחס לתת־קבוצות של $\mathbb{R}$. אין לבלבל אותה עם צפיפות הסדר, שהיא תכונה של יחס הסדר.

משפט 4.3.5 (צפיפות המספרים הרציונליים) $\mathbb{Q}$ צפופה ב־$\mathbb{R}$.

הוכחה יהיו $a < b$ ונראה שיש $r \in \mathbb{Q}$ עם $r \in (a,b)$. הרעיון הוא לנפח את הקטע מספר שלם של פעמים, עד שאורכו יותר מ־1, ולנצל את העובדה שאז יש בקטע מספר שלם. ואמנם לפי המשפט הארכימדי קיים n טבעי המקיים $n > \frac{1}{b-a}$. אם נגדיר $a' = na\,,\, b' = nb$ אז $a' < b'$ ומתקיים

$$b' - a' = nb - na = n(b-a) > \frac{b-a}{b-a} = 1$$

כלומר, הקטע (a',b') מאורך גדול מ־1. על סמך מסקנה 4.3.3 קיים מספר שלם m בקטע (a',b'). המספר m מקיים $na < m < nb$, ולאחר חלוקה ב־n מקבלים

$$a < \frac{m}{n} < b$$

ו־$\frac{m}{n}$ הוא המספר רציונלי כמבוקש. ■

מסקנה 4.3.6 בכל קטע לא ריק (a,b) יש אינסוף מספרים רציונליים.

הוכחה תהי C קבוצת המספרים הרציונליים ב־(a,b) ונניח בשלילה שהיא סופית. אז ל־C יש מינימום. נסמן $c = \min C$. לפי המשפט האחרון יש מספר רציונלי r בקטע (a,c), ובפרט מתקים $r < c$, בסתירה לכך ש־c הוא הרציונלי הקטן ביותר ב־(a,b). ■

תכונת הארכימדיות מאפשרת לנו להגדיר את פעולת העיגול של מספר. בהינתן מספר x שאינו בהכרח שלם פעולת העיגול נותנת מספר שלם קרוב ל־x.

הגדרה 4.3.7 יהי $x \in \mathbb{R}$. **החלק השלם** (integer part) של x, שמסומן $[x]$, הוא המספר השלם הגדול ביותר שאינו גדול מ־x, כלומר

$$[x] = \max\{n \in \mathbb{Z} : n \leq x\}$$

החלק השבור (fractional part) של x הוא המספר $x - [x]$ ומסומן ב־$\{x\}$.[8]

כדי להצדיק הגדרה זו יש להראות שיש לקבוצה $\{n \in \mathbb{Z} : n \leq x\}$ מקסימום. ההוכחה לכך היא כמו בהוכחת מסקנה 4.3.3.

שימו לב שהחלק השלם של x אינו נמצא בהכרח בין x ל־0. למשל, $[-\frac{1}{2}] = -1$. מאידך, החלק השבור של x מקיים תמיד $0 \leq \{x\} < 1$. האי־שוויון $\{x\} \geq 0$ מידי מההגדרה, והימני נובע מכך שאילו היה מתקיים $\{x\} \geq 1$ אז $x - [x] \geq 1$, ומכאן ש־ $[x] + 1$ הוא מספר טבעי שאינו גדול מ־x, בסתירה להגדרת הערך השלם.

אם רוצים לבחור באופן מפורש אם לעגל את x כלפי מעלה או כלפי מטה מסמנים ב־$\lfloor x \rfloor$ את המספר השלם הגדול ביותר שאינו גדול מ־x (דהיינו את $[x]$) וב־$\lceil x \rceil$ את המספר השלם הקטן ביותר שאינו קטן מ־x. ההוכחה שקיימים מספרים כאלה דומה להוכחה שקיים $[x]$, ומושארת כתרגיל.

תרגילים

1. נמקו בפירוט מדוע העובדה ש־$\mathbb{N}$ אינה חסומה מלעיל ב־$\mathbb{R}$ גוררת ש־$\mathbb{Z}$ ו־$\mathbb{Q}$ אינן חסומות מלעיל או מלרע ב־$\mathbb{R}$.
2. הוכיחו שלכל $x \in \mathbb{R}$ קיים מספר שלם בקטע $[x, x+1)$.
3. הוכיחו שאם $x \geq 0$ ואם לכל n טבעי מתקיים $x < \frac{1}{n}$ אז $x = 0$.
4. אילו מהקבוצות הבאות צפופות?
   (א) $\{\frac{1}{n} : n \in \mathbb{R}\}$.
   (ב) $\{\frac{k}{n} : k, n \in \mathbb{Z}\, ,\, n \neq 0\}$.
   (ג) $\{\pm\frac{1}{r} : r \in \mathbb{Q}\, ,\, 0 < r < 1\}$.
   (ד) קבוצת המספרים בעלי הצגה כשבר עשרוני (דהיינו מספרים מהצורה $\pm b_d \ldots b_1 b_0 . a_1 a_2 \ldots a_n$ כאשר b_i, a_j ספרות עשרוניות).
5. שאלה זו מתייחסת למושגים ערך שלם וערך שבור של מספר. הוכיחו או הפריכו:
   (א) לכל $x, y \in \mathbb{R}$ מתקיים $\{x+y\} = \{x\} + \{y\}$.
   (ב) לכל $x, y \in \mathbb{R}$ מתקיים $[x+y] = [x] + [y]$.

[8] הסימון $\{x\}$ לחלק השבור של מספר דומה לסימון של הקבוצה המכילה את האיבר היחיד x. לרוב אפשר להבדיל לפי ההקשר. במידת הצורך נציין במפורש למה הכוונה. יש ספרים בהם החלק השבור של x מסומן (x).

(ג) לכל $x \in \mathbb{R}$ מתקיים $[x+1] = [x] + 1$.

(ד) לכל k טבעי ולכל $x \geq 0$ מתקיים $[kx] \geq k[x]$.

6. יהיו $n, k \in \mathbb{N}$. בתרגיל (1) בעמוד 48 ראינו שיש מספרים טבעיים יחידים d, r כך ש־ $n = kd + r$ ו־ $0 \leq r < k$. הוכיחו ש־ $r = k \cdot \{\frac{n}{k}\}$ ו־ $d = [\frac{n}{k}]$ (כאן $\{x\}$ ו־$[x]$ מציינים את החלק השבור והחלק השלם של x).

7. הוכיחו שאם $A \subseteq \mathbb{R}$ היא קבוצה צפופה ב־$\mathbb{R}$ אז גם קבוצת ההפרשים $B = \{a' - a'' : a', a'' \in A\}$ צפופה.

4.4 חישוב של חסמים עליונים ותחתונים

בסעיף זה נדגים את החישוב של חסמים עליונים ותחתונים, ונפתח שיטות שיעזרו לנו בכך. נתחיל עם כמה דוגמאות פשוטות:

דוגמאות

1. יהי $A = \{\frac{1}{n} : n \in \mathbb{N}\}$, ונראה ש־ $\inf A = 0$. לכל $a \in A$ מתקיים $0 \leq a$, ולכן 0 הוא חסם מלרע של A. מצד שני, לכל $x > 0$ מסקנה **4.3.2** מבטיחה ש־x איננו חסם מלרע של A, מאחר שיש $n \in \mathbb{N}$ כך ש־ $\frac{1}{n} < x$ (והרי $\frac{1}{n} \in A$). לכן 0 הוא החסם מלרע הגדול ביותר ולכן $\inf A = 0$.[9]

2. יהי $A = \{\frac{1}{2^n} : n \in \mathbb{N}\}$ ונראה ש־ $\inf A = 0$. ואמנם כמו קודם 0 הוא בבירור חסם מלרע של A. כמו כן בהינתן $x > 0$ יש $n \geq 1$ טבעי כך ש־ $\frac{1}{n} < x$, וממילא $\frac{1}{2^n} \leq \frac{1}{n}$ (כי $2^n \geq n$ לכל n טבעי). לכן יש איבר ב־A שקטן מ־x, כלומר x אינו חסם מלרע של A. לכן $\inf A = 0$.

המשפט הבא הוא אפיון של החסם העליון. קיימת גרסה שלו לחסמים תחתונים, שאת ניסוחו והוכחתו אנו משאירים לכם (ראו תרגיל (1) למטה).

משפט 4.4.1 תהי $A \subseteq \mathbb{R}$ קבוצה לא ריקה ו־ $s \in \mathbb{R}$ חסם מלעיל של A. אז s הוא החסם העליון של A אמ"מ לכל מספר $\varepsilon > 0$, קטן ככל שיהיה, קיים איבר $x \in A$ שמקיים $s - \varepsilon < x \leq s$

הוכחה נניח ש־ $s = \sup A$. יהי $\varepsilon > 0$, ונראה שיש $x \in A$ כנדרש. שכן אחרת כל $x \in A$ מקיים $x \leq s - \varepsilon$, ולכן $s - \varepsilon$ חסם מלעיל של A. אבל $\varepsilon > 0$ ולכן $s - \varepsilon < s$, וזו סתירה לכך ש־s הוא החסם העליון של A, כנדרש.

להפך, נניח ש־s חסם מלעיל של A בעל התכונה שלכל $\varepsilon > 0$ יש $x \in A$ שמקיים $s - \varepsilon < x \leq s$. כדי להראות ש־$s$ החסם העליון של A, מספיק שנראה שאם $t < s$

[9] בנימוק זה השתמשנו בתכונת ארכימדיות, שהיא מסקנה מאקסיומת החסם העליון. על סמך אקסיומות השדה בלבד לא ניתן להוכיח ש־ $\inf A = 0$, אך לא נתעמק בנושא.

אז יש מספר $x \in A$ שמקיים $x > t$, כי אז t אינו חסם מלעיל של A. יהי אם כן $t < s$. אז קיים $\varepsilon > 0$ כך ש־ $t \leq s-\varepsilon < s$; זה נכון, למשל, עבור $\varepsilon = s-t$ (בדקו!). לפי ההנחה קיים $x \in A$ שמקיים $s-\varepsilon < x \leq s$, ולכן $x > s-\varepsilon \geq t$ ובפרט $x > t$. מכאן ש־t אינו יכול להיות חסם מלעיל של A, כנדרש. ■

הדרישה בתנאי המשפט ש־s יהיה חסם מלעיל של A אינה מיותרת. אם כל שידוע על s הוא שלכל $\varepsilon > 0$ יש $x \in A$ שמקיים $s-\varepsilon < x \leq s$, אז לא נובע ש־$\sup A = s$ (מצאו דוגמה נגדית!).

דוגמאות

1. יהי $A = (0,1) \cap \mathbb{Q} = \{x \in \mathbb{Q} \,:\, 0 < x < 1\}$, ונראה ש־ $\sup A = 1$. מההגדרה ברור ש־1 חסם מלעיל של A, ואילו אם $\varepsilon > 0$ אז יש מספרים רציונליים בקטע $(1-\varepsilon, 1)$ לפי תכונת הצפיפות של $\mathbb{Q}$ (משפט 4.3.5). לכן לפי המשפט האחרון מתקיים $\sup A = 1$.
2. תהי $A = \{\frac{n^2-n}{n^2} \,:\, n \in \mathbb{N}\}$ ונראה ש־ $\sup A = 1$ (ראשית, חישבו למה ניחשנו ש־1 הוא החסם העליון!). 1 הוא בבירור חסם מלעיל של A ולכן די שנראה שלכל $\varepsilon > 0$ יש n טבעי כך ש־ $\frac{n^2-n}{n^2} > 1-\varepsilon$, או באופן שקול ש־ $n > \frac{1}{\varepsilon}$. זו מסקנה מארכימדיות.

נביא עתה כמה משפטים המאפשרים לחשב חסם עליון (או תחתון) של קבוצה A במקרה ש־A מוגדרת על ידי צירוף מסוים של קבוצות אחרות, וזאת על ידי חישוב חסמים מתאימים של הקבוצות המרכיבות אותה. משפטים מסוג זה יחזרו שוב ושוב בהמשך, והם מכונים **משפטי אריתמטיקה**, או **משפטי תחשיב**.

ראשית נגדיר את הפעולות שיעסיקו אותנו:

הגדרה 4.4.2 יהיו $A, B \subseteq \mathbb{R}$ קבוצות. נגדיר קבוצות חדשות

$$\begin{aligned} -A &= \{-a \,:\, a \in A\} \\ A+B &= \{a+b \,:\, a \in A\,,\, b \in B\} \\ A \cdot B &= \{a \cdot b \,:\, a \in A\,,\, b \in B\} \end{aligned}$$

קל לראות מההגדרה שאם A, B קבוצות לא ריקות אז גם הקבוצות $-A$, $A+B$ ו־$A \cdot B$ אינן ריקות. יתר על כן יש קשר בין החסמים של A, B והחסמים של הקבוצות הנוצרות מהן. הקשר בין חסמים של A ושל $-A$ כבר נוסח בלמה 4.2.6. הלמה הבאה מתייחסת לפעולות האחרות:

למה 4.4.3 יהיו $A, B \subseteq \mathbb{R}$ ויהיו x חסם מלעיל של A ו־y חסם מלעיל של B. אז $x+y$ חסם מלעיל של $A+B$, ואם A, B מכילות רק מספרים אי־שליליים אז $x \cdot y$ הוא חסם מלעיל של AB.

הוכחה נוכיח את הטענה לגבי $A+B$, ואת האחרת נשאיר כתרגיל. יהי $c \in A+B$. עלינו להראות ש־ $c \le x+y$. מהגדרת $A+B$ קיימים $a \in A$, $b \in B$ כך ש־ $c = a+b$. אבל מההנחה $a \le x$, $b \le y$ ולכן $c = a+b \le x+y$. ■

משפט 4.4.4 (תחשיב של חסמים) יהיו $A, B \subseteq \mathbb{R}$ קבוצות לא ריקות.

1. אם A חסומה מלעיל אז $\inf(-A) = -\sup A$.
2. אם A, B חסומות מלעיל אז $\sup(A+B) = \sup A + \sup B$. ואם A, B חסומות מלרע אז $\inf(A+B) = \inf A + \inf B$.
3. אם האיברים של A, B אי־שליליים אז $\sup A \cdot B = \sup A \cdot \sup B$.

הוכחה חלק (1) אינו אלא מסקנה 4.2.7. אנו נוכיח להלן את חלק (2), ואת (3) נשאיר כתרגיל.

יהיו $A, B \subseteq \mathbb{R}$ חסומות מלעיל ולא ריקות. אז קיימים להן חסמים עליונים: נסמן $s = \sup A$ ו־ $t = \sup B$. מכיוון ש־ s, t חסמים מלעיל של A, B בהתאמה נסיק מהלמה ש־ $s+t$ הוא חסם מלעיל של $A+B$. בכדי להראות ש־ $s+t$ הוא חסם עליון של $A+B$ נשתמש באפיון שהוכח במשפט 4.4.1: נראה שלכל מספר $\varepsilon > 0$ יש $c \in A+B$ כך ש־ $c > (s+t) - \varepsilon$.

$A+B$ מכיל את כל האיברים מהצורה $a+b$, כאשר $a \in A$ ו־ $b \in B$. לכן בהינתן מספר $\varepsilon > 0$, מספיק למצוא $a \in A$, $b \in B$ כך ש־

$$a + b > (s+t) - \varepsilon$$

אם נמצא a, b כאלה כך ש־ $a > s - \frac{\varepsilon}{2}$ ו־ $b > t - \frac{\varepsilon}{2}$ אז יתקיים

$$a + b > (s - \frac{\varepsilon}{2}) + (t - \frac{\varepsilon}{2}) = (s+t) - \varepsilon$$

כנדרש. כל שנותר הוא לשים לב שמכיוון ש־ $\frac{\varepsilon}{2} > 0$, ומכיוון ש־ s, t חסמים עליונים של A, B בהתאמה, קיימים (לפי משפט 4.4.1) מספרים כאלה.

הטענה לגבי חסמים תחתונים נובעת ממה שהוכחנו ומהסעיף הראשון של המשפט. אם A, B לא ריקות וחסומות מלרע אז $-A, -B$ לא ריקות וחסומות מלעיל ולכן

$$\sup((-A) + (-B)) = \sup(-A) + \sup(-B) = -\inf A - \inf B$$

מצד שני $(-A) + (-B) = -(A+B)$ (וודאו זאת!), ולכן

$$\sup((-A) + (-B)) = \sup(-(A+B)) = -\inf(A+B)$$

וצירוף שני השוויונות וכפל ב־-1 נותן את השוויון $\inf(A+B) = \inf A + \inf B$, כנדרש. ■

לדוגמה, תהי

$$A = \{\frac{1}{2^n} + \frac{1}{3^m} \,:\, n, m \in \mathbb{N}\}$$

אז $A = A_2 + A_3$ כאשר $A_2 = \{\frac{1}{2^n} \,:\, n \in \mathbb{N}\}$ ו־ $A_3 = \{\frac{1}{3^n} \,:\, n \in \mathbb{N}\}$. ראינו קודם ש־ $\inf A_2 = 0$ ובאותה דרך מראים ש־ $\inf A_3 = 0$, ולכן לפי המשפט האחרון

$$\inf A = \inf(A_2 + A_3) = \inf A_2 + \inf A_3 = 0 + 0 = 0$$

נסיים את הסעיף עם משפט חשוב שיהיה שימושי בהמשך:

משפט 4.4.5 יהיו $A, B \subseteq \mathbb{R}$ קבוצות לא ריקות כך ש־ $a \leq b$ לכל $b \in B$, $a \in A$. אז מתקיים $\sup A \leq \inf B$, והתנאים הבאים שקולים:

1. $\sup A = \inf B$.
2. לכל מספר $\varepsilon > 0$ קיימים $a \in A, b \in B$ כך ש־ $0 \leq b - a < \varepsilon$.
3. יש בדיוק מספר ממשי אחד c כך ש־ $a \leq c \leq b$ לכל $a \in A, b \in B$.

הוכחה מהנתון, כל $b \in B$ הוא חסם מלעיל של A ולכן $\sup A \leq b$. מכאן נובע ש־$\sup A$ הוא חסם מלרע של B, ולכן $\sup A \leq \inf B$.

(1)⇐(2): יהי $\varepsilon > 0$. ממשפט 4.4.1, יש $a \in A$ עם $a > \sup A - \frac{\varepsilon}{2}$, ויש $b \in B$ עם $b < \inf B + \frac{\varepsilon}{2}$. לכן, מאחר ש־ $\inf B - \sup A = 0$ ואם $a \leq b$, מתקיים

$$0 \leq b - a < (\inf B + \frac{\varepsilon}{2}) - (\sup A - \frac{\varepsilon}{2}) = \varepsilon$$

(2)⇐(3): יהיו s, t מספרים כך ש־ $\sup A \leq s, t \leq \inf B$, ונראה ש־ $t = s$. יהי $\varepsilon > 0$ ויהיו $a \in A, b \in B$ כמו ב־(2). אז $a \leq \sup A \leq s, t \leq \inf B \leq b$, ולכן

$$|t - s| \leq b - a < \varepsilon$$

קיבלנו ש־ $|t - s| < \varepsilon$ לכל $\varepsilon > 0$, ולכן $s = t$.

(3)⇐(1): המספרים $s = \sup A$ ו־ $t = \inf B$ מקיימים $a \leq s, t \leq b$ לכל $a \in A, b \in B$ (למה?), ולכן לפי ההנחה מתקיים $s = t$, כנדרש. ■

תרגילים

1. הוכיחו שאם x חסם מלרע של $\emptyset \neq A \subseteq \mathbb{R}$, אז x החסם התחתון של A אמ"מ לכל $\varepsilon > 0$ יש $y \in A$ שמקיים $x \leq y < x + \varepsilon$ (זו גרסה עבור החסם התחתון של משפט 4.4.1).
2. יהיו $A, B \subseteq \mathbb{R}$ לא ריקות, ונניח שלכל $a \in A$ קיים $b \in B$ כך ש־ $a \leq b$, ולכל $b \in B$ קיים $a \in A$ כך ש־ $b \leq a$. הראו ש־A חסומה מלעיל אמ"מ B חסומה מלעיל, ואם הן חסומות מלעיל אז מתקיים $\sup A = \sup B$.
3. חשבו את החסמים העליונים והתחתונים של הקבוצות הבאות:
   (א) $\{\frac{1}{n} + \frac{1}{n^2} \,:\, n \in \mathbb{N}\}$.
   (ב) $\{\frac{1}{n} - \frac{1}{n^2} \,:\, n \in \mathbb{N}\}$.
   (ג) $\{\frac{1}{n} + \frac{1}{m^2} \,:\, n, m \in \mathbb{N}\}$.
   (ד) $\{\frac{1}{2^n} \,:\, n \in \mathbb{N}\}$.
   (ה) $\{\frac{m}{2^n} \,:\, m, n \in \mathbb{N}\}$.
   (ו) $\{r \in \mathbb{Q} \,:\, \exists n \in \mathbb{N}\ \frac{1}{2n+2} < r < \frac{1}{2n+1}\}$.
   (ז) $\{\frac{1}{2^n} \,:\, n \in \mathbb{N}\} \cup \{1 - \frac{1}{3^n} \,:\, n \in \mathbb{N}\}$.
4. הראו שמשפט 4.4.1 נשאר נכון אם מחליפים את התנאי $x - \varepsilon < y \leq x$ בתנאי $x - \varepsilon \leq y \leq x$, אך לא אם מחליפים אותו בתנאי $x - \varepsilon < y < x$ (במקרה האחרון יש למצוא דוגמה לקבוצה חסומה $\emptyset \neq A \subseteq \mathbb{R}$ כך שעבור $x = \sup A$ אין זה נכון שלכל $\varepsilon > 0$ יש $y \in A$ שמקיים $x - \varepsilon < y < x$).
5. תהי $A \subseteq \mathbb{R}$ קבוצה לא ריקה וחסומה מלעיל, ונניח שאין לה מקסימום. תהי B קבוצה סופית של מספרים. הראו ש־ $\sup A = \sup(A \setminus B)$.
6. יהיו $A, B \subseteq \mathbb{R}$. הוכיחו:
   (א) אם A, B לא ריקות וחסומות מלעיל אז כך גם $A \cup B$ לא ריקה וחסומה מלעיל, ומתקיים $\sup(A \cup B) = \max\{\sup A, \sup B\}$.
   (ב) אם A, B לא ריקות וחסומות מלעיל אז $A \cap B$ חסומה מלעיל, ואם היא לא ריקה אז $\sup(A \cap B) \leq \min\{\sup A, \sup B\}$.
7. יהיו $A, B \subseteq \mathbb{R}$ קבוצות לא ריקות וחסומות מלעיל של מספרים. הראו שאם $A + B = A$ אז $B = \{0\}$.
8. יהיו $A, B \subseteq \mathbb{R}$. נגדיר קבוצה $A - B$ על ידי

$$A - B = \{a - b \,:\, a \in A \,,\, b \in B\}$$

הוכיחו שאם A, B לא ריקות אז $A - B$ לא ריקה ושאם A חסומה מלעיל ו־B חסומה מלרע אז $A - B$ חסומה מלעיל ומתקיים

$$\sup(A - B) = \sup A - \inf B$$

נסחו והוכיחו טענה דומה עבור $\inf(A - B)$.

9. הוכיחו את סעיף (3) של משפט 4.4.4: אם $A, B \subseteq [0, \infty)$ לא ריקות וחסומות מלעיל אז $A \cdot B$ לא ריקה וחסומה מלעיל, ומתקיים

$$\sup(A \cdot B) = \sup A \cdot \sup B$$

הראו שהטענה אינה בהכרח נכונה אם A, B מכילות מספרים שליליים. נסחו והוכיחו טענה דומה עבור המקרה ששתיהן מכילות רק מספרים שליליים. (רמז: כדי להוכיח את השוויון חקו את ההוכחה של סעיף (2) במשפט 4.4.4).

4.5 המספר הממשי $\sqrt{2}$ והמספרים האי־רציונליים

בסעיף זה נראה שבמערכת המספרים הממשיים קיים מספר שריבועו 2. מספר זה מסומן $\sqrt{2}$ ונקרא **שורש ריבועי** של 2. קיומו מספק הוכחה לכך ש־ $\mathbb{Q} \neq \mathbb{R}$, כי ראינו בתחילת הפרק שאין מספר כזה ב־$\mathbb{Q}$.

משפט 4.5.1 קיים מספר חיובי r המקיים $r^2 = 2$.

הוכחה נגדיר קבוצה L על ידי

$$L = \{x \in \mathbb{R} \,:\, x \geq 0\,,\, x^2 \leq 2\}$$

בדיעבד יסתבר של־L יש מקסימום (הוא המספר המבוקש) אך איננו יודעים זאת מראש. L אינה ריקה כי $0 \in L$, והיא חסומה מלעיל על ידי 2, כי אם $x > 2$ אז $x^2 > 4 > 2$ וממילא $x \notin L$. לכן קיים ל־L חסם עליון. נסמן אותו ב־r. אנו טוענים ש־ $r^2 = 2$. כדי להוכיח זאת נוכיח את שני האי־שוויונות $r^2 \leq 2$ ו־ $r^2 \geq 2$.

נניח בשלילה ש־ $r^2 > 2$, ונראה שיש מספר חיובי קטן t כך ש־ $(r-t)^2 > 2$. מכאן ינבע שאין ב־L מספרים גדולים מ־$r - t$ (כי אם $s > r - t$ אז $s^2 > (r-t)^2 > 2$, ולכן $s \notin L$), ולכן $r - t$ הוא חסם מלעיל של L. אבל $r - t$ קטן יותר מ־r, וזו תהיה סתירה.

נותר להראות שתחת הנחת השלילה קיים t כזה. ואמנם, לכל t מתקיים

$$(r-t)^2 = r^2 - 2rt + t^2 > r^2 - 2rt$$

ולכן אם $0 < t < \frac{r^2-2}{2r}$ מתקיים $(r-t)^2 > 2$. שימו לב ש־ $\frac{r^2-2}{2r} > 0$ כי הנחנו $r^2 > 2$, ולכן יש t כזה, כנדרש.

בכיוון השני, נניח בשלילה ש־ $r^2 < 2$, ונראה שעבור מספר חיובי וקטן דיו t יתקיים גם $(r+t)^2 < 2$. זה לא ייתכן כי פירוש הדבר ש־ $r + t \in L$, בסתירה לכך ש־r חסם מלעיל של L.

ואמנם לכל $t > 0$ מתקיים

$$(r+t)^2 = r^2 + 2rt + t^2 = r^2 + t(2r+t)$$

אם נניח גם ש־ $0 < t < r$ אנו מסיקים ש־

$$(r+t)^2 < r^2 + t(2r+r) = r^2 + 3rt$$

ולכן אם יתקיים בנוסף $t < \frac{2-r^2}{3r}$, יתקיים $(r+t)^2 < 2$. אם כן אנו מחפשים מספר חיובי t שיקיים $t < r$ וגם $t < \frac{2-r^2}{3r}$. מכיוון שהנחנו ש־ $r^2 < 2$, שני המספרים באגף ימין של האי־שוויונות חיוביים, ולכן יש t כזה. ■

לא קשה להכליל את המשפט ולהראות שלכל $a > 0$ קיים מספר $r > 0$ כך ש־ $r^2 = a$ (כדאי להוכיח זאת!). בסעיף הבא נוכיח טענה כללית עוד יותר.

בסעיף 4.1 ראינו ש־$\sqrt{2}$ אינו רציונלי. לכן יש טעם בהגדרה הבאה:

הגדרה 4.5.2 מספר ממשי שאינו מספר רציונלי נקרא מספר **אי רציונלי** (irrational).

מסתבר שיש הרבה מאד מספרים אי־רציונליים. במובן מסוים, רוב המספרים הממשיים הם כאלה. ניתן לכך הוכחה בהמשך (סעיף 5.12 בפרק הבא). בינתיים נסתפק בהוכחה שהמספרים האי־רציונליים צפופים. לשם כך ניעזר באבחנה שאם $x \in \mathbb{R}$ הוא מספר אי רציונלי ו־ $0 \neq r \in \mathbb{Q}$ אז $x + r$ אי־רציונלי. ואמנם, נסמן $s = x + r$. אילו $s \in \mathbb{Q}$ היינו מקבלים $x = s - r$, כלומר x הוא הפרש של מספרים רציונליים ולכן הוא בעצמו רציונלי, בסתירה להנחה.

משפט 4.5.3 (צפיפות האי־רציונליים) המספרים האי־רציונליים צפופים ב־$\mathbb{R}$.

הוכחה די להראות שהקבוצה $A = \{\sqrt{2} + r \ : \ r \in \mathbb{Q}\}$ צפופה, כי מהדיון שקדם למשפט נובע שכל איברי A הם אי־רציונליים. יהיו $a < b$. עלינו למצוא $x \in A$ כך ש־ $a < x < b$, כלומר עלינו למצוא $r \in \mathbb{Q}$ כך ש־ $a < r + \sqrt{2} < b$. לשם כך די למצוא $r \in \mathbb{Q}$ כך ש־ $a - \sqrt{2} < r < b - \sqrt{2}$. אבל $a - \sqrt{2} < b - \sqrt{2}$, ולכן קיומו של r כזה נובע מצפיפות הרציונליים (משפט 4.3.5). ■

תרגילים

1. הוכיחו שקיים שורש ריבועי לכל מספר חיובי.
2. האם סכום של מספרים אי־רציונליים הוא אי־רציונלי? מה לגבי מכפלה?
3. הוכיחו שלמשוואה $ax^2 + bx + c = 0$ יש פתרון אמ"מ $b^2 - 4ac \geq 0$. אם יש שוויון אז הפתרון היחיד הוא $-\frac{b}{2a}$ ואחרת יש שני פתרונות, $\frac{-b+\sqrt{b^2-4ac}}{2a}$ ו־$\frac{-b-\sqrt{b^2-4ac}}{2a}$.

4. המספר $\sqrt{2} \in \mathbb{R}$ מראה ש־ $\mathbb{R} \neq \mathbb{Q}$. בשאלה זו נראה שיש מערכות ביניים, כלומר קבוצת מספרים F כך ש־ $\mathbb{Q} \subseteq F \subseteq \mathbb{R}$ ו־ $F \neq \mathbb{Q}, \mathbb{R}$, אשר מקיימות את אקסיומות השדה הסדור. נגדיר

$$F = \{s + t\sqrt{2} \,:\, s, t \in \mathbb{Q}\}$$

זו תת־קבוצה של $\mathbb{R}$ אשר מכילה את $\mathbb{Q}$ (למה?).

(א) הוכיחו שאם $x\,,\,y \in F$ אז גם $x + y\,,\,x \cdot y \in F$, וש־$F$ מקיימת את אקסיומות הנגדי וההפכי. הסיקו ש־F היא שדה סדור.

(ב) הוכיחו ש־$\sqrt{3} \notin F$, ולכן $F \neq \mathbb{R}$.

5. הוכיחו את **אי־שוויון קושי־שוורץ** (Cauchy-Schwartz inequality): לכל שתי n־יות מספרים, $a_1, \ldots, a_n$ ו־$b_1, \ldots, b_n$, מתקיים

$$\sum_{i=1}^{n} |a_i b_i| \leq \sqrt{\sum_{i=1}^{n} a_i^2} \cdot \sqrt{\sum_{i=1}^{n} b_i^2}$$

ויש שוויון אמ"מ $a_i = b_i$ לכל $i = 1, 2, \ldots, n$ (רמז: הראו שדי להוכיח זאת תחת ההנחה שהמספרים a_i, b_i אי־שליליים וש־ $\sum_{i=1}^{n} a_i^2 = \sum_{i=1}^{n} b_i^2 = 1$. תחת הנחה זו האי־שוויון הופך ל־ $\sum_{i=1}^{n} a_i b_i \leq 1$. כדי להוכיח אותו היעזרו בתרגיל (5(ג)) מעמוד 35. שימו לב שהטענה כאן מכלילה את תרגיל (5(ד)) מאותו עמוד).

4.6 פעולת החזקה

על המספרים הממשיים מוגדרות שתי פעולות בסיסיות: חיבור וכפל. בעזרת הנגדי וההפכי הגדרנו שתי פעולות נוספות: חיסור וחילוק (חילוק היא "פעולה חלקית" במובן זה שהיא אינה תמיד מוגדרת). בסעיף זה נשלים את התמונה ונגדיר את הפעולה (החלקית) העלאה בחזקה.

מטרתנו היא הוכחת המשפט הבא:

משפט 4.6.1 לכל מספר $a > 0$ ולכל $x \in \mathbb{R}$ קיים מספר ממשי a^x שנקרא a **בחזקת** x. המספר a נקרא **הבסיס**, והמספר x נקרא **המעריך**, של הביטוי a^x. הפעולה מוגדרת כך שלכל $a, b > 0$ ולכל $x, y \in \mathbb{R}$ מתקיים:

1. $a^x > 0$, $a^0 = 1$ ו־ $a^1 = a$.
2. $a^{x+y} = a^x \cdot a^y$.
3. $a^{xy} = (a^x)^y$.
4. $(ab)^x = a^x b^x$.

5. בהנחה $0<a<b$, אם $x>0$ אז $a^x<b^x$ ואם $x<0$ אז $a^x>b^x$.

6. בהנחה $x<y$, אם $a>1$ אז $a^x<a^y$ ואם $0<a<1$ אז $a^x>a^y$.

כמו־כן המספר 0^x מוגדר להיות 0 לכל $x>0$, ואילו $0^0=1$. סעיפים (2),(3) ו־(4) מתקיימים גם כאשר $a=0$ או $b=0$, בתנאי שכל החזקות בהן מוגדרות.

להוכחת המשפט יש שני חלקים. ראשית יש להגדיר את המספר a^x, ואז יש לבדוק שההגדרה מקיימת את התכונות האמורות. ההוכחה תתבצע בשבים. עבור $a>0$ נגדיר את הביטוי a^x תחילה כאשר x הוא מספר טבעי, לאחר מכן עבור x שלם, עבור המקרה $x=\frac{1}{n}$ (n טבעי), ולבסוף עבור x רציונלי כללי. את המקרה של מעריך ממשי כללי נדחה לפרק הבא, ועד שנגדיר אותו נעסוק רק בחזקות עם מעריך רציונלי.

חזקות עם מעריך טבעי

עבור $a\in\mathbb{R}$ ו־n טבעי, החזקה a^n מוגדרת בתור המכפלה של a עם עצמו n פעמים, כאשר $a^0=1$. הגדרה מדויקת ניתנה בהגדרה 3.6.1 והתכונות הוכחו בתרגיל (7) בעמוד 55. בשלב זה לא היה צורך להניח ש־a חיובי.

חזקות עם מעריך שלם

את החזקה a^n עם מעריך שלם נגדיר לא רק עבור בסיס חיובי אלא לכל בסיס $a\neq 0$. אפשר להגדיר את a^n על ידי הפרדה למקרים על פי הסימן של n, אך אנו נעדיף את ההגדרה הבאה. הזכרו בדיון שאחרי הגדרה 3.4.6, לפיו כל מספר שלם הוא הפרש של טבעיים.

הגדרה 4.6.2 יהי $a\neq 0$ ו־ $n\in\mathbb{Z}$, ונרשום $n=i-j$, כאשר $i,j\in\mathbb{N}$. המספר a^n מוגדר על ידי $a^n=\frac{a^i}{a^j}$, כשהמספרים a^i,a^j במונה ובמכנה מציינים חזקות טבעיות כפי שהוגדרו קודם.

שימו לב שהמכנה a^j בהגדרה שונה מאפס לפי תכונות החזקה עם מעריך טבעי.

לכל מספר שלם יש הצגות רבות כהפרש של טבעיים (למשל את 1 אפשר להציג בתור $1-0=2-1=3-2=4-3=\ldots$), ובהגדרה לא בחרנו הצגה מסוימת כזו. כדי שההגדרה תהיה חד משמעית עלינו להראות שהמספר $\frac{a^i}{a^j}$ אינו משתנה אם בוחרים זוג i,j אחר בעל אותו הפרש. ליתר דיוק:

למה 4.6.3 יהי $a\neq 0$ ו־i,j טבעיים. אז המנה $\frac{a^i}{a^j}$ תלויה רק בהפרש $i-j$ ולא במספרים i,j עצמם. בפרט, הגדרת החזקה השלמה היא חד משמעית.

הוכחה יהיו $i,j,i',j'\in\mathbb{N}$ ונניח $i-j=i'-j'$. עלינו להראות ש־ $a^i/a^j=a^{i'}/a^{j'}$, או, באופן שקול, ש־ $a^ia^{j'}=a^ja^{i'}$ (מדובר כאן בחזקות עם מעריך טבעי). מתכונות החזקה למעריכים טבעיים השוויון שקול לשוויון $a^{i+j'}=a^{j+i'}$, וזה נובע מכך ש־ $i+j'=j+i'$. ■

סוגיה נוספת שיש לתת עליה את הדעת היא שכל מספר טבעי הוא גם שלם, ולכן הסימן a^n מוגדר כעת בשתי דרכים: פעם כחזקה עם מעריך טבעי, ופעם כחזקה עם מעריך שלם. יש לבדוק ששתי ההגדרות מתלכדות, אחרת אין הצדקה להשתמש באותו סימן לשתיהן. ואמנם, אם $n\in\mathbb{N}$ אז אפשר לכתוב את n כהפרש של טבעיים בצורה $n=n-0$, ולכן החזקה השלמה a^n שווה למנה של חזקות טבעיות a^n/a^0, ומאחר ש־ $a^0=1$ אנו רואים שההגדרות באמת מתלכדות.

שימו לב שהסימן a^{-1}, שציין עד כה את האיבר ההפכי של a, קיבל גם הוא משמעות חדשה: a בחזקת -1. ואולם בפירוש החדש מתקיים $a^{-1}=a^{0-1}=a^0/a^1=1/a$, דהיינו שתי ההגדרות מתלכדות, ושוב אין צורך להבחין בין השניים.

טענה 4.6.4 משפט 4.6.1 מתקיים לחזקות שלמות.

הוכחה נוכיח למשל את השוויון $a^{m+n}=a^ma^n$. נרשום $m=i-j\,,\,n=i'-j'$ כאשר $i,j,i',j'\in\mathbb{N}$. אז מתקיים

$$a^m\cdot a^n=\frac{a^i}{a^j}\cdot\frac{a^{i'}}{a^{j'}}=\frac{a^i\cdot a^{i'}}{a^j\cdot a^{j'}}=\frac{a^{i+i'}}{a^{j+j'}}=a^{m+n}$$

(השוויון הראשון נובע מהגדרת החזקה עם מעריך שלם, השני מכללי החשבון של מנות, השלישי מתכונות החזקה עם מעריך טבעי, והרביעי שוב מהגדרת החזקה השלמה, כי $i+i',j+j'\in\mathbb{N}$ ומתקיים $m+n=(i+i')-(j+j')$)

יתר הסעיפים נובעים באופן דומה. הוכיחו אותם בעצמכם! ■

שורשים מסדר n

כבר הגדרנו שורשים ריבועיים והראינו את קיומם (סעיף 4.5). נטפל עתה בשורשים כלליים. בהינתן $a>0$ ו־ $n>0$ טבעי אנו מעוניינים למצוא מספר y כך ש־ $y^n=x$. לפני שנראה שיש מספר כזה נטפל בבעיית היחידות. יש מקרים שאין y יחיד כזה. למשל, קיימים שני מספרים y המקיימים $y^2=2$: המספר $\sqrt{2}$ שמצאנו בסעיף 4.5, והנגדי שלו, $-\sqrt{2}$. אולם אם נחפש שורש n־י חיובי, קיים לכל היותר מועמד אחד:

טענה 4.6.5 יהי $a>0$ ו־n טבעי. אם x,y מספרים חיוביים המקיימים $x^n=y^n=a$ אז $x=y$.

הוכחה אילו היה מתקיים $x\neq y$ אז או ש־ $x<y$ או ש־ $y<x$. לפי סעיף (5) של משפט 4.6.1 למעריכים טבעיים, במקרה הראשון מתקיים $x^n<y^n$ ובשני מתקיים $y^n<x^n$; ובכל מקרה $x^n\neq y^n$, בסתירה לנתון. לכן $x=y$. ■

ניגש עתה לשאלת הקיום של שורשים. ההוכחה דומה מאד להוכחה של קיום שורשים ריבועיים. נזדקק ללמה הטכנית הבאה:

למה 4.6.6 לכל $r > 1$ ו־ $n \in \mathbb{N}$ קיים $x > 1$ כך ש־ $x^n < r$.

הוכחה יהי $r > 1$. נשים לב שלכל $x > 1$ מתקיים $x^n < x^{(2^n)}$, וזאת משום ש־ $n < 2^n$ לכל n טבעי ולפי תכונות החזקה עם מעריך טבעי. לכן אם נראה שלכל n יש מספר $x > 1$ כך ש־ $x^{(2^n)} < r$ הרי שסיימנו.

את הטענה האחרונה נוכיח באינדוקציה, כלומר: נראה באינדוקציה על n שלכל $r > 1$ יש $x > 1$ כך ש־ $x^{(2^n)} < r$.

במקרה $n = 0$ הטענה היא שבהינתן $r > 1$ יש $x > 1$ עם $x^1 = x < r$, וזה מידי: נבחר למשל $x = \frac{1+r}{2}$.

נניח שהוכחנו את הטענה ל־n. יהי $r > 1$, ונרצה למצוא $x > 1$ עם $x^{(2^{n+1})} < r$. נשים לב ש־

$$x^{(2^{n+1})} = x^{(2^n \cdot 2)} = (x^{(2^n)})^2$$

ולכן אם נמצא $x > 1$ שמקיים $x^{(2^n)} < \sqrt{r}$ אז העלאת שני האגפים בריבוע ייתן $x^{(2^{n+1})} < r$, כנדרש. קיומו של x כזה נובע מהנחת האינדוקציה המופעלת על $\sqrt{r}$ במקום r. נציין שקיום המספר $\sqrt{r}$ מוכח כמו במשפט 4.5.1, ושמתקיים $\sqrt{r} > 1$ כי אחרת $0 < \sqrt{r} \leq 1$ ולכן $r^2 \leq 1^2 = 1$, בסתירה לנתון. ∎

מסקנה 4.6.7 אם $0 < r < 1$ ו־ $n \in \mathbb{N}$ קיים $0 < x < 1$ כך ש־ $x^n > r$.

הוכיחו את המסקנה!

משפט 4.6.8 לכל $a > 0$ ו־ $n > 0$ קיים מספר ממשי חיובי r כך ש־ $r^n = a$.

הוכחה ההוכחה שונה בפרטיה מהוכחת קיומו של שורש ריבועי ל־2, אך הרעיון דומה. כמו בהוכחת הקיום של שורש ריבועי, נתבונן בקבוצה

$$L = \{x \in \mathbb{R} : x \geq 0 \wedge x^n \leq a\}$$

L אינה ריקה והיא חסומה מלעיל, למשל על ידי $a + 1$ (הוכיחו!). לכן יש לה חסם עליון. נסמן אותו ב־r. אנו נראה ש־ $r^n = a$. לשם כך נראה שכל אחד מהאי־שוויונות $r^n > a$ ו־ $r^n < a$ מובילות לסתירה.

נניח בשלילה ש־ $r^n < a$. הסתירה תתקבל מכך שנראה שניתן להגדיל קצת את r ולקבל מספר $s > r$ שעדיין מקיים $s^n \leq a$, וזה בלתי אפשרי, כי אז $s \in L$ בסתירה להגדרת r. כדי למצוא s כזה די למצוא מספר $\theta > 1$ כך ש־ $(\theta r)^n \leq a$, כי אז $s = \theta r$ הוא המספר שייתן את הסתירה. כדי למצוא θ כזה, נשים לב שהאי־שוויון

$(\theta r)^n \le a$ שקול לאי־שוויון $\theta^n \le \frac{a}{r^n}$, ומההנחה $r^n < a$ נובע ש־ $\frac{a}{r^n} > 1$. לכן θ כזה קיים לפי למה 4.6.6, והגענו לסתירה.

נניח בשלילה ש־ $r^n > a$. הסתירה תתקבל מכך שנראה שניתן להקטין מעט את r ולקבל מספר $s < r$ שעדיין מקיים $s^n > a$, ומכאן שגם s חסם מלעיל של L (למה?), בניגוד להגדרת r. לשם כך די למצוא מספר $0 < \theta < 1$ כך ש־ $(\theta r)^n > a$, כי אז $s = \theta r$ הוא המספר שייתן את הסתירה. נותר להראות שקיים θ כזה. אנו רוצים שיתקיים $0 < \theta < 1$ וגם $(\theta r)^n > a$. האי־שוויון האחרון שקול ל־ $\theta^n > \frac{a}{r^n}$. לפי ההנחה, $r^n > a$ ולכן $\frac{a}{r^n} < 1$. לכן קיומו של θ נובע מהמסקנה 4.6.7, והגענו לסתירה.

בסיכום, קיבלנו $r^n = a$, כנדרש. ∎

יחד עם טענה 4.6.5, המשפט האחרון מצדיק את ההגדרה הבאה:

הגדרה 4.6.9 יהי $a > 0$ ו־ $n \in \mathbb{N}$. **השורש ה־n־י** (n-th root) **של** a הוא המספר החיובי היחיד x המקיים $x^n = a$, והוא מסומן $\sqrt[n]{a}$.

פעולת הוצאת השורש מקיימת תכונות אלגבריות דומות להעלאה בחזקה.

משפט 4.6.10 לכל $a, b > 0$ ו־$m, n > 0$ טבעיים מתקיים

1. $\sqrt[m]{\sqrt[n]{a}} = \sqrt[mn]{a}$.
2. $\sqrt[m]{ab} = \sqrt[m]{a} \cdot \sqrt[m]{b}$.
3. אם $a < b$ אז $\sqrt[n]{a} < \sqrt[n]{b}$.
4. אם $a > 1$ ו־ $m < n$ אז $\sqrt[m]{a} > \sqrt[n]{a}$. אם $0 < a < 1$ אז $\sqrt[m]{a} < \sqrt[n]{a}$.

הוכחה התכונה (1) נובעת כמעט מיד מהגדרת השורש. נסמן $x = \sqrt[m]{\sqrt[n]{a}}$. על מנת להראות ש־$x$ הוא שורש mn־י של a, מספיק להראות ש־ $x^{mn} = a$. ואמנם מתכונות החזקה השלמה מקבלים

$$x^{mn} = (x^m)^n = \left(\left(\sqrt[m]{\sqrt[n]{a}}\right)^m\right)^n =$$

מכיוון שלכל $z > 0$ מתקיים מההגדרה $(\sqrt[m]{z})^m = z$, זה נכון בפרט ל־ $z = \sqrt[n]{a}$, לכן

$$= (\sqrt[n]{a})^n = a$$

בשביל (2) מספיק לשים לב ש־

$$(\sqrt[m]{a} \cdot \sqrt[m]{b})^m = (\sqrt[m]{a})^m \cdot (\sqrt[m]{b})^m = a \cdot b$$

לגבי (3), אילו היה מתקיים $\sqrt[n]{b} \le \sqrt[n]{a}$ אז היינו מקבלים מתכונות החזקה השלמה ש־

$$b = (\sqrt[n]{b})^n \le (\sqrt[n]{a})^n = a$$

בסתירה להנחה $a < b$.

עבור (4), נוכיח את המקרה $a < 1$, ואת המקרה $a > 1$ נשאיר כתרגיל. נשים לב שהאי־שוויון $n < m$ גורר, לפי תכונות החזקה השלמה, את האי־שוויון $a^n < a^m$. לכן מסעיפים (2) ו־(1) נובע

$$\sqrt[m]{a} = \sqrt[m]{\sqrt[n]{a^n}} = \sqrt[mn]{a^n} < \sqrt[mn]{a^m} = \sqrt[n]{\sqrt[m]{a^m}} = \sqrt[n]{a}$$

כנדרש (הבהירו לעצמכם על מה מסתמך כל שלב באי־שוויון האחרון!). ■

חזקות עם מעריך רציונלי

הגדרה 4.6.11 יהי $a > 0$ ו־ $r \in \mathbb{Q}$. יהי m שלם ו־n טבעי כך ש־ $r = \frac{m}{n}$. אז המספר a^r מוגדר על ידי $a^r = (\sqrt[n]{a})^m$.

כמו בהגדרה של חזקה עם מעריך שלם, לפני שנקבל את ההגדרה עלינו לבדוק שהמספר a^r שהוגדר אינו תלוי בבחירת המונה והמכנה (שהרי לכל מספר רציונלי יש הצגות רבות כמנה של שלם וטבעי). ננסח זאת כטענה:

טענה 4.6.12 בהינתן $a > 0$, הביטוי $(\sqrt[n]{a})^m$ תלוי רק במספר הרציונלי $\frac{m}{n}$ ולא בזוג המספרים m, n, ולכן המספר $x^{m/n}$ מוגדר היטב (כלומר, בצורה חד־משמעית).

הוכחה נניח $r = \frac{m}{n}$ עבור m שלם ו־n טבעי. ראשית יהי k טבעי, כך שמתקיים גם $r = \frac{mk}{nk}$, ונראה ש־ $(\sqrt[n]{a})^m = (\sqrt[nk]{a})^{mk}$ (אלה הביטויים המגדירים את a^r כשמציגים את r בדרכים $\frac{m}{n}, \frac{mk}{nk}$). ואמנם, נשים לב שמתכונות השורש והחזקה השלמה מתקיים

$$(\sqrt[kn]{a})^{km} = (\sqrt[k]{\sqrt[n]{a}})^{km} = ((\sqrt[k]{\sqrt[n]{a}})^k)^m = (\sqrt[n]{a})^m$$

כאשר השוויון האחרון נובע מכך ש־ $(\sqrt[k]{y})^k = y$ לכל k טבעי, ובפרט עבור $y = \sqrt[n]{a}$ (הצדיקו את יתר השוויונות!).

כעת למקרה הכללי: יהיו $\frac{m}{n}$ ו־$\frac{m'}{n'}$ שתי הצגות של אותו מספר רציונלי r, ונרצה להראות ש־ $(\sqrt[n]{a})^m = (\sqrt[n']{a})^{m'}$. נעבור למכנה משותף ונכתוב $\frac{m}{n} = \frac{mn'}{nn'}$, $\frac{m'}{n'} = \frac{m'n}{n'n}$. מהאמור למעלה, מתקיים

$$\begin{aligned} (\sqrt[nn']{a})^{mn'} &= (\sqrt[n]{a})^m \\ (\sqrt[nn']{a})^{m'n} &= (\sqrt[n']{a})^{m'} \end{aligned}$$

מכיוון ש־ $\frac{m}{n}=\frac{m'}{n'}$ הרי ש־ $mn'=n'm$, ולכן אגפי שמאל של השוויונות האחרונים שווים. לכן גם אגפי ימין שווים, ומקבלים ש־ $(\sqrt[n']{a})^{m'}=(\sqrt[n]{a})^m$, שהוא השוויון המבוקש. ■

מספר שלם הוא בפרט רציונלי, ולכן אם n שלם אז a^n מוגדר כעת בשתי דרכים: כחזקה שלמה וכחזקה רציונלית. יש לוודא שההגדרות מתלכדות. ואמנם, אם $n\in\mathbb{Z}$ אז $n=n/1$, ולכן לפי ההגדרה של חזקה רציונלית,

$$a^{n/1}=(\sqrt[1]{a})^n=a^n$$

החזקה באגף ימין היא חזקה עם מעריך שלם במובן הישן, כפי שרצינו. שימו לב גם שהסימנים $a^{1/n}$ ו־$\sqrt[n]{a}$ מציינים את אותו מספר, כי

$$a^{1/n}=(\sqrt[n]{a})^1=\sqrt[n]{a}$$

מסקנה 4.6.13 לכל $a,b>0$ ו־$m,n>0$ טבעיים מתקיים

1. $(a^{1/m})^{1/n}=a^{1/mn}$
2. $(ab)^{1/m}=a^{1/m}\cdot b^{1/m}$
3. אם $a<b$ אז $a^{1/m}<b^{1/m}$
4. אם $a>1$ ו־ $m<n$ אז $a^{1/m}>a^{1/n}$. אם $0<a<1$ אז $a^{1/m}<a^{1/n}$

המשפט אינו אלא ניסוח מחדש של משפט 4.6.10, והוא כולל חלק מסעיפי משפט 4.6.1 למקרה של מעריכים מהצורה $1/k$. לגבי מעריכים רציונליים כלליים, נוכיח קודם למה.

למה 4.6.14 אם $a>0$ ו־ $n\in\mathbb{N}$, $m\in\mathbb{Z}$ אז $a^{m/n}=(a^m)^{1/n}$

הוכחה די להראות ש־ $\left(a^{m/n}\right)^n=a^m$. ואמנם

$$\left(a^{m/n}\right)^n=\left((a^{1/n})^m\right)^n=(a^{1/n})^{mn}=\left((a^{1/n})^n\right)^m=a^m$$

■

משפט 4.6.15 משפט 4.6.1 נכון למעריכים רציונליים.

הוכחה נוכיח למשל שלמספרים רציונליים r,s ול־ $a>0$ מתקיים $(a^r)^s=a^{rs}$. נבחר הצגה של $s=\frac{s_1}{s_2}$, $r=\frac{r_1}{r_2}$ כאשר $r_1,s_1\in\mathbb{Z}$ ו־ $r_2,s_2\in\mathbb{N}$ ונסמן $b=a^{1/r_2}$. כעת

$$(a^r)^s=(a^{r_1/r_2})^{s_1/s_2}=(b^{r_1})^{s_1/s_2}=\left((b^{r_1})^{1/s_2}\right)^{s_1}=$$

על סמך הלמה

$$= \left((b^{1/s_2})^{r_1}\right)^{s_1} = (b^{1/s_2})^{r_1 s_1}$$

וכעת אם נשים לב שלפי מסקנה 4.6.13 מתקיים $b^{1/s_2} = (a^{1/r_2})^{1/s_2} = a^{1/r_2 s_2}$ הרי שקיבלנו־

$$(a^r)^s = (a^{1/r_2 s_2})^{r_1 s_1} = a^{r_1 s_1 / r_2 s_2} = a^{rs}$$

כנדרש.

באותה דרך ניתן להוכיח את שאר הסעיפים. אנו משאירים זאת כתרגיל. ■

תרגילים

1. השלימו את הפרטים שהושמטו מהטענות במהלך הסעיף.
2. בשאלה זו נוכיח את **אי־שוויון הממוצעים**[10] (arithmetic-geometric in-equality): אם $a_1, \ldots, a_n$ מספרים חיוביים אז

$$\frac{n}{1/a_1 + 1/a_2 + \ldots + 1/a_n} \quad \le \quad \sqrt[n]{a_1 \cdot \ldots \cdot a_n} \quad \le \quad \frac{1}{n}(a_1 + \ldots + a_n)$$

והאי־שוויונות חזקים אלא אם $a_1 = a_2 = \ldots = a_n$.

(א) הוכיחו באינדוקציה על n שלכל $a_1, \ldots, a_n \ge 0$, אם $a_1 \cdot a_2 \ldots \cdot a_n = 1$ אז $a_1 + a_2 + \ldots + a_n \ge n$, ויש שוויון אמ"מ $a_1 = a_2 = \ldots = a_n = 1$.

(ב) הראו שמספיק להוכיח את האי־שוויון הימני באי־שוויון האריתמטי־גאומטרי תחת ההנחה ש־ $a_1 \cdot a_2 \cdot \ldots \cdot a_n = 1$, והסיקו אותו במקרה זה מהסעיף הקודם.

(ג) הראו שהאי־שוויון השמאלי נובע מהימני.

3. בסעיף זה ניתן הוכחה נוספת של אי־שוויון הממוצעים. כמו בשאלה הקודמת נסתמך על כך שמספיק להראות שלכל n מספרים חיוביים $a_1, \ldots, a_n$ מתקיים האי־שוויון $\sqrt[n]{a_1 \cdot \ldots \cdot a_n} \le \frac{1}{n}(a_1 + \ldots + a_n)$.

(א) הוכיחו באינדוקציה על k שהאי־שוויון נכון כאשר $n = 2^k$ (k טבעי). שימו לב שאפשר לפרק סדרת מספרים $a_1, a_2, \ldots, a_{2^{k+1}}$ לשתי תת־סדרות $a_1, \ldots, a_{2^k}$ ו־ $a_{2^k+1}, \ldots, a_{2^{k+1}}$, שכל אחת מהן באורך 2^k. היעזרו גם באי־שוויון $\sqrt{xy} \le \frac{1}{2}(x+y)$ (הוכיחו אותו!).

---

[10] אי־שוויון הממוצעים נקרא לפעמים גם **האי־שוויון האריתמטי־גאומטרי**, מכיוון שהוא משווה בין ממוצא אריתמטי $\frac{1}{n}(a_1 + \ldots + a_n)$ לממוצע גאומטרי $(a_1 \cdot \ldots \cdot a_n)^{1/n}$. נציין ש־ $n(a_1^{-1} + \ldots + a_n^{-1})^{-1}$ נקרא **ממוצע הרמוני**, והוא שווה לאחד חלקי הממוצע האריתמטי של המספרים $1/a_1, 1/a_2, \ldots, 1/a_n$.

(ב) הסיקו שהאי־שוויון נכון עבור n כללי. לשם כך בחרו k טבעי כך ש־ $2^k < n < 2^{k+1}$ (במקרה ש־ $n = 2^k$ ל־k כלשהו כבר טיפלנו בסעיף הקודם). יהיו $a_1, \ldots, a_n$ מספרים חיוביים. כמו בשאלה הקודמת אפשר להניח ש־ $a_1 \cdot a_2 \cdot \ldots \cdot a_n = 1$. יהי $A = \frac{1}{n}(a_1 + \ldots + a_n)$, ונגדיר

$$a_{n+1} = a_{n+2} = \ldots = a_{2^{k+1}} = A$$

אנו יודעים ש־ $\sqrt[2^{k+1}]{a_1 \cdot a_2 \ldots a_{2^{k+1}}} \leq \frac{1}{2^{k+1}}(a_1 + a_2 + \ldots + a_{2^{k+1}})$. הסיקו מכך את האי־שוויון המבוקש.

פרק 5

סדרות וגבולות

בפרק זה נחקור תכונות של סדרות של מספרים, ונתעניין במיוחד בסדרות ש"הולכות ומתקרבות" למספר מסוים ככל שמתקדמים בסדרה. סדרה כזאת נקראת סדרה מתכנסת, והמספר נקרא הגבול שלה. מושג הגבול הוא אחד המושגים המרכזיים בחשבון האינפיניטסימלי, ומופיע בצורה זו או אחרת כמעט בכל הפרקים הבאים.

רוב הפרק מוקדש להגדרת הגבול, חקר תכונותיו ופיתוח שיטות לחישובו. לאחר מכן נראה שימושים של סדרות. בין השאר נסיים את ההגדרה של פעולת החזקה, אשר התחלנו בה בפרק הקודם, ונגדיר את המספר החשוב e.

5.1 מושג הסדרה

הגדרה 5.1.1 סדרה (sequence) אינסופית של מספרים ממשיים היא התאמה בין המספרים הטבעיים החיוביים $\mathbb{N}\setminus\{0\}=\{1,2,3,\ldots\}$ והמספרים הממשיים $\mathbb{R}$. המספר המתאים ל־1 נקרא האיבר הראשון בסדרה, האיבר שמתאים ל־2 נקרא האיבר השני, וכן הלאה. נשתמש בסימון $(a_n)_{n=1}^{\infty}$ לציין את הסדרה שאיבריה הם $a_1,a_2,\ldots$. המספר n בסימון a_n נקרא **האינדקס** של האיבר a_n.

שתי סדרות $(a_n)_{n=1}^{\infty},(b_n)_{n=1}^{\infty}$ הן שוות אם $a_n=b_n$ לכל $n=1,2,3\ldots$.

באותו אופן ניתן להגדיר סדרות שמתחילות לא מהאינדקס 1 אלא מאינדקס אחר, למשל $(a_n)_{n=5}^{\infty}$ היא הסדרה $a_5,a_6,a_7,\ldots$. כאשר לא חשוב לנו להדגיש מאיזה אינדקס הסדרה מתחילה, נרשום פשוט (a_n). בהמשך נדבר גם על סדרות שהאיברים שלהן אינם מספרים ממשיים אלא איברים מסוג אחר. ההגדרה של סדרות כאלה דומה.

שימו לב שהתפקיד של האות n בסימון $(a_n)_{n=1}^{\infty}$ הוא כמשתנה סרק, כמו אינדקס הסכימה בסכום $\sum_{n=1}^{10}a_n$ (סעיף 3.6). אם $(a_n)_{n=1}^{\infty}$ היא סדרה אז $(a_i)_{i=1}^{\infty}$ היא

אותה סדרה בדיוק, כי האיבר ה־k בשתי הסדרות זהה לכל $k \in \mathbb{N}$. לעומת זאת אם נקבע $n = 1$ ו־ $i = 2$ אז באופן כללי יש הבדל בין האיבר a_n לאיבר a_i (אם כי במקרים מיוחדים ייתכן שוויון בין המספרים). שימו לב גם שהסימונים $(a_n)_{n=1}^{\infty}, (b_n)_{n=1}^{\infty}$ יכולים לציין סדרות שונות אף שהאינדקס בשניהם מסומן n.

לעתים נגדיר סדרה על ידי כתיבת כמה מאיבריה הראשונים. נעשה כך אם קל להסיק את הכלל לאיבר ה־n. למשל $2, 4, 6, 8, ...$ מסמן את סדרת המספרים הזוגיים לפי הסדר הטבעי שלהם. אם נקרא לסדרה הזו $(a_n)_{n=1}^{\infty}$, הנוסחה לאיבר ה־n היא $a_n = 2n$.

דוגמאות

1. עבור מספר $s \in \mathbb{R}$ הסדרה $s, s, s, s, s, \ldots$, נקראת **הסדרה הקבועה שערכה** s. אם נסמן את הסדרה $(a_n)_{n=1}^{\infty}$ אז $a_n = s$ לכל n.
2. הסדרה $1, \frac{1}{2}, \frac{1}{3}, \ldots$ נקראת **הסדרה ההרמונית**. אם נסמן את הסדרה הזו ב־ $(b_n)_{n=1}^{\infty}$ אז איברה הכללי נתון על ידי $b_n = 1/n$.
3. אם $a, d \in \mathbb{R}$ הסדרה $a, a+d, a+2d, \ldots$ נקראת **הסדרה החשבונית** המתחילה מ־a ובעלת הפרש d (כי ההפרש בין איברים עוקבים בסדרה הוא d). אם נסמן ב־c_n את האיבר ה־n בסדרה אז הוא נתון על ידי $c_n = a + d(n-1)$.
4. אם $s \in \mathbb{R}$ הסדרה $s, s^2, s^3, \ldots$ נקראת **הסדרה הגאומטרית** עם בסיס s. אם נסמן את הסדרה הזו ב־ $(d_n)_{n=1}^{\infty}$ אז איברה ה־n הוא $d_n = s^n$. לעתים קרובות מתחילים סדרה גאומטרית מהאינדקס אפס, כלומר מתבוננים בסדרה $(d_n)_{n=0}^{\infty}$, שאיבריה $1, s, s^2, s^3, \ldots$.

   מקרה מיוחד של סדרה גאומטרית הוא הסדרה $e_n = (-1)^n$. זוהי הסדרה שאיבריה הם ± 1 לסירוגין, דהיינו $1, -1, 1, -1, 1, -1, \ldots$.
5. אפשר להגדיר סדרות ברקורסיה. למשל תהי f_n הסדרה המוגדרת ברקורסיה על ידי $f_1 = 1$, $f_2 = 1$ ועבור $n > 2$ על ידי $f_n = f_{n-1} + f_{n-2}$. מתקבלת הסדרה

   $$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \ldots$$

   סדרה זו נקראת **סדרת פיבונאצ'י**.[1] כאשר מגדירים סדרה בצורה זו לא תמיד אפשר למצוא ביטוי פשוט לאיבר ה־n של הסדרה, אף שאיבריו מוגדרים היטב (במקרה של סדרת פיבונאצ'י ידועה נוסחה לאיבר הכללי בסדרה, ונראה איך למצוא אותה בסעיף 11.9).

בכל הדוגמאות למעלה נתנו תיאור פשוט של איברי הסדרה, למשל על ידי נוסחה. סדרות כאלה אינן הכלל, ויש סדרות רבות (למעשה, רוב הסדרות) שאינן ניתנות לתיאור על ידי נוסחה.

---

[1] Leonardo Fibonacci of Pisa, 1170-1250

חשוב להבדיל בין סדרה $(a_n)_{n=1}^{\infty}$ לקבוצת האיברים בסדרה $\{a_n : n \in \mathbb{N}\}$. הקבוצה אינה מכילה כל מידע על סדר האיברים או מספר האיברים השונים. כך למשל עבור שתי הסדרות

$$\begin{aligned}(a_n)_{n=1}^{\infty} &= 0,0,1,0,0,1,0,0,1,0,0,1,\ldots\\(b_n)_{n=1}^{\infty} &= 1,1,0,1,1,0,1,1,0,1,1,0,\ldots\end{aligned}$$

קבוצת האיברים היא $\{0,1\}$, על אף שהסדרות שונות (ואפילו מתקיים $a_n \neq b_n$ לכל n).

תרגילים

1. כתבו את עשרת האיברים הראשונים של בסדרים הבאות:
   (א) $a_n = n^2$.
   (ב) $b_n = 2^n$.
   (ג) $c_n = (-1)^n n$.
   (ד) $d_n = 0$ אם n מתחלק ב־3 ו־ $d_n = n$ אם n אינו מתחלק ב־3.
   (ה) e_n מוגדר ברקורסיה על ידי $e_0 = 0$ ולכל $n \geq 1$, $e_n = n^2 + e_{n-1}$.
2. להלן התחלות של כמה סדרות. לכל אחת מהן חפשו תיאור של הסדרה (נסו למצוא נוסחה לאיבר ה־n, או תיאור רקורסיבי).
   (א) $1,3,5,7,9,\ldots$.
   (ב) $1,1,2,2,3,3,4,4,5,5,\ldots$.
   (ג) $0,2,0,4,0,6,0,8,\ldots$.
   (ד) $1,3,7,15,31,63,127,\ldots$.
   (ה) $0,3,3,6,9,15,24,39,63,\ldots$.
3. תהי (a_n) סדרה. להלן רשומות כמה סדרות המוגדרות במונחים של (a_n). מצאו את כל השוויונות בין הסדרות (כלומר, שוויונות שנכונים ללא תלות ב־(a_n)).
   (א) $(a_{n+1})_{n=1}^{\infty}$.
   (ב) $(a_{2n})_{n=1}^{\infty}$.
   (ג) $(a_{j+1})_{j=2}^{\infty}$.
   (ד) $(a_{j+1})_{n=1}^{\infty}$.
   (ה) $(a_j)_{j=3}^{\infty}$.

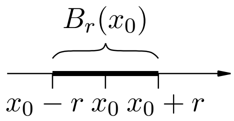
5.2 גבולות

ביטויים כמו "שואף לאפס" כבר השתרשו בשפת היום־יום. בעמודים הבאים נגדיר מושג זה במדויק.

אם (a_n) סדרה ו־a מספר, הביטוי "a_n שואף ל־a" נועד להביע את הרעיון שאיברי הסדרה נהיים קרובים ל־a "בסופו של דבר". דוגמה לסדרה כזאת היא הסדרה ההרמונית $a_n = 1/n$, ההולכת ומתקרבת ל־0 (אם כי היא לעולם אינה מגיעה, כי $a_n \neq 0$ לכל n). ישנן גם סדרות שלא נוכל לייחס להן התנהגות כזאת, כמו למשל הסדרה $1,-1,1,-1,\ldots$ שאיברה ה־n הוא $(-1)^{n+1}$. ערכי הסדרה הזו הם 1 ו־-1 לסירוגין, וכש־n גדל הסדרה מקפצת ביניהם ואינה מתקרבת באופן קבוע לאף אחד מהם, וגם לא לאף מספר אחר. ניסיון הגדרה ראשון יראה אם כן כך:

ניסיון הגדרה ראשון : סדרת מספרים (a_n) מתכנסת למספר a אם רוב איברי הסדרה קרובים ל־a.

שני מושגים כאן אינם ברורים: מהו "רוב", ומהו "קרוב". נתייחס קודם לבעיה השנייה. אם רוב איברי הסדרה "קרובים" ל־a, משמע שרובם נמצאים בסביבה קטנה של a. הנה ההגדרה של סביבה:



איור 5.2.1 סביבה ברדיוס r של x_0

הגדרה 5.2.1 יהי $x_0 \in \mathbb{R}$ ויהי $r > 0$ ממשי. **הסביבה** (neigborhood) של x_0 **ברדיוס** r מסומנת $B_r(x_0)$ ומוגדרת על ידי

$$B_r(x_0) = \{x \in \mathbb{R} \,:\, |x - x_0| < r\}$$

כדאי לחשוב על $B_r(x_0)$ כקבוצת הנקודות שמרחקם מ־x_0 קטן מ־r.[2] פירוש זה מתקבל מההגדרה על ידי פירוש הביטוי $|x-y|$ בתור המרחק בין x ל־y. נשוב ונשתמש בפירוש של $|x-y|$ כמרחק פעמים רבות בהמשך. אבל שימו לב גם ש־$y \in B_r(x_0)$ אמ"מ $-r < y - x_0 < r$, או באופן שקול $x_0 - r < y < x_0 + r$, וקיבלנו ש־$B_r(x_0)$ הוא קטע סימטרי פתוח סביב x_0:

$$B_r(x_0) = (x_0 - r, x_0 + r)$$

אם $y \in B_r(x)$ עבור מספר r "קטן", פירוש הדבר ש־y "קרוב" ל־x (כי המרחק ביניהם אינו עולה על r). מכאן שהגדרה אפשרית למושג השאיפה של סדרה (a_n) למספר a יכולה להיות:

---

[2] פירוש זה מראה את הקשר בין סביבה של נקודה לבין עיגול במישור. אם P היא נקודה במישור אז העיגול ברדיוס r סביב P היא קבוצת הנקודות במישור שמרחקן מ־P קטן מ־r. זו הסיבה לשימוש במילה "רדיוס" בהקשר הנוכחי.

ניסיון הגדרה שני : סדרת מספרים (a_n) שואפת ל־a אם לכל סביבה של a, קטנה ככל שתהיה, רוב איברי הסדרה a_n נמצאים באותה סביבה (באופן שקול: (a_n) שואפת ל־a אם לכל $r > 0$ רוב איברי הסדרה נמצאים בסביבה $B_r(a)$).

כעת עלינו לפרש את המושג "רוב איברי הסדרה". מהו רוב? היה אפשר לחשוב שלרוב איברי הסדרה יש תכונה מסוימת אם אינסוף מהם מקיימים אותה אך זו אינה בחירה מוצלחת. ההגדרה הנכונה נתונה להלן בשפה מעט יותר כללית:

הגדרה 5.2.2 לכל n טבעי תהי $P(n)$ טענה התלויה ב־n. נאמר ש־ $P(n)$ מתקיימת **לכל n גדול מספיק** (for all sufficiently large n) אם יש $N \in \mathbb{N}$ כך ש־$P(n)$ מתקיימת לכל $n > N$ טבעי. במקרה זה נאמר גם ש־$P(n)$ מתקיימת **החל ממקום מסוים**.[3]

טענה $P(n)$ מתקיימת לכל n גדול מספיק אמ"מ מספר ה־n־ים עבורם $P(n)$ אינה מתקיימת הוא סופי. השקילות נובעת מכך שאם $P(n)$ מתקיימת החל ממקום מסוים אז יש N כך ש־$P(n)$ מתקיימת לכל $n > N$, ואז ה־n־ים שלא מקיימים את P מוכלים בקבוצה הסופית $\{1, 2, \ldots, N\}$, ומהוות קבוצה סופית. מאידך אם $A \subseteq \mathbb{N}$ היא קבוצת ה־n־ים שלא מקיימים את P ואם A סופית אז נבחר $N = \max A$, והטענה $P(n)$ מתקיימת לכל $n > N$, ולכן $P(n)$ מתקיימת לכל n גדול מספיק.

דוגמאות

1. תהי $P(n)$ הטענה ש־ $\frac{1}{n} < 0.1$. אז $P(n)$ נכונה לכל $n > 10$ טבעי ולכן היא נכונה החל ממקום מסוים. נביע עובדה זו בקצרה על ידי כך שנאמר שאיברי הסדרה ההרמונית $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ קטנים מ־0.1 החל ממקום מסוים.

2. אין זה נכון שאיברי הסדרה $1, 2, 3, 4, \ldots$ זוגיים החל ממקום מסוים, כי יש אינסוף איברים אי־זוגיים.

   שימו לב שאף שלא נכון שהסדרה $1, 2, 3, 4 \ldots$ מכילה מספרים זוגיים החל ממקום מסוים, גם לא נכון שהחל ממקום מסוים המספרים אינם זוגיים, כי זה היה אומר שהחל ממקום מסוים המספרים אי־זוגיים, וזה לא נכון, כמובן.

3. לפעמים נתעניין בסדרות שאינן מתחילות מהאינדקס 1. דרך אחרת לומר זאת היא שנתעניין בסדרות המוגדרת החל ממקום מסוים.

   למשל, כאשר מגדירים סדרה בעזרת נוסחה לפעמים יש n־ים שעבורם הנוסחה אינה בעלת משמעות. אם יש רק מספר סופי של n־ים כאלה הרי שלא הצלחנו להגדיר את הסדרה לכל n טבעי, אך היא מוגדרת החל ממקום מסוים. לדוגמה, אם נגדיר את הסדרה שאיברה הכללי הוא $f_n = \sqrt{n-5}$ אז הסדרה מוגדרת החל מ־ $n = 5$.

[3] ישנם ספרים שבהם אומרים ש־$P(n)$ מתקיימת "כמעט לכל n" אם היא מתקיימת לכל n גדול מספיק.

כעת נוכל לחבר את כל המרכיבים:

הגדרה 5.2.3 תהי (a_n) סדרה של מספרים ממשיים ויהי $a \in \mathbb{R}$. נאמר כי a הוא **גבול** (limit) של הסדרה אם לכל סביבה של a, קטנה ככל שתהיה, החל ממקום מסוים כל איברי הסדרה נמצאים באותה סביבה. אם a גבול של הסדרה (a_n) נאמר ש־ a_n **שואפת** ל־a (כש־n שואף לאינסוף), או שהסדרה **מתכנסת ל־**a (converges to a), ונסמן זאת על ידי

$$\lim_{n\to\infty} a_n = a$$

לפעמים מסמנים גם $a \xrightarrow{n\to\infty} a_n$ או בקיצור $a_n \to a$. אם לסדרה יש גבול, נאמר שהיא **מתכנסת**.

אם בהגדרת הגבול נרשום בפירוט את הגדרת הסביבה ואת הפרוש של המונח "לכל n גדול מספיק", נקבל את האפיון הבא: סדרה (a_n) מתכנסת למספר a אמ"מ לכל $\varepsilon > 0$ קיים N (שתלוי, אולי, ב־ε) כך שלכל $n > N$ מתקיים $|a_n - a| < \varepsilon$. באופן תמציתי, אפשר לרשום זאת כך:[4]

$$\lim_{n\to\infty} a_n = a \qquad \Longleftrightarrow \qquad \forall\varepsilon > 0\ \exists N\ \forall n > N\ |a_n - a| < \varepsilon$$

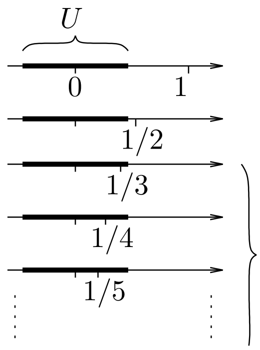
דוגמאות

1. יהי $s \in \mathbb{R}$. הסדרה הקבועה $a_n = s$ שואפת ל־s כש־n שואף לאינסוף, כי s נמצא בכל סביבה של s (כלומר, $s \in B_\varepsilon(s)$ לכל $\varepsilon > 0$) ולכן לכל $\varepsilon > 0$ מתקיים שהחל ממקום מסוים כל איברי הסדרה נמצאים בסביבה $B_\varepsilon(s)$ (בעצם, כל האיברים ממש נמצאים ב־$B_\varepsilon(s)$).

2. הסדרה ההרמונית $b_n = \frac{1}{n}$ שואפת ל־0. כדי להוכיח זאת אנו צריכים להראות שלכל $\varepsilon > 0$ קיים N טבעי עבורו לכל $n > N$ מתקיים $|b_n - 0| < \varepsilon$. מכיוון ש־ $b_n > 0$, התנאי $|b_n - 0| < \varepsilon$ שקול לאי־שוויון $b_n < \varepsilon$. לכן עלינו להראות שלכל $\varepsilon > 0$ קיים N כך שלכל $n > N$ מתקיים $b_n = \frac{1}{n} < \varepsilon$.

   יהי אם כן $\varepsilon > 0$. מתכונת הארכימדיות (משפט 4.3.1) קיים N כך ש־ $\frac{1}{N} < \varepsilon$, ולכל $n > N$ מתקיים $b_n = \frac{1}{n} < \frac{1}{N} < \varepsilon$, כפי שרצינו.

3. נתבונן בסדרה $a_n = \frac{1-(-1)^n}{2n}$. זו הסדרה $1, 0, \frac{1}{3}, 0, \frac{1}{5}, 0, \frac{1}{7}, 0, \ldots$ המתקבלת מהסדרה ההרמונית על ידי החלפת האיברים במקומות הזוגיים במספר 0. סדרה זו מתכנסת ל־0: כי בהינתן $\varepsilon > 0$ ומספר טבעי $n > \frac{1}{\varepsilon}$, או ש־$n$



איור 5.2.2 $\frac{1}{n}$ שייך לסביבה U של 0 לכל $n > 2$

[4] קיצרנו מעט את הכתיבה ולא ציינו במפורש ש־N, n הם מספרים טבעיים. יש בכך פגם מסוים, אך על מנת שלא לסרבל את הכתיבה ננהג כך גם בהמשך.

אי־זוגי, ואז $|a_n - 0| = \frac{1}{n} < \varepsilon$, או ש־$n$ זוגי, ואז $|a_n - 0| = 0 < \varepsilon$. בכל מקרה קיבלנו שאם $n > \frac{1}{\varepsilon}$ אז $|a_n - 0| < \varepsilon$. ואולם שימו לב שהסדרה אינה מתקרבת בהתמדה לגבול, אלא היא מתקרבת ומתרחקת: מרחקו מהגבול של איבר מהסוג a_{2n} הוא אפס, אבל האיבר הבא במרחק חיובי.

4. נחשב את הגבול של הסדרה $c_n = \frac{n^2+n-1}{n^2+1}$. נפשט את הביטוי של c_n ונקבל

$$c_n = \frac{(n^2+1)+(n-2)}{n^2+1} = 1 + \frac{n-2}{n^2+1}$$

נתבונן במחובר $\frac{n-2}{n^2+1}$ באגף ימין. מכיוון שעבור n גדול n^2 גדול בהרבה מ־n, נצפה שהמנה $\frac{n-2}{n^2+1}$ שואפת לאפס, ולכן נצפה ש־ $c_n \to 1$. כדי להראות זאת נעריך את הביטוי $|c_n - 1|$:

$$|c_n - 1| = |1 + \frac{n-2}{n^2+1} - 1| = \frac{|n-2|}{|n^2+1|} < \frac{n}{n^2} = \frac{1}{n}$$

(וודאי שהאי־שוויון נכון לכל $n \geq 1$!). יהי $\varepsilon > 0$. כבר ראינו שיש N כך שאם $n > N$ אז $\frac{1}{n} < \varepsilon$, ולכן עבור $n > N$ מתקיים $|c_n - 1| < \varepsilon$. מכאן נובע ש־ $c_n \to 1$, כפי שרצינו.

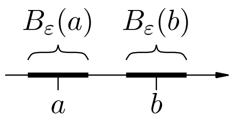
הערה בהגדרת הגבול דרשנו שלכל $\varepsilon > 0$ יהיה קיים מספר טבעי N עבורו מתקיימת תכונה מסוימת, וכשמוכיחים שסדרה מתכנסת לגבול מסוים, יש למצוא, עבור ε נתון, מספר N מתאים. אין חובה למצוא דווקא את המספר N המתאים הקטן ביותר. למשל בדוגמה האחרונה, בהינתן ε בחרנו את N כך שאם $n > N$ אז $|\frac{1}{n} - 0| < \varepsilon$, והראינו שה־$N$ הזה מתאים גם לצרכינו. ייתכן שהיה אפשר למצוא N קטן יותר, אך אין בכך צורך.

5. יהי $x \in \mathbb{R}$. קל לראות שיש סדרות שמתכנסות ל־x, כמו למשל הסדרה הקבועה $a_n = x$. נראה כעת שתמיד יש סדרה של מספרים רציונליים המתכנסת ל־x. ואמנם, לכל n יש מספרים רציונליים בסביבה $B_{1/n}(x)$, כי הסביבה היא קטע פתוח והרציונליים צפופים (משפט 4.3.5). לכן לכל n נוכל לבחור מספר רציונלי r_n כך ש־ $r_n \in B_{1/n}(x)$.

הגדרנו סדרה $(r_n)_{n=1}^{\infty}$ כך שלכל n מתקיים $r_n \in B_{1/n}(x)$, או באופן שקול: $|r_n - x| < 1/n$. אנו טוענים ש־ $r_n \to x$. ואמנם, בהינתן $\varepsilon > 0$ קיים N כך שאם $n > N$ אז $\frac{1}{n} < \varepsilon$, ואז לכל $n > N$ מתקיים $|r_n - x| < \frac{1}{n} < \varepsilon$. זה אומר בדיוק ש־ $r_n \to x$, כפי שרצינו.

הערה אותה ההוכחה נותנת תוצאה כללית יותר: אם A קבוצת מספרים צפופה ו־ $x \in \mathbb{R}$ אז יש סדרה שאיבריה מ־A המתכנסת ל־x.

האם ייתכן שלסדרה יהיו שני גבולות שונים? לא. ההוכחה מתבססת על הלמה הבאה:



איור 5.2.3

למה 5.2.4 לכל שני מספרים $a \neq b$ יש מספר $\varepsilon > 0$ כך שהסביבות $B_\varepsilon(a), B_\varepsilon(b)$ זרות (כלומר: אינן מכילות אף נקודה משותפת).

הוכחה עלינו למצוא $\varepsilon > 0$ כך שאין אף נקודה משותפת ל־$B_\varepsilon(a)$ ו־$B_\varepsilon(b)$. הנה הרעיון. אם יש נקודה $x \in B_\varepsilon(a) \cap B_\varepsilon(b)$. אז המרחקים בין a ל־x ובין x ל־b קטנים שניהם מ־ε, ולכן המרחק בין a ל־b קטן מ־2ε. אם נבחר את ε כך ש־2ε קטן יותר מהמרחק בין a ל־b נקבל סתירה, ונסיק שאין x כזה, כנדרש.

באופן הפורמלי, נבחר $\varepsilon = \frac{1}{2}|b-a|$. אם יש נקודה $x \in B_\varepsilon(a) \cap B_\varepsilon(b)$ פירוש הדבר ש־ $|x-a| < \varepsilon$ ו־ $|x-b| < \varepsilon$, ולכן

$$|a-b| = |(a-x)+(x-b)| \leq |x-a| + |x-b| < 2\varepsilon = |b-a|$$

וזו סתירה. לכן אין x משותף לשתי הסביבות. ■

שימו לב לאופן שבו פרשנו ביטויים מהסוג $|x-y|$ בתור המרחק בין x ל־y, ולתפקיד ששיחק כאן אי־שוויון המשולש. תמונה דומה תחזור פעמים אין־ספור בהמשך, וכדאי ללמוד אותה היטב כבר עכשיו.

משפט 5.2.5 תהי $(a_n)_{n=1}^{\infty}$ סדרה. אם a, b הם גבולות של הסדרה אז $a = b$.

הוכחה תהי (a_n) סדרה ויהיו $a \neq b$ מספרים. נניח בשלילה ש־ $a_n \to a$ וגם $a_n \to b$. נבחר על סמך הלמה $\varepsilon > 0$ כך שהסביבות $B_\varepsilon(a), B_\varepsilon(b)$ זרות. מכיוון ש־ $a_n \to a$ יש N_a כך שלכל $n > N_a$ מתקיים $a_n \in B_\varepsilon(a)$. מכיוון ש־ $a_n \to b$ יש N_b כך שלכל $n > N_b$ מתקיים $a_n \in B_\varepsilon(b)$. יהי $N = \max\{N_a, N_b\}$ ויהי $n = N+1$. אז $n > N_a$ וגם $n > N_b$ ולכן $a_n \in B_\varepsilon(a)$ וגם $a_n \in B_\varepsilon(b)$. אבל זה לא ייתכן, כי לשתי הסביבות האלה אין איברים משותפים, סתירה. ■

בזכות המשפט האחרון נוכל מעתה לדבר על הגבול של סדרה בהא הידיעה. כך למשל נוכל לומר שהגבול של הסדרה ההרמונית הוא 0, וגם להסיק מכך שהגבול שלה אינו 1 (או כל מספר אחר).

אם ננתח את ההוכחה האחרונה, נגלה שהיא התבססה על הטענה הבאה: אם $a_n \in B_\varepsilon(a)$ לכל n גדול מספיק וגם $a_n \in B_\varepsilon(b)$ לכל n גדול מספיק, אז לכל n גדול מספיק, מתקיים $a_n \in B_\varepsilon(a)$ וגם $a_n \in B_\varepsilon(b)$. מכיוון שטענות כאלה יופיעו לעתים תכופות מאד בהמשך ננסח טענה כללית בעניין:

למה 5.2.6 יהיו $P(n)$ ו־$Q(n)$ טענות התלויות ב־n. אם $P(n)$ מתקיימת לכל n גדול מספיק וגם $Q(n)$ מתקיימת לכל n גדול מספיק, אז לכל n גדול מספיק שתי הטענות מתקיימות יחד, כלומר, הטענה $P(n) \wedge Q(n)$ מתקיימת לכל n גדול מספיק.

ההוכחה כמו במשפט על יחידות הגבול, ומושארת כתרגיל.

עד כה לכל הסדרות שבדקנו היה גבול אך אין זה נכון באופן כללי: יש סדרות שאינן מתכנסות.

הגדרה 5.2.7 אם סדרה אינה מתכנסת (כלומר אם אף מספר ממשי אינו גבול שלה) נאמר שהיא **מתבדרת** (diverges).

כדי להבין מהי התבדרות עלינו לתת תיאור מפורט יותר של התופעה. נניח ש־a אינו הגבול של סדרה (a_n). פירוש הדבר שיש סביבה U של a שעבורה לא נכון ש־a_n שייך ל־U החל ממקום מסוים. כפי שראינו אחרי הגדרה 5.2.2, זה שקול לכך שיש אינסוף n־ים עבורם a_n לא שייכת ל־U, כלומר, לכל מספר N, גדול ככל שיהיה, יש $n > N$ כך ש־ $a_n \notin U$. אם נתרגם זאת לשפה כמותית, נקבל ש־(a_n) אינו מתכנס ל־a אמ"מ קיים $\varepsilon > 0$ כך שלכל N יש $n > N$ כך ש־ $|a_n - a| \geq \varepsilon$. בקצרה,

$$a_n \not\to a \qquad \iff \qquad \exists \varepsilon > 0 \, \forall N \, \exists n > N \; |a_n - a| \geq \varepsilon$$

אם כן, סדרה (a_n) מתבדרת אם לכל מספר a הסדרה (a_n) אינה מתכנסת ל־a. לכן (a_n) היא סדרה מתבדרת אמ"מ לכל $a \in \mathbb{R}$ קיים $\varepsilon > 0$ כך שלכל N קיים $n > N$ עבורו $|a_n - L| \geq \varepsilon$. בקצרה,

$$(a_n) \text{ מתבדרת} \qquad \iff \qquad \forall a \, \exists \varepsilon > 0 \, \forall N \, \exists n > N \; |a_n - a| \geq \varepsilon$$

וודאו שאתם מבינים כיצד אפיון זה התקבל מהגדרת ההתכנסות! (תוכלו להיעזר בדיון בשלילת פסוקים בעמוד 6).

דוגמאות

1. תהי $a_n = n$ ונראה שהסדרה הזו מתבדרת. הרעיון הוא שברגע שאיבר בסדרה גדול יותר ממספר a, כל יתר האיברים גדולים אף הם מ־a באותה מידה (ואפילו יותר), ולכן a אינו יכול להיות גבול הסדרה. באופן מדויק, לכל $a \in \mathbb{R}$ נשים לב שיש N טבעי שמקיים $a < N$ וממילא לכל $n > N$ מתקיים $a_n = n > a$. יתרה מזאת, אם $n > N$ אז

$$|a_n - a| = |n - a| = n - a \geq (N+1) - a > 1$$

קיבלנו שעבור הבחירה $\varepsilon = 1$ לכל $n > N$ מתקיים $|a_n - a| > \varepsilon$. לכן a אינו גבול של (a_n). המספר a היה שרירותי ולכן (a_n) מתבדרת.

2. נתבונן בסדרה $b_n = (-1)^n$ ונראה שהיא מתבדרת. יהי a מספר ממשי, ונראה ש־a אינו גבול של הסדרה. נניח תחילה ש־ $a \neq 1$. מכיוון ש־ $a_n = 1$ אינסוף פעמים, המרחק של a_n מ־a שווה ל־ $|a-1|$ אינסוף פעמים, ואם נסמן $\varepsilon = |a-1|$ אז $\varepsilon > 0$ ולאינסוף n־ים מתקיים $|a_n - a| \geq \varepsilon$. לכן a אינו גבול של a_n. כדי לראות ש־1 אינו גבול של הסדרה, נשים לב ש־a_n שווה ל־-1 אינסוף פעמים ולכן לאינסוף n־ים מתקיים $|a_n - 1| = |(-1) - 1| = 2$. לכן גם 1 אינו גבול של (a_n) וקיבלנו ש־(a_n) מתבדרת.

לגבול של סדרה יש התכונה שהוא אינו תלוי בהתנהגות ההתחלתית של הסדרה. למשל, אם נשנה את אלף האיברים הראשונים של הסדרה, או מיליון האיברים הראשונים, לא נשפיע על תכונות ההתכנסות. משום כך אומרים שגבול היא תכונה **אסימפטוטית** של סדרה. ליתר דיוק:

למה 5.2.8 (אי־רגישות הגבול לשינויים סופיים) תהיינה (a_n), (b_n) שתי סדרות. אם $a_n = b_n$ החל ממקום מסוים אז הגבול של (a_n) קיים אמ"מ הגבול של (b_n) קיים, ואם שניהם קיימים אז הם שווים.

הוכחה נניח ש־ $a_n \to a$ ותהי U סביבה של a. אנו יודעים ש־ a_n שייך ל־U החל ממקום מסוים, ואנו יודעים גם ש־ $a_n = b_n$ החל ממקום מסוים. לכן החל ממקום מסוים b_n שייך ל־U. מכיוון ש־U הייתה סביבה שרירותית של a קיבלנו ש־ $b_n \to a$. אותו נימוק (תוך החלפת התפקידים של (a_n) ו־(b_n)) מראה שאם $b_n \to a$ אז $a_n \to a$. בפרט קיבלנו שאם אחד הגבולות קיים אז השני קיים, כנדרש (היכן בהוכחה השתמשנו בלמה 5.2.6?). ■

באופן דומה, נניח ש־$(a_n)_{n=1}^{\infty}$ היא סדרה ו־k טבעי. אם מוחקים את k האיברים הראשונים של הסדרה מקבלים את הסדרה $b_1, b_2, \ldots = a_{k+1}, a_{k+2}, \ldots$. לדוגמה, הסדרה $\frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{8}, \ldots$ התקבלה מהסדרה ההרמונית $a_n = \frac{1}{n}$ על ידי מחיקת ארבעת האיברים הראשונים. ברור למדי שלסדרה (b_n) המתקבלת באופן זה יש אותן תכונות התכנסות כמו (a_n) המקורית. ואמנם,

למה 5.2.9 (אי־רגישות הגבול להזזה) תהיינה (a_n), (b_n) שתי סדרות. אם יש מספר טבעי k כך ש־ $b_n = a_{n+k}$ לכל n גדול מספיק אז הגבול של (a_n) קיים אמ"מ הגבול של (b_n) קיים, ואם בניהם קיימים אז הם שווים.

ההוכחה מושארת כתרגיל.

תרגילים

1. הוכיחו את התכונות הבאות של סביבות:

   (א) אם $0 < r_1 < r_2$ אז $B_{r_1}(x_0) \subseteq B_{r_2}(x_0)$.

   (ב) אם $y \in B_\varepsilon(x)$ אז $x \in B_\varepsilon(y)$.

(ג) אם $y \in B_\varepsilon(x)$ לכל $\varepsilon > 0$ אז $x = y$.

(ד) אם לכל $\varepsilon > 0$ מתקיים $B_\varepsilon(x) \cap B_\varepsilon(y) \neq \emptyset$ אז $x = y$.

(ה) אם $\varepsilon > 0$ ומתקיים $y \in B_\varepsilon(x)$ אז יש $\delta > 0$ כך ש־ $B_\delta(y) \subseteq B_\varepsilon(x)$.

(ו) אם $B_\varepsilon(x) \cap B_\delta(y) \neq \emptyset$ אז $|x - y| < \varepsilon + \delta$.

(ז) אם $x, y \in B_\varepsilon(a)$ אז $|x - y| < 2\varepsilon$.

(ח) אם $x < y$ אז יש $\varepsilon > 0$ כך שלכל $z \in B_\varepsilon(y)$ מתקיים $z > x$.

(ט) אם I קטע באורך קטן מ־r ואם $x \in I$ אז $I \subseteq B_r(x)$.

2. הוכיחו את למה 5.2.6.

3. יהיו $P(n), Q(n)$ טענות התלויות ב־n. אילו מהתנאים הבאים מבטיחים ש־$P(n) \wedge Q(n)$ מתקיימת לאינסוף n־ים? הוכיחו או תנו דוגמה נגדית!

(א) $P(n)$ מתקיימת לאינסוף n־ים ו־$Q(n)$ מתקיימת לכל n גדול מספיק.

(ב) $P(n)$ מתקיימת לאינסוף n־ים וגם $Q(n)$ מתקיימת לאינסוף n־ים.

4. נחשו את הגבול של הסדרות הבאות והוכיחו שהן מתכנסות אליו:

(א) $a_n = \frac{3}{n}$.

(ב) $a_n = -\frac{1}{n}$.

(ג) $a_n = 1 - \frac{1}{n}$.

(ד) $b_n = \frac{n-1}{n^2}$.

(ה) $c_n = \frac{1}{2^n}$.

(ו) $d_n = \frac{n^2+10}{n^3-10}$.

(ז) $e_n = \frac{1}{\alpha n}$ לקבוע $\alpha \neq 0$.

(ח) $f_n = \frac{1}{n^\alpha}$ לקבוע $\alpha > 1$.

5. הוכיחו שהסדרה $a_n = \frac{1+(-1)^n}{2}$ אינה מתכנסת ל־0. אחר־כך הוכיחו שהיא לא מתכנסת לשום גבול.

6. תהי $(a_n)_{n=1}^\infty$ סדרה ו־a מספר. אילו מהתנאים הבאים שקולים לכך שהסדרה מתכנסת ל־a?

(א) לכל $\varepsilon > 0$ קיים $N \in \mathbb{N}$ כך שאם $N \geq n$ אז $|a_n - a| < \varepsilon$. (ההבדל בין תנאי זה להגדרת הגבול הוא שכאן דרשנו $n \geq N$ במקום $n > N$).

(ב) קיים N כך שלכל $n > N$ אם $\varepsilon > 0$ אז $a_n \in B_\varepsilon(a)$.

(ג) לכל $\varepsilon > 0$ קיים N כך שלכל $n > N$ מתקיים $|a_n - a| < 100\varepsilon$.

(ד) קיים מספר $M > 0$ כך שלכל $\varepsilon > 0$ קיים N כך שלכל $n > N$ מתקיים $|a_n - a| < M\varepsilon$.

(ה) לכל $\varepsilon > 0$ קיים מספר $M > 0$ וקיים מספר N כך שלכל $n > N$ מתקיים $|a_n - a| < M\varepsilon$.

7. נגדיר את **הסביבה הסגורה של** a **ברדיוס** r על ידי

$$\overline{B}_r(a) = \{x \in \mathbb{R} \,:\, |x-a| \leq r\}$$

(שימו לב לאי־שוויון החלש שמופיע במקום האי־שוויון החזק בהגדרה של הסביבה $(B_r(a)$). אילו מהתנאים הבאים שקולים לכך ש־ $a_n \to a$?

(א) לכל $r > 0$ מתקיים $a_n \in \overline{B}_r(a)$ החל ממקום מסוים.

(ב) לכל $r \geq 0$ מתקיים $a_n \in \overline{B}_r(a)$ החל ממקום מסוים.

8. התבוננו בהוכחה הבאה, המראה שכל סדרה מתכנסת היא קבועה פרט למספר סופי של איברים. תהי (a_n) סדרה ונניח ש־ $a_n \to a$. אז לכל $\varepsilon > 0$ החל ממקום מסוים מתקיים $a_n \in B_\varepsilon(a)$. מכיוון שזה נכון לכל $\varepsilon > 0$ נובע ש־ $a_n = a$ (בתרגיל (1(ג)) למעלה ראינו שבאופן כללי אם $x \in B_\varepsilon(a)$ לכל $\varepsilon > 0$ אז $x = a$). קיבלנו אפוא ש־ $a_n = a$ החל ממקום מסוים, דהיינו הוכחנו שאם סדרה מתכנסת אז היא קבועה החל ממקום מסוים. היכן השגיאה?

9. הוכיחו את למה 5.2.9.

10. במשפט 5.2.5 הוכחנו את יחידות הגבול על דרך השלילה. הוכיחו אותה על דרך החיוב (רמז: הראו שאם $a_n \to a$ וגם $a_n \to b$ אז $|b-a| < \varepsilon$ לכל $\varepsilon > 0$ והסיקו $a = b$).

11. יהי $x \in \mathbb{R}$. הוכיחו שיש סדרה (r_n) של מספרים רציונליים כך ש־ $r_n < x$ ו־ $r_n \to x$. הסיקו אותו דבר אבל עם התנאי $r_n > x$.

12. תהי $A \subseteq \mathbb{R}$ קבוצה לא ריקה וחסומה מלעיל. הוכיחו ש־ $s = \sup A$ אמ"מ s חסם מלעיל של A ויש סדרת מספרים $a_n \in A$ כך ש־ $a_n \to s$.

13. נניח ש־ $a_n \to a$. הראו שהסדרה $b_n = \max\{a_n, a_{n+1}, \ldots, a_{n^2}\}$ מתכנסת גם־כן ל־a.

5.3 חסימות, תכונות סדר ומשפט הסנדוויץ'

סעיף זה עוסק בקשר בין אי־שוויונות שמקיימים האיברים של סדרות, לבין אי־שוויונות דומים שמקיימים הגבולות שלהם.

הגדרה 5.3.1 סדרה $(a_n)_{n=1}^{\infty}$ נקראת **חסומה מלעיל** (bounded from above) אם קיים מספר M עבורו לכל n טבעי מתקיים $a_n \leq M$. מספר M כזה נקרא **חסם מלעיל** (upper bound) של הסדרה. הסדרה נקראת **חסומה מלרע** (bounded from below) אם יש מספר M, הנקרא **חסם מלרע** (lower bound) של הסדרה, כך ש־ $a_n \geq M$ לכל n טבעי. סדרה נקראת **חסומה** (bounded) אם היא חסומה מלעיל ומלרע. תנאי זה שקול לקיומו של מספר M עבורו לכל $n \in \mathbb{N}$ מתקיים $|a_n| \leq M$. מספר M כזה נקרא **חסם** של הסדרה.

שימו לב שחסימות (בכל אחת מגרסותיה) של סדרה (a_n) שקולה לחסימות באותו מובן של קבוצת איברי הסדרה.

דוגמאות

1. הסדרה $(\frac{1}{n})_{n=1}^{\infty}$ חסומה מלעיל על ידי 1 ומלרע על ידי 0.
2. הסדרה $((-1)^n)_{n=1}^{\infty}$ חסומה מלעיל על ידי 1 ומלרע על ידי -1.
3. הסדרה $(n)_{n=1}^{\infty}$ חסומה מלרע על ידי 0 אך אינה חסומה מלעיל (זו תכונת הארכימדיות של הטבעיים).
4. הסדרה $((-1)^n n)_{n=1}^{\infty}$ אינה חסומה מלעיל או מלרע (הוכיחו!).

טענה 5.3.2 אם $(a_n)_{n=1}^{\infty}$ היא סדרה מתכנסת אז היא חסומה.

הוכחה יהי a הגבול של הסדרה. הרעיון של ההוכחה הוא לחלק את הסדרה לשני חלקים: "רוב" הסדרה (כל האיברים החל ממקום מסוים) מוכלים בסביבה של a (נוכל לבחור איזו סביבה שנרצה) ולכן חלק זה של הסדרה חסום. יתר האיברים מהווים קבוצה סופית, שהיא בוודאי חסומה. ומכאן שהסדרה כולה חסומה. זו כבר כמעט הוכחה מלאה, ונותר רק להשלים את הפרטים.

נבחר $\varepsilon = 1$ (כל בחירה אחרת טובה גם־כן). קיים N טבעי עבורו לכל $n > N$ מתקיים $|a_n - a| < \varepsilon = 1$, ובפרט $a_n < a + 1$. נסמן

$$M = \max\{a_1, a_2, \ldots, a_N, a+1\}$$

אז לכל n טבעי מתקיים $a_n \le M$, וקיבלנו שהסדרה חסומה מלעיל על ידי M. באותו אופן מוכיחים שהסדרה חסומה גם מלרע. $\blacksquare$

מהטענה נובע שאם סדרה אינה חסומה היא אינה מתכנסת. בסעיף הקודם ראינו שהסדרה $a_n = n$ אינה מתכנסת. כעת יש בידינו הוכחה קצרה יותר לעובדה זו: הסדרה $a_n = n$ אינה חסומה.

מה לגבי הכיוון ההפוך של הטענה, כלומר, האם חסימות של הסדרה מבטיחה התכנסות? התשובה שלילית: כבר ראינו שהסדרה $a_n = (-1)^n$ אינה מתכנסת על אף שהיא חסומה. בכל זאת, אם הסדרה מתכנסת יש קשר בין חסמים של הסדרה לגבול שלה.

טענה 5.3.3 תהי (a_n) סדרה מתכנסת, ונניח ש־ $a_n \le M$ החל ממקום מסוים. אז $\lim_{n\to\infty} a_n \le M$.

הוכחה נסמן $a = \lim_{n\to\infty} a_n$. אילו היה מתקיים $a > M$ אז היה אפשר לבחור סביבה U של a כך ש־ $x > M$ לכל $x \in U$ (למשל, $U = B_{a-M}(a)$ היא סביבה

כזאת). כיוון ש־ $a_n \to a$ החל ממקום מסוים מתקיים $a_n \in U$, ובפרט $a_n > M$ החל ממקום מסוים. יחד עם ההנחה היינו מקבלים ש־ $a_n > M$ החל ממקום מסוים וגם $a_n \le M$ החל ממקום מסוים, ולכן יש n כך ש־ $a_n \le M < a_n$, וזה לא ייתכן. ■

הטענה אינה נכונה אם מחליפים את האי־שוויונות החלשים באי־שוויונות חזקים. ייתכן בהחלט ש־ $a_n < M$ החל ממקום מסוים, ואפילו לכל n, אבל הגבול של a_n אינו קטן ממש מ־M. למשל $a_n = -\frac{1}{n}$ היא סדרה שכל איבריה קטנים ממש מ־0 אבל הגבול אינו קטן מ־0, אלא שווה לו.

קיימת כמובן גרסה של הטענה שבו האי־שוויונות הפוכים. אפשר להוכיח אותה ישירות אך היא נובעת גם מהטענה הבאה:

טענה 5.3.4 אם $(a_n),(b_n)$ סדרות המקיימות $a_n \le b_n$ החל ממקום מסוים, ואם $a_n \to a$ ו־ $b_n \to b$, אז $a \le b$.

הוכחה נניח בשלילה ש־ $a > b$. אז אפשר לבחור סביבה U_a של a וסביבה U_b של b כך שלכל $x \in U_a$ ולכל $y \in U_b$ מתקיים $x > y$ (למשל, זה נכון עבור הסביבות ברדיוס ε סביב a,b כאשר $\varepsilon = \frac{b-a}{2}$). לפי הנתון מתקיים $a_n \to a$ ו־ $b_n \to b$, ולכן החל ממקום מסוים $a_n \in U_a$ וגם $b_n \in U_b$. מכאן שהחל ממקום מסוים מתקיים $b_n < a_n$, בסתירה לכך שהחל ממקום מסוים $a_n \le b_n$. ■

בטענות האחרונות הנחנו את קיום הגבולות. המשפט הבא, לעומת זאת, מאפשר להסיק את קיום הגבול של סדרה, וזאת על ידי השוואת הסדרה עם סדרות אחרות החוסמות אותה מלמעלה ומלמטה. הוא גם מאפשר לחשב את הגבול. בהמשך ננצל זאת כדי לנתח סדרות מסובכות על ידי השוואתן עם סדרות פשוטות יותר.

משפט 5.3.5 (משפט הסנדוויץ') תהיינה (a_n),(b_n),(c_n) סדרות. נניח שהחל ממקום מסוים מתקיים $a_n \le b_n \le c_n$ ושהסדרות (a_n) ו־ (c_n) מתכנסות לגבול משותף L. אז גם (b_n) מתכנסת ל־L.

הערה אילו ידענו מראש שהסדרה (b_n) מתכנסת לגבול b היה נובע ממסקנה 5.3.4 ש־ $L \le b$ (כי $b_n \ge a_n$ החל ממקום מסוים) וגם $b \le L$ (כי $b_n \le c_n$ החל ממקום מסוים), ויחד היינו מקבלים $b = L$.

הוכחה הרעיון מאחורי ההוכחה פשוט: עבור n גדול מספיק המספרים a_n, c_n קרובים ל־L. אבל b_n נמצא בין a_n ו־c_n ולכן גם הוא לא יכול להיות רחוק מ־L.

הנה הפרטים. יהי $\varepsilon > 0$. החל ממקום מסוים מתקיים $a_n \in B_\varepsilon(L)$, ובפרט $a_n > L - \varepsilon$. החל ממקום מסוים מתקיים גם $b_n \ge a_n$, ולכן החל ממקום מסוים מתקיים $b_n > L - \varepsilon$.

באותו אופן, החל ממקום מסוים מתקיים $c_n \in B_\varepsilon(L)$ ולכן $c_n < L + \varepsilon$, והחל ממקום מסוים מתקיים $b_n \le c_n$. לכן החל ממקום מסוים מתקיים $b_n < L + \varepsilon$.

משתי הפסקאות האחרונות אנו מסיקים שהחל ממקום מסוים מתקיימים בו־זמנית $b_n > L - \varepsilon$ ו־ $b_n < L + \varepsilon$, ופירוש הדבר ש־ $b_n \in B_\varepsilon(L)$. מכיוון ש־ε היה שרירותי אנו מסיקים ש־ $b_n \to L$, כנדרש. ■

דוגמאות

1. יהי $\alpha > 1$ ויהי $a_n = \frac{1}{n^\alpha}$. כיוון שמתקיים $0 < \frac{1}{n^\alpha} \le \frac{1}{n}$ לכל n טבעי וכיוון שהסדרה ההרמונית וגם הסדרה הקבועה 0 מתכנסות שתיהן ל־0, נובע לפי כלל הסנדוויץ' כי $\lim_{n\to\infty} \frac{1}{n^\alpha} = 0$.

   גם אם $0 < \alpha < 1$ מתקיים $\frac{1}{n^\alpha} \to 0$. נוכיח זאת מאוחר יותר (זה נובע ממשפט 5.4.9 להלן).

2. תהי $a_n = \frac{1}{\sqrt{n}}(\sqrt{n+2} - \sqrt{n+1})$, ונראה ש־ $a_n \to 0$. אם נכפיל את המונה והמכנה במספר $\sqrt{(n+2)} + \sqrt{(n+1)}$ נקבל

$$\begin{aligned} a_n &= \frac{\sqrt{n+2} - \sqrt{n+1}}{\sqrt{n}} \cdot 1 \\ &= \frac{\sqrt{n+2} - \sqrt{n+1}}{\sqrt{n}} \cdot \frac{\sqrt{n+2} + \sqrt{n+1}}{\sqrt{n+2} + \sqrt{n+1}} \\ &= \frac{(n+2) - (n+1)}{\sqrt{n} \cdot (\sqrt{n+2} + \sqrt{n+1})} \\ &\le \frac{1}{\sqrt{n} \cdot 2\sqrt{n}} \\ &< \frac{1}{n} \end{aligned}$$

   קיבלנו את האי־שוויון $0 \le a_n \le \frac{1}{n}$, שבו צד ימין שואף ל־$0$ וכמובן כך גם צד שמאל (שהיא הסדרה הקבועה אפס). לכן מכלל הסנדוויץ', $a_n \to 0$.

3. נוכיח שהסדרה הגאומטרית $a_n = \frac{1}{2^n}$ שואפת ל־0. קל לבדוק באינדוקציה שמתקיים $2^n \ge n$ לכל n טבעי, ומכאן ש־ $0 \le \frac{1}{2^n} \le \frac{1}{n}$. מכיוון ש־ $\frac{1}{n} \to 0$ וגם הסדרה הקבועה 0 שואפת ל־0, כלל הסנדוויץ' גורר ש־ $\frac{1}{2^n} \to 0$.

4. נכליל את הדוגמה הקודמת לסדרות גאומטריות כלליות יותר. יהי $0 < s < 1$, ונראה ש־ $s^n \to 0$. כמו קודם, נרצה להראות שמהירות הגידול של $(\frac{1}{s})^n$ הוא לפחות כמו כפולה של n. ההנחה $0 < s < 1$ גוררת ש־ $\frac{1}{s} > 1$, ולכן נוכל לרשום $\frac{1}{s} = 1 + c$, כאשר c מספר חיובי. לפי אי־שוויון ברנולי מתקיים $(1+c)^n \ge 1 + cn$ (תרגיל (10) בעמוד 56). מכאן נובע ש־ $(\frac{1}{s})^n \ge cn$, או באופן שקול, $0 \le s^n \le \frac{1}{cn}$. כעת הגבול $s^n \to 0$ נובע מכלל הסנדוויץ' (ראו תרגיל (4)(ז)) בעמוד 97).

5. יהי $a > 0$ ונראה ש־ $\sqrt[n]{a} \to 1$. אם $a = 1$ הטענה טריביאלית כי $\sqrt[n]{1} = 1 \to 1$. נפריד לשני מקרים:

(א) נניח $0 < a < 1$ ונניח בשלילה כי $\sqrt[n]{a} \not\to 1$. נשים לב שלכל n מתקיים $0 < \sqrt[n]{a} < 1$ ולכן אם $\sqrt[n]{a} \not\to 1$ אז יש $0 < \varepsilon < 1$ ואינסוף n־ים כך ש־ $0 < \sqrt[n]{a} < 1 - \varepsilon$, כלומר, לאינסוף n־ים מתקיים $a < (1-\varepsilon)^n$. מצד שני, מהסעיף הקודם אנו יודעים ש־ $(1-\varepsilon)^n \to 0$, ולכן החל ממקום מסוים מתקיים $(1-\varepsilon)^n < a$, בסתירה לכך ש־ $a < (1-\varepsilon)^n$ לאינסוף n־ים.

(ב) יהי $a > 1$ ונראה ש־ $\sqrt[n]{a} \to 1$. יהי $\varepsilon > 0$. די להראות שהחל ממקום מסוים מתקיים $1 < \sqrt[n]{a} < 1 + \varepsilon$, או באופן שקול, שהחל ממקום מסוים מתקיים $\frac{1}{1+\varepsilon} < \frac{1}{\sqrt[n]{a}} < 1$. אבל מהסעיף הקודם אנו יודעים ש־ $\frac{1}{\sqrt[n]{a}} = \sqrt[n]{\frac{1}{a}} \to 1$, כי $0 < \frac{1}{a} < 1$, והמסקנה נובעת (וודאו שאתם מבינים מדוע!).

6. נראה ש־ $\sqrt[n]{\frac{n+1}{n-1}} \to 1$. נשים לב ש־ $1 < \frac{n+1}{n-1} < 2$ לכל $n \geq 4$ (הוכיחו זאת!) ולכן

$$\sqrt[n]{1} \leq \sqrt[n]{\frac{n+1}{n-1}} \leq \sqrt[n]{2}$$

אגף שמאל שווה ל־1 לכל n, ואילו $\sqrt[n]{2} \to 1$ לפי הדוגמה הקודמת. לכן מכלל הסנדוויץ' מקבלים $\sqrt[n]{\frac{n+1}{n-1}} \to 1$, כפי שרצינו.

מפתה לתת את ה"הוכחה" הבאה לגבול $\sqrt[n]{\frac{n+1}{n-1}} \to 1$ בדוגמה האחרונה. ראשית, מוכיחים ש־ $\frac{n+1}{n-1} \to 1$ (זה באמת נכון), ואז

$$\sqrt[n]{\frac{n+1}{n-1}} \to \sqrt[n]{1} \to 1$$

אבל זו שטות גמורה. לא ניתן להשאיף את n לאינסוף "בשלבים". באופן כללי, כשרושמים $a_n \to a$ האינדקס n אינו יכול להופיע באגף ימין. אם קבלתם שהגבול תלוי ב־n אז יש לכם טעות.

למשל, יהי $a_n = \sqrt[n]{\frac{1}{2^n}}$. ראינו ש־ $\frac{1}{2^n} \to 0$ ולכן מפתה לרשום

$$a_n = \sqrt[n]{\frac{1}{2^n}} \to \sqrt[n]{0} = 0$$

אבל כמובן זה לא נכון: $a_n = \frac{1}{2}$ לכל n ולכן $a_n \to \frac{1}{2}$.

תרגילים

1. הוכיחו שסכום של סדרות חסומות הוא חסום. האם מכפלה של סדרות חסומות מלעיל היא חסומה מלעיל?

2. עבור כל אחת מהסדרות הבאות קבעו אם היא חסומה (מלעיל או מלרע). אם היא חסומה בדקו האם היא מתכנסת, ואם כן חשבו את הגבול.
   (א) $a_n = \frac{n^2+3}{n+5}$.
   (ב) $a_n = \frac{1}{\sqrt{n^2+1}}$.
   (ג) $a_n = [\sqrt{n}]$.
   (ד) $a_n = (-1)^n(1+\frac{1}{n})$.
   (ה) $a_n = \sqrt{1+\frac{1}{n}}$.
   (ו) $a_n = \sqrt{n^2+2}-n$.
   (ז) $a_n = 2^{-\sqrt{n}}$.
3. הוכיחו או הפריכו:
   (א) אם $a_n \to a$ ו־ $b_n \to b$ ואם $a_n < b_n + 1$ לכל n, אז $a < b + 1$.
   (ב) אם $a_n \to a$ ו־$b_n \to b$ ואם $a < b$ אז $a_n < b_n$ החל ממקום מסוים.
   (ג) אם $a_n \to a$ ו־ $b_n \to b$, ומקיים $a_n > b_n$ ל־n זוגי ו־$a_n < b_n$ ל־n אי־זוגי, אז $a = b$.
   (ד) אם $a_n \to 0$ ואם $a_n < b_n < 2a_n$ אז $b_n \to 0$.
   (ה) אם $-1 < \frac{a_n}{b_n} < 1$ ואם (b_n) מתכנס למספר חיובי אז (a_n) מתכנס.
4. מה הגבול של $\sqrt[n]{2^n+3^n}$? באופן כללי יותר, אם $a, b > 0$ מה הגבול של $\sqrt[n]{a^n+b^n}$?
5. חשבו את הגבולות של הסדרות הבאות, אם הם קיימים, אחרת הוכיחו שהן מתבדרות.
   (א) $\frac{\{n\}}{2^n}$.
   (ב) $\frac{n^2+1}{n^3-2}$.
   (ג) $\frac{n-\sqrt{n^2+1}}{\sqrt{n}}$.
   (ד) $\sqrt[n]{\frac{n^2+5}{n^2+1}}$.
   (ה) $\frac{n-[\sqrt{n}]^2}{\sqrt{n}}$.
   (ו) $\frac{n-[\sqrt{n}]^2}{n}$.
   (ז) $\frac{x^n}{n!}$ עבור $x \in \mathbb{R}$ קבוע.
   (ח) $\frac{n^n}{x^n}$ עבור $0 \neq x \in \mathbb{R}$ קבוע.
   (ט) $\frac{n^k}{n!}$ עבור $k \in \mathbb{N}$ קבוע.
   (י) $\frac{n!}{n^n}$.
   (יא) $\frac{x^n+n}{n^n}$ עבור $x > 1$ קבוע.
6. נסחו והוכיחו גרסה של טענה 5.3.3 שבו כיוון האי־שוויון הוא הפוך.

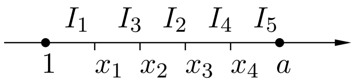
7. יהי x_0 מספר ממשי.

(א) יהי a_1 מספר שרירותי ונגדיר ברקורסיה $a_{n+1} = \frac{1}{2}(a_n + x_0)$. הוכיחו שלכל $r > 0$, מתקיים $a_n \in B_r(x_0)$ לכל n גדול מספיק, והסיקו שהסדרה (a_n) מתכנסת ל־x_0.

(ב) תהי (x_n) סדרה ונניח ש־ $x_n \to x_0$. יהי a_1 מספר שרירותי ונגדיר ברקורסיה $a_{n+1} = \frac{1}{2}(a_n + x_n)$. הוכיחו ש־ $a_n \to x_0$.

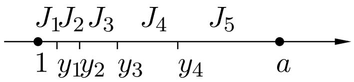
8. הנה הוכחה נוספת לעובדה שלכל $a > 1$ מתקיים $\sqrt[n]{a} \to 1$. יהי $n \in \mathbb{N}$ קבוע.

(א) מצאו סדרת נקודות $1 = x_0 < x_1 < x_2 < \ldots < x_n = a$ כך שהקטעים $I_i = [x_{i-1}, x_i]$ מחלקים את הקטע $[1, a]$ ל־n קטעים עוקבים שווי־אורך באורך $\frac{a-1}{n}$.



איור 5.3.1

(ב) מצאו סדרת נקודות $1 = y_0 < y_1 < y_2 < \ldots < y_n = a$ כך שהקטעים $J_i = [y_{i-1}, y_i]$ מחלקים את הקטע $[1, a]$ ל־n קטעים עוקבים כך שהאורך של J_i גדול פי $\sqrt[n]{a}$ מהאורך של J_{i-1}.



איור 5.3.2

(ג) שימו לב ש־J_1 הוא הקצר ביותר מבין הקטעים $J_1, \ldots J_n$. הסיקו מכך שהאורך של J_1 אינו עולה על האורך של I_1, והשתמשו באי־שוויון שקבלתם כדי להוכיח ש־ $\sqrt[n]{a} \to 1$.

9. בשאלה זו נוכיח ש־ $\sqrt[n]{n} \to 1$.

(א) היעזרו במשפט הבינום על מנת להוכיח שלכל $\varepsilon > 0$ יש קבוע חיובי c (התלוי ב־ε) כך ש־ $(1+\varepsilon)^n \geq cn^2$.

(ב) הסיקו $\sqrt[n]{n} \to 1$ (ההוכחה דומה להוכחה ש־ $\sqrt[n]{a} \to 1$ בדוגמה (5) בעמוד 101. הניחו בשלילה $\sqrt[n]{n} \not\to 1$, כלומר שיש $\varepsilon > 0$ כך ש־ $\sqrt[n]{n} > 1 + \varepsilon$ לאינסוף n־ים. כעת הגיעו לסתירה).

10. יהי $0 < s < 1$ ויהי $\alpha \in [0, \infty)$. הוכיחו ש־ $\lim_{n\to\infty} n^\alpha \cdot s^n = 0$ (רמז: שימו לב ש־ $n^\alpha s^n = ((\sqrt[n]{n})^\alpha s)^n$ והיעזרו עובדה ש־ $t^n \to 0$ לכל $0 < t < 1$ וש־ $\sqrt[n]{n} \to 1$).

5.4 אריתמטיקה של סדרות וגבולות

מה הגבול של הסדרה $c_n = \sqrt{(2+\frac{1}{n})(3-\frac{1}{n})}$? אם ניחשתם שהגבול הוא $\sqrt{6}$ הסיבה בוודאי הייתה השיקול הבא. כיוון ש־ $\frac{1}{n} \to 0$ נצפה ש־ $2 + \frac{1}{n} \to 2$, כיוון ש־ $\frac{1}{n} \to 0$ נצפה ש־ $3 - \frac{1}{n} \to 3$, ולכן נצפה שמכפלת הסדרות תשאף למכפלה $2 \cdot 3 = 6$ של הגבולות, ולבסוף מאחר ש־ $(2+\frac{1}{n})(1-\frac{1}{n}) \to 6$ נצפה ש־ $\sqrt{(2+\frac{1}{n})(3-\frac{1}{n})} \to \sqrt{6}$. ואמנם, אפשר להוכיח שהגבול הוא $\sqrt{6}$ באופן ישיר, אך המצב הטוב ביותר היה אילו יכולנו להצדיק את הנימוק האינטואיטיבי שנתנו למעלה. נעסוק בכך להלן.

כפי שרומזת הדוגמה למעלה, אנו נעסוק בסדרות המתקבלות מסדרות אחרות בעזרת הפעולות החשבוניות. אם $(a_n), (b_n)$ הן סדרות אז **הסכום** שלהן היא הסדרה (s_n) המוגדרת על ידי $s_n = a_n + b_n$. סדרה זו מסומנת בקיצור על ידי $(a_n + b_n)$. באופן דומה מקבלים את סדרת **ההפרש** שאיברה ה־n הוא $a_n - b_n$, סדרת **המכפלה** שאיברה ה־n הוא $a_n \cdot b_n$, **וסדרת המנה** שאיברה ה־n הוא $\frac{a_n}{b_n}$ (בתנאי ש־ $b_n \neq 0$). סדרת הערכים המוחלטים של (a_n) היא הסדרה $(|a_n|)$. אם $c \in \mathbb{R}$ נוכל להגדיר את הסדרה שאיברה ה־n הוא $c \cdot a_n$, ואם $r \in \mathbb{Q}$ נוכל להגדיר את הסדרה שאיברה ה־n הוא a_n^r (בתנאי ש־ $a_n > 0$). אפשר לשלב בין הפעולות ולקבל סדרות כמו $(a_n + \frac{b_n}{a_n})$, וכן הלאה.

שימו לב שעל מנת להגדיר את הסדרה $(\frac{a_n}{b_n})$ יש לדרוש ש־ $b_n \neq 0$ לכל n. עם זאת, אם $b_n \neq 0$ החל ממקום מסוים עדיין ניתן להגדיר סדרה $\frac{a_n}{b_n}$, אלא שהיא גם תהיה מוגדרת החל ממקום מסוים ולא בהכרח לכל n שעבורו a_n, b_n מוגדרים. הערה דומה נכונה לכל פעולה שאינה מוגדרת תמיד, כמו הוצאת שורש.

המטרה שלנו היא לחשב את הגבולות של סדרות מורכבות בעזרת הגבולות של הסדרות המרכיבות אותן. נתחיל בכמה אפיונים פשוטים אך חשובים של הגבול:

טענה 5.4.1 תהי $(a_n)_{n=1}^{\infty}$ סדרה ויהי $a \in \mathbb{R}$. התנאים הבאים שקולים:

1. $\lim_{n\to\infty} a_n = a$
2. $\lim_{n\to\infty}(a_n - a) = 0$
3. $\lim_{n\to\infty} |a_n - a| = 0$

בפרט, $a_n \to 0$ אמ"מ $|a_n| \to 0$.

הערה הגבול $\lim_{n\to\infty}(a_n - a)$ הוא הגבול של סדרת ההפרשים $(a_n - a)_{n=1}^{\infty}$, והגבול $\lim_{n\to\infty} |a_n - a|$ הוא הגבול של הסדרה $(|a_n - a|)_{n=1}^{\infty}$.

הוכחה אנו רוצים להוכיח את שקילות שלושת הטענות הבאות, שהן התרגום לכתיב פורמלי של שלושת התנאים בטענה:

1. $\forall \varepsilon > 0\ \exists N\ \forall n\ (n > N \implies |a_n - a| < \varepsilon)$.

2. $\forall\varepsilon > 0\ \exists N\ \forall n\ (n > N \implies |(a_n - a) - 0| < \varepsilon)$
3. $\forall\varepsilon > 0\ \exists N\ \forall n\ (n > N \implies ||a_n - a| - 0| < \varepsilon)$

השקילות נובעת באופן מידי מהאבחנה שלכל זוג מספרים x, y מתקיים

$$|x - y| = |(x - y) - 0| = ||x - y| - 0|$$

ובפרט כאשר $x = a_n$ ו־ $y = a$. לכן התנאים בצד ימין של שלוש הטענות שקולות, וממילא הטענות שקולות.

החלק האחרון נובע מיד על ידי הצבת $a = 0$ בסעיפים (1) ו־(3). ■

אנו מקבלים מאפיון זה את המשפט הראשון על אריתמטיקה של גבולות:

טענה 5.4.2 אם $a_n \to a$ אז $|a_n| \to |a|$.

הוכחה נניח ש־ $a_n \to a$. אז לפי הטענה הקודמת $|a_n - a| \to 0$. אבל מתכונות הערך המוחלט,

$$0 \leq ||a_n| - |a|| \leq |a_n - a|$$

וממשפט הסנדוויץ' אנו מסיקים ש־ $||a_n| - |a|| \to 0$. נפעיל שוב את הטענה ונקבל $|a_n| \to |a|$, כנדרש. ■

מומלץ לתת הוכחה ישירה למסקנה האחרונה. כמו כן שימו לב שההפך אינו נכון: ייתכן $|a_n| \to |a|$ אבל $a_n \not\to a$. מצאו דוגמה כזאת!

כעת אנו מגיעים למשפטי האריתמטיקה המרכזיים. משפטים אלה מאפשרים לחשב בדיוק את הגבול של סדרה מורכבת בעזרת הגבולות של המרכיבים שלה.

משפט 5.4.3 (כלל החיבור) אם (a_n), (b_n) סדרות המתכנסות לגבולות b, a בהתאמה אז מתקיים $\lim_{n\to\infty}(a_n + b_n) = a + b$.

הערה למשפט שתי טענות: ראשית, שהגבול של $(a_n + b_n)$ קיים, ושנית, שהוא שווה ל־ $a + b$. ניתן לרשום את מסקנת המשפט גם באופן הבא: אם (a_n) ו־(b_n) מתכנסים אז

$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$$

לכן אומרים לעתים שפעולת הגבול מתחלפת עם (או מכבדת את) פעולת החיבור.

הוכחה באופן גס, עלינו להראות שאם a_n קרוב ל־a ו־b_n קרוב ל־b אז $a_n + b_n$ קרוב ל־$a + b$. המפתח לניסוח הכמותי של רעיון זה הוא העובדה שאם המרחק בין a_n ל־a קטן מ־ε וגם המרחק בין b_n ל־b קטן מ־ε אז המרחק בין $a_n + b_n$ לבין

$a+b$ קטן מ־2ε. עובדה זו נובעת מאי־שוויון המשולש (ראו להלן). כעת כל שנותר הוא להחליף את ε ב־ $\varepsilon/2$ ולהשלים את הפרטים. הנה:

יהי $\varepsilon > 0$. עלינו להראות שהחל ממקום מסוים מתקיים

$$|(a_n + b_n) - (a+b)| < \varepsilon$$

אבל לפי אי־שוויון המשולש,

$$|(a_n + b_n) - (a+b)| \leq |a_n - a| + |b_n - b|$$

ולכן אם נראה שהחל ממקום מסוים מתקיימים שני האי־שוויונות

$$|a_n - a| < \frac{\varepsilon}{2} \quad , \quad |b_n - b| < \frac{\varepsilon}{2}$$

נסיים, כי מזה ינבע שהחל ממקום מסוים

$$|a_n + b_n - (a+b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

אבל מההתכנסות $a_n \to a$ ו־ $b_n \to b$ אנו יודעים שהחל ממקום מסוים $|a_n - a| < \frac{\varepsilon}{2}$ ושהחל ממקום מסוים מתקיים $|b_n - b| < \frac{\varepsilon}{2}$ (זה נובע ישירות מהגדרת הגבול על ידי התבוננות בסביבות $B_{\varepsilon/2}(a)$ ו־$B_{\varepsilon/2}(b)$), ולכן החל ממקום מסוים שניהם מתקיימים, כפי שרצינו. ■

דוגמאות

1. ראינו ש־ $\lim_{n\to\infty} \sqrt[n]{4} = 1$ וכן $\lim_{n\to\infty} \frac{1}{n} = 0$. לכן נובע מכלל החיבור ש־

$$\lim_{n\to\infty} (\frac{1}{n} + \sqrt[n]{4}) = 0 + 1 = 1$$

2. נתבונן בסדרה $a_n = (-1)^n - \sqrt[n]{4}$. אילו היה לסדרה גבול אז לפי כלל החיבור גם הסדרה המתקבלת על ידי חיבור $(a_n)_{n=1}^\infty$ עם הסדרה $(\sqrt[n]{4})_{n=1}^\infty$ הייתה מתכנסת, כי $(\sqrt[n]{4})$ מתכנסת גם־כן. אבל סדרת הסכום היא $((-1)^n)_{n=1}^\infty$, וזו סדרה מתבדרת. לכן הסדרה $(-1)^n - \sqrt[n]{4}$ מתבדרת.

הערות

1. כלל הסכום היא גרירה בכיוון אחד בלבד: התכנסות הסדרות (a_n) ו־(b_n) גוררת התכנסות (a_n+b_n). ההפך אינו נכון, כלומר, התכנסות של (a_n+b_n) אינה גוררת את התכנסות של (a_n) ו־(b_n). למשל אם $a_n = (-1)^n$ ו־ $b_n = (-1)^{n+1}$, אז שתי הסדרות מתבדרות, אבל $a_n + b_n = 0$ לכל n ולכן סדרת הסכום מתכנסת.

2. כלל החיבור תקף גם עבור לסכומים של יותר משתי סדרות. למשל יהיו $(a_n),(b_n),(c_n)$ סדרות המתכנסות ל־ a,b,c בהתאמה ונגדיר את סדרת הסכום, $x_n = a_n + b_n + c_n$. אז אם נפעיל את המשפט פעמיים נקבל $x_n \to a+b+c$. באופן מדויק, נגדיר סדרת ביניים $y_n = a_n + b_n$. אז לפי המשפט $y_n \to a+b$. כמו־כן $x_n = y_n + c_n$ ולכן שוב לפי המשפט, $x_n \to (a+b)+c$, כפי שרצינו. התוצאה נכונה גם לסכומים של יותר משלוש סדרות (נסחו והוכיחו את הטענה הכללית!)

כפי שציינו, כלל הסכום חל על סדרה המתקבלת כסכום של מספר סופי של סדרות, אך לא כל סדרה שאיבריה מוצגים כסכומים מתקבלת בצורה זו. הנה דוגמה לשימוש שגוי במשפט: נתבונן בסדרה $(a_n)_{n=1}^{\infty}$ המוגדרת על ידי הסכום

$$a_n \;=\; \underbrace{\frac{1}{n}+\frac{1}{n}+\ldots+\frac{1}{n}}_{n \text{ מחוברים}}$$

מכיוון שהסדרה ההרמונית מקיימת $\frac{1}{n} \to 0$ ו־a_n הוא סכום של סדרות הרמוניות ניתן לכאורה להסיק ממשפט האריתמטיקה של גבולות ש־ $a_n \to 0+0+\ldots+0 = 0$. כמובן שזה לא נכון! הרי $a_n = 1$ לכל n, ולכן $a_n \to 1$. השגיאה כאן היא ש־a_n מוצג כסכום של מספר הולך וגדל של מחוברים, ולא כסכום של שתי סדרות, כמו במשפט, או מספר סופי אך קבוע של סדרות, כמו בהערה הקודמת. אפשר לכתוב את הנימוק השגוי גם באופן הבא: נתונה הסדרה ש־ $a_n = \sum_{k=1}^{n} \frac{1}{n}$, ולכן אנו טוענים ש־

$$\lim_{n\to\infty} a_n = \sum_{k=1}^{n} \lim_{n\to\infty} \frac{1}{n} = \sum_{k=1}^{n} 0 = 0$$

זו כמובן שגיאה, כי האינדקס n אותו משאיפים לאינסוף "ברח" ומופיע מחוץ לגבול כבר בשוויון הראשון. עיינו בדיון בעמוד 102.

במקביל לכלל החיבור ישנו גם כלל חיסור. כדי להוכיח אותו אין צורך לחזור על הוכחת כלל החיבור, אלא אפשר להתבסס עליו:

משפט 5.4.4 (כלל החיסור) יהיו $(a_n)_{n=1}^{\infty}$, $(b_n)_{n=1}^{\infty}$ סדרות המתכנסות לגבולות b,a בהתאמה אז

$$\lim_{n\to\infty}(a_n - b_n) = a - b \qquad , \qquad \lim_{n\to\infty}(-b_n) = -b$$

הוכחה לפי טענה 5.4.1, כדי להסיק ש־ $-b_n \to -b$ די שנראה ש־ $|-b_n-(-b)| \to 0$. לכל n מתקיים $|-b_n-(-b)| = |b_n - b|$, ולכן די שנראה שמתקיים $|b_n - b| \to 0$. אבל זה נובע מהנתון $b_n \to b$.

כעת אם נשים לב ש־ $a_n - b_n = a_n + (-b_n)$ נסיק מכלל החיבור ש־

$$\lim_{n\to\infty}(a_n - b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty}(-b_n) = a + (-b) = a - b$$

כנדרש. ■

נעבור לדון במכפלות של סדרות. לצורך כך ניעזר בטענה השימושית הבאה:

טענה 5.4.5 אם $a_n \to 0$ ואם (b_n) סדרה חסומה אז $a_n \cdot b_n \to 0$.

הוכחה לפי טענה 5.4.1 די להוכיח ש־ $|a_n b_n| \to 0$. יהי $M > 0$ חסם של (b_n). אז $|b_n| \le M$ ולכן

$$|a_n b_n| = |a_n| \cdot |b_n| \le M \cdot |a_n|$$

יהי $\varepsilon > 0$. החל ממקום מסוים מתקיים $|a_n| < \frac{\varepsilon}{M}$, ולכן החל ממקום מסוים מתקיים

$$|a_n b_n| \le \varepsilon$$

וזה מספיק להסיק $|a_n b_n| \to 0$. ■

למשל, נוכל להסיק בקלות ש־ $\lim_{n\to\infty} \frac{(-1)^n}{n} = 0$, כי הסדרה $\left(\frac{(-1)^n}{n}\right)$ היא מכפלה של הסדרה ההרמונית בסדרה החסומה $((-1)^n)$ (כמובן, אפשר גם להוכיח זאת ישירות).

משפט 5.4.6 (כלל המכפלה) יהיו (a_n),(b_n) סדרות עם גבולות b, a בהתאמה. אז

$$\lim_{n\to\infty}(a_n \cdot b_n) = a \cdot b$$

הוכחה מספיק שנוכיח כי

$$\lim_{n\to\infty} |a_n \cdot b_n - a \cdot b| = 0$$

ניעזר בזהות

$$xy - uv = (x-u)y + (y-v)u$$

אשר מתקיימת לכל ארבעה מספרים $x, y, u, v \in \mathbb{R}$ (בדקו על ידי פתיחת סוגריים!). נקבל

$$\begin{aligned} |a_n b_n - ab| &= |(a_n - a)b_n + (b_n - b)a| \\ &\le |a_n - a| \cdot |b_n| + |b_n - b| \cdot |a| \end{aligned}$$

אם נראה שכל אחד מהמחוברים בסכום האחרון שואף ל־0 נקבל מכלל החיבור שאגף ימין כולו שואף ל־0, ואז $|a_nb_n-ab|$ היא סדרה אי־שלילית שחסומה מלמעלה על ידי סדרה ששואפת לאפס, ומכלל הסנדוויץ נסיק ש־ $|a_nb_n-ab| \to 0$, כפי שאנו רוצים.

כל מחובר בצד ימין מורכב ממכפלה של שתי סדרות. לפי טענה **5.4.5**, מכפלה של סדרה חסומה בסדרה השואפת ל־0 שואפת גם היא ל־0 לכן מספיק להראות שכל מחובר באגף ימין הוא מכפלה של סדרות כאלה.

ואמנם, המחובר הראשון הוא מכפלה של הסדרה $|a_n-a|$, השואפת ל־0 כי $a_n \to a$, בסדרה $|b_n|$, שהיא סדרה חסומה כי היא מתכנסת. כמו־כן המחובר השני הוא מכפלת הסדרה $|b_n-b|$, השואפת ל־0, בסדרה הקבועה $|a|$, שהיא בוודאי סדרה חסומה. ■

מסקנה 5.4.7 תהי (a_n) סדרה ו־ $c \in \mathbb{R}$. אם $a_n \to a$ אז $ca_n \to ca$.

הוכחה אם חושבים על c כסדרה קבועה המתכנסת ל־c זה נובע מכלל המכפלה. ■

ההערות המופיעות אחרי ההוכחה של כלל הסכום (משפט **5.4.3**) חלות גם על מכפלות, לאחר השינויים המתבקשים בניסוח.

לדוגמה, נתבונן בסדרה $(3-\frac{1}{n})\cdot(2+\frac{1}{n})$. כיוון ש־ $\frac{1}{n} \to 0$, נובע מכלל הסכום ש־ $2+\frac{1}{n} \to 2$, ומכלל החיסור $3-\frac{1}{n} \to 3$. לכן לפי כלל המכפלה $(2+\frac{1}{n})\cdot(3-\frac{1}{n}) \to 6$.

מתבקש להוכיח גם משפט בנוסח "גבול של מנה הוא מנת הגבולות". משפט כזה אכן קיים אך יש להיזהר בניסוחו, שכן אם הגבול של הסדרה במכנה הוא אפס לא נוכל לחלק בו.

משפט 5.4.8 (כלל המנה) יהיו (a_n),(b_n) סדרות המתכנסות לגבולות a, b בהתאמה. אם $b \neq 0$ אז

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$$

הערה המנה $\frac{a_n}{b_n}$ מוגדרת רק כאשר $b_n \neq 0$. ההנחה $b \neq 0$ מבטיחה ש־ $b_n \neq 0$ החל ממקום מסוים (למה?), ולכן גם הסדרה $(\frac{a_n}{b_n})$ מוגדרת החל ממקום מסוים. זו הסדרה שלגבולה המשפט מתייחס. אם רוצים שהסדרה $\frac{a_n}{b_n}$ תהיה מוגדרת לכל n יש לדרוש במפורש ש־ $b_n \neq 0$ לכל n.

הוכחה נשים לב שמנה $\frac{x}{y}$ היא בעצם מכפלה $x \cdot \frac{1}{y}$, ולכן אם נדע שבתנאי המשפט מתקיים $\lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{b}$ נקבל לפי כלל המכפלה ש־

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty}(a_n \cdot \frac{1}{b_n}) = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} \frac{1}{b_n} = a \cdot \frac{1}{b} = \frac{a}{b}$$

אם כן מספיק שנוכיח כי בתנאים הנ"ל $\lim_{n\to\infty}\frac{1}{b_n}=\frac{1}{b}$. לשם כך מספיק שנוכיח כי $\lim_{n\to\infty}|\frac{1}{b_n}-\frac{1}{b}|=0$. נעבור למכנה משותף ונקבל

$$|\frac{1}{b_n}-\frac{1}{b}|=|\frac{b-b_n}{b\cdot b_n}|=\frac{1}{|b\cdot b_n|}\cdot|b-b_n|$$

לכן די שנראה שאגף ימין הוא סדרה השואפת ל־0. אגף ימין הוא מכפלה של שתי סדרות, ודי להראות שאחת שואפת ל־0 והשנייה חסומה. מאחר ש־ $b_n\to b$ מתקיים $|b_n-b|\to 0$, ונותר להראות שהסדרה $\frac{1}{|b\cdot b_n|}$ חסומה. זו סדרה אי־שלילית וממילא חסומה מלרע, ולכן די להוכיח שהיא חסומה מלעיל. לשם כך נתבונן בסדרה $|b\cdot b_n|$. לפי כלל המכפלה וטענה 5.4.1 סדרה זו מתכנסת ל־ b^2. מכיוון ש־ $b^2>0$ אנו יודעים ש־ $|b\cdot b_n|>\frac{b^2}{2}$ החל ממקום מסוים, ולכן $\frac{1}{|b\cdot b_n|}<\frac{2}{b^2}$ החל ממקום מסוים וממילא היא חסומה (מוכיחים זאת כמו בהוכחה של טענה 5.3.2). ■

נחשב לדוגמה את הגבול

$$\lim_{n\to\infty}\frac{n^3+6\cdot n^2-6}{3\cdot n^3+5\cdot n+10}$$

תוך שימוש באריתמטיקה של גבולות. נחלק את המונה והמכנה ב־n^3 ונקבל ביטוי שקול

$$\lim_{n\to\infty}\frac{1+\frac{6}{n}-\frac{6}{n^3}}{3+\frac{5}{n^2}+\frac{10}{n^3}}$$

(כאן לא השתמשנו עדיין במשפט האריתמטיקה. השוויון נובע מכך שהסדרות בתוך סימן הגבול שוות). אנו יודעים כי גבולות מהצורה $\lim\frac{1}{n^\alpha}$ עבור $\alpha\geq 1$ קיימות ושוות ל־0 (ראו דוגמה (1) בעמוד 101), ולכן נוכל לפי כלל המכפלה (או לפי טענה 5.4.5) להסיק

$$\lim_{n\to\infty}\frac{6}{n}=\lim_{n\to\infty}\frac{6}{n^3}=0$$

ומכלל הסכום נקבל

$$\lim_{n\to\infty}(1+\frac{6}{n}-\frac{6}{n^3})=1+0+0=1$$

באופן דומה נקבל

$$\lim_{n\to\infty}(3+\frac{5}{n^2}+\frac{10}{n^3})=3+0+0=3$$

ואז, לפי כלל המנה, הגבול המבוקש הוא מנת הגבולות האלה, דהיינו $\frac{1}{3}$. כפי שאנו רואים, דרך ההוכחה הזו מסורבלת ולכן בדרך כלל מדלגים על רוב השלבים ופשוט כותבים

$$\lim_{n\to\infty}\frac{1+\frac{6}{n}-\frac{10}{n^3}}{3+\frac{5}{n^2}+\frac{10}{n^3}}=\frac{1+\lim_{n\to\infty}\frac{6}{n}-\lim_{n\to\infty}\frac{10}{n^3}}{3+\lim_{n\to\infty}\frac{5}{n^2}+\lim_{n\to\infty}\frac{10}{n^3}}=\frac{1+0-0}{3+0+0}=\frac{1}{3}$$

נדגיש כי חישובים כאלה מוצדקים "מהסוף להתחלה". אריתמטיקה של גבולות דורשת קודם כל לדעת את קיום הגבולות הפשוטים ומתוך כך להסיק את קיומו וערכו של הגבול המורכב (ראו הערה (1) בעמוד 107). במילים אחרות, קיום הגבול מימין מוכיח את קיום הגבול משמאל.

משפט 5.4.9 (כלל החזקה) יהי $r \in \mathbb{Q}$ ו־ $a_n \to a$. נניח שהחזקה a^r מוגדרת וש־a_n^r מוגדרת החל ממקום מסוים. אז $a_n^r \to a^r$.

הוכחה נוכיח את המשפט בשלבים, תחילה למעריך טבעי, אח"כ למעריך שלם, אח"כ לשורשים שלמים ולבסוף למקרה הכללי.

1. **מעריך טבעי:** יהי $k \in \mathbb{N}$. העובדה ש־ $\lim_{n\to\infty} a_n^k = a^k$ נובעת באינדוקציה על k מכלל המכפלה (עבור $k = 2$ זהו בדיוק כלל המכפלה במקרה ששתי הסדרות שוות).
2. **מעריך שלם:** יהי $k \in \mathbb{Z}$. אם $k \geq 0$ זה המקרה הקודם. אם $k < 0$ אז לפי המקרה של מעריך טבעי מתקיים $\lim_{n\to\infty} a_n^{-k} = a^{-k}$. מההנחה ש־$a^k$ מוגדר אנו מקבלים $a \neq 0$ ולכן $a^{-k} \neq 0$, ואז לפי כלל המנה,
$$\lim_{n\to\infty} a_n^k = \lim_{n\to\infty} \frac{1}{a_n^{-k}} = \frac{1}{a^{-k}} = a^k$$
3. **שורשים:** יהי $k \in \mathbb{N}$ ונניח $a_n \to a$. נניח גם ש־ $a \geq 0$ וש־ $a_n > 0$ החל ממקום מסוים. נראה ש־ $\sqrt[k]{a_n} \to \sqrt[k]{a}$. שכן נניח שלא: אז יש $\varepsilon > 0$ כך שלאינסוף n־ים מתקיים
$$|\sqrt[k]{a_n} - \sqrt[k]{a}| \geq \varepsilon$$
מכאן שאו שלאינסוף n־ים מתקיים $\sqrt[k]{a_n} \geq \sqrt[k]{a} + \varepsilon$, או שלאינסוף n־ים מתקיים $\sqrt[k]{a_n} \leq \sqrt[k]{a} - \varepsilon$. נראה שבמקרה הראשון מקבלים סתירה. המקרה השני מושאר כתרגיל.

עבור כל n שמקיים $\sqrt[k]{a_n} \geq \sqrt[k]{a} + \varepsilon$ נקבל על ידי העלאה של האי־שוויון בחזקת k ש־
$$a_n \geq (\sqrt[k]{a} + \varepsilon)^k$$
כיוון ש־ $\sqrt[k]{a} + \varepsilon > \sqrt[k]{a}$ הרי ש־ $(\sqrt[k]{a} + \varepsilon)^k > a$, ואפשר למצוא $\delta > 0$ כך שיתקיים $(\sqrt[k]{a} + \varepsilon)^k = a + \delta$. קיבלנו ש־ $a_n \geq a + \delta$ לאינסוף n־ים, ומכיוון ש־ $\delta > 0$ זו סתירה לכך ש־ $a_n \to a$.
4. **מעריך רציונלי:** יהי $r = \frac{p}{q}$ עם p, q שלמים ו־ $q > 0$. אם $a_n \to a$ אז מהסעיף הקודם מקבלים $a_n^{1/q} \to a^{1/q}$ ולכן מכיוון ש־ p שלם מקבלים על סמך המקרה

של מעריך שלם ש־

$$\begin{aligned}\lim_{n\to\infty} a_n^r &= \lim_{n\to\infty}(a_n^{1/q})^p\\ &= (\lim_{n\to\infty} a_n^{1/q})^p\\ &= (a^{1/q})^p\\ &= a^{p/q}\end{aligned}$$

כנדרש (הצדיקו כל שוויון!). ■

הערות

1. בהנחת המשפט דרשנו במפורש ש־a^r יהיה מוגדר כי הדבר אינו נובע מכך ש־a_n^r מוגדר ו־ $a_n \to a$. למשל תהי $a_n = \frac{1}{n}$. אז a_n^{-1} מוגדרים אבל $a_n \to 0$ וכמובן 0^{-1} אינו מוגדר.

2. לאחר שנגדיר חזקות למעריך ממשי נוכיח שהמשפט תקף גם במקרה זה.

נעבור לנושא האחרון של הסעיף. אם $x_1, \ldots, x_n$ מספרים אז **הממוצע החשבוני** (average או arithmetic mean) שלהם הוא $\frac{1}{n}(x_1 + \ldots + x_n)$. בהינתן סדרה (a_n) ניתן ליצור את **סדרת הממוצעים** שלה

$$a_1, \frac{1}{2}(a_1+a_2), \frac{1}{3}(a_1+a_2+a_3), \ldots$$

האיבר ה־n של הסדרה החדשה הוא

$$s_n = \frac{1}{n}(a_1 + a_2 + \ldots + a_n) = \frac{1}{n}\sum_{k=1}^{n} a_k$$

האם יש קשר בין הגבול של (a_n) לגבול של סדרת הממוצעים (s_n)? בכיוון אחד יש:

משפט 5.4.10 (משפט צ'זארו)[5] תהי $(a_n)_{n=1}^{\infty}$ סדרה מתכנסת ו־ $s_n = \frac{1}{n}\sum_{k=1}^{n} a_k$ סדרת הממוצעים שלה. אם $a_n \to a$ אז גם $s_n \to a$.

הוכחה נניח תחילה $a_n \to 0$. די שנוכיח $|s_n| \to 0$. יהי C חסם של הסדרה (a_n) (יש כזה כי היא מתכנסת) ויהי $\varepsilon > 0$. קיים N כך שלכל $n > N$ מתקיים $|a_n| < \varepsilon$. לכל

[5] Ernesto Cesaro, 1859־1906.

$n>N$ אפשר לפרק את הסכום $a_1+\ldots+a_n$ לסכום של N המחוברים הראשונים, ועוד $n-N$ המחוברים הנותרים, ונקבל

$$\begin{aligned} |s_n| &= |\frac{a_1+\ldots+a_N+a_{N+1}+\ldots+a_n}{n}| \\ &\le \frac{1}{n}(|a_1|+\ldots+|a_N|)+\frac{1}{n}(|a_{N+1}|+\ldots+|a_n|) \\ &\le \frac{NC}{n}+\frac{1}{n}(n-N)\cdot\varepsilon \end{aligned}$$

האי־שוויון האחרון נובע מכך ש־ $|a_i|\le C$ לכל i ובסכום $|a_1|+\ldots+|a_N|$ מופיעים N מחוברים, ומכך ש־ $|a_i|<\varepsilon$ לכל $i>N$ ובסכום $|a_{N+1}|+\ldots+|a_n|$ מופיעים $n-N$ מחוברים. קיבלנו ש־

$$|s_n|\le\frac{NC}{n}+\varepsilon$$

לכל $n>N$. מכיוון ש־ $\lim_{n\to\infty}\frac{NC}{n}=0$ יש M כך שאם $n>M$ אז $NC/n<\varepsilon$. כעת אם $n>\max\{M,N\}$ אז מתקיים $|s_n|\le\varepsilon+\varepsilon=2\varepsilon$. קיבלנו שלכל ε, החל ממקום מסוים מתקיים $|s_n|<2\varepsilon$, ולכן $s_n\to 0$.

במקרה הכללי, נניח שמתקיים $a_n\to a$. נסמן $b_n=a_n-a$ ונגדיר את הממוצע $t_n=\frac{1}{n}(b_1+\ldots+b_n)$. מכיוון ש־ $b_n\to 0$, מהמקרה הקודם אנו יודעים ש־ $t_n\to 0$. מצד שני אם נסמן ב־(s_n) את סדרת הממוצעים של (a_n), מתקיים

$$t_n=\frac{1}{n}\sum_{k=1}^{n}b_k=\frac{1}{n}\sum_{k=1}^{n}(a_k-a)=\frac{1}{n}\sum_{k=1}^{n}a_k-\frac{1}{n}\sum_{k=1}^{n}a=s_n-a$$

ולכן הגבול $t_n\to 0$ פירושו $(s_n-a)\to 0$, וזה גורר $s_n\to a$, כנדרש. ■

ישנן פעולות אחרות שיכולות להיחשב מיצוע. אם $x<y$ מספרים חיוביים אז המספר $\sqrt{xy}$ נמצא ביניהם ויכול להיחשב סוג של ממוצע שלהם. באופן כללי יותר, אם $x_1,\ldots,x_n$ מספרים חיוביים אפשר לחשוב על המספר $\sqrt[n]{x_1\cdot x_2\cdot\ldots\cdot x_n}$ כעל מעין ממוצע שלהם. מספר זה נקרא **הממוצע הגאומטרי** או **ההנדסי** (geometric mean) של $x_1,\ldots,x_n$. נשאלת השאלה אם $a_n\to a$ גורר שהממוצעים הגאומטריים $g_n=\sqrt[n]{a_1\cdot a_2\cdot\ldots\cdot a_n}$ של a_n מתכנסים גם הם ל־a. התשובה חיובית:

מסקנה 5.4.11 אם (a_n) סדרה חיובית ו־ $a_n\to a$ אז הסדרה $g_n=\sqrt[n]{a_1\cdot\ldots\cdot a_n}$ מתכנסת גם־כן ל־a.

הוכחה נניח תחילה שהגבול a שונה מ־0. אז $1/a_n\to 1/a$. לפי אי־שוויון הממוצעים (תרגיל (2) בעמוד 85) מתקיים

$$\frac{1}{\frac{1}{n}(\frac{1}{a_1}+\ldots+\frac{1}{a_n})}\le\sqrt[n]{a_1\cdot\ldots\cdot a_n}\le\frac{1}{n}(a_1+\ldots+a_n)$$

צד ימין שואף ל־a לפי משפט צ'זארו. המכנה של צד שמאל שואף ל־ $1/a$ מאותה סיבה (וודאו שאתם מבינים מדוע!), ולכן צד שמאל גם־כן שואף ל־a. ממשפט הסנדוויץ' נובע ש־ $g_n \to a$ כנדרש.

אם $a = 0$, אז לא צריך את החצי השמאלי של האי־שוויון: מספיק לנו האי־שוויון

$$0 \leq \sqrt[n]{a_1 \cdot \ldots \cdot a_n} \leq \frac{1}{n}(a_1 + \ldots + a_n)$$

כדי להסיק $g_n \to 0$. ■

המסקנה האחרונה פותחת אפשרויות רבות לחישוב של גבולות שקודם לא יכולנו לחשב, או שחישובן היה קשה. הנה מסקנה שימושית:

למה 5.4.12 תהי $(a_n)_{n=1}^{\infty}$ סדרה של מספרים חיוביים כך ש־ $\frac{a_{n+1}}{a_n} \to a$. אז $\sqrt[n]{a_n} \to a$.

הוכחה נסמן $b_1 = a_1$ ועבור $n > 1$ נגדיר $b_n = \frac{a_n}{a_{n-1}}$. לפי הנתון $b_n \to a$, ולכן לפי המשפט הקודם

$$\sqrt[n]{a_n} = \sqrt[n]{\frac{a_n}{a_{n-1}} \cdot \frac{a_{n-1}}{a_{n-2}} \cdot \ldots \cdot \frac{a_2}{a_1} \cdot a_1} = \sqrt[n]{b_n \cdot b_{n-1} \cdot \ldots \cdot b_2 \cdot b_1} \to a$$

כנדרש. ■

מכאן אנו יכולים להסיק, למשל, ש־ $\sqrt[n]{n} \to 1$, כי $\frac{n+1}{n} \to 1$.

תרגילים

1. הוכיחו או הפריכו:
   (א) אם $|a_n| \to |a|$ אז $a_n \to a$.
   (ב) אם $(|a_n| - a) \to 0$ אז $|a_n| \to |a|$.
   (ג) אם $|a_n| \to a$ וגם $|b_n| \to b$ אז $(|a_n| - |b_n|) \to |a| - |b|$.
2. תנו הוכחה ישירה (כלומר, מהגדרת הגבול) לכך שאם $a_n \to a$ אז $|a_n| \to |a|$.
3. חשבו את הגבולות של הסדרות הבאות, או הוכיחו שהגבול אינו קיים:
   (א) $(1 + \frac{1}{n})(1 - \frac{1}{2^n})$.
   (ב) $\sqrt[n]{2} + \frac{n+1}{n^2+1}$.
   (ג) $\frac{1}{n} + \frac{1+(-1)^n}{\sqrt[n]{2}}$.
   (ד) $\frac{1}{n} \cdot \frac{1+(-1)^n}{\sqrt[n]{2}}$.
   (ה) $\frac{1/n^2}{1/n+1/n^2}$.

(ו) $\frac{1/n^2}{1/n^3+1/n^4}$.
(ז) $\frac{1/n^3}{1/n+1/n^2}$.
(ח) $\frac{1-\sqrt[2n]{3}}{1-\sqrt[n]{3}}$.
(ט) $\sqrt{\frac{1}{n}+\frac{1}{2^n}}$.

4. יהיו $(a_n),(b_n)$ סדרות.
   (א) הוכיחו שאם $a_n \to a$ ואם $(b_n-a_n)\to 0$ אז $b_n \to a$.
   (ב) הוכיחו שאם $a_n \to a \neq 0$ ואם $(\frac{b_n}{a_n})\to 1$ אז $b_n \to a$. הוכיחו זאת גם במקרה ש־ $a=0$.
5. הוכיחו שאם $a_n \to a$ אז $a_n \cdot a_{n+1} \to a^2$.
6. יהיו $x,y\in\mathbb{R}$.
   (א) הוכיחו שלכל $\varepsilon>0$ קיים δ כך ש־ $B_\delta(x)+B_\delta(y)\subseteq B_\varepsilon(x+y)$ (כאן מדובר בסכום של קבוצות: דהיינו $A+B=\{a+b\,:\,a\in A\,,\,b\in B\}$). השתמשו בעובדה זו כדי לתת הוכחה נוספת לכלל הסכום.
   (ב) הוכיחו שלכל $\varepsilon>0$ קיים $\delta>0$ כך ש־ $B_\delta(x)\cdot B_\delta(y)\subseteq B_\varepsilon(xy)$ (כאן מדובר במכפלה של קבוצות, דהיינו $A\cdot B=\{ab\,:\,a\in A\,,\,b\in B\}$). השתמשו בעובדה זו כדי לתת הוכחה נוספת לכלל המכפלה.
7. הוכיחו שאין סדרה חיובית המתכנסת ל־0 הכי מהר או הכי לאט. ליתר דיוק, הראו שאם (a_n) סדרה של מספרים חיוביים כך ש־ $a_n\to 0$, אז קיימות סדרות $(b_n),(c_n)$ המקיימות $|b_n|<|a_n|<|c_n|$ אבל $b_n\to 0$ וגם $c_n\to 0$.
8. תהי (a_n) סדרה כך ש־ $a_n\to 1$. הוכיחו או הפריכו:
   (א) $\frac{a_{n+1}}{a_n}\to 1$.
   (ב) $(a_n)^n\to 1$.
   (ג) $\sqrt[n]{a_n}\to 1$.
9. תהי (a_n) סדרה כך ש־ $a_n\to 0$. הוכיחו או הפריכו:
   (א) $\frac{a_{n+1}}{a_n}\to 0$.
   (ב) $a_n^n\to 0$.
   (ג) $\sqrt[n]{a_n}\to 0$.
10. השלימו את המקרה החסר בשלב (3) של משפט 5.4.9.
11. מצאו דוגמה לסדרה (a_n) כך שסדרת הממוצעים $s_n=\frac{1}{n}(a_1+\ldots+a_n)$ מתכנסת, אבל (a_n) מתבדרת.
12. יהיו $x_1,\ldots,x_n$ מספרים חיוביים. הוכיחו ש־

$$\min\{x_1,\ldots,x_n\}\leq\sqrt[n]{x_1\cdot\ldots\cdot x_n}\leq\max\{x_1,\ldots,x_n\}$$

ויש שוויון אמ"מ $x_1 = \ldots = x_n$.

13. תנו הוכחה ישירה ללמה 5.4.12.

14. הוכיחו שאם $a_n \to a$ אז הסדרה $t_n = \frac{1}{n^2}(a_1 + 2a_2 + 3a_3 + \ldots + na_n)$ מתכנסת ל־$\frac{1}{2}a$.

15. חשבו את $\lim_{n\to\infty} \sqrt[n]{\binom{2n}{n}}$.

16. הוכיחו שלכל $a > 1$ מתקיים $\frac{1}{2} \to {}^{1/n}(\sqrt[2n]{a} - 1)$.

17. (*) תהי (a_n) סדרה המקיימת $\lim_{n\to\infty} \frac{a_n}{n^a} = 0$ לכל $\alpha > 0$. האם נובע ש־$a_n \to 0$?

5.5 גבולות במובן הרחב

כאשר סדרה גדלה "בלי סוף" היא אינה מתכנסת (כי אינה חסומה) אך עדיין יש לה גבול במובן מסוים. למשל הסדרה $a_n = n$ היא כזאת: אין לה גבול אבל כש־n גדל היא גדלה בלי סוף.

הגדרה 5.5.1 תהי $(a_n)_{n=1}^{\infty}$ סדרה. נאמר שהיא **מתכנסת לאינסוף** (ל־∞), או שהיא שואפת לאינסוף, ונסמן זאת על ידי $\lim_{n\to\infty} a_n = \infty$ או $a_n \to \infty$, אם לכל מספר M מתקיים $a_n > M$ החל ממקום מסוים. באופן מפורט, $a_n \to \infty$ אם לכל M ממשי קיים N טבעי כך שלכל $n > N$ מתקיים $a_n > M$.[6]

נזכיר שוב ש־∞ אינו מספר ואין להתייחס אליו ככזה. בפרט אין לבצע בו פעולות חשבוניות.

דוגמאות

1. $\lim_{n\to\infty} n^2 = \infty$, כי לכל M ממשי ניתן לבחור מספר טבעי N עם $N^2 > M$ (למשל נבחר $N > \max\{M, 1\}$) ואז לכל $n > N$ מתקיים $n^2 > N^2 \geq M$.

2. יהי $s > 1$. נראה ש־ $s^n \to \infty$. ואמנם אפשר לרשום $s = 1 + c$ עבור מספר c חיובי, ואז מאי־שוויון ברנולי $s^n = (1+c)^n \geq cn$. לכן לכל M, אם $n > \frac{|M|}{c}$ אז $s^n \geq cn > |M| \geq M$ ולכן $s^n \to \infty$

באופן דומה נגדיר התכנסות במובן הרחב ל־$(-\infty)$:

---

[6] ניתן לקבל צורה שקולה של ההגדרה הזו שדומה יותר להגדרה הרגילה של הגבול בעזרת התכסיס הבא. נגדיר **סביבה של אינסוף** להיות קרן מהצורה (a, ∞). אז ההגדרה למעלה אומרת ש־ $a_n \to \infty$ אם לכל סביבה U של אינסוף מתקיים $a_n \in U$ החל ממקום מסוים.

הגדרה 5.5.2 תהי $(a_n)_{n=1}^{\infty}$ סדרה. נאמר כי הסדרה **מתכנסת ל־$-\infty$**, או שהיא שואפת למינוס־אינסוף, ונסמן זאת על ידי $\lim_{n\to\infty} a_n = -\infty$ או $a_n \to -\infty$, אם לכל M ממשי מתקיים $a_n < M$ החל ממקום מסוים.

כעת יש בידינו שני מושגים שונים שאנו מכנים התכנסות, ועלינו לדייק בלשוננו. כשנאמר ש־$(a_n)_{n=1}^{\infty}$ מתכנסת נתכוון תמיד שהגבול סופי, אלא אם צוין אחרת במפורש. אם נרצה להדגיש במיוחד שהגבול סופי, נאמר שהסדרה מתכנסת **במובן הצר**. כדי שנוכל לדבר על שני המושגים יחד, נכניס את המושג הבא:

הגדרה 5.5.3 נאמר שסדרה (a_n) **מתכנסת במובן הרחב** אם יש לה גבול ממשי או שהיא מתכנסת ל־∞ או ל־$-\infty$. נאמר שהיא **אינה מתכנסת במובן הרחב** אם אין לה גבול ממשי וגם $\pm\infty$ אינם גבולות שלה.

שימו לב שסדרה מתבדרת היא סדרה שאין לה גבול ממשי (הגדרה 5.2.7) ולכן סדרה המתכנסת ל־∞ היא סדרה מתבדרת (קיום גבול אינסופי שולל אפשרות של קיום גבול ממשי, כפי שנוכיח מיד).

ראינו שסדרה מתכנסת במובן הצר היא חסומה. באופן דומה אפשר להראות שסדרה השואפת לאינסוף חסומה מלרע (כמובן שהיא אינה חסומה מלעיל). אנו משאירים את ההוכחה לכך כתרגיל.

טענה 5.5.4 (יחידות הגבול במובן הרחב) אם סדרה מתכנסת במובן הרחב לגבולות a, b (ממשים או אינסופיים) אז $a = b$.

הוכחה יהיו a, b גבולות של (a_n) במובן הרחב.

אם $a_n \to a$ ו־$a \in \mathbb{R}$ אז (a_n) חסומה, ולכן אינה מתכנסת ל־$\pm\infty$. מכאן ש־ $b \in \mathbb{R}$, ולפי משפט היחידות של גבולות סופיים מסיקים $a = b$.

אם $a = \infty$ אז אין לסדרה גבול סופי כי הסדרה אינה חסומה (ולכן $b \notin \mathbb{R}$), ומאידך היא חסומה מלרע (ולכן $b \neq -\infty$). נותרה רק האפשרות $b = \infty$, כלומר $b = a$.

במקרה $a = -\infty$ מטפלים בצורה דומה. ■

נעיר שיש סדרות שאינן חסומות מלעיל ובכל־זאת אינן שואפות ל־∞ או ל־$-\infty$. למשל הסדרה

$$1, 0, 2, 0, 3, 0, 4, 0, 5, 0, \ldots$$

אינה חסומה אך היא אינה שואפת לאינסוף כי אין זה נכון שהיא גדולה מ־0 החל ממקום מסוים.

משפט 5.5.5 (מבחן ההשוואה) יהיו (a_n), (b_n) סדרות המקיימות $b_n \geq a_n$ לכל n החל ממקום מסוים. אם $a_n \to \infty$ אז גם $b_n \to \infty$, ואם $b_n \to -\infty$ אז גם $a_n \to -\infty$

הוכחה נניח כי $\lim_{n\to\infty} a_n = \infty$. יהי M ממשי. מהנתון מתקיים $a_n > M$ החל ממקום מסוים. כיוון ש־ $b_n \geq a_n$ לכל n גדול מספיק נובע ש־ $b_n > M$ החל ממקום מסוים. מכיוון ש־M היה שרירותי, $\lim_{n\to\infty} b_n = \infty$. המקרה $b_n \to -\infty$ דומה. ■

נרחיב עתה את משפטי האריתמטיקה של גבולות למקרה בו אחד מהגבולות הוא אינסופי. יש לשנות את ניסוחי המשפטים במקרה זה: אם $a_n \to a$ ו־ $b_n \to \infty$ לא נוכל להסיק ש־ $a_n + b_n \to a + \infty$, כי אין שום משמעות לביטוי מהצורה $a + \infty$. מצד שני, נראה שניתן להחליש קצת את ההנחות.

טענה 5.5.6 (כלל הסכום) אם $\lim_{n\to\infty} a_n = \infty$ ואם (b_n) סדרה חסומה מלרע אז $\lim_{n\to\infty}(a_n + b_n) = \infty$.

הוכחה יהי K חסם מלרע של הסדרה $(b_n)_{n=1}^{\infty}$. אז $a_n + b_n \geq a_n + K$, ולפי משפט ההשוואה לגבולות אינסופיים די להראות ש־ $a_n + K \to \infty$. יהי M ממשי. כל שעלינו להראות הוא ש־ $a_n + K > M$ החל ממקום מסוים, כלומר ש־ $a_n > M - K$ החל ממקום מסוים. אבל זה נכון כי $a_n \to \infty$, כפי שרצינו. ■

מסקנה 5.5.7 אם $\lim_{n\to\infty} a_n = \infty$ ואם (b_n) מתכנסת במובן הרחב לגבול השונה מ־ $-\infty$ אז מתקיים $\lim_{n\to\infty}(a_n + b_n) = \infty$

טענה 5.5.8 (כלל המכפלה) אם $\lim_{n\to\infty} a_n = \infty$ ויש $K > 0$ כך ש־ $b_n \geq K$ החל ממקום מסוים אז מתקיים $\lim_{n\to\infty} a_n \cdot b_n = \infty$. כמו־כן אם יש $K < 0$ כך ש־ $b_n \leq K$ החל ממקום מסוים אז $\lim_{n\to\infty} a_n \cdot b_n \to -\infty$.

הוכחה יהי $K > 0$ כאמור. הסדרה a_n חיובית החל ממקום מסוים (כי היא שואפת לאינסוף) ולכן $a_n b_n \geq K a_n$ החל ממקום מסוים. לכן די להראות ש־ $K a_n \to \infty$: יתר הפרטים דומים לסוף הוכחת טענה 5.5.6. החלק השני של הטענה מושאר כתרגיל. ■

מהטענה נובע ש־ $a_n \to \infty$ אמ"מ $-a_n \to -\infty$.

טענה 5.5.9 (כלל המנה) אם $a_n \to \pm\infty$ אז $\frac{1}{a_n} \to 0$. מאידך, אם $a_n \to 0$ ו־ $a_n \neq 0$ החל ממקום מסוים, אז $\frac{1}{|a_n|} \to \infty$.

הוכחה נוכיח את המקרה $a_n \to \infty$. ראשית נעיר שההנחה $a_n \to \infty$ גוררת ש־$a \neq 0$ החל ממקום מסוים ולכן $\frac{1}{a_n}$ מוגדר החל ממקום מסוים.

יהי $\varepsilon > 0$. עבור $M = \frac{1}{\varepsilon}$ מתקיים $a_n > M$ החל ממקום מסוים. לכן החל ממקום מסוים מתקיים $0 < \frac{1}{a_n} < \varepsilon$, ולכן $\lim_{n\to\infty} \frac{1}{a_n} = 0$.

ההוכחה של החלק השני מושארת כתרגיל. ■

הערות

1. בכלל המכפלה, הדרישה ש־(b_n) חסומה מלרע על ידי מספר חיובי ממש היא הכרחית, ולא מספיק שכל ה־b_n־ים יהיו חיוביים. לדוגמה, יהי $a_n = n$ ו־ $b_n = 1/n$. אז $a_n \to \infty$ ו־ $b_n > 0$ לכל n אבל $a_n b_n = 1 \to 1$ וממילא $a_n b_n \not\to \infty$.
2. ייתכן ש־ $a_n \to 0$ ובכל זאת $\frac{1}{a_n} \not\to \pm\infty$. למשל יהי $a_n = \frac{(-1)^n}{n}$. סדרה זו שואפת ל־0 אבל $\frac{1}{a_n} = (-1)^n \cdot n$, וזו סדרה שאינה חסומה מלעיל או מלרע וממילא אינה מתכנסת ל־∞ או ל־$-\infty$.

תרגילים

1. חשבו את הגבולות במובן הרחב של הסדרות הבאות, או הוכיחו שהגבול לא קיים:
   (א) $\frac{n^3+n^{-3}}{n^2-n^{-2}}$.
   (ב) $\frac{(\sqrt{n}+1)^2}{n^2+1}$.
   (ג) $\frac{(\sqrt{n}+1)(\sqrt{n}+(-1)^n)}{\sqrt{n}}$.
   (ד) $\sqrt[100]{n}$.
   (ה) $\frac{n^n}{n!}$.
   (ו) $\frac{2^n}{n}$.
   (ז) $2^{\sqrt{n}}$.
2. השלימו את החלקים החסרים של ההוכחות של טענות 5.5.8, 5.5.9.
3. נניח $a_n - b_n \to c$ עם $c \in \mathbb{R}$ ו־ $a_n \to \infty$. האם (b_n) מתכנס, ולאן?
4. נניח $\frac{a_n}{b_n} \to c$ עם $c \in \mathbb{R}$. אם $a_n \to \infty$, מה ניתן לומר על (b_n)? אם $b_n \to \infty$ מה ניתן לומר על (a_n)?
5. תהי (a_n) כך ש־ $a_n \to c > 0$. נגדיר $s_n = a_1 + \ldots + a_n$. הוכיחו ש־ $s_n \to \infty$.
6. נרצה לסדר את הסדרות בסדר לפי מהירות השאיפה שלהן לאינסוף. לשם כך נאמר ש־(a_n) שואפת לאינסוף מהר יותר מ־(b_n) אם $\frac{a_n}{b_n} \to \infty$. במקרה כזה נסמן $(a_n) \succ (b_n)$.
   (א) הוכיחו שהיחס $\succ$ הוא תורשתי במובן של אקסיומה VIII מסעיף 3.3. כלומר, אם $(a_n) \succ (b_n)$ ואם $(b_n) \succ (c_n)$ אז $(a_n) \succ (c_n)$.
   (ב) מצאו דוגמה לסדרות שונות השואפות לאינסוף שאף אחת אינה שואפת לאינסוף מהר יותר מהשנייה. מכאן שהיחס שהגדרנו אינו מקיים את תכונת ההשוואה (אקסיומה VII). בעקבות התכונות האלה אומרים ש־$\succ$ הוא **יחס סדר חלקי**.

(ג) נניח ש־$(a_n) \succ (b_n)$ וגם $(a_n) \succ (c_n)$. תהי (d_n) סדרה כלשהי השואפת לאינסוף. הוכיחו כי $(a_n) \succ (b_n + c_n)$, $(a_n + d_n) \succ (b_n)$ ו־ $(d_n \cdot a_n) \succ (b_n)$.

(ד) דרגו את הסדרות הבאות לפי הסדר החלקי $\succ$: $n!$, 2^n,, n^2, $\sqrt{n}$, n^n, n, $2^{\sqrt{n}}$.

(ה) שני הסעיפים האחרונים מאפשרים בקלות לקבוע לגבי סדרות רבות האם הן מתכנסות ל־∞ או ל־0 על ידי בדיקה של המרכיבים שלהן. למשל, הסדרה $\frac{n^n+2^n}{n!+n^k}$ מתכנסת ל־∞, כי המרכיב ה"שולט" הוא n^n והוא נמצא במונה. הוכיחו זאת באופן מדויק על סמך הסעיפים הקודמים!

7. לפעמים רוצים לדעת האם סדרה (a_n) שואפת ל־∞ בקצב גאומטרי (מעריכי או אקספוננציאלי), כלומר, האם ניתן לומר ש־ (a_n) גדלה "מהר כמו s^n" לאיזשהו מספר $s > 1$? שאלה זו טבעית בהקשרים רבים שכן תופעות רבות מהעולם האמתי ניתנות לתיאור על ידי סדרה עם גידול גאומטרי, כמו למשל מספר החיידקים במושבה (בתנאים אידאליים), ההחזר על השקעה צוברת ריבית, ומספר האתרים באינטרנט שניתן להגיע אליהם תוך n קישורים (לפחות כש־n לא גדול מדי).

(א) יהיו $s, t > 0$ ונניח $(t^n) \succ (a_n) \succ (s^n)$ במובן של הסעיף הקודם. הוכיחו שאם $x < s$ ו־ $y > t$ אז $(y^n) \succ (a_n) \succ (x^n)$.

(ב) תהי (a_n) סדרה. נגדיר

$$\sigma = \sup\{s \,:\, (a_n) \succ (s^n)\}\ ,\ \tau = \inf\{t \,:\, (t^n) \succ (a_n)\}$$

בהנחה שהקבוצות לא ריקות וחסומות. הוכיחו ש־ $\sigma \leq \tau$.

אם $\sigma = \tau$ נאמר ש**קצב הגידול** של (a_n) הוא **גאומטרי** (או מעריכי או אקספוננציאלי) עם בסיס σ.

(ג) חשבו את σ ו־τ לסדרות n^2, 2^n, $n2^n$, $\frac{2^n}{n}$, 2^{n^2}, $3^{\sqrt{n}}$.

(ד) תנו דוגמה לסדרה עבורה $0 < \sigma < \tau < \infty$.

(ה) הראו שיש סדרות $(a_n), (b_n)$ ששתיהן בעלות גידול אקספוננציאלי עם אותו בסיס אבל $(a_n) \succ (b_n)$.

8. יהיו $(a_n), (b_n)$ סדרות כך ש־ (b_n) עולה ממש ושואפת ל־∞ ונניח שמתקיים

$$\lim_{n\to\infty} \frac{a_{n+1} - a_n}{b_{n+1} - b_n} = L$$

הוכיחו שגם $\frac{a_n}{b_n} \to L$.

9. תהי $(a_n)_{n=1}^{\infty}$ סדרה המתכנסת ל־∞. תהי $s_n = \frac{1}{n}(a_1 + \ldots + a_n)$ סדרת הממוצעים האריתמטיים שלה. הוכיחו ש־ $s_n \to \infty$.

10. הוכיחו את ההכללה הבאה של משפט צ'זארו (משפט 5.4.10). תהי $(a_n)_{n=1}^{\infty}$ המתכנסת במובן הרחב ל־a. תהי $(\omega_n)_{n=1}^{\infty}$ סדרה אי־שלילית ונניח שהסדרה $\sigma_n = \omega_1 + \ldots + \omega_n$ מקיימת $\sigma_n \to \infty$. נגדיר

$$t_n = \frac{\omega_1 a_1 + \omega_2 a_2 + \ldots + \omega_n a_n}{\omega_1 + \ldots + \omega_n}$$

הוכיחו ש־ $t_n \to a$. שימו לב שאם $\omega_n = 1$ מקבלים בדיוק את משפט צ'זארו הרגיל.

5.6 סדרות מונוטוניות והלמה של קנטור

בסעיף זה נעסוק במחלקה מצומצמת של סדרות: הסדרות המונוטוניות. אלה סדרות שיש להן מגמה קבועה, כלומר, הן גדלות בהתמדה, או קטנות בהתמדה.

הגדרה 5.6.1 סדרה $(a_n)_{n=1}^{\infty}$ נקראת **סדרה עולה** (increasing) אם לכל n טבעי מתקיים $a_{n+1} \geq a_n$ ונקראת **סדרה יורדת** (decreasing) אם לכל n טבעי מתקיים $a_{n+1} \leq a_n$. סדרה נקראת **מונוטונית** (monotone) אם היא עולה או יורדת. אם בנוסף האי־שוויונות הם חזקים, נאמר כי הסדרה **עולה ממש** (strictly increasing) או **יורדת ממש** (strictly decreasing), או **מונוטונית ממש**[7] (strictly monotone) בהתאמה. אם רוצים להדגיש שסדרה היא מונוטונית אך אינה בהכרח מונוטונית ממש, אומרים שהיא מונוטונית **חלש**.

קל לראות שאם (a_n) עולה אז $a_m \geq a_n$ לכל $m \geq n$, כי מתקיימת שרשרת האי־שוויונות

$$a_m \geq a_{m-1} \geq \ldots \geq a_{n+1} \geq a_n$$

(באופן פורמלי מוכיחים זאת באינדוקציה על ההפרש $m-n$). אם הסדרה עולה ממש אלה הם אי־שוויונות חזקים, ונקבל שאם $m > n$ אז $a_m > a_n$. הערה דומה חלה על סדרות יורדות.

דוגמאות

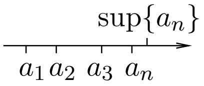
1. הסדרה $a_n = n$ עולה ממש.
2. הסדרה ההרמונית $\frac{1}{n}$ יורדת ממש.
3. הסדרה $(-1)^n$ אינה עולה ואינה יורדת.

[7] לפעמים אומרים שסדרה היא **מונוטונית חזק** במקום מונוטונית ממש.

4. הסדרות הקבועות הן גם עולות וגם יורדות, ואלה הסדרות היחידות בעלות תכונה זו (הוכיחו זאת!).

שימו לב שסדרה עולה היא בהכרח חסומה מלרע, ובעצם האיבר הראשון בסדרה הוא האיבר המינימלי שלה כי לכל $n \in \mathbb{N}$ מתקיים $a_n \geq a_1$. באותו אופן, האיבר הראשון בסדרה יורדת הוא המקסימום של הסדרה.

החשיבות של סדרות מונוטוניות נעוצה במשפט הבא:



איור 5.6.1

משפט 5.6.2 תהי $(a_n)_{n=1}^{\infty}$ סדרה עולה. אם (a_n) חסומה מלעיל אז הסדרה מתכנסת לגבול סופי ומתקיים $\lim_{n\to\infty} a_n = \sup\{a_n : n \in \mathbb{N}\}$. אחרת (כלומר אם (a_n) אינה חסומה מלעיל) מתקיים $a_n \to \infty$.

הוכחה ההוכחה למקרה ש־ (a_n) אינה חסומה קלה במיוחד, ונתחיל בה. בהינתן $M \in \mathbb{R}$ יש N כך ש־ $a_N > M$, כי הסדרה אינה חסומה מלעיל, וממונוטוניות נובע שלכל $n > N$ מתקיים $a_n \geq a_N > M$. מכאן ש־ $a_n \to \infty$.

נעבור למקרה ש־(a_n) חסומה. נסמן $a = \sup\{a_n : n \in \mathbb{N}\}$, ויהי $\varepsilon > 0$. מתכונות החסם העליון קיים N טבעי עבורו $a_N > a - \varepsilon$. היות והסדרה מונוטונית עולה, לכל $n > N$ מתקיים $a_n \geq a_N > a - \varepsilon$. מאידך, היות ו־$a$ חסם מלעיל של הסדרה הרי שלכל n, ובפרט ל־ $n > N$, מתקיים $a_n \leq a$, ולכן לכל $n > N$ מתקיים $|a_n - a| < \varepsilon$. מכיוון ש־ε היה שרירותי, נסיק ש־ $\lim_{n\to\infty} a_n = a$, כנדרש. ∎

להלן המשפט המקביל לסדרות מונוטוניות יורדות. ההוכחות מושארות כתרגיל.

משפט 5.6.3 תהי $(a_n)_{n=1}^{\infty}$ סדרה מונוטונית יורדת. אם (a_n) חסומה מלרע אז הסדרה מתכנסת לגבול סופי ומתקיים $\lim_{n\to\infty} a_n = \inf\{a_n : n \in \mathbb{N}\}$. אחרת (כלומר אם (a_n) אינה חסומה מלרע) מתקיים $a_n \to -\infty$.

מסקנה 5.6.4 כל סדרה מונוטונית היא סדרה מתכנסת במובן הרחב.

אפשר להגדיר מושג של סדרה "מונוטונית החל ממקום מסוים", במובן שיש N כך שהסדרה $(a_n)_{n=N}^{\infty}$ מונוטונית. סדרה כזאת גם־כן מתכנסת במובן הרחב (למה?) אבל הנוסחה $\lim a_n = \sup\{a_n : n \in \mathbb{N}\}$ אינה בהכרח נכונה (מה הנוסחה הנכונה?).

דוגמאות

1. הנה הוכחה נוספת לכך שאם $0 < s < 1$ אז $s^n \to 0$. ראשית, $s^n > 0$ לכל n, לכן (s^n) חסומה מלרע. מכיוון ש־ $s < 1$ מתקיים $s^{n+1} = s \cdot s^n < s^n$ ולכן (s^n) היא סדרה יורדת (אפילו יורדת ממש). לכן יש לה גבול. נסמן $L = \lim s^n$. כעת נשים לב שלפי משפט האריתמטיקה של גבולות מתקיים

$$s \cdot L = s \cdot \lim_{n\to\infty} s^n = \lim_{n\to\infty} s \cdot s^n = \lim_{n\to\infty} s^{n+1} = L$$

וקיבלנו $sL = L$. מכיוון ש־ $s \neq 1$ זה מחייב $L = 0$, כפי שרצינו.

2. ניתן הוכחה נוספת לכך ש־ $\sqrt[n]{a} \to 1$ לכל $a > 0$. נראה זאת למשל עבור $a > 1$. מתכונות השורש הסדרה $a_n = \sqrt[n]{a}$ יורדת ממש, והיא חסומה מלרע על ידי 1 וממילא מתכנסת. כדי להוכיח $a_n \to 1$ די אם כן להוכיח ש־ $\inf\{\sqrt[n]{a} : n \in \mathbb{N}\} = 1$.

נסמן $x = \inf\{\sqrt[n]{a} : n \in \mathbb{N}\}$. אם $x \neq 1$ אז $x > 1$ (למה?). לכן גם $\sqrt{x} > 1$, ונוכל לבחור מספר y המקיים

$$\sqrt{x} < y < x$$

(למשל $y = \frac{1}{2}(\sqrt{x} + x)$). נעלה אי־שוויון זה בריבוע ונקבל $x < y^2 < x^2$.

מכיוון ש־ $x < y^2$ ו־x הוא החסם התחתון של $\{\sqrt[n]{a} : n \in \mathbb{N}\}$, קיים n כך ש־ $\sqrt[n]{a} < y^2$. נוציא שורש ריבועי ונקבל ש־

$$\sqrt[2n]{a} < y$$

ובשילוב עם האי־שוויון $y < x$ קיבלנו $\sqrt[2n]{a} < x$. אבל זו סתירה להגדרת x, וסיימנו.

3. יהי x מספר חיובי. בדוגמה זו נתאר סדרה מפורשות $(a_n)_{n=1}^{\infty}$ המתכנסת ל־$\sqrt{x}$.

נבחר מספר חיובי a_1 כלשהו, שנחשוב עליו כעל ניחוש התחלתי לערך של $\sqrt{x}$. אפשר, למשל, לבחור $a_1 = 1$ או $a_1 = x$. נשים לב שאם $a_1 \leq \sqrt{x}$ אז $\frac{x}{a_1} \geq \sqrt{x}$, ושאם $a_1 \geq \sqrt{x}$ אז $\frac{x}{a_1} \leq \sqrt{x}$. בכל מקרה אנו רואים ש־ $\sqrt{x}$ נמצא בין a_1 לבין x/a_1, ולכן ניתן לשער שהממוצע בין שני המספרים a_1 ו־ x/a_1 יהיה קרוב יותר ל־$\sqrt{x}$ מאשר הניחוש הראשון a_1. לכן נגדיר $a_2 = \frac{1}{2}(a_1 + \frac{x}{a_1})$. את אותו השיקול נוכל להפעיל על הקירוב a_2, וכן הלאה. נקבל סדרה המוגדרת ברקורסיה על ידי

$$a_{n+1} = \frac{1}{2}(a_n + \frac{x}{a_n})$$

קל להראות באינדוקציה ש־ $a_n > 0$ לכל n ולכן המחובר $\frac{x}{a_n}$ בהגדרה של a_{n+1} מוגדר. אנו נראה שהסדרה (a_n) מתכנסת ל־$\sqrt{x}$. ראשית נוכיח שהסדרה מתכנסת, ואח"כ נברר מהו הגבול. כדי להראות ש־(a_n) מתכנסת מספיק שנוכיח שהיא חסומה מלרע ומונוטונית יורדת החל ממקום מסוים.

ואמנם, הסדרה חסומה מלרע כי היא חיובית, ונותר להוכיח שהיא מונוטונית. אנו רוצים של־n מספיק גדול יתקיים $a_n - a_{n+1} \geq 0$. ואכן:

$$\begin{aligned} a_n - a_{n+1} \geq 0 \quad &\iff \quad a_n - \frac{a_n + x/a_n}{2} \geq 0 \\ &\iff \quad 2 \cdot a_n - a_n - \frac{x}{a_n} \geq 0 \\ &\iff \quad a_n^2 - x \geq 0 \end{aligned}$$

עבור $n > 1$ נוכל להציב את ההגדרה הרקורסיבית של a_n ונקבל

$$\begin{aligned} a_n^2 - x \geq 0 \quad &\iff \quad \frac{1}{4}(a_{n-1} + \frac{x}{a_{n-1}})^2 \geq x \\ &\iff \quad a_{n-1} + \frac{x}{a_{n-1}} \geq 2\sqrt{x} \\ &\iff \quad a_{n-1}^2 - 2a_{n-1}\sqrt{x} + x \geq 0 \\ &\iff \quad (a_{n-1} - \sqrt{x})^2 \geq 0 \end{aligned}$$

והאי־שוויון האחרון כמובן נכון, כי ריבוע של כל מספר הוא אי־שלילי. אם כן, הוכחנו שהסדרה (a_n) חסומה מלרע ויורדת החל מהמקום השני, ולכן היא מתכנסת לגבול סופי a.

נותר רק לחשב את הגבול a. נשים לב ש־a מקיים

$$a = \lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \frac{a_n + \frac{x}{a_n}}{2} = \frac{a + \frac{x}{a}}{2}$$

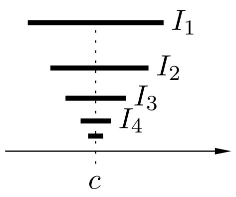
(השוויון האחרון נובע מאריתמטיקה של גבולות. הצדיקו את יתר השוויונות!) קיבלנו שהגבול מקיים את המשוואה $a = \frac{1}{2}(a+\frac{x}{a})$, השקולה למשוואה $a^2 = x$, ולכן $a = \pm\sqrt{x}$. לגבי הסימן, מכיוון שהסדרה חיובית לא ייתכן שגבול הסדרה יהיה שלילי, ולכן נסיק $a = \sqrt{x}$ כפי שרצינו.

למשפט הבא שנתאר יש טעם גאומטרי. נניח שנתונה סדרה של קטעים I_n בישר כך שכל אחד מוכל בקודם, כלומר $I_{n+1} \subseteq I_n$. במצב כזה אפשר לצפות שיהיו נקודות משותפות לכל הקטעים או במלים אחרות, שהחיתוך של הקטעים יהיה קבוצה לא ריקה (ראו איור 5.6.2 להלן). ללא הנחות נוספות, מסקנה זו אינה נכונה.

דוגמאות

1. עבור הקטעים $I_n = [0, \frac{1}{n}]$ קל לראות שחיתוך כל הקטעים הוא $\{0\}$ (למה?). גם עבור בו $J_n = [-\frac{1}{n}, \frac{1}{n}]$ החיתוך הוא $\{0\}$.
2. כאשר בודקים מה קורה אם הקטעים אינם סגורים מגלים שהמצב עלול להיות שונה. למשל, עבור $I'_n = (0, \frac{1}{n})$ לא קשה להוכיח כי החיתוך הוא ריק. מצד שני עבור $J'_n = (-\frac{1}{n}, \frac{1}{n})$ החיתוך הוא שוב $\{0\}$.
3. החיתוך של סדרת קטעים כזו יכול גם להכיל יותר מנקודה אחת. למשל נתבונן בקטעים $K_n = [-1 - \frac{1}{n}, 1 + \frac{1}{n}]$. גם כאן הקטעים מוכלים אחד בשני (וההכלה היא הכלה ממש, כלומר, $K_{n+1} \neq K_n$) אבל הפעם החיתוך הוא כל הקטע $[-1, 1]$, המכיל אינסוף נקודות.

נזכיר שאם $I = [a, b]$ קטע אז $|I|$ מסמן את אורך הקטע, דהיינו את $|b - a|$.



איור 5.6.2

משפט 5.6.5 (הלמה של קנטור[8]) תהי $(I_n)_{n=1}^{\infty}$ סדרת קטעים סגורים וחסומים בישר הממשי בעלי התכונה ש־ $I_{n+1} \subseteq I_n$ לכל n, ונניח $\lim_{n\to\infty} |I_n| = 0$. אז קיים מספר יחיד c השייך לכל הקטעים, כלומר

$$\bigcap_{n=1}^{\infty} I_n = \{c\}$$

הוכחה נסמן $I_n = [a_n, b_n]$. התנאי $I_{n+1} \subseteq I_n$ גורר שהסדרה (a_n) עולה (במובן החלש), והסדרה $(b_n)_{n=1}^{\infty}$ יורדת (במובן החלש). כמו־כן הנתון ש־ $|I_n| \to 0$ שקול לכך ש־ $|b_n - a_n| \to 0$.

עלינו להראות שיש נקודה יחידה x השייכת לכל הקטעים, או באופן שקול שיש x יחיד כך ש־ $a_n \le x \le b_n$ לכל n. נטפל תחילה בשאלת היחידות. אם x, y מספרים כאלה אז

$$0 \le |x - y| \le |b_n - a_n|$$

לכל n. האגפים הקיצוניים של האי־שוויון שואפים לאפס, ולכן לפי משפט הסנדוויץ' מקבלים $|x - y| = 0$, דהיינו $x = y$.

נראה כעת שיש x כזה. נשים לב שלכל n מתקיים $a_n \le b_n \le b_1$ ולכן b_1 חסם מלעיל של הסדרה (a_i). לכן הקבוצה $\{a_n : n \in \mathbb{N}\}$ חסומה מלעיל ואפשר להגדיר

$$c = \sup\{a_n : n \in \mathbb{N}\}$$

מההגדרה ברור ש־ $c \ge a_n$ לכל n. נותר להראות ש־ $c \le b_n$ לכל n, כי אז $c \in [a, b]$ (כאן אנו משתמשים בכך שהקטעים סגורים!). לשם כך די להראות שלכל m המספר b_m הוא חסם מלעיל של הקבוצה $\{a_n : n \in \mathbb{N}\}$. ואמנם אם $n \ge m$ אז $a_n \le b_n \le b_m$ (האי־שוויון הימני כי הסדרה (b_i) יורדת), ומצד שני אם $n \le m$ אז $a_n \le a_m \le b_m$ (האי־שוויון השמאלי כי (a_i) סדרה עולה), כפי שרצינו. ∎

שימו לב שהשתמשנו בהנחה $|I_n| \to 0$ רק כדי להוכיח את יחידות הנקודה בחיתוך הקטעים. אותה הוכחה נותנת את התוצאה הבאה, המושארת כתרגיל:

טענה 5.6.6 תהי $(I_n)_{n=1}^{\infty}$ סדרת קטעים סגורים וחסומים כך ש־ $I_n \subseteq I_{n+1}$ לכל n. אז החיתוך $\cap_{n=1}^{\infty} I_n$ הוא קטע סגור לא ריק.

[8] Georg Cantor, 1845‏-1918.

תרגילים

1. אילו מהסדרות הבאות מונוטוניות?
   (א) 2^n.
   (ב) $2^n - n$.
   (ג) $\frac{(-1)^n}{n}$.
   (ד) $n + \frac{(-1)^n}{n}$.
   (ה) $n!$.
   (ו) $\frac{n!}{n^n}$.
   (ז) $\sqrt[n]{n}$.

2. הראו שסדרה (a_n) היא עולה או עולה ממש אם ורק אם הסדרה $(-a_n)$ היא יורדת או יורדת ממש, בהתאמה. מתי עלייה של (a_n) גוררת עליה של $(\frac{1}{a_n})$?

3. הראו שלכל $x \in \mathbb{R}$ קיימת סדרה מונוטונית עולה ממש של מספרים רציונליים (r_n) כך ש־ $r_n \to x$.

4. נסחו והוכיחו את הגרסה של משפט 5.6.2 המתאימה לסדרות יורדות.

5. הוכיחו שאם $a_n \to a$ אז הסדרה $b_n = \max\{a_1, \ldots, a_n\}$ מתכנסת. מתי $\lim a_n = \lim b_n$?

6. בשאלה זו תתנו פירוש לסימן

$$\sqrt{2\sqrt{2\sqrt{2\sqrt{2}}}}$$

יהי $a_1 = \sqrt{2}$ ונגדיר $a_{n+1} = \sqrt{2a_n}$. הראו ש־a_n מתכנסת וחשבו את גבולה.

7. יהי x מספר, ונגדיר $a_1 = x$ ו־ $a_{n+1} = \frac{1}{2}(a_n^2 + 1)$. בדקו את התכנסות הסדרה עבור הערכים $x = \frac{1}{2}, x = 2$ וחשבו את הגבולות (אם הם קיימים).

8. יהיו $x, y > 0$ ונגדיר $a_1 = 1$, וברקורסיה $a_{n+1} = \sqrt{x + a_n \cdot y}$. הוכיחו שהסדרה מתכנסת וחשבו את גבולה.

9. יהיו $a_1, b_1 > 0$. נגדיר ברקורסיה $a_{n+1} = \frac{1}{2}(a_n + b_n)$ ו־ $b_{n+1} = \sqrt{a_n \cdot b_n}$. הוכיחו שהסדרות $(a_n), (b_n)$ מתכנסות, ולאותו גבול. לגבול קוראים **הממוצע האריתמטי־גאומטרי** (arithmetic-geometric mean) של a_1, b_1. (רמז: הראו ש־ $a_n \le b_n$ לכל $n \ge 2$ וש־ $|b_{n+1} - a_{n+1}| < \frac{1}{2}|b_n - a_n|$).

10. יהיו $I_n = (a_n, b_n)$ סדרה יורדת של קטעים פתוחים, כלומר $I_{n+1} \subseteq I_n$ לכל n. נניח שהסדרה (a_n) עולה ממש והסדרה (b_n) יורדת ממש. האם יש נקודה משותפת לכל הקטעים?

11. הוכיחו בפירוט את ההכללה של משפט קנטור (טענה 5.6.6).

12. בתרגיל זה נפתח שיטה לייצג במפורש כל מספר ממשי. נגדיר ברקורסיה על n כלל המתאים לכל סדרה סופית $a_0 \dots a_n$ של מספרים טבעיים מספר רציונלי המסומן $[a_0, \dots, a_n]$ ונקרא **שבר משולב** (continued fraction). אם $n = 0$ נגדיר $[a_0] = a_0$, וברקורסיה

$$[a_0, \dots, a_n] = a_0 + \frac{1}{[a_1, \dots, a_n]}$$

למשל,

$$[1, 2, 3, 4] = 1 + \frac{1}{[2, 3, 4]} = 1 + \frac{1}{2 + \frac{1}{[3,4]}} = 1 + \frac{1}{2 + \frac{1}{3+1/4}}$$

חישוב פשוט על ידי מעבר למכנה משותף מראה ש־ $[1, 2, 3, 4] = \frac{43}{30}$.

(א) תהי $(a_n)_{n=0}^{\infty}$ סדרה של מספרים טבעיים חיוביים, למעט a_0 שיכול להיות 0. הוכיחו שהסדרה $f_n = [a_0, \dots, a_n]$ מתכנסת. הגבול שלה מסומן $[a_0, a_1, \dots]$ או $[a_n]_{n=0}^{\infty}$ ונקרא **שבר משולב אינסופי** (רמז: הראו שהסדרות

$$A_n = [a_0, \dots, a_{2n}] \qquad , \qquad B_n = [a_0, \dots a_{2n+1}]$$

מקיימות $A_n \le A_{n+1} \le B_{n+1} \le B_n$).

(ב) יהי $x > 0$. נגדיר סדרה (x_n) ברקורסיה: יהי $x_0 = x$. בשלב ה־n אם x_n שלם נפסיק, אחרת נגדיר $x_{n+1} = (x_n - \lfloor x_n \rfloor)^{-1}$. לכל n עבורו הגדרנו את x_n יהי $a_n = \lfloor x_n \rfloor$. הוכיחו שאם התהליך נעצר אחרי מספר סופי N של צעדים אז x רציונלי ומתקיים $x = [a_0, \dots, a_N]$ ואחרת $x = [a_0, a_1, \dots]$.

(ג) הראו שלמספר רציונלי חיובי יש בדיוק שתי הצגות כשבר משולב סופי, ושלמספר אי־רציונאלי יש בדיוק הצגה אחת.

(ד) מצאו נוסחות אלגבריות $[1, 1, 1, 1, \dots]$ ו־ $[1, 2, 2, 2, 2, \dots]$.

13. (*) סדרה אי־שלילית $(a_n)_{n=1}^{\infty}$ נקראת **תת אדיטיבית** (sub-additive) אם היא המקיימת $a_{n+m} \le a_n + a_m$ לכל $m, n \in \mathbb{N}$. הוכיחו שהגבול $\lim_{n\to\infty} \frac{a_n}{n}$ של סדרה תת־אדיטיבית קיים במובן הרחב ומתקיים

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf\{\frac{a_n}{n} \ : \ n \in \mathbb{N}\}$$

כאשר מפרשים את החסם התחתון של קבוצה שאינה חסומה מלרע בתור $-\infty$ (רמז: הראו ש־ $\frac{a_{kn}}{kn} \le \frac{a_m}{m}$).

14. (*) הוכיחו את משפט היינה־בורל (Heine-Borel): תהי Ω קבוצת אינדקסים ולכל $\omega \in \Omega$ נניח שנתון קטע פתוח I_ω. יהי $[a,b]$ קטע סגור ונניח שלכל $x \in [a,b]$ יש $\omega \in \Omega$ כך ש־ $x \in I_\omega$, דהיינו $[a,b] \subseteq \cup_{\omega\in\Omega} I_\omega$ (אומרים אז

שהאוסף $\{I_\omega : \omega \in \Omega\}$ הוא **כיסוי פתוח** של $[a,b]$). הוכיחו שיש קבוצת אינדקסים סופית $\{\omega_1,\dots,\omega_n\} \subseteq \Omega$ כך שהאוסף הסופי $\{I_{\omega_1},\dots,I_{\omega_n}\}$ כבר מכסה את $[a,b]$, דהיינו $[a,b] \subseteq \cup_{i=1}^n I_{\omega_i}$ (רמז: הראו שיש מספר $t \in [a,b]$ גדול ביותר שעבורו קיים תת־אוסף סופי של קטעים שמכסה את $[a,t]$, והראו ש־ $t = b$).

5.7 תת־סדרות וגבולות חלקיים

נסמן ב־$(a_n)_{n=1}^\infty$ את הסדרה ההרמונית $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$. אם נמחק חלק מאיבריה, ואם נשארו בכל־זאת אינסוף מאיברי הסדרה המקורית, ניתן לחשוב על האיברים הנותרים כאיבריה של סדרה חדשה. למשל, אם נמחק את האיברים עם אינדקסים אי־זוגיים מהסדרה ההרמונית נקבל את הסדרה $\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \dots$. אז נקרא לסדרה זו $(b_k)_{k=1}^\infty$ אז איבריה נתונים על ידי $b_k = \frac{1}{2k}$. נשים לב שניתן לכתוב את b_k גם במונחים של (a_n), על ידי $b_k = a_{2k}$. כלומר, b_k הוא האיבר ה־k מבין האיברים שנשארו מהסדרה (a_n) לאחר מחיקת האיברים האי־זוגיים, שהוא האיבר ה־$2k$ של הסדרה המקורית (a_n).

a_1	
a_2	b_1
a_3	
a_4	b_2
a_5	b_3
a_6	
a_7	
a_8	b_4
$\vdots$	$\vdots$

איור 5.7.1

באופן כללי, תת־סדרה של סדרה (a_n) מתקבלת על ידי בחירה של חלק מאיברי הסדרה המקורית, או באופן שקול, מחיקת האיברים האחרים, כפי שמודגם באיור 5.7.1. אם האיברים שבחרנו הם אלה שבאינדקסים $n_! < n_2 < n_3 <$, אז הסדרה (b_k) המתקבלת מוגדרת על ידי $b_k = a_{n_k}$.

הגדרה 5.7.1 תהי $(a_n)_{n=1}^\infty$ סדרה ותהי $(n_k)_{k=1}^\infty$ סדרה עולה ממש של מספרים טבעיים חיוביים. הסדרה $(b_k)_{k=1}^\infty$ המוגדרת על ידי $b_k = a_{n_k}$ נקראת **תת־סדרה** (subsequence) של $(a_n)_{n=1}^\infty$. אומרים אז ש־(b_n) התקבלה מ־(a_n) בעזרת סדרת האינדקסים $(n_k)_{k=1}^\infty$.[9]

כפי שהערנו בתחילת הפרק, האות המייצגת אינדקס של סדרה היא במידה רבה שרירותית: אם $(a_n)_{n=1}^\infty$ היא סדרה אז היא זהה לסדרה $(a_i)_{i=1}^\infty$. כאשר עוברים לתת־סדרות המצב משתנה בהתאם. למשל תהי $a_n = \frac{1}{n}$ הסדרה ההרמונית, ונגדיר שתי תת־סדרות של (a_n) בעזרת סדרות האינדקסים $n_k = 2k\,,\ m_k = 2k+1$. התת־סדרות $(a_{n_k})_{k=1}^\infty$ ו־ $(a_{m_k})_{k=1}^\infty$ שונות זו מזו לחלוטין! (הציבו כמה ערכים של k כדי להשתכנע). לעומת זאת, $(a_{n_k})_{k=1}^\infty$ שווה לסדרה $(a_{n_j})_{j=1}^\infty$, כי כעת j,k הם האינדקסים החופשיים.

דוגמאות

1. הסדרה $b_k = \frac{1}{2k}$ היא תת־סדרה של הסדרה ההרמונית $a_n = \frac{1}{n}$ כי עבור סדרת האינדקסים $n_k = 2k$ מתקיים $b_k = a_{n_k}$. מצד שני, (a_n) אינה תת־

[9] לפעמים קוראים לתת־סדרה גם **סדרה חלקית**.

סדרה של (b_n) כי אין אף אינדקס k כך ש־ $b_k = a_3 = \frac{1}{3}$.

2. כל סדרה היא תת־סדרה של עצמה (המתאימה לסדרת האינדקסים $n_k = k$). ייתכנו גם שתי סדרות שכל אחת היא תת־סדרה של השנייה, כמו למשל עבור

$$\begin{aligned} a_n &= 0,1,0,1,0,1,0,1,0,1,0,1\ldots \\ b_n &= 0,0,1,1,0,0,1,1,0,0,1,1,\ldots \end{aligned}$$

(מצאו סדרות אינדקסים כך ש־ $a_k = b_{n_k}$ ו־ $b_k = a_{m_k}$!).

3. הסדרה הקבועה 1 היא תת־סדרה של הסדרה $(a_n)_{n=1}^{\infty} = ((-1)^n)_{n=1}^{\infty}$ אך נשים לב שיש דרכים רבות לקבל אותה. למשל ניתן לקבל אותה על ידי בחירת האינדקסים $n_k = 2k$ (המקומות הזוגיים) וגם על ידי בחירת האינדקסים $m_k = k\cdot(k+1)$ (שימו לב ש־ m_k תמיד זוגי, אך (m_k) אינה כוללת את כל המספרים הזוגיים). אף שסדרות האינדקסים שונות, שתי התת־סדרות האלה שוות: $(a_{n_k})_{k=1}^{\infty} = (a_{m_k})_{k=1}^{\infty}$.

למה 5.7.2 (תכונות של סדרות אינדקסים ושל תת־סדרות)

1. אם $(n_k)_{k=1}^{\infty}$ סדרה עולה ממש של מספרים טבעיים חיוביים, אז $n_k \geq k$ לכל $k \in \mathbb{N}$.
2. תהי $(n_k)_{k=1}^{\infty}$ סדרה עולה ממש של מספרים טבעיים ותהי $P(m)$ טענה התלויה ב־m.
   (א) אם $P(m)$ מתקיימת לכל m גדול מספיק אז $P(n_k)$ מתקיימת לכל k גדול מספיק.
   (ב) אם $P(n_k)$ מתקיימת לכל k גדול מספיק אז $P(m)$ מתקיימת לאינסוף m־ים.
3. תהי (a_n) סדרה. כל תת־סדרה של תת־סדרה של (a_n) היא בעצמה תת־סדרה של (a_n).
4. אם $A \subseteq \mathbb{N}$ קבוצה אינסופית של הטבעיים אז יש סדרת אינדקסים עולה ממש $(n_k)_{k=1}^{\infty}$ כך ש־ $A = \{n_k \,:\, k \in \mathbb{N}\}$.

הוכחה רוב התכונות ברורות, ונוכיח רק את הסעיף האחרון. נגדיר $n_1 = \min A$ וברקורסיה נגדיר את n_k להיות האיבר המינימלי של $A \setminus \{n_1,\ldots,n_{k-1}\}$. יש איבר כזה כי A אינסופית ולכן $A \setminus \{n_1,\ldots,n_{k-1}\}$ קבוצה לא ריקה של מספרים טבעיים. קל לוודא של־(n_k) יש את התכונות המבוקשות. ■

הגדרה 5.7.3 מספר ממשי a נקרא **גבול חלקי**[10] (accumulation points) של הסדרה (a_n) אם קיימת תת־סדרה $(a_{n_k})_{k=1}^{\infty}$ של (a_n) עבורה $\lim_{k\to\infty} a_{n_k} = a$. כמו־כן נאמר ש־$\infty$ הוא גבול חלקי של הסדרה אם יש תת־סדרה $(a_{n_k})_{n=1}^{\infty}$ של (a_n) כך ש־ $\lim_{k\to\infty} a_{n_k} = \infty$. באופן דומה מגדירים מתי $-\infty$ הוא גבול חלקי של הסדרה.

[10] לפעמים קוראים לגבול חלקי גם **נקודת הצטברות**.

כרגיל, כשנאמר ש־a גבול חלקי של (a_n) נתכוון תמיד ש־a ממשי אלא אם צוין אחרת. אם נאמר ש־a גבול חלקי **במובן הרחב** נתכוון שהוא או ממשי או $\pm\infty$.

דוגמאות

1. לסדרה הקבועה $a_n = c$ יש רק גבול חלקי אחד, c, כי כל תת־סדרה של (a_n) היא גם הסדרה הקבועה c, ולכן מתכנסת ל־c.
2. נתבונן בסדרה $((-1)^n)_{n=1}^{\infty}$. תת־הסדרה המתקבלת מהמקומות הזוגיים היא הסדרה הקבועה שערכה 1, ותת הסדרה המתקבלת מהמקומות האי־זוגיים היא הסדרה הקבועה שערכה -1. לכן $1, -1$ הם גבולות חלקיים של הסדרה המקורית. למעשה אלה הגבולות החלקיים היחידים. שכן אם (a_n) סדרה המכילה את 1 ואת -1 אינסוף פעמים אז היא אינה מתכנסת, ולכן תת־סדרה מתכנסת של הסדרה $a_n = (-1)^n$ חייבת להיות קבועה החל ממקום מסוים, עם ערך 1 או -1. זה גורר שהגבול של התת־סדרה הוא 1 או -1, בהתאמה.
3. ייתכן שלסדרה יהיו אינסוף גבולות חלקיים. נתבונן לדוגמה בסדרה

$$1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 4, 5, 1, 2, 3, 4, 5, 6, 1, 2, 3, 4, 5, 6, 7, \ldots$$

בסדרה זו כל מספר טבעי מופיע אינסוף פעמים. לכן כל מספר טבעי הוא גבול חלקי שלה. לא קשה למצוא סדרה חסומה עם אינסוף גבולות חלקיים (בנו סדרה כזו!).

4. אם $a_n \to a$ אז a גבול חלקי של הסדרה (a_n), כי (a_n) תת־סדרה של עצמה.

הדוגמה האחרונה היא מקרה פרטי של הטענה הכללית הבאה:

טענה 5.7.4 אם $(a_n)_{n=1}^{\infty}$ סדרה המתכנסת במובן הרחב לגבול a אז כל סדרה חלקית שלה מתכנסת גם היא לאותו גבול.

הוכחה נניח שהגבול של (a_n) סופי ושווה למספר ממשי a. תהי $(a_{n_k})_{k=1}^{\infty}$ תת־סדרה של $(a_n)_{n=1}^{\infty}$. עלינו להראות שלכל סביבה $B_\varepsilon(a)$ של a, מתקיים $a_{n_k} \in B_\varepsilon(a)$ עבור k גדול מספיק. אבל לכל n גדול מספיק מתקיים $a_n \in B_\varepsilon(a)$ ולכן לפי למה 5.7.2 לכל k גדול מספיק מתקיים $a_{n_k} \in B_\varepsilon(a)$, כנדרש.

ההוכחה למקרה בו $a = \pm\infty$ דומה ומושארת כתרגיל. ■

מסקנה 5.7.5 אם (a_n) סדרה בעלת שני גבולות חלקיים שונים (במובן הרחב) אז היא אינה מתכנסת, אפילו במובן הרחב.

כבר ראינו שהסדרה $a_n = (-1)^n$ אינה מתכנסת. כעת נוכל להוכיח עובדה זו ביתר קלות: כל שיש לשים לב הוא שלסדרה זו יש שני גבולות חלקיים שונים, 1 ו־ -1, ולכן אינה מתכנסת.

מסתבר שגם הכיוון ההפוך של הטענה האחרונה נכון: אם סדרה אינה מתכנסת, יש לה לפחות שני גבולות חלקיים שונים במובן הרחב. נוכיח זאת בהמשך (טענה 5.7.10 להלן).

הטענה הבאה מאפיינת גבולות חלקיים בלי להשתמש בתת־סדרות:

טענה 5.7.6 מספר ממשי a הוא גבול חלקי של הסדרה $(a_n)_{n=1}^{\infty}$ אם ורק אם כל סביבה של a מכילה אינסוף מאיברי הסדרה.

הערה כשאומרים שאינסוף מאיברי הסדרה (a_n) שייכים לסביבה U הכוונה שיש אינסוף n־ים כך ש־ $a_n \in U$. בהחלט ייתכן שהערכים של ה־a_n־ים האלה שווים. למשל אם (a_n) היא הסדרה הקבועה 0 ו־U סביבה של 0 אז אינסוף מאיברי הסדרה שייכים ל־U, אף שכל איברי הסדרה הם בעלי אותו ערך 0.

הוכחה נניח כי a הוא גבול חלקי ותהי U סביבה של a. מהגדרת הגבול החלקי, קיימת תת־סדרה (a_{n_k}) של (a_n) כך ש־ $a_{n_k} \to a$. מהגדרת הגבול נובע ש־ $a_{n_k} \in U$ החל ממקום מסוים, ולכן אינסוף מאיברי הסדרה שייכים ל־U.

בכיוון השני, נניח שכל סביבה של a מכילה אינסוף מאיברי הסדרה ונבנה ברקורסיה תת־סדרה המתכנסת ל־a. כיוון שכל סביבה של a מכילה אינסוף מאיברי הסדרה, בפרט יש איברים מהסדרה בסביבה $B_1(a)$ ולכן נוכל לבחור אינדקס n_1 עבורו $a_{n_1} \in B_1(a)$. נניח שבחרנו אינדקסים $n_1 < n_2 < n_3 < \cdots < n_k$ כך שלכל $1 \le j \le k$ מתקיים $a_{n_j} \in B_{1/j}(a)$, ונתבונן בסביבה $B_{1/(k+1)}(a)$. מכיוון שיש אינסוף איברים של הסדרה (a_n) בסביבה זו, בפרט קיים $n \in \mathbb{N}$ המקיים $n > n_k$ וגם $a_n \in B_{1/(k+1)}(a)$. נגדיר אז $n_{k+1} = n$.

הסדרה (n_k) עולה ממש (זה מידי מהגדרתה) ולכן $(a_{n_k})_{k=1}^{\infty}$ היא תת־סדרה של (a_n). נראה שהיא מתכנסת ל־a. ואמנם, הסדרה נבחרה כך ש־ $a_{n_k} \in B_{1/k}(a)$, ולכן $0 \le |a_{n_k} - a| < \frac{1}{k}$. מכלל הסנדוויץ' אנו מסיקים ש־ $\lim_{k\to\infty} |a_{n_k} - a| = 0$, דהיינו $a_{n_k} \to a$; כנדרש. ∎

שימו לב שאם (a_n) סדרה ובכל סביבה של a יש איברים מהסדרה אין זה מבטיח ש־a הוא גבול חלקי שלה (בדקו היכן בהוכחה השתמשנו בהנחה שיש בכל סביבה אינסוף איברים!). למשל אם $a_n = \frac{1}{n}$ אז בכל סביבה של 1 יש איברים של הסדרה (האיבר a_1) אבל $a_n \to 0$ וממילא 1 אינו גבול חלקי שלה. מצד שני אם בכל סביבה של a יש איברים של הסדרה **השונים מ־**a קל לראות שבכל סביבה יש אינסוף איברים, ואז לפי הטענה a גבול חלקי של (a_n) (הוכיחו זאת!).

הטענה הבאה משלימה את קודמתה ומאפיינת מתי $\pm\infty$ הוא גבול חלקי של סדרה:

טענה 5.7.7 ∞ הוא גבול חלקי של סדרה אמ"מ הסדרה אינה חסומה מלעיל. כמו כן $-\infty$ הוא גבול חלקי של סדרה אם ורק אם הסדרה אינה חסומה מלרע.[11]

[11]כדי לראות את הקשר עם טענה 5.7.6, היזכרו בהערה בעמוד 117, שם הגדרנו סביבה של אינסוף

ההוכחה דומה להוכחה הקודמת, ומושארת כתרגיל.

המשפט הבא מבטיח קיום של תת־סדרה מתכנסת, והוא ישחק תפקיד מרכזי בהמשך.

משפט 5.7.8 (משפט בולצאנו[12]-וויירשטראס[13]) אם $(a_k)_{k=1}^{\infty}$ סדרה חסומה אז יש לה תת־סדרה המתכנסת לגבול סופי.

הוכחה הרעיון פשוט. נניח שאיברי הסדרה שייכים כולם לקטע $I=[-M,M]$. אם נחצה את הקטע לשני תת־קטעים שווים, שכל אחד באורך מחצית אורכו של הקטע המקורי, אז אחד החצאים יכיל בהכרח אינסוף מאיברי הסדרה. את אותו חצי קטע נחצה שוב לשניים. אחד משני הקטעים שיתקבלו שוב מכיל אינסוף מאיברי הסדרה. כך נמשיך ונקבל סדרה יורדת של קטעים שאורכיהם שואפים לאפס. מהלמה של קנטור נסיק שיש מספר (יחיד) x בחיתוך הקטעים, ונוכיח ש־x גבול חלקי של הסדרה.

נעבור לפרטי ההוכחה. יהי M חסם של הסדרה $(a_k)_{k=1}^{\infty}$. נבנה ברקורסיה סדרת קטעים סגורים $I_1,I_2,...,I_n$ בעלי התכונות הבאות:

1. האורך של I_k הוא $M\cdot 2^{-k+2}$ לכל $1\leq k\leq n$.
2. $I_k\subseteq I_{k-1}$ לכל $1<k\leq n$.
3. כל אחד מהקטעים $I_1,\ldots,I_n$ מכיל אינסוף מאיברי הסדרה $(a_k)_{k=1}^{\infty}$.

עבור $n=1$ נגדיר $I_1=[-M,M]$. אז כל איברי הסדרה (ובפרט אינסוף מאיברי הסדרה) שייכים לקטע I_1 וקל לבדוק שהתכונות האחרות מתקיימות. כדי להגדיר את I_{n+1} נניח שנתונים הקטעים $I_1,\ldots,I_n$ והם מקיימים את התכונות האמורות. נסמן $I_n=[a,b]$ ונחלק את I_n לשני חלקים שווי־אורך,

$$I_n'=[a,\frac{a+b}{2}]\ ,\ I_n''=[\frac{a+b}{2},b]$$

כיוון שאינסוף מאיברי הסדרה נמצאים בקטע I_n ומכיוון ש־ $I_n=I_n''\cup I_n''$ נובע שלפחות אחד מבין הקטעים I_n'',I_n'' מכיל אינסוף מאיברי הסדרה. נסמן קטע זה ב־I_{n+1} (אם בשניהם יש אינסוף מאברי הסדרה נבחר שרירותית אחד מהם). אז מתקיים $I_{n+1}\subseteq I_n$, וב־I_{n+1} יש אינסוף מאיברי הסדרה. כמו־כן האורך של הקטעים I_n',I_n'' הוא מחצית האורך של I_n ולכן

$$|I_{n+1}|=\frac{1}{2}|I_n|=\frac{1}{2}\cdot M2^{-n+2}=M\cdot 2^{-(n+1)+2}$$

להיות קרן מהצורה (a,∞). שימו לב שסדרה (a_n) אינה חסומה מלעיל אמ"מ יש אינסוף מאיברי הסדרה בכל סביבה של אינסוף. מכאן שטענה 5.7.7 אומרת ש־∞ הוא גבול חלקי של (a_n) אמ"מ בכל סביבה של אינסוף יש אינסוף מאיברי הסדרה.

[12] (Bernard Bolzano, 1781-1848).

[13] (Karl Weierstrass, 1815-1897).

וכל התנאים מתקיימים.

אם כן הגדרנו ברקורסיה סדרת קטעים סגורים $(I_n)_{n=1}^{\infty}$ כאמור. מהתכונות $I_n \supseteq I_{n+1}$ ומכיוון ש־ $|I_n| = M2^{-n+2} \to 0$, מהלמה של קנטור יש נקודה x בחיתוך שלהן. נוכיח ש־x היא גבול חלקי של הסדרה $(a_k)_{k=1}^{\infty}$. לשם כך כל שעלינו לעשות הוא להראות שבכל סביבה של x יש אינסוף מאיברי הסדרה. תהי אם כן $B_r(x)$ סביבה של x עם רדיוס $r > 0$. כיוון שסדרת אורכי הקטעים שהגדרנו לעיל שואפים ל־0 נובע שקיים $n \in \mathbb{N}$ עבורו $|I_n| < r$. כיוון ש־ $x \in I_n$ וכיוון שאורכו של I_n קטן מ־r נובע ש־ $I_n \subseteq B_r(x)$ (למה?). מכיוון שב־I_n יש אינסוף מאיברי הסדרה גם בסביבה $B_r(x)$ יש אינסוף מאיברי הסדרה, כפי שרצינו. ■

גם לסדרות שאינן חסומות תקף משפט זה בגרסה הבאה:

מסקנה 5.7.9 לכל סדרה יש תת־סדרה מתכנסת במובן הרחב.

הוכחה תהי (a_n) סדרה. אם היא חסומה יש לה גבול חלקי ממשי לפי משפט בולצאנו־ויירשטראס. אחרת או שאינה חסומה מלעיל או שאינה חסומה מלרע, ואז (לפי טענה 5.7.7) יש לה גבול חלקי ∞ או $-\infty$ בהתאמה. ■

כעת יש בידינו כלים המאפשרים לאפיין התכנסות של סדרה באמצעות הגבולות החלקיים שלה ותכונות של התת־סדרות שלה.

למה 5.7.10 תהי (a_n) סדרה שאינה מתכנסת במובן הרחב. אז יש לה לפחות שני גבולות חלקיים שונים במובן הרחב.

הוכחה יהי a גבול חלקי של (a_n). נניח תחילה ש־a ממשי. מההנחה $a_n \not\to a$ יש $\varepsilon > 0$ ואינסוף אינדקסים n כך ש־ $a_n \notin B_\varepsilon(a)$. בפרט ניתן לבנות תת־סדרה $(a_{n_k})_{k=1}^{\infty}$ של (a_n) כך ש־ $a_{n_k} \notin B_\varepsilon(a)$. יהי b גבול חלקי במובן הרחב של הסדרה (a_{n_k}), ולכן גם של (a_n). אז $b \neq a$ כי אף איבר של (a_{n_k}) אינו שייך ל־ $B_\varepsilon(a)$.

אם $a = \infty$ ההוכחה דומה: מכיוון ש־ $a_n \not\to \infty$ יש M כך שלאינסוף n־ים מתקיים $a_n \leq M$. נבנה תת־סדרה (a_{n_k}) כך שלכל k מתקיים $a_{n_k} \leq M$. אם b גבול חלקי (במובן הרחב) של הסדרה (a_{n_k}) אז $b = -\infty$ או b ממשי ומקיים $b \leq M$ ובכל מקרה $b \neq a$.

המקרה $a = -\infty$ דומה למקרה $a = \infty$, ומושאר כתרגיל. ■

כפי שראינו ישנן סדרות עם בדיוק שני גבולות חלקיים (למשל הסדרה $(-1)^n$), ולכן לא ניתן לחזק את הטענה האחרונה.

מסקנה 5.7.11 תהי (a_n) סדרה ויהי $a \in \mathbb{R}$ או $a = \pm\infty$. אז $a_n \to a$ אמ"מ a הוא הגבול החלקי היחיד של (a_n).

ההוכחה מידית מהלמה הקודמת ומטענה 5.7.4.

נסיים את הסעיף בהכללה שימושית. כזכור, הגדרנו תת־סדרה של סדרה (a_n) להיות סדרה מהצורה $(a_{n_k})_{k=1}^{\infty}$ כאשר n_k סדרה עולה ממש של אינדקסים. אפשר לתת הגדרה כללית יותר שבה מוותרים על דרישת המונוטוניות ודורשים רק ש־ $n_k \to \infty$. הסדרה (a_{n_k}) המתקבלת נקראת **תת־סדרה מוכללת** של (a_n). ברור שכל תת־סדרה של (a_n) היא גם תת־סדרה מוכללת, אבל ההפך אינו נכון: למשל הסדרה $1,1,2,2,3,3,4,4,5,5,\dots$ היא תת־סדרה מוכללת של $1,2,3,4,5,\dots$, אבל אינה תת־סדרה שלה.

כל המשפטים שהוכחנו על תת־סדרות נכונים לתת־סדרות מוכללות, למעט הסעיף הראשון בלמה 5.7.2, אך אם תבדקו היכן השתמשנו בסעיף זה תראו שאפשר למצוא לו תחליף. נזכיר בפרט את העובדה ש־ $a_n \to a$ אמ"מ $a_{n_k} \to a$ לכל תת־סדרה מוכללת של (a_n). תוצאה זו תהיה שימושית בהמשך.

תרגילים

1. לכל זוג סדרות מהרשימה הבאה, קבעו האם אפשר לכתוב אחת כתת־סדרה של השנייה, ואם כן הראו כיצד. למשל, (b_n) היא תת־סדרה של (a_n) ונתונה על ידי $b_n = a_{n^2}$.

   (א) $a_n = n$.

   (ב) $b_n = n^2$.

   (ג) $c_n = [\sqrt{n}]$.

   (ד) $d_n = 2^n$.

   (ה) $e_n = 2^{2n}$.

2. מצאו את כל הגבולות החלקיים של הסדרות הבאות ($[x]$ מציין את הערך השלם של x):

   (א) $a_n = \frac{1}{n}$ אם n זוגי ו־ $a_n = 1 - \frac{1}{n}$ אם n אי־זוגי.

   (ב) $[n]$.

   (ג) $[(\frac{n}{3})^2]$.

   (ד) $2^{(-1)^n/n}$.

   (ה) $\frac{1}{1+n-[n/5]}$.

   (ו) $n \cdot (1 + (-1)^n)$.

   (ז) $\frac{n-[\sqrt{n}]^2}{[\sqrt{n}]+1}$

   (ח) (a_n) מוגדרת ברקורסיה על ידי $a_n = 0$, ו־ $a_n = \frac{1}{2}(a_n + (-1)^{[\sqrt{n}]})$.

3. נניח $a_n \to a$ ותהי (n_k) סדרה של מספרים טבעיים חיוביים המקיימת $n_k \to \infty$. הוכיחו ש־ $\lim_{k\to\infty} a_{n_k} = a$ (שימו לב ש־(a_{n_k}) אינה תת־סדרה אלא תת־סדרה מוכללת כי (n_k) לא בהכרח עולה ממש).

4. תהי $(a_n)_{n=1}^{\infty}$ סדרה.

(א) נניח ש־ $a_{2n} \to a$ וגם $a_{2n+1} \to a$. הוכיחו ש־ $a_n \to a$.

(ב) באופן כללי יותר, יהיו $(m_k)_{k=2}^{\infty}$, $(n_k)_{k=1}^{\infty}$ סדרות עולות ממש של אינדקסים שכוללות יחד את $1,2,3\ldots$, דהיינו $\mathbb{N} = \{n_k, m_k \,:\, k \in \mathbb{N}\}$. הראו שאם $a_{n_k} \to a$ וגם $a_{m_k} \to a$ אז $a_n \to a$.

5. לכל $n \in \mathbb{N}$ בנו סדרה עם בדיוק n גבולות חלקיים.

6. תהי (a_n) סדרה ותהי (b_n) סדרה המסכימה עם (a_n) החל ממקום מסוים. האם כל גבול חלקי של (b_n) הוא גבול חלקי של (a_n)?

7. הוכיחו את טענה 5.7.7.

8. יהיו $(a_n),(b_n)$ סדרות ויהיו A,B קבוצות הגבולות החלקיים שלהן במובן הצר, בהתאמה. תהי $c_n = a_n + b_n$, ותהי C קבוצת הגבולות החלקיים של (c_n). הוכיחו או הפריכו את היחסים הבאים:

(א) $C \subseteq A$.

(ב) $C \subseteq A \cup B$.

(ג) $C \subseteq A \cap B$.

(ד) $C = \{a+b \,:\, a \in A\,,\, b \in B\}$.

9. בשאלה זו ניתן הוכחה נוספת ש־ $\sqrt[n]{n} \to 1$. נסמן $a_n = \sqrt[n]{n}$.

(א) התבוננו בתת־סדרה (a_{2^k}) וחשבו את הגבול שלה.

(ב) היעזרו בעובדה שלכל n קיים k כך ש־ $2^k \leq n < 2^{k+1}$ על מנת להוכיח ש־ $a_n \to 1$.

(ג) באופן כללי יותר, הראו שלכל $0 < m \in \mathbb{N}$ מתקיים $n^{(1/n^{1/m})} \to 1$.

10. בשאלה זו נוכיח בדרך אחרת את משפט בולצאנו־ויירשטראס, ללא שימוש בלמה של קנטור. תהי (a_n) סדרה חסומה. מספיק להוכיח של־(a_n) יש תת־סדרה מונוטונית (למה?). תהי $P(n)$ הטענה ש־a_n הוא האיבר המקסימלי של קבוצת האיברים $\{a_i \,:\, i \geq n\}$. הוכיחו שאם $P(n)$ מתקיימת לאינסוף n־ים אז (a_n) מכילה תת־סדרה יורדת, ואחרת היא מכילה תת־סדרה עולה. הסיקו את משפט בולצאנו־ויירשטראס.

11. השלימו את המקרה $-\infty$ בהוכחה של למה 5.7.10.

12. תהי (a_n) סדרה חסומה, ונניח שעבור כל תת־סדרה (a_{n_k}) של (a_n), קבוצת הגבולות החלקים של (a_{n_k}) שווה לקבוצת הגבולות החלקיים של (a_n). האם נובע מכך ש־(a_n) מתכנסת?

13. תהי (a_n) סדרה ולכל k יהי x_k גבול חלקי ממשי של (a_n). נניח $x_k \to x$. הוכיחו ש־x גבול חלקי של (a_n). הוכיחו זאת גם במקרה ש־ $x = \pm\infty$.

14. **הצפיפות** (relative density) של קבוצה $A \subseteq \mathbb{N}$ מוגדרת על ידי

$$d(A) = \lim_{n\to\infty} \frac{|A \cap \{1,2,\ldots,n\}|}{n}$$

אם הגבול קיים., ואחרת נאמר של־A אין צפיפות.

(א) הראו שלקבוצה $\{2,4,6\dots\}$ יש צפיפות $\frac{1}{2}$, שלקבוצה $\{1^2,2^2,3^2,\dots\}$ יש צפיפות 0, שלקבוצה $\{n\in\mathbb{N}\,:\,\forall k\in\mathbb{N}\, n\neq 2^k\}$ יש צפיפות 1, ושלקבוצה

$$\{n\in\mathbb{N}\,:\,\exists k\in\mathbb{N}\, 2^{2k}\leq n<2^{2k+1}\}$$

אין צפיפות.

(ב) הוכיחו שאם A,B קבוצות עם צפיפות 1 אז ל־ $A\cup B$ ול־ $A\cap B$ גם־כן יש צפיפות 1, ואם לשתיהן צפיפות אפס אז ל־$A\cup B$ ול־$A\cap B$ יש צפיפות אפס.

(ג) הוכיחו שאם A,B קבוצות זרות עם צפיפות a,b בהתאמה אז ל־ $A\cup B$ יש צפיפות $a+b$.

15. כשהגדרנו את מושג הגבול בחרנו לתרגם את התכונה "לכל סביבה של a רוב איברי (a_n) נמצאים בסביבה" לשפה מדויקת על ידי שפרשנו "רוב" בתור "החל ממקום מסוים" (עמוד 91). נגדיר כעת מושג דומה: נאמר שסדרה (a_n) **מתכנסת ל־a בצפיפות**, ונסמן $\operatorname{dlim} a_n=a$, אם לכל $\varepsilon>0$ יש קבוצה $A\subseteq\mathbb{N}$ עם צפיפות 1 כך שלכל $n\in A$ מתקיים $a_n\in B_\varepsilon(a)$ (מושג הצפיפות של קבוצה הוגדר בשאלה הקודמת).

(א) הוכיחו שאם $a_n\to a$ אז $\operatorname{dlim} a_n=a$, אבל לא להפך (מסקנה: מושגי ההתכנסות האלה באמת שונים).

(ב) בדקו אילו מהמשפטים על סדרות, כמו למשל משפטי האריתמטיקה, תקפים בגרסאות לגבולות בצפיפות (כדאי לשים לב באילו תכונות של המושג "החל ממקום מסוים" השתמשנו בהוכחות השונות, ולבדוק שאם במקומם מופיע הביטוי "לקבוצת אינדקסים שצפיפותה אחד" ההוכחה עדיין עובדת).

(ג) תהי $(a_n)_{n=1}^\infty$ סדרה חסומה. נסמן $s_n=\frac{1}{n}(a_1+\ldots+a_n)$. הוכיחו שאם $\operatorname{dlim} a_n=a$ אז $s_n\to a$. האם ההפך נכון?

5.8 גבולות עליונים וגבולות תחתונים

התבוננו בסדרה ההרמונית $1,\frac{1}{2},\frac{1}{3},\frac{1}{4},\dots$. החסם העליון של הסדרה הוא 1, אך הסיבה שאין חסם מלעיל קטן יותר נעוצה רק באיברים הראשונים בסדרה. אם נמחק את האיבר הראשון נשאר עם הסדרה $\frac{1}{2},\frac{1}{3},\frac{1}{4},\dots$, והמספר $\frac{1}{2}$ הוא חסם מלעיל שלה. למעשה נוכל לקבל חסם מלעיל חיובי קטן כרצוננו אם רק נסכים למחוק מספר גדול מספיק (אך סופי) של איברים. מאידך אף מספר שלילי אינו חסם מלעיל של סדרה המתקבלת באופן זה. באופן לא פורמלי, נוכל לומר ש־0 הוא מעין "חסם עליון אסימפטוטי" של הסדרה.

נעבור לדיון פורמלי יותר. החסם העליון של קבוצה A הוא החסם מלעיל הקטן ביותר שלה, דהיינו

$$\sup A = \inf\{M \;:\; a \le M \text{ לכל } a \in A\}$$

(למעשה אפשר לרשום min במקום inf). השוו להגדרה זו את ההגדרה הבאה:

הגדרה 5.8.1 תהי (a_n) סדרה חסומה. **הגבול העליון** (upper limit) של (a_n) הוא המספר המסומן $\limsup_{n\to\infty} a_n$[14] ומוגדר על ידי

$$\limsup_{n\to\infty} a_n = \inf\{M \;:\; a_n \le M \text{ החל ממקום מסוים}\}$$

באופן דומה, **הגבול התחתון** (lower limit) של (a_n) מסומן $\liminf_{n\to\infty} a_n$ ומוגדר על ידי

$$\liminf_{n\to\infty} a_n = \sup\{L \;:\; a_n \ge L \text{ החל ממקום מסוים}\}$$

כדי לא לסרבל את הכתיבה, אם אין חשש לאי הבנה לעיתים קרובות נכתוב $\limsup a_n$ במקום $\limsup_{n\to\infty} a_n$, וכך גם לגבול התחתון.

טענה 5.8.2 תהי (a_n) סדרה חסומה. אז הגבול העליון והתחתון של (a_n) מוגדרים היטב, ו־

$$\liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n$$

הוכחה תהי (a_n) סדרה חסומה ותהי A קבוצת כל המספרים M כך ש־ $a_n \le M$ החל ממקום מסוים. תהי B קבוצת מספרים L כך ש־ $a_n \ge L$ החל ממקום מסוים. הקבוצה A מכילה כל חסם מלעיל של הסדרה ולכן לפי הנתון אינה ריקה. באותו אופן B לא ריקה כי היא מכילה את החסמים מלרע של הסדרה.

אם $M \in A$ ו־ $L \in B$ אז יש n כך ש־ $L \le a_n \le M$ (למה יש n כזה?) ולכן בפרט $L \le M$. מכאן שכל איבר של A הוא חסם מלעיל של B, וכל איבר של B הוא חסם מלרע של A. לכן ממשפט 4.4.5, $\inf A$ ו־$\sup B$ קיימים ומקיימים

$$\liminf_{n\to\infty} b_n = \sup B \le \inf A = \limsup_{n\to\infty} A_n$$

כנדרש. ■

[14] לפעמים מסמנים את הגבול העליון ב־$\overline{\lim}_{n\to\infty} a_n$ ואת הגבול התחתון ב־$\underline{\lim}_{n\to\infty} a_n$.

דוגמאות

1. אם $a_n = \frac{1}{n}$ ו־ $M > 0$ אז $a_n \leq M$ לכל $n > \frac{1}{M}$. מצד שני אם $M \leq 0$ אז כל איברי הסדרה גדולים ממש מ־M ולכן אין זה נכון ש־ $a_n \leq M$ החל ממקום מסוים. לכן $\limsup a_n = \inf\{M \,:\, M > 0\} = 0$. שימו לב שהחסם התחתון אינו מתקבל, כלומר אי אפשר לרשום במקומו מינימום.

   קל גם לבדוק ש־0 הוא גם הגבול התחתון של (a_n): הרי 0 הוא חסם מלרע של הסדרה, ואילו לכל $L > 0$ אינסוף מאיברי הסדרה מקיימים $a_n < L$ (כי הסדרה שואפת לאפס). מכאן ש־ $\liminf a_n = \sup\{L \,:\, L \leq 0\} = 0$.

2. תהי $a_n = (-1)^n$. קל לבדוק ש־1 הוא חסם מלעיל של הסדרה ואילו לכל $M < 1$ יש אינסוף איברים בסדרה שגדולים מ־M ולכן אין זה נכון ש־M גדול מאיברי הסדרה החל ממקום מסוים. לכן $\limsup a_n = 1$. נימוק דומה מראה $\liminf a_n = -1$.

הדוגמה האחרונה מראה שהגבולות העליונים והתחתונים אינם מושג שקול לגבול. יש בהם יתרון מסוים כיוון שהם קיימים גם במקרים שהגבול אינו קיים, כמו למשל בדוגמה האחרונה.

ישנו קשר הדוק בין גבולות עליונים ותחתונים לגבולות חלקיים.

למה 5.8.3 תהי (a_n) סדרה חסומה ו־ $s = \limsup a_n$. אם $t > s$ אז $a_n \leq t$ לכל n גדול מספיק.

הוכחה יהי $t > s$. מהגדרת ה־$\limsup$, קיים $s \leq M < t$ כך ש־ $a_n \leq M$ לכל n גדול מספיק. כיוון ש־ $M < t$ יוצא ש־ $a_n < t$ לכל n גדול מספיק. ■

משפט 5.8.4 תהי (a_n) סדרה חסומה ויהי $s = \limsup a_n$ הגבול העליון שלה. אז s הוא הגבול החלקי המקסימלי של (a_n). באופן דומה, הגבול התחתון הוא הגבול החלקי המינימלי.

הערה המשפט מוכיח בפרט שלסדרה חסומה קיים גבול חלקי מקסימלי ומינימלי, דבר שלא ידענו קודם (אפשר להסיק זאת גם מתרגיל (13) בעמוד 136).

הוכחה נוכיח רק את הטענה על הגבול העליון. הוכחת החלק השני מושארת כתרגיל. תחילה נראה שאין גבול חלקי גדול מ־s. ואמנם, יהי a גבול חלקי של הסדרה. לפי הלמה לכל $t > s$ מתקיים $a_n < t$ החל ממקום מסוים. מכאן ש־ $a \leq t$. מאחר שזה נכון לכל $t > s$ הרי $a \leq s$.

נותר להראות ש־s הוא גבול חלקי של הסדרה. לשם כך די להראות שבכל סביבה של s יש אינסוף איברים מהסדרה (טענה 5.7.6). יהי $\varepsilon > 0$, ונראה שמתקיים $s - \varepsilon \leq a_n \leq s + \varepsilon$ לאינסוף n־ים. ואמנם, מכיוון ש־ $s + \varepsilon > s$ הרי שמשלמה הקודמת, $a_n < s + \varepsilon$ החל ממקום מסוים. מצד שני, אילולא היו אינסוף n־ים כך

ש־$a_n > s - \varepsilon$, היה מתקיים $a_n \leq s - \varepsilon$ החל ממקום מסוים, וכיוון ש־s הגבול העליון של הסדרה היה יוצא ש־ $s \leq s - \varepsilon$, סתירה. נובע משתי המסקנות האלה שלאינסוף n־ים מתקיים $s - \varepsilon < a_n \leq s + \varepsilon$, כפי שרצינו. ■

מסקנה 5.8.5 סדרה חסומה (a_n) מתכנסת אמ"מ $\limsup a_n = \liminf a_n$, ובמקרה זה $\limsup a_n = \liminf a_n = \lim a_n$.

הוכחה אם (a_n) מתכנסת למספר a אז כל גבול חלקי שלה שווה ל־a, ולכן לפי המשפט מתקיים $\limsup a_n = \liminf a_n = a$.

להפך, נניח $\limsup a_n = \liminf a_n = a$. לפי המשפט, כל גבול חלקי של (a_n) נמצא בין $\limsup a_n$ ו־ $\liminf a_n$ ולכן כל הגבולות החלקיים שווים ל־a. מכאן וממסקנה 5.7.11 נובע שהסדרה מתכנסת והגבול שווה ל־a. ■

נפנה למשפטי תחשיב על גבולות עליונים ותחתונים. בכל הקשור לאי־שוויונות, אין הפתעות מיוחדות:

טענה 5.8.6 אם $(a_n), (b_n)$ סדרות חסומות המקיימות $a_n \leq b_n$ החל ממקום מסוים, אז

$$\begin{aligned}\limsup_{n\to\infty} a_n &\leq \limsup_{n\to\infty} b_n\\ \liminf_{n\to\infty} a_n &\leq \liminf_{n\to\infty} b_n\end{aligned}$$

הוכחה כיוון ש־ $a_n \leq b_n$ החל ממקום מסוים, כל מספר M שעבורו $b_n \leq M$ החל ממקום מסוים מקיים גם $a_n \leq M$ החל ממקום מסוים, כלומר

$$\{M\ :\ \text{החל ממקום מסוים } b_n \leq M\} \subseteq \{M\ :\ \text{החל ממקום מסוים } a_n \leq M\}$$

מכאן שהחסם התחתון של הקבוצה משמאל אינו קטן מהחסם התחתון של הקבוצה מימין, כלומר: $\limsup b_n \geq \limsup a_n$. הטענה בדבר הגבולות התחתונים מושארת כתרגיל. ■

לעומת זאת, ההתנהגות של גבולות עליונים ותחתונים תחת פעולות החשבון היא פחות טובה משל גבולות:

משפט 5.8.7 יהיו $(a_n), (b_n)$ סדרות חסומות.

1. מתקיים

$$\begin{aligned}\limsup_{n\to\infty}(-a_n) &= -\liminf_{n\to\infty} a_n\\ \liminf_{n\to\infty}(-a_n) &= -\limsup_{n\to\infty} a_n\end{aligned}$$

2. מתקיים

$$\begin{aligned}\limsup_{n\to\infty}(a_n+b_n) &\le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \\ \liminf_{n\to\infty}(a_n+b_b) &\ge \liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n\end{aligned}$$

3. אם בנוסף הסדרות אי־שליליות, אז

$$\begin{aligned}\limsup_{n\to\infty}(a_n\cdot b_n) &\le \limsup_{n\to\infty} a_n \cdot \limsup b_n \\ \liminf_{n\to\infty}(a_n\cdot b_n) &\ge \liminf_{n\to\infty} a_n \cdot \liminf b_n\end{aligned}$$

הוכחה נוכיח חלק מהסעיפים. הנותרים מושארים כתרגיל.

את (1) קל ביותר להוכיח בעזרת גבולות חלקיים: אם a גבול חלקי של (a_n) אז $-a$ גבול חלקי של $(-a_n)$, ומכאן ש־s הוא הגבול החלקי המקסימלי של (a_n) אמ"מ $-s$ הוא הגבול החלקי המינימלי של $(-a_n)$ (הבהירו לעצמכם שאתם מבינים מדוע!). מכאן ש־

$$-\limsup_{n\to\infty} a_n = \liminf_{n\to\infty}(-a_n)$$

השוויון השני מוכח בצורה דומה.

לגבי (2), קל לוודא שאם $a_n \le L$ החל ממקום מסוים וגם $b_n \le M$ החל ממקום מסוים אז $a_n+b_n \le L+M$ החל ממקום מסוים, ולכן אם נסמן

$$\begin{aligned}A &= \{L : \text{החל ממקום מסוים } a_n \le L\} \\ B &= \{M : \text{החל ממקום מסוים } b_n \le M\} \\ C &= \{R : \text{החל ממקום מסוים } a_n+b_n \le R\}\end{aligned}$$

אז

$$C \supseteq \{a+b : a\in A\,,\, b\in B\} = A+B$$

ולכן

$$\begin{aligned}\limsup_{n\to\infty}(a_n+b_n) &= \inf C \\ &\le \inf(A+B) \\ &= \inf A + \inf B \\ &= \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n\end{aligned}$$

כאשר השוויון $\inf(A+B)=\inf A+\inf B$ נובע ממשפט 4.4.4 (הצדיקו את יתר השוויונות!).

סעיף (3) מושאר כתרגיל. ■

באופן כללי אי אפשר לחזק את האי שוויונית במשפט ולקבל שוויונות, כפי שאפשר לראות על ידי התבוננות בסדרות $a_n=(-1)^n$ ו־ $b_n=(-1)^{n+1}$. כיוון ש־ $a_n+b_n=0$ לכל n הרי ש־ $\limsup a_n+b_n=0$, ואילו $\limsup a_n=\limsup b_n=1$. מכאן שהשוויון $\limsup(a_n+b_n)=\limsup a_n+\limsup b_n$ אינו נכון באופן כללי.

יש מקרה אחד חשוב שבו מובטח שוויון במשפטי התחשיב:

משפט 5.8.8 יהיו $(a_n),(b_n)$ סדרות חסומות ונניח $a_n\to a$. אז מתקיים

$$\begin{aligned}\limsup_{n\to\infty}(a_n+b_n) &= a+\limsup_{n\to\infty} b_n\\ \liminf_{n\to\infty}(a_n+b_n) &= a+\liminf_{n\to\infty} b_n\end{aligned}$$

אם בנוסף $a_n,b_n\geq 0$ החל ממקום מסוים אז

$$\begin{aligned}\limsup_{n\to\infty}(a_n\cdot b_n) &= a\cdot\limsup_{n\to\infty} b_n\\ \liminf_{n\to\infty}(a_n\cdot b_n) &= a\cdot\liminf_{n\to\infty} b_n\end{aligned}$$

הוכחה נוכיח למשל $\limsup(a_n\cdot b_n)=a\cdot\limsup b_n$ תחת ההנחה $a_n,b_n\geq 0$. אנו כבר יודעים מהמשפט הקודם שמתקיים אי־שוויון $\limsup(a_n\cdot b_n)\leq a\cdot\limsup b_n$ ולכן די להראות שאם b גבול חלקי של (b_n) אז ab הוא גבול חלקי של (a_nb_n), כי אז נבחר $b=\limsup b_n$. אבל זה נובע מיד מכללי האריתמטיקה של גבולות, כי אם $b_{n_k}\to b$, אז מכיוון ש־ $a_n\to a$ הרי גם $a_{n_k}\to a$, ולכן $a_{n_k}b_{n_k}\to ab$. ■

כדוגמה לשימוש בגבולות עליונים ותחתונים נראה שוב ש־ $s^n\to 0$ לכל $0<s<1$. די להראות ש־ $\liminf s^n=\limsup s^n=0$. מאחר ש־0 הוא חסם מלרע של s^n הרי $\liminf s^n\geq 0$. מאידך, נשים לב ש־ $s^n=s\cdot s^{n-1}$. לכן

$$\limsup_{n\to\infty} s^n=\limsup_{n\to\infty} s\cdot s^{n-1}=s\cdot\limsup_{n\to\infty} s^{n-1}=s\cdot\limsup_{n\to\infty} s^n$$

(השוויון האמצעי היא מסקנה של המשפט האחרון. למה השוויון האחרון נכון?). קיבלנו ש־ $\limsup s^n=s\cdot\limsup s^n$, אבל לפי ההנחה $0<s<1$, ולכן השוויון אפשרי רק אם $\limsup s^n=0$. קיבלנו אפוא $0\leq\liminf s^n\leq\limsup s^n\leq 0$ וזה מספיק.

שימו לב שאילו ידענו ש־s^n מתכנסת אז היינו יכולים בצורה דומה לחשב את הגבול a בעזרת אריתמטיקה של גבולות:

$$a=\lim_{n\to\infty} s^n=\lim_{n\to\infty} s\cdot s^{n-1}=s\cdot\lim_{n\to\infty} s^n=s\cdot a$$

ולכן $a=0$ (השוו הוכחה זו עם ההוכחה מדוגמה 1 בעמוד 123). הנקודה היא שכאן לא הנחנו שהסדרה מתכנסת, ובכל זאת הצלחנו להפעיל את השיקול הזה בזכות קיומם של הגבולות העליונים והתחתונים ותכונות אריתמטיות שלהם.

נסיים את הסעיף בהכללה הבאה:

הגדרה 5.8.9 אם (a_n) סדרה שאינה חסומה מלעיל נכתוב $\limsup a_n=\infty$, ואם היא אינה חסומה מלרע נכתוב $\liminf a_n=-\infty$.

הגדרה זו נותנת סימון חדש למושג האי־חסימות של סדרה אבל היא משתלבת היטב עם ההגדרה הקודמת של גבול עליון ותחתון. למשל, נשים לב ש־ $\limsup a_n=\infty$ אמ"מ ∞ הוא גבול חלקי של הסדרה, ולכן גם במקרה זה הגבול העליון שווה ל"מקסימום" של הגבולות החלקיים במובן הרחב.

תרגילים

1. חשבו את הגבולות העליונים והתחתונים של הסדרות בשאלה (2) בעמוד 135.
2. הוכיחו ישירות מהגדרת הגבול העליון והתחתון את הטענה שאם $a_n\to a$ אז $\limsup a_n=\liminf a_n=a$.
3. אילו מהתנאים הבאים שקולים ל־ $s=\limsup a_n$?
   (א) $s=\inf\{x\ :\ \text{החל ממקום מסוים}\ a_n<x\}$.
   (ב) $s=\inf\{x\ :\ \text{החל ממקום מסוים}\ a_n\le x\}$.
   (ג) $s=\sup\{x\ :\ \text{אינסוף פעמים}\ a_n>x\}$.
   (ד) $s=\sup\{x\ :\ \text{אינסוף פעמים}\ a_n\ge x\}$.
4. אילו מהתנאים הבאים שקולים לאי־שוויון $\limsup a_n\ge s$?
   (א) לכל $\varepsilon>0$ יש N כך ש־ $a_n>s-\varepsilon$ לכל $n>N$.
   (ב) לכל $\varepsilon>0$ ולכל N יש $n>N$ עם $a_n>s-\varepsilon$.
   (ג) לכל $\varepsilon>0$ ולכל N יש $n>N$ עם $s-\varepsilon<a_n<s$.
   (ד) לכל $\varepsilon>0$ יש רק מספר סופי של n־ים כך ש־ $a_n>s+\varepsilon$.
   (ה) לכל $\varepsilon>0$ יש רק מספר סופי של n־ים כך ש־ $a_n<s-\varepsilon$.
   (ו) הקבוצה $\{n\ :\ a_n\ge s\}$ אינה חסומה מלעיל.
5. תהי (a_n) סדרה חסומה ולכל k נסמן

$$A_k=\{a_n\ :\ n\ge k\}=\{a_k,a_{k+1},a_{k+2},\ldots\}$$

ברור ש־ $A_1\supseteq A_2\supseteq\ldots\supseteq A_k\supseteq\ldots$ ושכל הקבוצות האלה חסומות ולא ריקות. לכן אפשר להגדיר

$$s_k=\sup A_k\qquad,\qquad i_k=\inf A_k$$

(א) הראו ש־ (s_k) סדרה יורדת וש־(i_k) סדרה עולה, וששתיהן חסומות.

(ב) נגדיר $s = \lim s_k\,,\ i = \lim i_k$ (הגבולות קיימים לפי הסעיף הקודם). הראו ש־$s = \limsup a_n$ וש־ $i = \liminf a_n$.

6. הוכיחו את סעיף (2) ממשפט 5.8.7 בעזרת משפט 5.8.4.

7. נסחו והוכיחו כלל תחשיב המקשר בין הגבולות העליונים והתחתונים של סדרות (a_n) ו־ (ca_n), כאשר c קבוע.

8. הוכיחו שאם (a_n) סדרה חיובית וחסומה ואם $\limsup a_n = a$ אז לכל $r \in \mathbb{Q}$ מתקיים $\limsup \sqrt[r]{a_n} = \sqrt[r]{\limsup a_n}$.

9. יהיו $(a_n), (b_n), (c_n)$ סדרות חסומות המקיימות $a_n \le b_n \le c_n$ החל ממקום מסוים, ונניח ש־ $\liminf a_n = \limsup c_n$. הראו שאז כל שלוש הסדרות מתכנסות לאותו גבול (זו הכללה של משפט הסנדוויץ').

10. נקבע $a_1 \in [0,1]$ ונגדיר ברקורסיה $a_{n+1} = \sqrt{1-a_n^2}$. קל לוודא ש־(a_n) סדרה חסומה (בדקו!). נסמן $a_+ = \limsup a_n$ ו־ $a_- = \liminf a_n$. אז

$$a_+ = \limsup \sqrt{1-a_{n-1}^2} = \sqrt{1-(\limsup a_{n-1})^2} = \sqrt{1-(a_+)^2}$$

וחילוץ של a_+ מהמשוואה נותן $a_+ = \frac{1}{\sqrt{2}}$. אותו חישוב מראה ש־ $a_- = \frac{1}{\sqrt{2}}$, ולכן הסדרה כולה מתכנסת ל־$\sqrt{2}$. לעומת זאת בדיקה קלה מראה שאם קובעים $a_1 = 0$ אז $a_n = \frac{1}{2}(1+(-1)^n)$, וסדרה זו אינה מתכנסת. היכן כאן השגיאה?

11. תהי (a_n) סדרה חסומה, ונסמן $x = \liminf a_n\,,\ y = \limsup a_n$. נניח ש־ $a_{n+1} - a_n \to 0$. הוכיחו שכל מספר $c \in [x,y]$ הוא גבול חלקי של הסדרה.

12. בנו סדרה כך ש־ $a_{n+1} - a_n \to 0$ ו־ $\liminf a_n < \limsup a_n$.

13. הראו ש־ $\limsup a_n = \infty$ אמ"מ יש ל־(a_n) תת־סדרה ששואפת לאינסוף.

14. נאמר שסדרה (a_n) היא **בי־מונוטונית** אם לכל n מתקיימת אחת האפשרויות הבאות: (א) $a_k \le a_n$ לכל $k \ge n$, או (ב) $a_k \ge a_n$ לכל $k \ge n$.

(א) הראו שלסדרה בי־מונוטונית וחסומה יש לכל היותר שני גבולות חלקיים שונים.

(ב) הוכיחו ש־(a_n) היא סדרה בי־מונוטונית חסומה אמ"מ יש מספר c כך שסדרת האיברים שקטנים או שווים ל־c עולה וסדרת האיברים שגדולים או שווים ל־c יורדת. הסיקו שסדרה בי־מונוטונית וחסומה (a_n) עם יותר מגבול חלקי אחד מתפרקת לשתי תת־סדרות $(a_{i_k}), (a_{j_k})$, כך ש־(a_{i_k}) יורדת ומקיימת $a_{i_k} \to \limsup a_n$, ואילו (a_{j_k}) עולה ומקיימת $a_{j_k} \to \liminf a_n$.

15. תהי (a_n) סדרה ויהי (x, y) קטע. נאמר שהסדרה חוצה את (x, y) אינסוף פעמים אם יש אינדקסים $\ldots < i_1 < j_1 < i_2 < j_2 < \ldots < i_k < j_k$ כך ש־$a_{i_m} < x$ ו־$a_{j_m} > y$ לכל m, ואחרת נאמר שהיא חוצה את הקטע רק מספר סופי של פעמים. הוכיחו שאם (a_n) חוצה כל קטע לא מנוון (x, y) רק מספר סופי של פעמים אז (a_n) מתכנסת במובן הרחב.

16. נסחו והוכיחו גרסה של משפט האריתמטיקה של גבולות עליונים ותחתונים כאשר אחת או שתיים מהסדרות המעורבות אינה חסומה.

5.9 תנאי קושי

כדי להכריע התכנסות של סדרה (a_n) יש לבדוק האם קיים מספר a בעל תכונה מסוימת. זו דוגמה למה שמכונה **תנאי חיצוני**, כי ההתכנסות אינה נקבעת לפי תכונה של הסדרה (a_n) עצמה אלא על ידי תכונה של האובייקט החיצוני a, שאינו חלק מהסדרה. היה רצוי למצוא **תנאי פנימי** שיאפשר לבדוק אם סדרה מתכנסת רק על סמך תכונות של איברי הסדרה, מבלי להזדקק לאובייקטים חיצוניים.

מטרת הסעיף הנוכחי היא לתת אפיון כזה. המוטיבציה מאחוריה היא האבחנה הבאה: אם (a_n) מתכנסת למספר a אז איברי הסדרה הולכים ומצטופפים לקראת הגבול, כלומר, לא רק שעבור n־ים גדולים האיברים קרובים לגבול a אלא שהם קרובים גם אחד לשני. קרבה ל־a היא תנאי חיצוני, אך הקרבה של איברי הסדרה זה לזה היא תנאי פנימי.

נכניס את ההגדרה הבאה:

הגדרה 5.9.1 תהי $P(n, m)$ טענה התלויה ב־m, n. נאמר ש־$P(n, m)$ מתקיימת לכל m, n מספיק גדולים (או שהיא מתקיימת החל ממקום מסוים) אם יש N כל שלכל $m, n > N$ טבעיים הטענה $P(m, n)$ מתקיימת.

כעת כדי להביע את הטענה שהאיברים של סדרה מתכנסת מצטופפים יותר ויותר, נוכל לומר שהחל ממקום מסוים הם קרובים זה לזה. ליתר דיוק,

הגדרה 5.9.2 (תנאי קושי[15]) תהי $(a_n)_{n=1}^{\infty}$ סדרה. נאמר שהיא **סדרת קושי** (Cauchy sequence) אם לכל $\varepsilon > 0$, לכל m, n גדולים מספיק מתקיים $|a_n - a_m| < \varepsilon$. לפעמים נאמר במקרה כזה שהסדרה מקיימת את **תנאי קושי**.

ביתר פירוט, (a_n) היא סדרת קושי אמ"מ לכל $\varepsilon > 0$ יש N כך שלכל $m, n > N$ טבעיים מתקיים $|a_n - a_m| < \varepsilon$. דרך שקולה לומר זאת הוא שלכל $\varepsilon > 0$ יש N כך שלכל $n > N$ ולכל k טבעי מתקיים $|a_n - a_{n+k}| < \varepsilon$.

[15] (Augustine Cauchy, 1789־1857).

כפי שציינו לפני ההגדרה, אינטואיטיבית ברור שאם סדרה מתכנסת אז איבריה מצטופפים יותר ויותר ולכך תנאי קושי מתקיים. ואמנם,

טענה 5.9.3 אם $(a_n)_{n=1}^{\infty}$ סדרה מתכנסת אז היא סדרת קושי.

הוכחה נניח כי $\lim_{n\to\infty} a_n = a$ ויהי $\varepsilon > 0$. אז החל ממקום מסוים איברי הסביבה נמצאים כולם בסביבה $B_{\varepsilon/2}(a)$, ולכן רחוקים אחד מהשני לכל היותר מרחק ε. ביתר פירוט, קיים N טבעי עבורו לכל $n > N$ מתקיים $|a_n - a| < \frac{\varepsilon}{2}$ ולכן לכל $n, m > N$ מתקיים

$$|a_n - a_m| = |a_n - a + a - a_m| \le |a_n - a| + |a - a_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

ותנאי קושי מתקיים. ■

הכיוון ההפוך של הטענה, הקובעת שכל סדרת קושי מתכנסת לגבול סופי, מעט יותר מורכב. כדי להוכיח שסדרה (a_n) מתכנסת, יש למצוא מועמד לגבול. בדיעבד, אילו ידענו ש־(a_n) מתכנסת אז הגבול שלה שווה לגבול של כל תת־סדרה מתכנסת. לכן האסטרטגיה שלנו תהיה להוכיח שיש ל־(a_n) תת־סדרה מתכנסת, ואז להראות שאותו גבול חלקי הוא למעשה גבול של הסדרה כולה. כדי להבטיח שיש גבול חלקי נוכיח תחילה שסדרת קושי היא חסומה, ואז נפעיל את משפט בולצאנו־ויירשטראס. לאחר מכן, משמצאנו תת־סדרה מתכנסת, נראה שתנאי קושי מבטיח שהסדרה המקורית מתכנסת לאותו גבול כמו התת־סדרה.

למה 5.9.4 סדרת קושי היא חסומה.

הוכחה תהי (a_n) סדרת קושי. מתנאי קושי, עבור $\varepsilon = 1$ קיים N טבעי כך שלכל $n, m > N$ מתקיים $|a_n - a_m| < 1$ ובפרט לכל $m > N$ מתקיים $|a_{N+1} - a_m| < 1$, כלומר

$$a_{N+1} - 1 \le a_m \le a_{N+1} + 1$$

לכן המקסימום

$$M = \max\{|a_{N+1} - 1|, |a_{N+1} + 1|\}$$

הוא חסם של הסדרה $(a_n)_{n=N+1}^{\infty}$. סדרה זו נבדלת מהסדרה המקורית רק במספר סופי של איברים בתחילת הסדרה, ולכן הסדרה המקורית חסומה. ■

משפט 5.9.5 (תנאי קושי) סדרה (a_n) מתכנסת אמ"מ היא סדרת קושי.

הוכחה ראינו כבר שסדרה מתכנסת היא סדרת קושי. נראה את הכיוון השני.

נניח ש־(a_n) סדרת קושי. מהלמה אנו יודעים ש־(a_n) חסומה. ממשפט בולצאנו־ויירשטראס נסיק שקיימת תת־סדרה (a_{n_k}) המתכנסת למספר a. עלינו להראות

שהסדרה (a_n) כולה מתכנסת ל־a. הרעיון הוא שלפי תנאי קושי, החל ממקום מסוים כל איברי (a_n) קרובים זה לזה ולכן כדי שכולם יהיו קרובים ל־a מספיק שאחד מהם יהיה קרוב ל־a. מצד שני, בין האיברים האלה מופיעים אינסוף האיברים מהתת־סדרה (a_{n_k}) (למעשה כולם, למעט אולי מספר סופי), ולכן יש ביניהם איברים קרובים ל־a.

באופן מדויק, נקבע $\varepsilon > 0$. עלינו להראות שיש N כך שלכל $m > N$ מתקיים $|a_m - a| < \varepsilon$. תנאי קושי מבטיח שקיים N טבעי עבורו לכל $k, m > N$ מתקיים $|a_k - a_m| < \varepsilon$. כמו־כן לכל i גדול מספיק מתקיים $|a_{n_i} - a| < \varepsilon$ (כי $a_{n_i} \to a$) ולכן קיים i כך ש־$n_i > N$ וגם $|a_{n_i} - a| < \varepsilon$. כעת בהינתן $m > N$ מתקיים $|a_m - a_{n_i}| < \varepsilon$, ולכן

$$|a_m - a| = |a_m - a_{n_i} + a_{n_i} - a| \leq |a_m - a_{n_i}| + |a_{n_i} - a| < \varepsilon + \varepsilon = 2\varepsilon$$

ראינו שבהינתן $\varepsilon > 0$ קיים N כך שלכל $m > N$ מתקיים $|a_m - a| < 2\varepsilon$, וזה כמובן מספיק על מנת להסיק ש־ $a_n \to a$. ■

תרגילים

1. תהי (a_n) סדרה. הוכיחו או הפריכו:

   (א) נניח שלכל N יש $\varepsilon > 0$ כך שלכל $m, n > N$ מתקיים $|a_n - a_m| < \varepsilon$. אז (a_n) מתכנסת.

   (ב) נניח שלכל $\varepsilon > 0$ יש N, M כך שלכל $n > N$ ולכל $m > M$ מתקיים $|a_n - a_{n+m}| < \varepsilon$. אז (a_n) מתכנסת.

   (ג) נניח שלכל $\varepsilon > 0$ ולכל k קיים N כך ש־ $|a_n - a_{n+k}| < \varepsilon$ לכל $n > N$. אז (a_n) מתכנסת.

   (ד) נניח ש־(a_n) חסומה מלעיל וחסומה מלרע על ידי מספר חיובי, ולכל $\varepsilon > 0$ קיים N כך שלכל $m, n > N$ מתקיים $1 - \varepsilon < |\frac{a_n}{a_m}| < 1 + \varepsilon$. אז (a_n) מתכנסת.

   (ה) נניח שלכל $M > 0$ יש N כך שלכל $n, m > N$ מתקיים $|a_n - a_m| > M$. אז (a_n) מתכנסת במובן הרחב.

2. תנו הוכחה חדשה, בעזרת תנאי קושי, לעובדה שסדרה מונוטונית וחסומה מתכנסת.

3. הראו ישירות מההגדרה שסכום ומכפלה של סדרות קושי היא סדרת קושי.

4. (*) מתחילת הפרק ועד כה הוכחנו סדרה של משפטי התכנסות חשובים. בעזרת אקסיומת החסם העליון הוכחנו שלכל סדרה מונוטונית וחסומה יש גבול. ממשפט זה הוכחנו את הלמה של קנטור. הלמה של קנטור ותכונת

הארכימדיות גררו את משפט בולצאנו־ויירשטראס (היכן השתמשנו בהוכחה בארכימדיות?). ובסעיף הנוכחי הראינו שממשפט בולצאנו־ויירשטראס נובע שכל סדרת קושי מתכנסת.

מתברר שאפשר לסגור את המעגל. הוכיחו שתחת ההנחה ש־$\mathbb{R}$ מקיימת את אקסיומות השדה הסדור ואת תכונת הארכימדיות, ותחת ההנחה שכל סדרת קושי מתכנסת, נובע שלכל קבוצה חסומה מלעיל ולא ריקה יש חסם עליון (רמז: אם $A \subseteq \mathbb{R}$ קבוצה לא ריקה וחסומה, קבעו $a_1 \in A$ ו־ b_1 חסם מלעיל של A והגדירו ברקורסיה סדרה עולה (a_n) וסדרה יורדת (b_n) כך ש־ $a_n \in A$ ו־b_n חסם מלעיל של A, ומתקיים $|a_n - b_n| < 2^{-n}|a_1 - b_1|$. הוכיחו שאלה סדרות קושי עם גבול משותף ושהגבול הוא החסם העליון של A).

5.10 חזקות עם מעריך ממשי

הגדרנו בפרק 4 חזקות בעלות מעריך רציונלי. בפרק זה נסיים את המלאכה ונגדיר חזקה עם מעריך ממשי כללי, כלומר נגדיר את a^x כאשר $a > 0$ ו־ $x \in \mathbb{R}$. הרעיון הוא להגדיר את a^x להיות הערך אליו מתקרב a^r כאשר r הוא מספר רציונלי קרוב ל־x. בשפה של גבולות רעיון זה ניתן לניסוח כדלקמן:

הגדרה 5.10.1 יהי $a > 0$ ויהי $x \in \mathbb{R}$. המספר a^x מוגדר על ידי

$$a^x = \lim_{n\to\infty} a^{r_n}$$

כאשר (r_n) היא כל סדרה של מספרים רציונליים כך ש־ $r_n \to x$, והביטוי a^{r_n} הוא החזקה ה־r_n של a לפי ההגדרה של חזקות רציונליות (הגדרה 4.6.11).

כמו בשלבי ההגדרה הקודמים של החזקה עלינו לוודא שההגדרה החדשה חד־משמעית ומתיישבת עם ההגדרות הקודמות. ראשית, נציין שלכל $x \in \mathbb{R}$ קיימות סדרות של מספרים רציונליים המתכנסות ל־x (דוגמה (5) בעמוד 93). ולכן ההגדרה אינה מתקיימת באופן ריק. נותר לוודא שהגבול $\lim_{n\to\infty} a^{r_n}$ קיים, ושאינו תלוי בבחירת הסדרה (r_n).

למה 5.10.2 תהי (t_n) סדרה של מספרים רציונליים השואפת ל־0 ויהי $a > 0$. אז $a^{t_n} \to 1$.

הוכחה נוכיח את המקרה $a > 1$. (ההוכחה במקרה $a < 1$ נובעת ממקרה זה בעזרת אריתמטיקה של גבולות). יהי $\varepsilon > 0$. מכיוון ש־ $a^{1/n} \to 1$ וגם $a^{-1/n} \to 1$ (דוגמה (5) בעמוד 101), קיים n כך ש־ $1 - \varepsilon < a^{-1/n} < a^{1/n} < 1 + \varepsilon$. מכיוון ש־ $t_k \to 0$, לכל k גדול מספיק מתקיים $-1/n < t_k < 1/n$ ולכן לכל k גדול מספיק מתקיים

$$1 - \varepsilon < a^{-1/n} < a^{t_k} < a^{1/n} < 1 + \varepsilon$$

וקיבלנו $a^{t_k} \to 1$ (היכן השתמשנו בהנחה $a > 1$?). ■

נוכיח כעת שאם הגבול בהגדרת החזקה קיים אז הוא אינו תלוי בסדרה (r_n):

למה 5.10.3 אם $(r_n), (s_n)$ סדרות של מספרים רציונליים השואפות לנקודה $x \in \mathbb{R}$, ואם $a > 0$, אז

$$\lim_{n\to\infty} a^{r_n} = \lim_{n\to\infty} a^{s_n}$$

בהנחה ששני הגבולות קיימים.

הוכחה נסמן $R = \lim a^{r_n}$, $S = \lim a^{s_n}$. נשים לב שקיום הגבולות האלה גורר בפרט ששניהם חיוביים ממש. כדי לראות זאת, נשים לב שקיים מספר חיובי M כך ש־ $-M \leq r_n, s_n \leq M$ לכל n, ולכן כל האיברים a^{r_n} , a^{s_n} חסומים על־ידי המספרים a^{-M}, a^M (מי מהם הגדול ומי הקטן תלוי ביחס בין a ו־1). לכן גם $\lim a^{r_n}$, $\lim a^{s_n}$ חסומים על־ידי a^{-M}, a^M, ובפרט $R, S > 0$. אם כן די שנראה ש־ $R/S = 1$. אבל מאריתמטיקה של גבולות מתקיים

$$\frac{R}{S} = \lim_{n\to\infty} \frac{a^{r_n}}{a^{s_n}} = \lim_{n\to\infty} a^{r_n - s_n}$$

ואנו יודעים מהנתון ש־ $r_n - s_n \to 0$, וממילא $r_n - s_n$ היא סדרה רציונלית. לכן מהלמה הקודמת נובע $a^{r_n - s_n} \to 1$, לכן $R/S = 1$, כנדרש. ■

מהלמה האחרונה נובע גם קיום הגבולות בהגדרת החזקה הממשית:

מסקנה 5.10.4 אם $a > 0$ ו־(r_n) סדרה מתכנסת של מספרים רציונליים, אז (a^{r_n}) גם־כן סדרה מתכנסת.

הוכחה נסמן $x = \lim r_n$. כמו בהוכחה הקודמת העובדה ש־(r_n) חסומה גורר ש־(a^{r_n}) חסומה. לכן כדי להוכיח ש־(a^{r_n}) מתכנסת למספר ממשי די להראות שכל שתי תת־סדרות מתכנסות של (a^{r_n}) מתכנסות לאותו גבול (זו מסקנה 5.7.11). אבל זה נובע מהלמה הקודמת, כי כל שתי תת־סדרות של (r_n) מתכנסות לאותו הגבול x. ■

אם כן, הגדרה 5.10.1 היא חד משמעית ומגדירה בצורה יחידה את המספר a^x. כעת עלינו לבדוק אם ההגדרה החדשה של a^x מתלכדת עם ההגדרה הישנה במקרה שבו x רציונלי. זה באמת המצב, וההוכחה קלה: אם $x \in \mathbb{Q}$ אז הסדרה הקבועה $r_n = x$ היא סדרה של מספרים רציונליים המתכנסת ל־x, ולכן a^x לפי ההגדרה החדשה שווה ל־ $\lim a^{r_n}$, כאשר a^{r_n} היא החזקה לפי ההגדרה הישנה. אבל $r_n = x$ לכל n ולכן ההגדרה הישנה והחדשה של a^x מתלכדות.

המטרה הבאה שלנו היא להראות שכללי החזקה מתקיימים בהגדרה החדשה. נעזר בתוצאה הטכנית הבאה:

למה 5.10.5 יהי $a>0$ ו־ $x\in\mathbb{R}$.

1. לכל $\varepsilon>0$ קיים מספר רציונלי r כך ש־ $|x-r|<\varepsilon$ וגם $|a^x-a^r|<\varepsilon$.
2. אם $x_n\to x$ אז $a^{x_n}\to a^x$.

הוכחה החלק הראשון נובע בקלות מההגדרה: יהי $\varepsilon>0$ ונבחר סדרה (r_n) של מספרים רציונליים שמקיימת $r_n\to x$. אז $a^{r_n}\to a^x$, ולכן לכל n מספיק גדול הבחירה $r=r_n$ תהיה טובה.

כדי להוכיח את החלק השני, לכל n נבחר r_n רציונלי המקיים $|r_n-x_n|<\frac{1}{n}$ וגם $|a^{x_n}-a^{r_n}|<\frac{1}{n}$ (ניתן לבחור r_n כזה לפי החלק הראשון). מכיוון ש־ $x_n\to x$ נובע מ־ $|r_n-x_n|<\frac{1}{n}$ שגם $r_n\to x$, ולכן לפי הגדרת החזקה, $a^{r_n}\to a^x$. מצד שני, התכונה $|a^{x_n}-a^{r_n}|<\frac{1}{n}$ מבטיחה ש־ a^{x_n} מתכנס לאותו גבול כמו a^{r_n}, כלומר: $a^{x_n}\to a^x$, כנדרש. ■

משפט 5.10.6 החזקות הממשיות מקיימות את כללי החזקה שבמשפט 4.6.1.

הוכחה נוכיח את הסעיפים לפי הסדר הנוח ביותר ולא לפי הסדר המקורי שלהם.

נראה שאם $x<y$ ו־ $a>1$ אז $a^x<a^y$. אפשר למצוא סדרות $(r_n),(s_n)$ של מספרים רציונליים כך ש־ $r_n\to x\,,\ s_n\to y$ ו־ $r_n\le s_n$ לכל n (הוכיחו שיש סדרות כאלה!). מהתכונות של חזקה עם מעריך רציונלי מקבלים $a^{r_n}\le a^{s_n}$ לכל n, ולכן

$$a^x=\lim_{n\to\infty}a^{r_n}\le\lim_{n\to\infty}a^{s_n}=a^y$$

כנדרש.

נראה שאם $a,b>0$ אז $(ab)^x=a^xb^x$. תהי (r_n) סדרה רציונלית כך ש־ $r_n\to x$. אז

$$(ab)^x=\lim_{n\to\infty}(ab)^{r_n}=$$

על סמך תכונות החזקה עם מעריך רציונלי מתקיים $(ab)^{r_n}=a^{r_n}\cdot b^{r_n}$, ולכן

$$=\lim_{n\to\infty}a^{r_n}b^{r_n}=$$

לפי אריתמטיקה של גבולות

$$=\lim_{n\to\infty}a^{r_n}\cdot\lim_{n\to\infty}b^{r_n}=a^x\cdot b^x$$

כנדרש.

נראה שלכל $a>0$ ו־ $x,y\in\mathbb{R}$ מתקיים $a^{x+y}=a^x\cdot a^y$. יהיו $(r_n),(s_n)$ סדרות של מספרים רציונליים כך ש־ $r_n\to x\,,\ s_n\to y$. לפי משפט האריתמטיקה של גבולות,

הסדרה $(t_n) = (r_n + s_n)$ היא סדרה של מספרים רציונליים המתכנסת ל־ $x+y$.
לכן

$$a^{x+y} = \lim_{n\to\infty} a^{t_n} = \lim_{n\to\infty} a^{r_n+s_n} =$$

על סמך תכונות החזקות הרציונליות ואריתמטיקה של גבולות

$$= \lim_{n\to\infty} a^{r_n} \cdot a^{s_n} = \lim_{n\to\infty} a^{r_n} \cdot \lim_{n\to\infty} a^{s_n} = a^x \cdot a^y$$

נראה כי לכל $a > 0$ ולכל $x, y \in \mathbb{R}$ מתקיים $(a^x)^y = a^{xy}$. ראשית נשים לב שאם y רציונלי הטענה מתקיימת. ואמנם, לפי כלל החזקה עם מעריך רציונלי (משפט 5.4.9) אם $s \in \mathbb{Q}$ ואם (a_n) סדרה מתכנסת אז $(\lim a_n)^s = \lim(a_n)^s$. לכן עבור $s \in \mathbb{Q}$ ועבור a, x כמו במשפט, אם נבחר סדרה רציונלית $r_n \to x$ אז

$$(a^x)^s = (\lim_{n\to\infty} a^{r_n})^s = \lim_{n\to\infty}((a^{r_n})^s) = \lim_{n\to\infty}(a^{r_n \cdot s}) = a^{x\cdot s}$$

(השוויון האחרון כי $r_n s \to xs$ ו־$(r_n s)$ היא סדרה רציונלית). כעת אם $a > 0$, $x, y \in \mathbb{R}$ ו־(s_n) סדרה של מספרים רציונליים המקיימת $s_n \to y$, אז

$$(a^x)^y = \lim_{n\to\infty}((a^x)^{s_n}) = \lim_{n\to\infty} a^{x\cdot s_n} = a^{x\cdot y}$$

כאן השוויון השמאלי נובע מכך ש־ $b^y = \lim b^{s_n}$ לכל b ובפרט ל־ $b = a^x$. השוויון השני כי s_n רציונלי והראינו קודם שכאשר s רציונלי מתקיים $(a^x)^s = a^{xs}$. השוויון האחרון כי $x \cdot s_n \to x \cdot y$ ועל סמך למה 5.10.5 (אי אפשר להסתמך כאן על הגדרת החזקה הממשית כי $x \cdot s_n$ אינם רציונליים).

יתר התכונות מוכחות באופן דומה, ומושארות כתרגיל. ■

אפשר היה להגדיר את החזקה ללא שימוש בסדרות, בשיטה הדומה לשיטה בה הגדרנו שורשים שלמים. הרעיון הוא להגדיר את a^x עבור $a > 1$ להיות החסם העליון של כל המספרים a^r כאשר $r < x$ רציונלי. הטענה הבאה מראה שהגדרה זו וההגדרה שלנו שקולות:

טענה 5.10.7 יהי $a > 1$ ו־ $x \in \mathbb{R}$. אז

$$a^x = \sup\{a^r \,:\, r < x\,,\, r \in \mathbb{Q}\} = \inf\{a^r \,:\, r > x\,,\, r \in \mathbb{Q}\}$$

הוכחה תהי $A = \{a^r \,:\, r < x\,,\, r \in \mathbb{Q}\}$ ונוכיח את השוויון $a^x = \sup A$. השוויון השני מוכח בצורה דומה. אם $r < x$ אז $a^r < a^x$ (כי $a > 1$) ולכן a^x חסם מלעיל של A. מאידך, תהי (r_n) סדרה של מספרים רציונליים כך ש־ $r_n < x$ ו־ $r_n \to x$. אז $a^{r_n} \in A$ ו־ $a^{r_n} \to a^x$, ונובע ש־ a^x החסם העליון של A. ■

נסיים את הסעיף עם משפט אריתמטיקה אחרון:

משפט 5.10.8 (כלל החזקה) אם (a_n) סדרה ו־ $a_n \to a > 0$ אז לכל $x \in \mathbb{R}$ מתקיים $a_n^x \to a^x$.

הוכחה נניח ש־ $a > 1$ (המקרה $a < 1$ נובע ממנו – איך?). נקבע מספר רציונלי $r > x$. החל ממקום מסוים $a_n > 1$, ואז $a_n^x < a_n^r$. מכאן ש־

$$\limsup a_n^x \le \limsup a_n^r = a^r$$

(בשוויון האחרון השתמשנו בכלל החזקה למעריכים רציונליים). אי־שוויון זה נכון לכל $r > x$ רציונלי, ולכן לפי הטענה הקודמת,

$$\limsup a_n^x \le \inf\{a^r \,:\, r > x\,,\, r \in \mathbb{Q}\} = a^x$$

באותו אופן מראים ש־ $\liminf a_n^x \ge a^x$. קיבלנו בסיכום ש־

$$a^x \le \liminf a_n^x \le \limsup a_n^x \le a^x$$

וכיוון שבקצוות מופיע אותו מספר a^x, כל האי־שוויונות הם שוויונות, וממסקנה 5.8.5 נסיק ש־ $a_n^x \to a^x$. ■

תרגילים

1. השלימו את הוכחת התכונות של החזקה הממשית.
2. הוכיחו שאם $a_n \to a > 0$ ו־ $b_n \to b$ אז $(a_n)^{b_n} \to a^b$.
3. השלימו את הוכחת המקרה של inf בטענה 5.10.7.

5.11 המספר e והחזקות e^x

בסעיף זה נגדיר את המספר החשוב e ונמצא לכל מספר ממשי x סדרה פשוטה המתכנסת לחזקה e^x. החשיבות של e וחזקותיו תתבהר בהמשך.

משפט 5.11.1 לכל $x \in \mathbb{R}$ קיים הגבול $\lim_{n\to\infty}(1+\frac{x}{n})^n$.

למרבה הצער לא נוכל לתת בשלב זה את המוטיבציה מאחורי ההגדרה של הסדרה $(1+\frac{x}{n})^n$. בינתיים נקבל את ההגדרה כמו שהיא ונתרכז בחקירת התכונות שלה. תוך כדי כך יתברר שלנוסחה $(1+\frac{x}{n})^n$ יש תכונות אלגבריות מעניינות, שבעזרתן נוכיח את המשפט.

נתחיל בהוכחת התכנסות הסדרה $(1+\frac{1}{n})^n$. מסתבר שזו סדרה מונוטונית, אף שבמבט ראשון זה כלל לא ברור: כשמגדילים את n הבסיס $1+\frac{1}{n}$ קטן, אבל המעריך גדל, כלומר יש כוחות הפועלים בו זמנית להקטין ולהגדיל את הביטוי. כדי להעריך את $(1+\frac{1}{n})^n$ נפתח את הסוגריים בעזרת במשפט הבינום (משפט 3.6.7) ונקבל

$$(1+\frac{1}{n})^n = \sum_{k=0}^{n}\binom{n}{k}\cdot(\frac{1}{n})^k\cdot 1^{n-k} = \sum_{k=0}^{n}\binom{n}{k}\cdot\frac{1}{n^k}$$

כאשר $\binom{n}{k}$ הוא המספר

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{k!}$$

כשמציבים את ההגדרה של $\binom{n}{k}$ ומסדרים מחדש, מקבלים את השוויון

$$(1+\frac{1}{n})^n = \sum_{k=0}^{n}\frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{n^k}\cdot\frac{1}{k!} = \sum_{k=0}^{n}\frac{1}{k!}a_{n,k}$$

כאשר $a_{n,k}$ מסמן את המספר

$$a_{n,k} = \frac{n}{n}\cdot\frac{n-1}{n}\cdot\ldots\cdot\frac{n-k+1}{n}$$

ברור ש־ $0\leq a_{n,k}\leq 1$ לכל $0\leq k\leq n$ טבעיים, כי הוא מכפלה של מספרים בין אפס לאחת. יתרה מזאת, לכל $k\in\mathbb{N}$ ולכל $n\geq k$ מתקיים $a_{n+1,k}\geq a_{n,k}$: כדי לראות זאת, כופלים זה בזה את האי־שוויונות $\frac{n+1-m}{n+1}\geq\frac{n-m}{n}$ עבור $m=0,1,\ldots,k-1$ (בדקו!).

טענה 5.11.2 הסדרה $((1+\frac{1}{n})^n)_{n=1}^{\infty}$ מתכנסת.

הוכחה די להראות שהסדרה עולה וחסומה מלעיל. נוכיח תחילה שהסדרה חסומה מלעיל. בסימונים הקודמים מקבלים

$$(1+\frac{1}{n})^n = \sum_{k=0}^{n}a_{n,k}\cdot\frac{1}{k!} \leq \sum_{k=0}^{n}\frac{1}{k!}$$

כי $a_{n,k}\leq 1$. נשים לב שלכל $k\geq 2$ מתקיים

$$\frac{1}{k!} = \frac{1}{k\cdot(k-1)\cdot\ldots\cdot 2\cdot 1} \leq \frac{1}{k(k-1)} = \frac{1}{k-1}-\frac{1}{k}$$

שימו לב ש־

$$\sum_{k=2}^{n}(\frac{1}{k-1}-\frac{1}{k}) = 1-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\frac{1}{3}-\frac{1}{4}+\ldots\frac{1}{n-1}-\frac{1}{n} = 1-\frac{1}{n}$$

כי פרט למחובר הראשון והאחרון שאר המחוברים מתחלקים לזוגות שמצמצמים אחד את השני. ולכן

$$\sum_{k=0}^{n}\frac{1}{k!}=1+1+\sum_{k=2}^{n}\frac{1}{k!}\le 2+\sum_{k=2}^{n}\left(\frac{1}{k-1}-\frac{1}{k}\right)=2+\left(1-\frac{1}{n}\right)\le 3$$

וקיבלנו בסיכום שלכל $n\in\mathbb{N}$ מתקיים $(1+\frac{1}{n})^n\le 3$.

כדי לראות ש־$(1+\frac{1}{n})^n$ סדרה עולה מספיק להראות ש־ $(1+\frac{1}{n+1})^{n+1}-(1+\frac{1}{n})^n\ge 0$ לכל n. ואמנם,

$$\begin{aligned}(1+\frac{1}{n+1})^{n+1}-(1+\frac{1}{n})^n &= \sum_{k=0}^{n+1}a_{n+1,k}\cdot\frac{1}{k!}-\sum_{k=0}^{n}a_{n,k}\cdot\frac{1}{k!}\\ &= \frac{1}{(n+1)!}a_{n+1,n+1}+\sum_{k=0}^{n}(a_{n+1,k}-a_{n,k})\frac{1}{k!}\end{aligned}$$

ראינו קודם ש־ $a_{n+1,k}-a_{n,k}\ge 0$, וממילא $a_{n+1,n+1}\ge 0$, ולכן כל המחוברים בסכום האחרון חיוביים, כפי שרצינו להראות. ■

הגדרה 5.11.3 הגבול $\lim_{n\to\infty}(1+\frac{1}{n})^n$ מסומן ב־e.

ראינו כבר ש־3 חסם מלעיל של הסדרה $(1+\frac{1}{n})^n$ ולכן $e=\lim(1+\frac{1}{n})^n\le 3$. מצד שני $2<\frac{9}{8}=(1+\frac{1}{2})^2$ ומכיוון שהסדרה עולה לגבול e קיבלנו $e>2$. בפרקים מאוחרים יותר נמצא דרכים להעריך את e בדיוק רב יותר. נציין כאן שהוא שווה בקירוב ל־ $2.7182818284\ldots$. בהמשך נראה גם ש־e אינו רציונלי.

מטרתנו הבאה היא להראות שלכל $x\in\mathbb{R}$ הגבול $\lim_{n\to\infty}(1+\frac{x}{n})^n$ קיים ושווה ל־ e^x. עבור $x>0$ אפשר להוכיח את ההתכנסות של $(1+\frac{x}{n})^n$ בצורה דומה להוכחת ההתכנסות במקרה $x=1$, אך אנו ניתן הוכחה אחרת לעובדה זו. לשם כך נזדקק ללמה הבאה:

למה 5.11.4 תהי (a_n) סדרה כך ש־ $a_n\to\infty$. אז $(1+\frac{1}{a_n})^{a_n}\to e$

הוכחה המקרה $a_n=n$ היא המשפט הקודם, ואם (a_n) סדרה עולה ממש של מספרים טבעיים זה נובע מהמשפט הקודם כי אז הסדרה $((1+\frac{1}{a_n})^{a_n})$ היא תת־סדרה מוכללת של הסדרה $((1+\frac{1}{n})^n)$, ומתכנסת לאותו גבול. ההוכחה לסדרה (a_n) כללית מבוססת על הערכה לפיה $(1+\frac{1}{a_n})^{a_n}$ שווה בקירוב ל־$(1+\frac{1}{k})^k$ עבור k טבעי מתאים.

בלי הגבלת הכלליות נוכל להניח ש־ $a_n\ge 1$. לכל x ממשי הערך השלם $[x]$ של x מקיים $[x]\le x\le [x]+1$, ונובע ש־

$$(1+\frac{1}{[a_n]+1})^{[a_n]}\le(1+\frac{1}{a_n})^{a_n}\le(1+\frac{1}{[a_n]})^{[a_n]+1}$$

נוכל לכתוב אי־שוויון זה גם באופן הבא:

$$(1+\frac{1}{[a_n]+1})^{-1}\cdot(1+\frac{1}{[a_n]+1})^{[a_n]+1}\leq(1+\frac{1}{a_n})^{a_n}\leq(1+\frac{1}{[a_n]})\cdot(1+\frac{1}{[a_n]})^{[a_n]}$$

מכיוון ש־ $a_n\to\infty$ גם $[a_n]\to\infty$ ולכן $(1+\frac{1}{[a_n]})\to 1$. וגם $1+\frac{1}{[a_n]+1}\to 1$. מצד שני, הסדרות $(1+\frac{1}{[a_n]})^{[a_n]}$ ו־ $(1+\cdot\frac{1}{[a_n]+1})^{[a_n]+1}$ הן תת־סדרות מוכללות של $(1+\frac{1}{n})^n$ ולכן

$$\lim_{n\to\infty}(1+\frac{1}{[a_n]})^{[a_n]}=\lim_{n\to\infty}(1+\frac{1}{[a_n]+1})^{[a_n]+1}=e$$

מכאן שהצלחנו לחסום את $(1+\frac{1}{a_n})^{a_n}$ בין סדרות המתכנסות ל־e, ומכלל הסנדוויץ' מקבלים $(1+\frac{1}{a_n})^{a_n}\to e$. ■

משפט 5.11.5 עבור $x\geq 0$ מתקיים $\lim_{n\to\infty}(1+\frac{x}{n})^n=e^x$ (ובפרט הגבול קיים).

הוכחה המקרה $x=0$ מידי. נוכיח כעת את המקרה ש־ $x>0$, המקרה השלילי ידון בהמשך. נניח אם־כן ש־ $x>0$. נשים לב שאפשר לכתוב

$$(1+\frac{x}{n})^n=((1+\frac{1}{n/x})^{n/x})^x$$

מכיוון ש־ $\lim_{n\to\infty}\frac{n}{x}\to\infty$ אנו מקבלים מהטענה הקודמת ש־

$$(1+\frac{1}{n/x})^{n/x}\to e$$

ומאריתמטיקה של גבולות (כלל החזקה, משפט 5.10.8) מסיקים

$$(1+\frac{x}{n})^n=((1+\frac{1}{n/x})^{n/x})^x\to e^x$$

כנדרש. ■

נותר להוכיח את המשפט עבור $x<0$. ניעזר בלמה הבאה, שהיא בעלת עניין בפני עצמה. תהי (ε_n) סדרה אי־שלילית השואפת לאפס ונתבונן בסדרה $(1+\varepsilon_n)^n$. ישנם שני כוחות הפועלים על איברי הסדרה: השאיפה לאפס של ε_n מנסה להקטין, אבל השאיפה לאינסוף של n מנסה להגדיל. התוצאה נקבעת על ידי יחס הכוחות. ראינו כבר שאם בוחרים $\varepsilon_n=\frac{x}{n}$ אז $(1+\varepsilon_n)^n\to e^x$. הלמה הבאה אומרת שאם ε_n קטן יותר, כלומר אם קצב השאיפה שלו לאפס מהיר יותר מכל כפולה של הסדרה ההרמונית, אז $(1+\varepsilon_n)^n\to 1$:

למה 5.11.6 תהי $(\varepsilon_n)_{n=1}^{\infty}$ סדרה אי־שלילית כך ש־ $n\varepsilon_n\to 0$. אז $(1+\varepsilon_n)^n\to 1$.

הוכחה מספיק להוכיח זאת תחת ההנחה $\varepsilon_n > 0$ (למה?). נשים לב שאז $\frac{1}{\varepsilon_n} \to \infty$ ולכן מלמה 5.11.4 מתקיים $(1+\varepsilon_n)^{1/\varepsilon_n} \to e$. בפרט אנו מקבלים את האי־שוויון

$$1 \le (1+\varepsilon_n)^{1/\varepsilon_n} \le 3$$

לכל n גדול מספיק. נעלה את שני צדי האי־שוויון בחזקה $n\varepsilon_n$ ונקבל שלכל n גדול מספיק

$$1 \le (1+\varepsilon_n)^n \le 3^{n\varepsilon_n}$$

אבל $n\varepsilon_n \to 0$ ולכן אגף ימין שואף ל־1 (למה?), ולכן מכלל הסנדוויץ' קיבלנו $(1+\varepsilon_n)^n \to 1$ כפי שרצינו. ■

מסקנה 5.11.7 תהי $(\varepsilon_n)_{n=1}^\infty$ סדרה אי־שלילית כך ש־ $n\varepsilon_n \to 0$. אז $(1-\varepsilon_n)^n \to 1$.

הוכחה ניעזר בזהות $1-x = (1+\frac{x}{1-x})^{-1}$ ונקבל

$$(1-\varepsilon_n)^n = \frac{1}{(1+\frac{\varepsilon_n}{1-\varepsilon_n})^n} \to \frac{1}{1} = 1$$

כאשר הגבול נובע מאריתמטיקה של גבולות ומהלמה הקודמת, מאחר שמתקיים $n \cdot \frac{\varepsilon_n}{1-\varepsilon_n} \to 0$. ■

טענה 5.11.8 עבור $x < 0$ מתקיים $\lim_{n\to\infty}(1+\frac{x}{n})^n = e^x$ (ובפרט הגבול קיים).

הוכחה יהי $x < 0$. נשתמש בזהות $1+a = \frac{1-a^2}{1-a}$ על מנת להסי

$$\lim_{n\to\infty}(1+\frac{x}{n})^n = \lim_{n\to\infty}\frac{(1-\frac{x^2}{n^2})^n}{(1+\frac{(-x)}{n})^n} = \frac{\lim_{n\to\infty}(1-\frac{x^2}{n^2})^n}{\lim_{n\to\infty}(1+\frac{(-x)}{n})^n} = \frac{1}{e^{-x}}$$

כאן השתמשנו בעובדה ש־ $-x > 0$ כדי להסיק ש־ $(1+\frac{(-x)}{n})^n \to e^{-x}$, והשתמשנו במסקנה 5.11.7 כדי להסיק שהמונה שואף ל־1. ■

בכך סיימנו את ההוכחה שלכל $x \in \mathbb{R}$ מתקיים $\lim_{n\to\infty}(1+\frac{x}{n})^n = e^x$, כמובטח.

תרגילים

1. מצאו את השגיאה בנימוק הבא: אנו יודעים ש־ $1+\frac{1}{n} \to 1$ ומתקיים $(1+\frac{1}{n})^n = \prod_{i=1}^n(1+\frac{1}{n})$. לכן מכלל המכפלה נובע $(1+\frac{1}{n})^n \to 1\cdot1\cdot\ldots\cdot1 = 1$.
2. הוכיחו ש־ $e < 3$ (אנו הוכחנו אי־שוויון חלש). נסו לשפר עוד את ההערכה שלכם.

3. חשבו את הגבולות של הסדרות הבאות:
4. $(1+\frac{n+1}{n^2+1})^{3n+1}$.
5. $(1+\frac{n+1}{n^2+1})^{3n^2+1}$.
6. $(1+\frac{1}{n^2+1})^{3n+1}$.
7. הוכיחו שלכל x מתקיים $1+x \le e^x$ ושאם $x > -1$ גם מתקיים $e^{\frac{x}{1+x}} \le 1+x$ (בשביל האי־שוויון הראשון היעזרו באי־שוויון ברנולי, ראו תרגיל (10) בעמוד 56. האי־שוויון השני נובע מהראשון).
8. הוכיחו שלכל $x > 0$ הסדרה $(1+\frac{x}{n})^n$ עולה, והראו שהיא מתכנסת. הסיקו מכך ש־ $\lim_{n\to\infty} n^\alpha e^{-n} = 0$ לכל α.
9. הוכיחו שאם $n\varepsilon_n \to 0$ אז $(1+\varepsilon_n)^n \to 1$, מבלי להניח ש־ $\varepsilon_n > 0$.

5.12 קבוצות בנות מניה ועוצמת הממשיים

בסעיף 3.5 הבחנו בין קבוצות סופיות וקבוצות אינסופיות. אפשר להשוות את גודלן של קבוצות הסופיות על ידי השוואת מספר האיברים בהן. נשאלת השאלה האם ניתן למיין בצורה דומה את הקבוצות האינסופיות, כלומר, האם יש קבוצות אינסופיות גדולות יותר וגדולות פחות? את העובדה שמיון כזה אפשרי גילה גאורג קנטור בחצי השני של המאה ה־19 כאשר הראה שיש קבוצות אינסופיות גדולות יותר וגדולות פחות. התגלית באה בעקבות מחקר שעשה על קבוצות של מספרים ממשיים. במהרה הסתבר שישנה היררכיה שלמה של גדלים אינסופיים. יש אינסוף אינסופים שונים, אבל אין אינסוף גדול ביותר (אם כי יש אינסוף קטן ביותר!).

כפי שגדלים סופיים נקראים מספרים, גדלים אינסופיים מכונים **עוצמות** (cardinals). לא נגדיר במדויק את מושג העוצמה אלא נפתח את אותו חלק של התורה שמעניין בהקשר שלנו: נראה ש־$\mathbb{R}$ היא מעוצמה גדולה יותר מ־$\mathbb{N}$.

הקבוצות האינסופיות הקטנות ביותר הן הקבוצות בנות־המניה:

הגדרה 5.12.1 נאמר על קבוצה A שהיא **קבוצה בת מניה** (countable set) אם יש סדרה $(a_n)_{n=1}^\infty$ כך שכל איבר של a מופיע בסדרה, כלומר, לכל $a \in A$ יש $n \in \mathbb{N}$ כך ש־ $a_n = a$ (ייתכן ש־a מופיע בסדרה יותר מפעם אחת).

כל קבוצה סופית היא בת מניה כי אם $A = \{a_1, \ldots, a_n\}$ אז בוודאי יש סדרה שמתחילה באיברים $a_1, a_2, \ldots, a_n$. מאותה הסיבה הקבוצה הריקה בת־מניה.[16]

[16] ישנם ספרים בהם המונח בן־מניה שמור רק לקבוצות אינסופיות. על קבוצה כזאת אומרים שהיא מעוצמה $\aleph_0$. נציין שאם A קבוצה בת־מניה אינסופית אפשר למצוא סדרה (a_n) שכל איבריה הם מ־A וכך שכל איבר מ־A מופיע בסדרה בדיוק פעם אחת. תתבקשו להוכיח זאת בתרגיל (4) בסוף הסעיף.

הקבוצה $\mathbb{N}$ בת־מניה כי איבריה הם איברי הסדרה $a_n = n$ וגם $\mathbb{Z}$ היא בת־מניה, כי הסדרה

$$0, 1, -1, 2, -2, 3, -3, \dots$$

מכילה את כל השלמים (אפשר לתת נוסחה לסדרה זו, למשל $a_n = [\frac{n+1}{2}]\cdot(-1)^{n+1}$). באופן מפתיע גם $\mathbb{Q}$ היא קבוצה בת מניה. על מנת להוכיח זאת נעזר בלמה הכללית הבאה:

למה 5.12.2 לכל n טבעי תהי A_n קבוצה סופית. תהי $A = \bigcup_{n=1}^{\infty} A_n$ איחוד הקבוצות $A_1, A_2, \dots$. אז A בת־מניה.

הוכחה נסמן ב־ $k_n = |A_n|$ את מספר האיברים ב־A_n, ונסמן את האיברים ב־ $A_n = \{a_{n,1}, \dots, a_{n,k_n}\}$. נבנה את הסדרה

$$a_{1,1}, \dots, a_{1,k_1}, a_{2,1}, \dots, a_{2,k_2}, \dots, a_{i,1}, \dots, a_{i,k_i}, \dots$$

באופן מדויק, נשים לב שלכל $i \in \mathbb{N}$ יש j, m יחידים כך ש־ $1 \le j \le k_m$ ו־ $i = k_1 + \ldots + k_{m-1} + j$ (הוכיחו זאת!), ולכן נוכל להגדיר $x_i = a_{j,m}$ לכל שלשה i, j, m כזו. ברור שכל איבר של A מופיע בסדרה (x_i). ■

טענה 5.12.3 $\mathbb{Q}$ בת־מניה.

הוכחה לכל $0 < n \in \mathbb{N}$ תהי $A_n \subseteq \mathbb{Q}$ קבוצת כל המספרים $\frac{p}{q}$ עם p, q שלמים המקיימים $|p|, |q| \le n$. ברור ש־ A_n סופית (קל לראות למשל שיש ב־A_n לכל היותר $(2n+1)^2$ איברים). כמו־כן $\mathbb{Q} = \bigcup_{n=1}^{\infty} A_n$, ולכן לפי הלמה $\mathbb{Q}$ בת מניה. ■

העובדה המפתיעה היא שלא כל קבוצה היא בת־מניה. אנו נסתפק בדוגמה אחת לקבוצה כזאת (נראה דוגמה נוספת בתרגילים בסוף הסעיף):

משפט 5.12.4 (קנטור, 1887): $\mathbb{R}$ אינה בת מניה.[17]

הוכחה עלינו להראות שאין סדרה $(a_n)_{n=1}^{\infty}$ של מספרים שכל מספר ממשי מופיע בה (מובן שאין צורך לדון בסדרות שחלק מאיבריהן אינם מספרים), דהיינו שלכל סדרה כזו יש מספר $x \in [0,1]$ שאינו מופיע בה.

תהי $(a_n)_{n=1}^{\infty}$ סדרה של מספרים. נבנה ברקורסיה סדרת קטעים סגורים $(I_n)_{n=1}^{\infty}$ שכולם מוכלים ב־$[0,1]$ ובעלת התכונות הבאות:

1. $I_{n+1} \subseteq I_n$
2. $0 < |I_{n+1}| < \frac{1}{2}|I_n|$

[17] אומרים שהקבוצה $\mathbb{R}$ היא **מעוצמת הרצף** (cardinality of the continuum) ומסמנים את עוצמת הרצף ב־א או ב־$\mathfrak{c}$.

3. $a_n \notin I_n$

נשתמש בעובדה הפשוטה הבאה, שהוכחתה מושארת כתרגיל: אם $[u,v]$ קטע בעל אורך חיובי ואם $w \in \mathbb{R}$ אז יש תת־קטע $[u',v'] \subseteq [u,v]$, גם־כן בעל אורך חיובי, שאינו מכיל את w, ואפשר לבחור את $[u',v']$ כך שאורך הקטע קטן כרצוננו. בעזרת הטענה נוכל לבחור את I_1 להיות תת־קטע של $[0,1]$ כך ש־ $a_1 \notin I_1$, ובהנחה שבחרנו את הקטעים $I_1,\ldots,I_n$ נבחר תת־קטע I_{n+1} של I_n שאורכו פחות ממחצית אורך I_n ושאינו מכיל את a_{n+1}.

מהלמה של קנטור קיים מספר $x \in \bigcap_{n=1}^{\infty} I_n$. אם היה n טבעי כך ש־ $a_n = x$ היינו מקבלים $a_n \in \bigcap_{n=1}^{\infty} I_n$ ובפרט $a_n \in I_n$. אבל זו סתירה, ולכן x לא מופיע בסדרה (a_n), כפי שרצינו. ∎

לא קשה להכליל את המשפט ולהראות שכל קטע לא מנוון אינו בן־מניה.

מדוע בעצם אי אפשר לסדר את איברי הממשיים בסדרה, למשל על ידי בחירת המספרים אחד־אחד באופן שעוברים על כולם? ובכן, קבוצת הממשיים פשוט גדולה מדי. מהמשפט נובע שבכל ניסיון לבחור סדרת איברים מקבוצה בת מניה יישארו בסוף איברים שלא נבחרו. ויותר מכך: מה שיישאר יהיה עצמו קבוצה גדולה. ואמנם,

למה 5.12.5 אם A,B קבוצות בנות מניה אז $A \cup B$ בת מניה.

הוכחה יהיו $(a_n),(b_n)$ סדרות המכילות את כל איברי A,B בהתאמה. תהי (c_n) הסדרה $a_1,b_1,a_2,b_2,a_3,b_3,\ldots$. אז $\{c_n : n \in \mathbb{N}\}$ מכילה את כל איברי $A\cup B$ ולכן $A \cup B$ בת מניה. ∎

מסקנה 5.12.6 אם A קבוצה שאינה בת מניה ואם $A_0 \subseteq A$ בת מניה אז $A \setminus A_0$ שוב אינה בת מניה.

הוכחה אילו $A \setminus A_0$ הייתה בת מניה, מהלמה הקודמת $A = A_0 \cup (A \setminus A_0)$ הייתה איחוד של שתי קבוצות בנות מניה ולכן בעצמה בת־מניה, בסתירה לנתון. ∎

אפשר לפרש את המסקנה באופן הבא: קבוצה שאינה בת מניה היא קבוצה כל כך גדולה ביחס לקבוצות בנות מניה שאם נוציא ממנה קבוצה בת־מניה, אפילו שהיא אינסופית, לא נצליח להקטין במידה ניכרת את הקבוצה המקורית.

אחד השימושים של מושג העוצמה היא כדי להוכיח את קיומם של איברים מסוג מסוים. השיטה עובדת באופן הבא. נניח שרוצים להראות שבקבוצה A יש איברים עם תכונה P, נניח ש־A אינה בת־מניה. אם אפשר להראות שקבוצת האיברים ב־A שלא מקיימים P היא קבוצה בת־מניה, אז נוכל על סמך המסקנה הקודמת להסיק שיש ב־A איברים שמקיימים את P.[18] הנה יישום טיפוסי של שיטה זו:

[18] הוכחות מסוג זה נקראות **הוכחות קיום לא קונסטרוקטיביות**, כי במקום להצביע על איבר מסוים עם התכונה המבוקשת מראים שיש הרבה כאלה אך מבלי למצוא דוגמה קונקרטית. האפשרות לתת

מסקנה 5.12.7 קיימים מספרים אי־רציונליים. יתרה מזאת, קבוצת המספרים האי־רציונליים היא תת־קבוצה שאינה בת מניה של $\mathbb{R}$.

הוכחה $\mathbb{Q}$ בת מניה ו־$\mathbb{R}$ אינה בת־מניה, ולכן לפי הלמה הקודמת $\mathbb{R} \setminus \mathbb{Q}$ אינה בת מניה ובפרט אינה ריקה. ■

הנה שימוש נוסף של אותה השיטה, הנותנת מסקנה מפתיעה ומעוררת ענווה. לצורך הדיון נסכים שמשפט בשפה העברית הוא סדרה סופית של סמלים המכילה אותיות, ספרות, רווחים, וסימני פיסוק. אנו נתעניין במספרים הניתנים לתיאור בשפה העברית.[19] יש מספרים רבים כאלה, ביניהם כל המספרים שהצלחנו לתאר בספר עד כה. למשל, את המספרים הטבעיים אפשר לכתוב בייצוג עשרוני, את המספר $\sqrt{2}$ אפשר לתאר כאותו מספר חיובי שהכפלתו עם עצמו נותן 2, המספר e הוא הגבול של הסדרה $(1+\frac{1}{n})^n$, וכן הלאה. כפי שרואים מהדוגמאות השפה העברית (כמו שפות מדוברות אחרות) היא בעלת כושר הבעה רב, ונשאלת השאלה אם קיימים בכלל מספרים שאינם ניתנים לתיאור, או שמא את כולם אפשר לתאר. התשובה היא שקיימים מספרים כאלה, ויתרה מזאת, רוב המספרים אינם ניתנים לתיאור!

נסמן ב־A_n את קבוצת המשפטים באורך n בשפה העברית. כיוון שיש בשפה רק מספר סופי של סמלים הקבוצה A_n היא סופית, ולכן קבוצת כל המשפטים בעברית, כלומר $A = \bigcup_{n=1}^{\infty} A_n$, היא קבוצה בת־מניה (זה נובע מלמה 5.12.2). לכן קיימת סדרה $(a_n)_{n=1}^{\infty}$ שבין איבריה מופיעים כל המשפטים בעברית. נגדיר כעת סדרה $(b_n)_{n=1}^{\infty}$ באופן הבא: עבור אינדקס n, אם a_n הוא משפט שמתאר מספר x, אז $b_n = x$, ואחרת $b_n = 0$. כך קיבלנו סדרה (b_n) שבין איבריה מופיעים כל המספרים הניתנים לתיאור, ולכן המספרים הניתנים לתיאור מהווים קבוצה בת־מניה. מכאן אנו מסיקים שקבוצת המספרים הממשיים שאינם ניתנים לתיאור היא קבוצה לא בת־מניה, ובפרט לא ריקה!

משמעות הדבר הוא שאנו חוקרים עולם של מספרים שאת רובם אין לנו אפשרות לתאר וגם לעולם לא תהיה לנו אפשרות לתאר! אבל אל דאגה, עולם המספרים שאנו יכולים לתאר עשיר מעבר לכל דמיון, ואין סכנה שנבין לחלוטין אפילו אותו.

תורת העוצמות האינסופיות שייכת לתורת הקבוצות. תוכלו ללמוד עליה ב־[12],[13].

תרגילים

1. הראו שתת־קבוצה של קבוצה בת־מניה היא בת־מניה.

הוכחות קיום כאלה עוררה התנגדות נמרצת בקרב חלק מהמתמטיקאים, והובילה בסוף המאה ה־19 ותחילת המאה ה־20 לכמה תנועות ששללו את השימוש בהוכחות כאלה. לרוב תנועות אלה ערערו גם על יסודות נוספים של המתמטיקה או של הלוגיקה. כיום רוב המתמטיקאים אינם מקבלים התנגדויות האלה, והוכחות קיום לא קונסטרוקטיביות הן כלי מקובל וחשוב המתמטיקה.

[19] לא ניכנס כאן לשאלה מה זה אומר לתאר מספר, כי זו שאלה קשה. במקום זאת נניח שקיים בורר חכם שמכריע אם משפט מתאר מספר ואם כן אז הוא יודע לחשב אותו.

2. יהיו A,B קבוצות בנות מניה. הוכיחו שגם קבוצת המכפלה $A\times B$ (כלומר אוסף כל הזוגות $A\times B=\{(a,b)\,:\,a\in A\,,\,b\in B\}$) היא בת־מניה. זו הכללה של טענה 5.12.3, ותוכלו להיעזר ברעיון שמופיע בהוכחה שלה.

3. הראו שאם $A_1,A_2,A_3\ldots$ הן קבוצות בנות מניה אז הקבוצה $A=\cup_{n=1}^{\infty}A_n$, המורכבת מכל הנקודות ששייכות לאחת הקבוצות A_n, היא בת־מניה (זו הכללה של למה 5.12.2).

4. תהי A קבוצה בת מניה אינסופית. הוכיחו שיש סדרה $(a_n)_{n=1}^{\infty}$ שכל איבריה לקוחים מ־A ומכילה כל איבר של A בדיוק פעם אחת (שימו לב שבהגדרה של קבוצה בת־מניה לא דרשנו שכל איבר בסדרה יהיה איבר ב־A, אלא רק שכל איבר של A יופיע בסדרה, וגם הרשינו לסדרה להכיל חזרות).

5. תהי A קבוצה ונניח שלכל $a\in A$ נתון קטע לא טריביאלי I_a. נניח ש־$I_a\cap I_b=\emptyset$ לכל $a\neq b$ מתוך A. הוכיחו ש־A בת־מניה. (רמז: התאימו לכל $a\in A$ מספר רציונלי $r_a\in I_a$, הראו שאם $a\neq b$ אז $r_a\neq r_b$, וזכרו ש־$\mathbb{Q}$ בת מניה).

6. נסמן ב־Σ את אוסף כל הסדרות שאיבריהם 0 או 1. למשל הסדרות הקבועות $(0,0,0,\ldots)$ ו־$(1,1,1,\ldots)$ שייכות ל־Σ, וגם הסדרה $(0,1,0,1,0\ldots)$, וכן הלאה. בשאלה זו נראה ש־Σ אינה קבוצה בת־מניה.

ברור ש־Σ אינסופית ולכן אילו הייתה בת־מניה היה אפשר לרשום אותה כ־$\Sigma=\{a^{(n)}\,:\,n=1,2,3\ldots\}$ כאשר $a^{(n)}=(a_1^{(n)},a_2^{(n)},a_3^{(n)},\ldots)$ היא סדרה של אפסים ואחדות. נגדיר את הסדרה $(b_k)_{k=1}^{\infty}$ על ידי

$$b_k=\begin{cases}0 & a_k^{(k)}=1\\ 1 & a_k^{(k)}=0\end{cases}$$

הראו ש־ $(b_k)\in\Sigma$ אבל אין אף n כך ש־ $(b_k)_{k=1}^{\infty}=(a_k^{(n)})_{k=1}^{\infty}$, וזו סתירה.

פרק 6

טורים

בפרק זה נגדיר "סכום אינסופי", כלומר סכום עם מספר אינסופי של מחוברים, הנקרא טור. סכום כזה אינו מוגדר כחלק מהפעולות הבסיסיות בין מספרים, והוא למעשה מקרה פרטי של מושג הגבול של סדרה. ישנן הקבלות רבות בין טורים לסכומים רגילים, אך יש גם הבדלים מפתיעים: למשל, תכונות החילוף והקיבוץ לא תמיד מתקיימות. נחקור עניינים אלה לעומק. במהלך הפרק נראה שימושים שונים של טורים, ובין השאר נגדיר בעזרתם את ייצוג המספרים בעזרת שברים עשרוניים אינסופיים.

6.1 טורים

חיבור של מספרים ממשיים מוגדר תחילה בין זוגות של מספרים. סכומים סופיים של יותר משני איברים מוגדרים על ידי פעולה חוזרת של הפעולה הבסיסית של חיבור זוגות. כך ניתן להגדיר סכום של מספר סופי כלשהו של מספרים. לעומת זאת בעזרת הפעולות האלה בלבד לא ניתן להגדיר סכום של סדרת מספרים אינסופית. מסתבר שהכלי הרלוונטי להגדרה של סכום אינסופי, הנקרא טור, הוא מושג הגבול של סדרה.

נקדים דוגמה להגדרה. נניח שאנו רוצים לסכם את כל המספרים מהצורה $\frac{1}{2^n}$ ל־n טבעי, כלומר לחשב את ה"סכום"

$$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}+\ldots+\frac{1}{2^n}+\ldots$$

אין בידינו הגדרה לסכום של אינסוף מספרים, אבל טבעי בכל־זאת להתחיל לחבר ולראות מה קורה. כך נקבל סדרה של תוצאות ביניים של החישוב, שאליהן אנו מוסיפים בכל שלב את האיבר הבא מסדרת המחוברים כדי לקבל את תוצאת הביניים הבאה. אפשר לקוות שהסדרה הזו תתכנס לגבול שיוכל להיקרא הסכום.

כדי לנסח זאת במפורש, נתבונן בסדרה $(\frac{1}{2^n})_{n=1}^{\infty}$, ולכל $N \in \mathbb{N}$ נסמן ב־S_N את סכום N האיברים הראשונים בסדרה, כלומר

$$S_N = \frac{1}{2^1} + \frac{1}{2^2} + \ldots + \frac{1}{2^N}$$

ראו איור 6.1.1. לפי הנוסחה לסכום של טור הנדסי, נקבל שלכל N מתקיים $S_N = 1 - \frac{1}{2^N}$, ומכאן נובע $S_N \to 1$. אפשר לחשוב על הסדרה $(S_N)_{N=1}^{\infty}$ כעל סדרת תוצאות הביניים של חישוב הסכום האינסופי $\frac{1}{2} + \frac{1}{2^2} + \ldots$, ומתקבל על הדעת להתייחס לגבול של סדרה זו (במידה והיא קיימת) כאילו הוא הערך של הסכום אינסופי אותו אנו רוצים להגדיר. יהיה זה לכן טבעי לומר שה"סכום האינסופי" $\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots$ שווה ל־1.

באופן כללי יותר, נגדיר:

הגדרה 6.1.1 תהי $(a_n)_{n=1}^{\infty}$ סדרת מספרים ממשיים ולכל N יהי

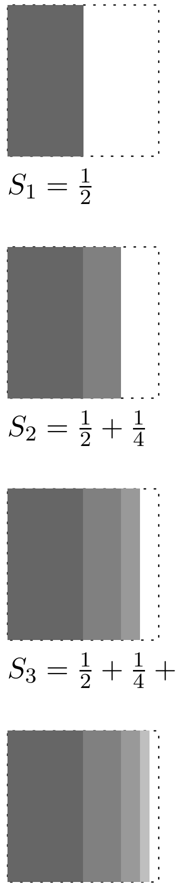
$$S_N = \sum_{n=1}^{N} a_n = a_1 + a_2 + \ldots + a_N$$

הסדרה (S_N) נקראת **סדרת הסכומים החלקיים** של הסדרה (a_n), והסימן $\sum_{n=1}^{\infty} a_n$ נקרא **הטור** (series) המתאים לסדרה (a_n). אם (S_N) מתכנסת למספר ממשי S, נאמר שהטור $\sum_{n=1}^{\infty} a_n$ **מתכנס**, נאמר ש**הסכום** שלו הוא S, ונסמן זאת על ידי $\sum_{n=1}^{\infty} a_n = S$. אם (S_N) אינה מתכנסת, נאמר שהטור $\sum_{n=1}^{\infty} a_n$ **מתבדר**.

באופן דומה, אם $S_N \to \infty$ נאמר שהטור $\sum_{n=1}^{\infty} a_n$ **מתכנס לאינסוף** (ל־∞) ונכתוב $\sum_{n=1}^{\infty} a_n = \infty$. התכנסות ל־$-\infty$ מוגדר באופן דומה. אם $\sum_{n=1}^{\infty} a_n$ מתכנס למספר ממשי או ל־$\pm\infty$ נאמר שהוא **מתכנס במובן הרחב**.

ניתן לרשום את הסכום של טור $\sum_{n=1}^{\infty} a_n$ מבלי להזכיר מפורש את סדרת הסכומים החלקיים על ידי

$$\sum_{n=1}^{\infty} a_n = \lim_{N\to\infty} \sum_{n=1}^{N} a_n$$



$S_1 = \frac{1}{2}$

$S_2 = \frac{1}{2} + \frac{1}{4}$

$S_3 = \frac{1}{2} + \frac{1}{4} + \frac{1}{8}$

$S_4 = \frac{1}{2} + \ldots + \frac{1}{16}$

איור 6.1.1 כמה סכומים חלקיים של הטור $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots$. שטח המלבנים מייצג את ערך המחוברים.

בהינתן טור $\sum_{n=1}^{\infty} a_n$, הסדרה (a_n) נקראת **סדרת איברי הטור**. לעתים קרובות לא נדבר מפורשות על הסדרה (a_n), אלא נציג מיד את הטור $\sum_{n=1}^{\infty} a_n$, גם כשהוא אינו מתכנס. כך הסימן $\sum_{n=1}^{\infty} a_n$ מקבל שתי משמעויות: לפעמים הוא נועד להציג את הסדרה אותה אנו רוצים לסכום, ובמקרים שהטור מתכנס הוא גם מסמל את הערך של הסכום.

לעתים יהיה צורך לדבר על טורים המתאימים לסדרה $(a_n)_{n=k}^{\infty}$ שאינה מתחילה מהאינדקס 1. נשתמש אז בסימון $\sum_{n=k}^{\infty} a_n$ ונתכוון לגבול של סדרת הסכומים החלקיים $S_N = \sum_{n=k}^{N} a_n$.

לעיתים נבחר בכתיב מקוצר $\sum a_n$ כדי ליצג את הטור במקום הצורה המלאה $\sum_{n=1}^{\infty} a_n$. בכך השמטנו את האינדקס ממנו מתחילה הסכימה, ולכן צורת הכתיבה לא מתאימה ליצג את ערך הטור, אשר עשוי להשתנות אם משנים את גבולות הסכימה. אבל תכונות אחרות של הטור, כמו עצם התכנסותו, אינם משתנים אם משמיטים איברים בתחילתו, ואז הכתיב המקוצר נוח ואינו פוגם בהבנה.

אפשר לפתח מושג מקביל של מכפלות אינסופיות. נדון בכך בסעיף 6.6.

דוגמאות

1. יהיו $a_1, \ldots, a_k$ מספרים ונגדיר $a_i = 0$ לכל $i > k$. אז $\sum_{i=1}^{\infty} a_i = \sum_{i=1}^{k} a_i$, כי לכל $n > k$ מתקיים

$$S_n = a_1 + \ldots + a_k + a_{k+1} + \ldots + a_n = (a_1 + \ldots + a_k) + 0 + \ldots + 0 = S_k$$

ולכן $\lim_{n\to\infty} S_n = S_k$.

2. יהי $c \in \mathbb{R}$ ונגדיר $a_n = c$ לכל n. אז $S_n = n \cdot c$ ומכאן ש־ $\sum_{n=1}^{\infty} a_n = 0$ אם $c = 0$, ושאם $c > 0$ אז $\sum_{n=1}^{\infty} a_n = \infty$ ואם $c < 0$ אז $\sum_{n=1}^{\infty} a_n = -\infty$.

3. מהדוגמה בתחילת הסעיף רואים ש־ $\sum_{k=1}^{\infty} 2^{-k} = 1$. באופן כללי יותר יהי $q \in \mathbb{R}$ ונתבונן בסדרה ההנדסית $(a_n)_{n=0}^{\infty}$ המוגדרת על ידי $a_n = q^n$, כלומר הסדרה $1, q, q^2, q^3, \ldots$. תהי $S_N = \sum_{k=0}^{N} q^k$ סדרת הסכומים החלקיים (שימו לב שהתחלנו את הטור מהאינדקס 0). אם $q \neq 1$ אז לפי הנוסחה לסכום של טור הנדסי נקבל ש־ $S_N = \frac{1-q^{N+1}}{1-q}$, ואם $q = 1$ אז $S_N = N$. לכן

   (א) אם $q = 1$ אז $S_N = N + 1 \to \infty$ ולכן $\sum_{n=0}^{\infty} a_n = \infty$.

   (ב) אם $|q| < 1$ נקבל ש־ $S_N = \frac{1-q^{N+1}}{1-q} \to \frac{1}{1-q}$, ולכן $\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$.

   (ג) אם $q > 1$ מקבלים $S_N \to \infty$ ולכן $\sum_{n=0}^{\infty} q^n = \infty$.

   (ד) אם $q \leq -1$ הטור אינו מתכנס, גם לא במובן הרחב. ואמנם, כיוון ש־ $q^{2n} \to \infty$ ואילו $q^{2n+1} \to -\infty$, הרי שמהנוסחה $S_N = \frac{1-q^{N+1}}{1-q}$ אנו מקבלים ש־ $S_{2N} \to -\infty$ ו־ $S_{2N+1} \to \infty$. בפרט, $\pm\infty$ הם גבולות חלקיים של הסדרה $(S_N)_{N=1}^{\infty}$, וממילא היא אינה מתכנסת במובן הרחב.

4. מקרה פרטי של הסעיף האחרון הוא המקרה $q = -1$. נקבל את הטור המתבדר

$$\sum_{n=0}^{\infty} (-1)^n = 1 - 1 + 1 - 1 + 1 - 1 + \ldots$$

שסכומיו החלקיים הם 1 ו־0 לסירוגין.

5. נתבונן בסדרה $a_n = \frac{1}{n(n+1)}$. נשים לב ש־$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, ולכן נוכל לחשב:

$$\begin{aligned} S_N &= \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \ldots + \frac{1}{N(N+1)} \\ &= (\frac{1}{1} - \frac{1}{2}) + (\frac{1}{2} - \frac{1}{3}) + (\frac{1}{3} - \frac{1}{4}) + \ldots + (\frac{1}{N} - \frac{1}{N+1}) \end{aligned}$$

פרט לסוגריים הימניים ביותר, בכל סוגריים מבטל האיבר הימני את האיבר השמאלי שבסוגריים שלימינו, ונקבל ש־$S_N = 1 - \frac{1}{N+1}$. לכן $S_N \to 1$, ולכן $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$.

6. כל בעיה על התכנסות של סדרה שקולה לבעיה על התכנסות של טור מתאים. ואמנם, בהינתן סדרה $(a_n)_{n=1}^{\infty}$ נתבונן בטור $\sum_{n=1}^{\infty} b_n$ שאיבריו מוגדרים על ידי $b_1 = a_1$ ו־ $b_n = a_n - a_{n-1}$ לכל $n \geq 2$. קל לראות שאם (S_n) היא סדרת הסכומים החלקיים של (b_n) אז $S_n = a_n$, ולכן התכנסות הסדרה (a_n) שקולה להתכנסות הטור $\sum_{n=1}^{\infty} b_n$, ואם הם מתכנסים הם שווים.

על מנת שסדרת הסכומים הסופיים של טור תתכנס, התוספת בכל שלב חייבת להיות קטנה יותר ויותר, וסביר לצפות שסדרת איברי הטור שואפת ל־0. ואמנם:

משפט 6.1.2 אם הטור $\sum_{n=1}^{\infty} a_n$ מתכנס למספר ממשי אז $a_n \to 0$.

הוכחה נניח שהטור מתכנס ל־S, כלומר שסדרת הסכומים החלקיים $(S_N)_{N=1}^{\infty}$ שלו מתכנסת ל־S. מכך נובע בעזרת אריתמטיקה של גבולות

$$\lim_{N\to\infty} (S_N - S_{N-1}) = S - S = 0$$

אבל $S_N - S_{N-1} = a_N$, וקיבלנו $a_n \to 0$. ■

התנאי $a_n \to 0$ הכרחי להתכנסות של הטור $\sum a_n$ אך אינו מספיק. הדוגמה הנגדית הפשוטה ביותר היא **הטור ההרמוני** $\sum_{n=1}^{\infty} \frac{1}{n}$: אמנם מתקיים $\frac{1}{n} \to 0$, אבל הטור $\sum_{n=1}^{\infty} \frac{1}{n}$ מתבדר, אם כי דרוש תחכום מסוים כדי להוכיח זאת. נשים לב שלכל $k \in \mathbb{N}$, הסכום של k האיברים בסדרה ההרמונית החל מהאיבר ה־k ועד האיבר ה־$2k-1$ מקיים

$$\frac{1}{k} + \frac{1}{k+1} + \ldots + \frac{1}{2k-1} \geq \frac{1}{2k-1} + \ldots + \frac{1}{2k-1} \geq \frac{k}{2k-1} \geq \frac{1}{2}$$

אפשר לפרק את הטור ההרמוני לרצפים מסוג זה:

$$\underbrace{1}_{>\frac{1}{2}} + \underbrace{\frac{1}{2} + \frac{1}{3}}_{>\frac{1}{2}} + \underbrace{\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}}_{>\frac{1}{2}} + \underbrace{\frac{1}{8} + \frac{1}{9} + \ldots + \frac{1}{15}}_{>\frac{1}{2}} + \ldots$$

ולכן הטור אינו יכול להתכנס, ואפילו שואף לאינסוף. באופן מדויק, יהי T_k הסכום של איברי הסדרה ההרמונית מהאיבר ה־2^k עד האיבר $2\cdot 2^k-1$:

$$T_k = \frac{1}{2^k}+\frac{1}{2^k+1}+\ldots+\frac{1}{2\cdot 2^k-1} = \sum_{n=2^k}^{2^{k+1}-1}\frac{1}{n}$$

לפי מה שראינו $T_k \geq \frac{1}{2}$ ולכן לכל n מתקיים

$$S_{2^n-1} = T_0+T_1+\ldots+T_{n-1} \geq \frac{1}{2}+\frac{1}{2}+\ldots+\frac{1}{2} = \frac{n}{2} \to \infty$$

לכן ל־(S_N) יש תת־סדרה המתכנסת ל־∞ וממילא (S_N) אינה מתכנסת, כלומר $\sum_{n=1}^{\infty}\frac{1}{n}$ אינו מתכנס. לא קשה לראות ש־ $S_N \to \infty$.

פעם נוספת אנו רואים שהאינטואיציה שלנו עלולה להטעותנו. עולם הטורים מלא בדוגמאות כאלה. בסעיפים הבאים ננסה לעשות מעט סדר בדברים.

תרגילים

1. אילו מהטורים הבאים מתכנסים (במובן הצר או במובן הרחב):
   (א) $\sum_{n=1}^{\infty}\sqrt{n}$.
   (ב) $\sum_{n=1}^{\infty}(-1)^n[\sqrt{\frac{n}{2}}]$.
   (ג) $\sum_{n=2}^{\infty}a_n$ כאשר $a_n = \frac{(-1)^n}{[n/2]}$ ($[x]$ מציין את הערך השלם של x).
   (ד) $\sum_{n=1}^{\infty}\frac{2n+1}{n^2(n+1)^2}$ (היעזרו בחישוב של $\sum_{n=1}^{\infty}\frac{1}{n(n+1)}$ בדוגמה 5 בעמוד 166).
   (ה) $\sum_{n=1}^{\infty}a_n$ כאשר a_n מוגדרת על ידי $a_n = (\frac{1}{2})^k$ עבור $2^k \leq n < 2^{k+1}$.
   (ו) $\sum_{n=1}^{\infty}a_n$ כאשר a_n מוגדרת על ידי $a_n = (\frac{1}{3})^k$ עבור $2^k \leq n < 2^{k+1}$.
2. יהי $\sum_{n=1}^{\infty}a_n$ טור. נגדיר סדרה b_n על ידי $b_{2n-1}=a_n$ ו־ $b_{2n}=0$, כלומר

   $$(b_n) = a_1, 0, a_2, 0, a_3, 0, \ldots$$

   הוכיחו ש־ $\sum_{n=1}^{\infty}b_n$ מתכנס אמ"מ $\sum_{n=1}^{\infty}a_n$ מתכנס, ואם הם מתכנסים אז הגבול זהה.

   באופן כללי יותר, יהי $\sum_{n=1}^{\infty}a_n$ טור מתכנס. תהי (n_k) סדרה עולה ממש של מספרים טבעיים ונגדיר $c_{n_k}=a_k$ ו־ $c_n=0$ לכל $n \notin \{n_k\ :\ k\in\mathbb{N}\}$, כלומר (c_n) היא הסדרה שבה a_k מופיע במקום ה־ n_k ושאר איבריה אפסים. הוכיחו ש־ $\sum_{n=1}^{\infty}c_n$ מתכנס במובן הרחב אמ"מ $\sum_{n=1}^{\infty}c_n$ מתכנס במובן הרחב, ומתקיים $\sum_{n=1}^{\infty}c_n = \sum_{n=1}^{\infty}a_n$.
3. הראו ש־ $\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}=\infty$ על ידי הוכחת האי־שוויון $S_N \geq \sqrt{N}$ כאשר S_N סדרת הסכומים החלקיים.

4. הוכיחו ש־ $\sum_{n=1}^{\infty}\frac{1}{n}=\infty$.
5. הפילוסוף היווני זנון[1] הציע את הפרדוקס הבא, שידוע בתור **הפרדוקס של אכילס והצב**. אכילס, כידוע, היה אצן בעל שיעור קומה. יום אחד הצב, שהיה אטי מאכילס בהרבה, הציע לערוך תחרות ריצה ביניהם, בתנאי שיינתן לו (לצב) יתרון התחלתי קטן. אכילס כמובן הסכים. אהה! קרא הצב, לעולם לא תוכל לנצח בתחרות! והנה הסיבה: מסלול הריצה ישר, ובזמן t_0 אכילס נמצא בנקודה x_0 והצב ב־x_1 ומתקיים $x_1>x_0$. בזמן t_0 התחרות מתחילה ושניהם מתחילים לרוץ. לפני שאכילס יתפוס את הצב עליו לעבור דרך הנקודה x_0, נניח בזמן t_1. אבל בזמן t_1 הצב כבר נמצא בנקודה $x_2>x_1$. כעת לפני שיוכל לעקוף את הצב חייב אכילס לעבור בנקודה x_2, נניח בזמן t_2, אבל אז הצב כבר יהיה בנקודה $x_3>x_2$, וכן הלאה. אכילס לכן לעולם לא ישיג את הצב.

   מה הפגם בטיעון הזה, וכיצד הוא קשור לטורים? תנו הסבר מפורט. הניחו שהמהירות של אכילס היא v ושל הצב היא w (אנו מניחים שהמהירויות קבועות, וכמובן $v>w>0$). כדאי לחשב את הזמנים t_n והנקודות x_n.
6. אדם יוצא מנקודה X ונע במהירות קבועה, וכעבור שעה מגיע לנקודה Y. ברגע הגעתו האדם מחליף כיוון ומכפיל את מהירותו, כך שהוא מגיע חזרה ל־X כעבור חצי שעה; אז הוא שוב מחליף כיוון ומכפיל מהירות, וכן הלאה. איזה מרחק יעבור האיש בתום שעתיים? תוכלו גם להשתעשע בשאלה, היכן נמצא האדם כעבור שעתיים?

6.2 תנאי קושי ותכונות בסיסיות

ניתן להפיק מידע על טורים בעזרת הכלים שפיתחנו בפרק על סדרות. אין זה מפתיע, שכן התכנסות של טור הוגדרה במונחים של התכנסות סדרות.

נתחיל בתנאי הכרחי ומספיק להתכנסות של טור, שהוא תוצאה מידית של קריטריון קושי להתכנסות סדרות:

משפט 6.2.1 (תנאי קושי להתכנסות טורים) הטור $\sum_{n=1}^{\infty}a_n$ מתכנס אמ"מ לכל $\varepsilon>0$ קיים N כך שלכל $n>N$ ולכל k מתקיים $|a_{n+1}+a_{n+2}+\ldots+a_{n+k}|<\varepsilon$.

הוכחה לפי ההגדרה התכנסות הטור שקולה להתכנסות סדרת הסכומים החלקיים $S_N=a_1+a_2+\ldots+a_N$, ולכן די להראות שהתנאי האמור שקול להתכנסות הסדרה (S_N). נשים לב שלכל n,k מתקיים

$$|S_{n+k}-S_n|=|(a_1+\ldots+a_{n+k})-(a_1+\ldots+a_n)|=|a_{n+1}+\ldots+a_{n+k}|$$

[1] Zeno of Elea, ‏490־425 לסנה"ס.

כעת, מתנאי קושי להתכנסות סדרות, אנו יודעים ש־(S_N) סדרה מתכנסת אמ"מ לכל ε קיים N כך שלכל $n > N$ ולכל k טבעי מתקיים $|S_n - S_{n+k}| < \varepsilon$, כלומר לכל $n > N$ ולכל $k \in \mathbb{N}$ מתקיים $|a_{n+1} + \ldots + a_{n+k}| < \varepsilon$, כפי שרצינו להוכיח. ■

לפעמים מנסחים את תנאי קושי גם כך: לכל $\varepsilon > 0$ קיים N כך שלכל $n \geq m > N$ מתקיים $|a_m + \ldots + a_n| < \varepsilon$. שקילות שני הניסוחים מידית.

דוגמאות

1. נחזור על ההוכחה שהטור ההרמוני $\sum_{n=1}^{\infty} \frac{1}{n}$ מתבדר, הפעם תוך שימוש בתנאי קושי. נניח בשלילה שהטור מתכנס, ולכן מקיים את תנאי קושי. בפרט עבור $\varepsilon = \frac{1}{2}$ קיים N עבורו לכל $n > N$ ולכל k מתקיים $\frac{1}{n+1} + \ldots + \frac{1}{n+k} < \frac{1}{2}$ (אין צורך בערך המוחלט כי הסכום חיובי). בפרט עבור $k = n = N + 1$ מתקיים

$$\frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n} < \frac{1}{2}$$

אבל בסכום מופיעים n מחוברים שכולם גדולים או שווים ל־$\frac{1}{2n}$, ולכן הסכום הוא לפחות $n \cdot \frac{1}{2n} = \frac{1}{2}$. זו סתירה, ונובע שהטור מתבדר.

2. נוכיח שהטור $\sum_{n=1}^{\infty} \frac{1}{n^2}$ מתכנס, על ידי שנוכיח שמתקיים תנאי קושי. תהי (S_n) סדרת הסכומים החלקיים שלה. לכל n, k מתקיים

$$\begin{aligned}
|S_{n+k} - S_n| &= |\frac{1}{(n+1)^2} + \frac{1}{(n+2)^2} + \ldots + \frac{1}{(n+k)^2}| \\
&= \frac{1}{(n+1)^2} + \frac{1}{(n+2)^2} + \ldots + \frac{1}{(n+k)^2} \\
&\leq \frac{1}{n(n+1)} + \frac{1}{(n+1)(n+2)} + \ldots + \frac{1}{(n+k-1)(n+k)} \\
&= (\frac{1}{n} - \frac{1}{n+1}) + \\
&\qquad + (\frac{1}{n+1} - \frac{1}{n+2}) + \ldots + (\frac{1}{n+k-1} - \frac{1}{n+k}) \\
&= \frac{1}{n} - \frac{1}{n+k} \\
&< \frac{1}{n}
\end{aligned}$$

לכן לכל N אם $n > N$ אז $|\frac{1}{(n+1)^2} + \ldots + \frac{1}{(n+k)^2}| < \frac{1}{n} < \frac{1}{N}$. בהינתן $\varepsilon > 0$, אם נבחר N כך ש־ $\frac{1}{N} < \varepsilon$ אז לכל $n > N$ ולכל k טבעי יתקיים $|\frac{1}{(n+1)^2} + \ldots + \frac{1}{(n+k)^2}| < \varepsilon$, ותנאי קושי מתקיים.

בתורת הסדרות ראינו שהתכנסות של סדרה היא תכונה אסימפטוטית, כלומר היא אינה תלויה בהתנהגות האיברים הראשונים בסדרה. ננסה לתרגם עובדה זו לתוצאה

על טורים. השמטת m האיברים הראשונים של טור מגדירה טור חדש שבו הסכימה מתחילה מהמקום ה־$m+1$. באופן מדויק,

הגדרה 6.2.2 יהי נתון טור $\sum_{n=1}^{\infty} a_n$ (לאו דווקא מתכנס). ויהי $m \in \mathbb{N}$. הטור $\sum_{n=m+1}^{\infty} a_n$ נקרא **ה־m־זנב** של $\sum_{n=1}^{\infty} a_n$ (זהו הטור המקורי לאחר שקוצצו m איבריו הראשונים). אם ה־m־זנב מתכנס, נסמן את סכומו ב־r_m.

משפט 6.2.3 יהי $\sum_{n=1}^{\infty} a_n$ טור ותהי (S_N) סדרת הסכומים החלקיים שלו.

1. $\sum_{n=1}^{\infty} a_n$ מתכנס אמ"מ כל זנב שלו מתכנס, ואז אם $\sum_{n=1}^{\infty} a_n = S$ אז $r_m = S - S_m$
2. $\sum_{n=1}^{\infty} a_n$ מתכנס ל־∞ אמ"מ כל זנב שלו מתכנס ל־∞

הוכחה נוכיח את המקרה שהטור מתכנס במובן הצר. המקרה השני מושאר כתרגיל.

יהי $m \in \mathbb{N}$ קבוע. תהי $(\widetilde{S}_k)$ סדרת הסכומים החלקיים של ה־m־זנב של $\sum_{n=1}^{\infty} a_n$, כלומר

$$\widetilde{S}_k = a_{m+1} + a_{m+2} + \ldots + a_{m+k}$$

נשים לב שלכל k מתקיים $\widetilde{S}_k = S_{m+k} - S_m$. מכיוון שהמספר S_m בצד ימין קבוע (לא תלוי ב־k) אנו מקבלים לפי אריתמטיקה של גבולות ש־ $(\widetilde{S}_k)_{k=1}^{\infty}$ מתכנס אמ"מ $(S_{m+k})_{k=1}^{\infty}$ מתכנס, כלומר אמ"מ (S_k) מתכנס. יתרה מזאת, קיבלנו שאם הגבול קיים אז

$$r_m = \lim_{k\to\infty} \widetilde{S}_k = \lim_{k\to\infty} (S_{m+k} - S_m) = S - S_m$$

כאשר $S = \sum_{n=1}^{\infty} a_n$. ■

מסקנה 6.2.4 אם $\sum_{n=1}^{\infty} a_n$ ו־ $\sum_{n=1}^{\infty} b_n$ הם טורים הנבדלים במספר סופי של איברים אז $\sum a_n$ מתכנס אמ"מ $\sum b_n$ מתכנס.

הוכחה מההנחה, לטורים יש זנבות משותפים, ולכן זו מסקנה מידית של המשפט. ■

המסקנה נכונה גם אם מחליפים את הנחת ההתכנסות בהתכנסות במובן הרחב.

מסקנה 6.2.5 אם הטור $\sum_{n=1}^{\infty} a_n$ מתכנס ו־r_m הוא ה־m־זנב של הטור אז $r_m \to 0$.

הוכחה יהי $S = \sum_{n=1}^{\infty} a_n$. ראינו ש־ $r_m = S - S_m$. אבל $S_m \to S$, ולכן מאריתמטיקה של גבולות $r_m = S - S_m \to S - S = 0$. ■

גם משפט האריתמטיקה הבא נובע מהמשפט המקביל לסדרות:

משפט 6.2.6 (כללי תחשיב לטורים) אם הטורים $\sum_{n=1}^{\infty} a_n$ ו־ $\sum_{n=1}^{\infty} b_n$ מתכנסים ל־S ול־T בהתאמה, אז הטור $\sum_{n=1}^{\infty} (a_n + b_n)$ מתכנס וסכומו הוא $S + T$. אם $c \in \mathbb{R}$ ואם $\sum_{n=1}^{\infty} a_n$ מתכנס אז $\sum_{n=1}^{\infty} c \cdot a_n$ מתכנס וסכומו $c \cdot S$.

הוכחה נוכיח את החלק הראשון. נסמן ב־ (S_N) את סדרת הסכומים החלקיים של $\sum_{n=1}^{\infty} a_n$, ב־ (T_N) את סדרת הסכומים החלקיים של $\sum_{n=1}^{\infty} b_n$ וב־ (R_N) את סדרת הסכומים החלקיים של $\sum_{n=1}^{\infty}(a_n + b_n)$. מכללי הסכימה של סכומים סופיים מתקיים

$$\begin{aligned} R_N &= (a_1 + b_1) + \ldots + (a_N + b_N) \\ &= (a_1 + \ldots + a_N) + (b_1 + \ldots + b_N) \\ &= S_N + T_N \end{aligned}$$

ההנחה שלנו היא ש־ $S_N \to S$ וש־ $T_N \to T$, ולכן מאריתמטיקה של גבולות, נקבל ש־ $R_N \to S + T$, ולכן $\sum_{n=1}^{\infty}(a_n + b_n) = S + T$, כפי שרצינו.

הוכחת הטענה על כפל הטור בקבוע מושארת כתרגיל. ■

כמו במקרה של סדרות, הכיוון ההפוך אינו נכון: אם $\sum(a_n + b_n)$ מתכנס לא נובע בהכרח ש־ $\sum a_n, \sum b_n$ מתכנסים. תנו דוגמה!

אפשר לנסח משפטי אריתמטיקה גם לטורים שערכם $\pm\infty$. אלה נובעים בקלות ממשפטי האריתמטיקה לגבולות במובן הרחב שפגשנו בסעיף 5.5

מה בנוגע לכלל המכפלה? באופן כללי אין לערך של הטור $\sum a_n b_n$ קשר פשוט לטורים $\sum a_n, \sum b_n$ (ואין סיבה לצפות שיהיה, כי גם במקרה של סכומים סופיים אין קשר פשוט). נשוב לשאלות אלה בסעיף 6.6 להלן.

תרגילים

1. נכון או לא נכון:
   (א) אם קיים N כך שלכל $n > N$ מתקיים $|S_{n+1000} - S_n| < \varepsilon$ אז $\sum a_n$ מתכנס.
   (ב) אם $\sum a_n$ מתכנס אז לכל $\varepsilon > 0$ קיים N כך שלכל $n > N$ ולכל $k = 1, 2, \ldots, n$ מתקיים $|S_n - S_{n+k}| < \varepsilon$.
   (ג) (*) אם לכל $\varepsilon > 0$ קיים N כך שלכל $n > N$ ולכל $k = 1, 2, \ldots, n$ מתקיים $|S_n - S_{n+k}| < \varepsilon$ אז $\sum a_n$ מתכנס.
2. תנו דוגמה לטורים $\sum a_n$ ו־$\sum b_n$ שאינם מתכנסים אבל כך הטור $\sum(a_n + b_n)$ מתכנס.
3. יהיו $\sum_{n=1}^{\infty} a_n$ ו־$\sum_{n=1}^{\infty} b_n$ טורים מתכנסים. נגדיר סדרה (c_n) ע"י שילוב הסדרות: $(c_n) = a_1, b_1, a_2, b_2, \ldots$. הוכיחו ש־ $\sum_{n=1}^{\infty} c_n$ מתכנס ומתקיים $\sum_{n=1}^{\infty} c_n = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n$.
4. הראו שלכל טור מתכנס $\sum a_n$ יש טור מתכנס $\sum b_n$ כך ש־ $b_n > a_n$ לכל n.
5. הוכיחו את סעיף (2) של משפט 6.2.3.

6. תנו הוכחה נוספת למשפט 6.1.2 בעזרת קריטריון קושי (כתבו את התנאי למקרה $k=1$).
7. תנו הוכחה נוספת למסקנה 6.2.5 בעזרת קריטריון קושי.
8. הוכיחו את הסעיף על כפל בקבוע במשפט 6.2.6. נסחו והוכיחו גרסה של אותו משפט כאשר אחד או שניים מהטורים מתכנסים ל־$\pm\infty$.
9. הוכיחו את משפט הסנדוויץ' הבא: יהיו $\sum a_n, \sum b_n, \sum c_n$ טורים כך ש־ $a_n \leq b_n \leq c_n$ לכל n מספיק גדול. אז אם הטורים $\sum a_n, \sum c_n$ מתכנסים גם $\sum b_n$ מתכנס (רמז: תנאי קושי).

6.3 טורים חיוביים

הגדרה 6.3.1 טור $\sum_{n=1}^{\infty} a_n$ נקרא **טור חיובי** אם $a_n \geq 0$ לכל n.

שימו לב שמדובר בטורים שאיבריהם אי־שליליים. למרות זאת נשתמש במונח המקובל (והקצר יותר) של טור חיובי. טור שאיבריו חיוביים ממש נקרא **טור חיובי ממש**. באופן דומה מגדירים טור שלילי וטור שלילי ממש, וכל האמור בסעיף זה על טורים חיוביים נכון לאחר שינויים מתאימים בניסוח גם לטורים שליליים.

הסיבה שטורים חיוביים נוחים לניתוח היא שעבור טור חיובי $\sum a_n$ סדרת הסכומים החלקיים (S_N) היא סדרה מונוטונית עולה, כי

$$S_{N+1} = S_N + a_{N+1} \geq S_N$$

ולכן לפי משפט 5.6.2, (S_N) מתכנסת במובן הרחב לגבול סופי או לאינסוף. אנו מקבלים את המשפט הבא:

משפט 6.3.2 אם $\sum_{n=1}^{\infty} a_n$ הוא טור חיובי הוא מתכנס במובן הרחב. אם סדרת הסכומים החלקיים חסומה מלעיל אז הטור מתכנס למספר אי־שלילי, אחרת (אם סדרת הסכומים החלקיים אינה חסומה מלעיל) אז $\sum_{n=1}^{\infty} a_n = \infty$.

לאור שתי האפשרויות במשפט, עבור טור חיובי מסמנים לפעמים $\sum_{n=1}^{\infty} a_n < \infty$ כדי לציין שהטור מתכנס למספר ממשי.

המשפט הבא הוא הכלי החשוב ביותר שיש לנו לבדיקת ההתכנסות של טור:

משפט 6.3.3 (מבחן ההשוואה לטורים חיוביים) יהיו $\sum_{n=1}^{\infty} a_n$ ו־ $\sum_{n=1}^{\infty} b_n$ טורים חיוביים, ונניח ש־ $a_n \leq b_n$ לכל n. אז

1. אם הטור $\sum_{n=1}^{\infty} b_n$ מתכנס, אז גם הטור $\sum_{n=1}^{\infty} a_n$ מתכנס, והם מקיימים את האי־שוויון $\sum_{n=1}^{\infty} a_n \leq \sum_{n=1}^{\infty} b_n$.

2. אם הטור $\sum_{n=1}^{\infty} a_n$ מתבדר (כלומר מתכנס ל־∞), אז גם $\sum_{n=1}^{\infty} b_n = \infty$.

הוכחה ראשית נוכיח את סעיף (2). נסמן את סדרת הסכומים החלקיים של $\sum_{n=1}^{\infty} a_n$ ב־ (S_N), ואת סדרת הסכומים החלקיים של $\sum_{n=1}^{\infty} b_n$ ב־ (T_N). לפי ההנחות שלנו מתקיים $S_N \leq T_N$. שתי סדרות אלה עולות ולכן מתכנסות במובן הרחב, והגבול שלהן סופי אמ"מ הן חסומות מלעיל. אם $S_N \to \infty$ אז (S_N) אינה חסומה מלעיל ומכיוון ש־ $S_N \leq T_N$ לכל N הרי שגם (T_N) אינה חסומה מלעיל ומתקיים $T_N \to \infty$.

לגבי (1), אם (T_N) מתכנסת לגבול סופי T אז היא חסומה מלעיל על ידי T (למעשה $T = \sup T_N$ לפי משפט 5.6.2) ולכן גם (S_N) חסומה מלעיל על ידי T, כי $S_N \leq T_N \leq T$ לכל N. מכאן נובע ש־(S_N) מתכנסת לגבול סופי ומתקיים $\lim S_N \leq T = \lim T_N$, כנדרש. ■

במצב המתואר במשפט אומרים ש־ $\sum_{n=1}^{\infty} b_n$ **שולט** (dominates או majorizes) על $\sum_{n=1}^{\infty} a_n$, וש־ $\sum_{n=1}^{\infty} a_n$ **נשלט** על ידי $\sum_{n=1}^{\infty} b_n$. נציין שבתור מבחן התכנסות המשפט נכון גם אם מניחים שהאי־שוויון בין איברי הטורים מתקיים רק החל ממקום מסוים, אם כי אז אי אפשר להסיק את האי־שוויון בין סכומי הטורים.

דוגמאות

1. לכל $\alpha \geq 2$ הטור $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ מתכנס, כי הוא טור חיובי ונשלט על ידי הטור החיובי $\sum \frac{1}{n^2}$, שהוא טור מתכנס (דוגמה (2) מעמוד 169).
2. לכל $0 < \alpha \leq 1$ הטור $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ מתבדר, כי הוא שולט על הטור ההרמוני $\sum \frac{1}{n}$, שהוא טור מתבדר (דוגמה (1) מעמוד 169).
3. נתבונן בטור $\sum_{n=1}^{\infty} \frac{1+(-1)^n/n^2}{n}$. נשים לב ש־ $1+\frac{(-1)^n}{n^2} > \frac{1}{2}$ לכל $n > 1$, ולכן הטור $\sum_{n=1}^{\infty} \frac{1}{2n}$ נשלט על ידי הטור המקורי. כיוון שהטור $\sum \frac{1}{2n}$ מתבדר, גם הטור המקורי מתבדר.

בדוגמאות יישבנו את שאלת ההתכנסות של טורים מהצורה $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ עבור $\alpha \geq 2$ ועבור $\alpha \leq 1$. בהמשך נראה שהטור מתכנס גם כאשר $1 < \alpha < 2$, אך ההוכחה תדרוש כלים מעט יותר מתוחכמים (ראו משפט 6.3.9 להלן והדוגמה שאחריו).

המשפט הבא נובע בקלות ממבחן ההשוואה, ומושאר כתרגיל:

משפט 6.3.4 יהיו $\sum_{n=1}^{\infty} a_n$ ו־ $\sum_{n=1}^{\infty} b_n$ טורים חיוביים. נניח שקיים מספר חיובי u כך ש־ $\frac{a_n}{b_n} \leq u$. אז אם $\sum_{n=1}^{\infty} b_n$ מתכנס גם $\sum_{n=1}^{\infty} a_n$ מתכנס. בפרט אם קיימים מספרים $u, v > 0$ כך ש־ $u \leq \frac{a_n}{b_n} \leq v$ החל ממקום מסוים (למשל, אם הגבול $\lim \frac{a_n}{b_n}$ קיים וחיובי) אז הטורים מתכנסים ומתבדרים יחד, כלומר האחד מתכנס אמ"מ השני מתכנס.

למשל, הטור $\sum_{n=1}^{\infty}(n^2-n)/(n^4-10)$ מתכנס כי אם נסמן $a_n = (n^2-n)/(n^4-10)$ ו־ $b_n = 1/n^2$ אז $a_n/b_n \to 1$ ו־ $\sum b_n$ מתכנס. זו דרך מדויקת לבטא את העובדה של־$(n^2-n)/(n^4-10)$ ול־$1/n^2$ יש קצב דומה של שאיפה לאפס.

בשני מבחני ההשוואה האחרונים הסקנו התכנסות או התבדרות של טור אחד על ידי השוואתו עם טור אחר. במשפטים הבאים נראה תנאים המבטיחים התכנסות על ידי התבוננות בתכונות של איברי הטור בלבד. עם זאת, על אף שלכאורה אלה אינם מבחני השוואה, ההוכחה מתבססת בכל־זאת על השוואה של הטור עם טור גאומטרי.

משפט 6.3.5 (מבחן השורש של קושי) יהי $\sum_{n=1}^{\infty} a_n$ טור חיובי.

1. אם קיים $0 < q < 1$ כך ש־ $\sqrt[n]{a_n} \leq q$ החל ממקום מסוים אז הטור מתכנס.
2. אם $\sqrt[n]{a_n} \geq 1$ עבור אינסוף n־ים אז הטור מתבדר.

הוכחה כאמור, נראה שאפשר להשוות את הטור לטור גאומטרי. במקרה הראשון, מכיוון שהתכנסות או אי־התכנסות היא תכונה של הזנב של הטור, אפשר להניח ש־ $\sqrt[n]{a_n} \leq q$ לכל n, כלומר ש־ $a_n \leq q^n$ לכל n. כיוון שהטור $\sum_{n=1}^{\infty} q^n$ מתכנס, נובע ממבחן ההשוואה ש־ $\sum_{n=1}^{\infty} a_n$ מתכנס.

במקרה השני, אם לאינסוף n־ים מתקיים $\sqrt[n]{a_n} \geq 1$ אז גם $a_n \geq 1$ ובפרט $a_n \not\to 0$, ולכן הטור מתבדר. ■

המסקנה הבאה מושארת כתרגיל:

מסקנה 6.3.6 יהי $\sum_{n=1}^{\infty} a_n$ טור חיובי. אם $\limsup \sqrt[n]{a_n} < 1$, אז הטור מתכנס, ואם $\liminf \sqrt[n]{a_n} > 1$, אז הטור מתבדר. בפרט, אם $q = \lim \sqrt[n]{a_n}$ קיים אז אם $q < 1$ הטור מתכנס ואם $q > 1$ הטור מתבדר.

דוגמאות

1. נתבונן בטור $\sum_{n=1}^{\infty} \frac{3^n}{2^{n^2}}$. נחשב:

$$\sqrt[n]{\frac{3^n}{2^{n^2}}} = \frac{\sqrt[n]{3^n}}{\sqrt[n]{2^{n^2}}} = \frac{3}{2^{n^2/n}} = \frac{3}{2^n} \to 0$$

   לכן לפי המסקנה הטור מתכנס.

2. באופן כללי מסקנה 6.3.6 חלשה יותר ממבחן השורש. למשל, מהמשפט נובע שהטור $\sum_{n=1}^{\infty} 1$ מתבדר, בעוד שמהמסקנה אי אפשר להסיק אפילו שהטור $\sum_{n=1}^{\infty} n$ מתבדר, כי $\lim \sqrt[n]{1} = \lim \sqrt[n]{n} = 1$.

נציין שמבחן השורש (וכמובן המסקנה שלו) אינו מאפיין התכנסות של טורים חיוביים. כפי שהזכרנו, הוא מחביא בתוכו השוואה עם טור גאומטרי, ולא יוכל לקבוע התכנסות או התבדרות של טור שאינה נשלט על ידי טור גאומטרי. למשל, הטורים $\sum_{n=1}^{\infty} 1/n$ ו־ $\sum_{n=1}^{\infty} 1/n^2$ מתבדרים ומתכנסים בהתאמה, אך אי אפשר להסיק זאת ממבחן השורש. אגב, בשני הדוגמאות האלו $\sqrt[n]{a_n} \to 1$, ומכאן שבמסקנה 6.3.6 אי אפשר להחליף את האי־שוויון החזק באי־שוויון חלש.

משפט 6.3.7 (מבחן המנה של דלאמבר)[2] יהי $\sum_{n=1}^{\infty} a_n$ טור חיובי ממש. אם קיים $0 < q < 1$ כך ש־ $\frac{a_{n+1}}{a_n} \leq q$ החל ממקום מסוים אז הטור מתכנס. מאידך, אם קיים $\frac{a_{n+1}}{a_n} \geq 1$ החל ממקום מסוים אז הטור מתבדר.

הוכחה שוב נערוך השוואה לטור גאומטרי. כרגיל, נוכל להניח שהתנאים מתקיימים לכל n. בחלק הראשון, התנאי $\frac{a_{n+1}}{a_n} \leq q < 1$ מבטיח שאיברי הטור הולכים וקטנים ביחס קבוע, ומתקיים

$$a_2 \leq qa_1 \quad , \quad a_3 \leq qa_2 \leq q^2 a_1 \quad , \quad \ldots$$

והוכחה קלה באינדוקציה מראה ש־ $0 \leq a_n \leq q^{n-1} a_1$. הטור $\sum_{n=1}^{\infty} q^{n-1}$ מתכנס כי $0 < q < 1$ ולכן גם $\sum_{n=1}^{\infty} q^{n-1} a_1$ מתכנס, וממבחן ההשוואה מסיקים ש־ $\sum_{n=1}^{\infty} a_n$ מתכנס.

לגבי החלק השני, מההנחה $a_{n+1} \geq a_n$ לכל n נובע ש־ $a_n \geq a_1 > 0$, ולכן $a_n \not\to 0$ והטור מתבדר. ■

הוכחת המסקנה הבאה מושארת כתרגיל:

מסקנה 6.3.8 יהי $\sum_{n=1}^{\infty} a_n$ טור חיובי ממש. אם $\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} < 1$ אז הטור מתכנס, ואם $\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} > 1$ אז הטור מתבדר. בפרט אם קיים הגבול $q = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}$ אז אם $q < 1$ הטור מתכנס ואם $1 < q$ הטור מתבדר.

דוגמאות

1. נתבונן בטור $\sum_{n=1}^{\infty} \frac{2^n n!}{n^n}$. נחשב מנות של איברים עוקבים:

$$\begin{aligned}
\frac{2^{n+1}(n+1)!/(n+1)^{n+1}}{2^n n!/n^n} &= 2 \cdot \frac{(n+1)!}{n!} \cdot \frac{n^n}{(n+1)^n \cdot (n+1)} \\
&= 2 \cdot (n+1)(1+\frac{1}{n})^{-n} \cdot \frac{1}{n+1} \\
&= 2 \cdot (1+\frac{1}{n})^{-n} \\
&\to \frac{2}{e}
\end{aligned}$$

(כאן e הוא המספר שהוגדר בסעיף 5.11). מכיוון ש־ $e > 2$ הגבול הזה קטן מ־1 וממבחן המנה הטור מתכנס. שימו לב שקשה מאד להוכיח את התכנסות הטור ממבחן השורש (נסו!).

[2] Jean d'Alambert, 1717־1783.

2. נתבונן בטור $1+\frac{1}{2}+\frac{1}{2}+\frac{1}{4}+\frac{1}{4}+\dots$ הנתון על ידי $a_n=2^{-[n/2]}$. אז $\sum a_n$ מתכנס. קל להוכיח זאת ישירות ואף לחשב את גבולו, אך הנה דרך אחרת. נשים לב ש־ $\frac{[n/2]}{n}\to\frac{1}{2}$, ולכן ממבחן השורש

$$\sqrt[n]{a_n}=2^{-[n/2]/n}\to 2^{-1/2}<1$$

מצד שני לכל n מתקיים $\frac{a_{2n+1}}{a_{2n}}=1$, ולכן אי אפשר להסיק דבר על התכנסות הטור ממבחן המנה.

אם אפשר להכריע התכנסות של טור בעזרת מבחן המנה אז אפשר להכריעו בעזרת מבחן השורש. הסיבה היא שאם $\frac{a_{n+1}}{a_n}\le q<1$ אז $a_n\le q^{n-1}a_1$ ולכן

$$\sqrt[n]{a_n}\le q^{(n-1)/n}\cdot\sqrt[n]{a_1}=q\cdot\sqrt[n]{qa_1}\to q<1$$

כלומר גם מבחן השורש יגלה את דבר ההתכנסות. חישוב דומה חל על תנאי ההתבדרות של המשפטים. לכן לא מפתיע שגם מבחן המנה אינו תנאי הכרחי להתכנסות, ואינו מסוגל להכריע התכנסות של טורים מהסוג $\sum\frac{1}{n^2}$, $\sum\frac{1}{n}$.

עם זאת, כפי שראינו בדוגמה של הטור $\sum\frac{2^n n!}{n^n}$, לעתים קרובות קל יותר לחשב מנות מאשר שורשים ולכן מבחן המנה הוא בעל ערך מעשי רב.

ישנן כמה תוצאות מיוחדות לטורים שסדרת איבריהם מונוטונית. רוב הדוגמאות שפגשנו עד כה היו כאלה: למשל הטורים $\sum_{n=1}^\infty\frac{1}{n^\alpha}$ והטורים הגאומטריים $\sum_{n=1}^\infty q^n$.

נתחיל במבחן התכנסות על טורים מונוטוניים. יהי $\sum a_n$ טור חיובי שבה הסדרה (a_n) יורדת. אם $m<n$ נשים לב שמתקיים

$$(n-m)a_{n-1}\le\sum_{k=m}^{n-1}a_k\le(n-m)a_m$$

כאן השתמשנו במונוטוניות ובכך שהסכום מכיל $n-m$ מחוברים שהגדול מביניהם הוא a_m והקטן הוא a_{n-1}. בפרט עבור $m=2^{k-1},n=2^k-1$ מתקיים

$$2^{k-1}a_{2^k}\le(2^k-2^{k-1})a_{2^k-1}\le\sum_{i=2^{k-1}}^{2^k}a_i\le(2^k-2^{k-1})a_{2^{k-1}}=2^{k-1}a_{2^{k-1}}$$

ולכן אם (S_N) הם הסכומים החלקיים של הטור $\sum_{n=1}^\infty a_n$ אז

$$\begin{aligned}S_{2^k-1}&=a_1+(a_2+a_3)+\ldots+(a_{2^{k-1}}+\ldots+a_{2^k-1})\\&=\sum_{j=1}^k(\sum_{i=2^{j-1}}^{2^j-1}a_i)\end{aligned}$$

ואנו מסיקים ש־

$$\sum_{j=1}^k 2^{j-1}a_{2^j}\le S_{2^k-1}\le\sum_{j=1}^k 2^{j-1}a_{2^{j-1}}$$

או באופן שקול, אם נסמן $b_j = 2^j a_{2^j}$ קיבלנו

$$\frac{1}{2}\sum_{j=1}^{k} b_j \le S_{2^k-1} \le a_1 + \sum_{j=1}^{k-1} b_j$$

בשני האגפים הקיצוניים מופיעים כעת סכומים חלקיים של אותו טור חיובי $\sum b_j$. אנו מסיקים:

משפט 6.3.9 (מבחן העיבוי) יהי $\sum_{n=1}^{\infty} a_n$ טור חיובי ומונוטוני ונסמן $b_n = 2^n a_{2^n}$. אז $\sum a_n$ מתכנס אמ"מ $\sum b_n$ מתכנס.

הוכחה יהיו S_N הסכומים החלקיים של $\sum a_n$ ו־ (T_N) הסכומים החלקיים של $\sum b_n$. עלינו להראות ש־(S_N) מתכנסת אמ"מ (T_N) מתכנסת.

נניח ש־(T_N) חסומה. הדיון למעלה מראה ש־ $S_{2^n-1} \le a_1 + T_{n-1}$. מכאן שחסימות מלעיל של (T_N) גוררת חסימות מלעיל של (S_{2^n-1}) , ומכיוון ש־(S_N) היא סדרה עולה הדבר גורר חסימות של (S_N) והתכנסות של הטור $\sum a_n$ (כי זהו טור חיובי).

מאידך, אם (S_N) מתכנסת אז (S_N) חסומה מלעיל, ובפרט (S_{2^n-1}) חסומה מלעיל. הדיון למעלה נותן את האי־שוויון $\frac{1}{2}T_n \le S_{2^n-1}$, ומכאן ש־$(T_N)$ חסומה מלעיל ולכן $\sum b_n$ מתכנס. ∎

בעזרת מבחן העיבוי נוכל ליישב שאלה מציקה: עבור $1 < \alpha < 2$, האם $\sum_{n=1}^{\infty} \frac{1}{n^a}$ מתכנס או מתבדר? נסמן $a_n = \frac{1}{n^a}$. מתקיים

$$2^n a_{2^n} = \frac{2^n}{2^{\alpha n}} = 2^{-(\alpha-1)n}$$

ולכן $\sum_{n=1}^{\infty} 2^n a_{2^n} = \sum_{n=1}^{\infty} (2^{-\alpha+1})^n$ הוא טור מתכנס, כי זהו טור גאומטרי עם בסיס קטן מ־1 (הנחנו $\alpha > 1$). לכן $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ מתכנס. יחד עם התוצאות מעמוד 173 הוכחנו את הטענה הבאה:

טענה 6.3.10 הטור $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ מתכנס עבור $\alpha > 1$ ומתבדר עבור $\alpha \le 1$.

בהמשך נראה הוכחות נוספות לטענה זו.

ישנם מבחני התכנסות נוספים לטורים בעלי סדרת איברים מונוטונית. ראו למשל תרגיל (13) בסוף הסעיף.

נסיים בתוצאה נאה על קצב השאיפה לאפס של האיברים של טור מתכנס. ראינו שאם $\sum a_n$ הוא טור מתכנס (לאו דווקא חיובי), אז $a_n \to 0$. כמה אטית יכולה להיות השאיפה? ראינו שהטור ההרמוני $\sum_{n=1}^{\infty} \frac{1}{n}$ מתבדר. לכן ניתן לצפות שכל טור חיובי מתכנס ישאף לאפס מהר יותר מהסדרה ההרמונית. אם ננסה לנסח זאת

במדויק נקבל את ההשערה שאם $\sum a_n$ טור מתכנס אז $\frac{a_n}{1/n} = na_n \to 0$. אולם השערה זו אינה נכונה, כפי שמדגימה הסדרה

$$a_n = \begin{cases} 1/n & n = k^2 \\ 0 & \text{אחרת} \end{cases}$$

הטור $\sum a_n$ מתכנס (למה?) אבל $na_n = 1$ לכל ריבוע n, וממילא $na_n \not\to 0$.

עם זאת ההשערה נכונה אם מחזקים מעט את ההנחות על הטור. הנה הניסוח וההוכחה:

משפט 6.3.11 תהי $(a_n)_{n=1}^{\infty}$ סדרה חיובית ומונוטונית יורדת. אם הטור $\sum a_n$ מתכנס, אז $na_n \to 0$.

הוכחה די להראות

$$\lim_{n\to\infty} \frac{na_n}{2} = 0$$

נשים לב ש־

$$0 \le \frac{n}{2}a_n \le \sum_{\lfloor \frac{n}{2} \rfloor \le k \le n} a_n \le \sum_{\lfloor \frac{n}{2} \rfloor \le k \le n} a_k \le \sum_{k=\lfloor \frac{n}{2} \rfloor}^{\infty} a_k$$

(האי־שוויון האמצעי נובע מכך ש־(a_n) סדרה יורדת: הצדיקו את האי־שוויונות האחרים!). צד ימין הוא $[n/2]$־זנב של הטור, ומכיוון שהטור מתכנס צד ימין שואף לאפס כש־n שואף לאינסוף. כעת מכלל הסנדוויץ' נובע ש־ $na_n \to 0$. ■

אנו מקבלים מכך הוכחה נוספת לכך ש־$\sum_{n=1}^{\infty} n^{\alpha}$ מתבדר אם $-1 \le \alpha < 0$, שכן אז סדרת איברי הטור מונוטונית יורדת אבל $n \cdot n^{\alpha} = n^{1-\alpha} \ge 1$.

תרגילים

1. בדקו את התכנסות הטורים הבאים:
   (א) $\sum \frac{a^n}{n!}$ ($0 < a$ קבוע).
   (ב) $\sum \frac{1}{(n+1)^2-1}$.
   (ג) $\sum \frac{n}{n+2^n}$.
   (ד) $\sum \frac{1}{\sqrt{n(n+1)}}$.
   (ה) $\sum \frac{1}{n^{1+1/n}}$.
   (ו) $\sum \frac{1}{n \cdot \sqrt[n]{n}}$.
   (ז) $\sum \frac{a^n \cdot n!}{n^n}$ עבור $a \in \mathbb{R}$.

(ח) $\sum \frac{(n!)^n}{(2n)!}$.

(ט) $\sum n^{(2+(-1)^n)/2}$.

(י) $\sum(1+\frac{1}{n})^{n^2}$.

(יא) $\sum a_n$ כאשר $a_n = \frac{1}{n\cdot k^\alpha}$ עבור $2^k \le n < 2^{k+1}$ (α קבוע).

(יב) $\sum a_n$ כאשר $a_n = \frac{1}{k^n}$ עבור $2^k \le n < 2^{k+1}$.

2. הוכיחו את מסקנות 6.3.6 ו־6.3.8.

3. יהי $0 < q < 1$ ותהי $(a_n)_{n=1}^\infty$ הסדרה המוגדרת על ידי $a_n = q^{n-1}$ ל־n זוגי, ו־ $a_n = q^{n+1}$ ל־n אי־זוגי. הוכיחו בעזרת מסקנה 6.3.6 שהטור $\sum_{n=1}^\infty a_n$ מתכנס, והראו שממשפט 6.3.7 לא ניתן להסיק דבר על התכנסות הטור.

4. תהי (a_n) סדרה חיובית ונניח ש־ $\sqrt[n]{a_n} \le r < 1$. יהי $s = \sum_{n=1}^\infty a_n$ (הטור מתכנס ממשפט קושי). יהי $S_N = a_1+\ldots+a_N$. הוכיחו ש־ $|s-S_N| < \frac{r^{N+1}}{1-r}$.

5. יהיו $\sum_{n=1}^\infty a_n$ ו־ $\sum_{n=1}^\infty b_n$ טורים חיוביים, ונניח $\frac{a_{n+1}}{a_n} \le \frac{b_{n+1}}{b_n}$ החל ממקום מסוים. הוכיחו

(א) אם $\sum b_n$ מתכנס אז $\sum a_n$ מתכנס.

(ב) אם $\sum a_n$ מתבדר אז $\sum b_n$ מתבדר.

6. הוכיחו שטור חיובי $\sum a_n$ מתכנס אמ"מ $\sum \frac{a_n}{1+a_n}$ מתכנס.

7. התבוננו בסדרה (a_n) שהוגדרה לפני משפט 6.3.11 והשלימו את ההוכחה שהטור $\sum a_n$ מתבדר.

8. הוכיחו שעבור $\alpha \le 2$, הטור $\sum(\sqrt{n+1}-\sqrt{n})^\alpha$ מתבדר, ושעבור $\alpha \ge 4$ הוא מתכנס.

9. לאיזה $a > 0$ מתכנס הטור $\sum_{n=1}^\infty((\sum_{k=1}^n \frac{1}{k})a^n)$?

10. הוכיחו שאם $S = \sum_{n=1}^\infty \frac{1}{n^2}$, אז $S\frac{3}{4} = \ldots + \frac{1}{7^2} + \frac{1}{5^2} + \frac{1}{3^2} + 1$.

11. חשבו את סכומי הטורים $\sum_{n=1}^\infty \frac{1}{n(n+2)}$ ו־ $\sum_{n=1}^\infty \frac{1}{n(n+1)(n+2)}$.

12. הראו שאין טור חיובי המתבדר "הכי לאט", על ידי שתוכיחו שאם טור חיובי $\sum a_n$ מתבדר אז יש טור חיובי מתבדר $\sum b_n$ כך ש־ $\frac{b_n}{a_n} \to 0$.

13. הוכיחו את **מבחן ראבה**:[3] אם (a_n) טור חיובי, ואם הגבול

$$r = \lim_{n\to\infty} n(1-\frac{a_{n+1}}{a_n})$$

קיים ומקיים $r > 1$, אז הטור מתכנס. לעומת זאת אם המנה $n(1-\frac{a_{n+1}}{a_n}) \le 1$ החל ממקום מסוים אז הטור מתבדר (רמז: הוכיחו שתחת כל אחת מההנחות הסדרה $b_n = na_{n+1}$ מונוטונית החל ממקום מסוים. הגדירו $c_n = b_n - b_{n+1}$ והראו ש־ $\sum c_n$ מתכנס במובן הרחב. כעת היעזרו בטענה 6.3.4).

[3] Joseph Raabe, 1801־1859.

הראו שאם אפשר להכריע התכנסות של טור חיובי בעזרת מבחן השורש אז אפשר להכריע אותו בעזרת מבחן ראבה, ומצאו טור שניתן להכריע את התכנסותו בעזרת מבחן ראבה אך לא בעזרת מבחן השורש.

6.4 טורים עם סימנים משתנים

בסעיף זה נפתח תנאים מספיקים להתכנסות של טורים לא חיוביים מסוימים.

נתחיל בשיטה גסה יחסית אך שימושית. לפעמים ניתן ללמוד על טור שאינו חיובי על ידי התבוננות בטור המתקבל ממנו על ידי התעלמות מהסימן של איבריו, שהוא טור חיובי:

הגדרה 6.4.1 נאמר שהטור $\sum_{n=1}^{\infty} a_n$ הוא **מתכנס בהחלט** (converges absolutely) אם הטור $\sum_{n=1}^{\infty} |a_n|$ מתכנס. אם הטור מתכנס אבל לא מתכנס בהחלט, נאמר שהטור **מתכנס בתנאי** (converges conditionally).

משפט 6.4.2 אם הטור $\sum_{n=1}^{\infty} a_n$ מתכנס בהחלט אז הוא מתכנס.

הערה ברור שטור חיובי מתכנס אמ"מ הוא מתכנס בהחלט. לכן המשפט הוא בעל עניין רק לטורים שאינם חיוביים.

הוכחה נשתמש בתנאי קושי. יהי $\varepsilon > 0$. לפי הנתון ומשפט 6.2.1 קיים N עבורו לכל $n > N$ ולכל k מתקיים

$$||a_{n+1}| + |a_{n+2}| + \ldots + |a_{n+k}|| = |a_{n+1}| + |a_{n+2}| + \ldots + |a_{n+k}| < \varepsilon$$

כעת מאי־שוויון המשולש נובע שלכל $n > N$ ולכל k מתקיים

$$|a_{n+1} + a_{n+2} + \ldots + a_{n+k}| \leq |a_{n+1}| + |a_{n+2}| + \ldots + |a_{n+k}| < \varepsilon$$

ולכן גם הטור $\sum a_n$ מקיים את תנאי קושי, ולכן מתכנס. ■

מסקנה מעשית מהמשפט היא שבבואנו להכריע האם טור מסוים $\sum a_n$ מתכנס, כדאי תחילה לבדוק אם הטור מתכנס בהחלט, כי אז הוא ודאי מתכנס.

דוגמאות

1. ראינו שהטור $\sum_{n=1}^{\infty} \frac{1}{n^2}$ מתכנס. לכן הטור $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$ מתכנס, כי הוא מתכנס בהחלט.

2. הנה דוגמה מעט יותר מסובכת המבוססת על אותו עיקרון: נסמן ב־$\{x\}$ את החלק השבור של x. נתבונן בטור $\sum_{n=1}^{\infty}\frac{1-2\{\sqrt{n}\}}{n^2}$. האיבר ה־$n$ בטור הוא אי־שלילי אם $\{\sqrt{n}\}\leq\frac{1}{2}$ ואחרת הוא שלילי. קשה לתאר בדיוק את סימני האיברים או לחשב במפורש את הסכומים החלקיים, אבל הערך המוחלט של האיבר ה־n בטור קטן מ־$\frac{1}{n^2}$ ולכן הטור מתכנס בהחלט.

3. הכנסות והתכנסות בהחלט אינם מושגים שקולים, כלומר קיימים טורים שמתכנסים בתנאי. למשל, נתבונן בסדרה $(a_n)=1-1+\frac{1}{2}-\frac{1}{2}+\frac{1}{3}-\frac{1}{3}+\ldots$ ובטור $\sum_{n=1}^{\infty}a_n$. אז

$$a_{2n}+a_{2n+1}=\frac{1}{n}-\frac{1}{n}=0$$

ומכאן שאם (S_N) היא סדרת הסכומים החלקיים אז

$$S_{2N+1}=(a_2+a_3)+(a_4+a_5)+\ldots+(a_{2N}+a_{2N+1})=0+0+\ldots+0=0$$

ואילו־

$$S_{2N}=S_{2N-1}+a_{2N}=0+a_{2N}=a_{2N}$$

כיוון ש־ $a_n\to 0$ נובע ש־ $S_N\to 0$ ולכן הטור מתכנס (לאפס). מאידך, $|a_n|=\frac{1}{[n/2]}\geq\frac{2}{n}$ ולכן לפי מבחן ההשוואה הטור $\sum|a_n|$ מתבדר.

מחלקת הטורים המתכנסים בהחלט סגורה תחת סכום וכפל בסקלר:

משפט 6.4.3 אם $\sum a_n$ טור מתכנס בהחלט אז $\sum ca_n$ מתכנס בהחלט לכל $c\in\mathbb{R}$. אם גם $\sum b_n$ מתכנס בהחלט אז $\sum(a_n+b_n)$ מתכנס בהחלט.

ההוכחה אינה קשה, ומושארת כתרגיל.

נעבור עתה לטפל בטורים שאינם מתכנסים בהחלט. הרעיון יהיה למצוא תנאים בהם האיברים החיוביים והשליליים של הטור מבטלים זה את זה במידה מספקת. ראינו טור כזה בדוגמה (3). דוגמה מורכבת יותר היא הטור

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\ldots$$

כאן לאחר כל תוספת המגדילה את הסכום החלקי בא חיסור המקטין אותו, אך לא ברור אם יש "מספיק" הצטמצמות כדי שהטור יתכנס (הוא הרי אינו מתכנס בהחלט). עצם העובדה שכל איבר הוא בעל סימן הפוך מהאיבר הקודם לו אינו מבטיח התכנסות, כי למשל תופעה כזו קיימת גם בטור

$$1-1+1-1+1-\ldots$$

אבל טור זה אינו מתכנס אפילו במובן הרחב. כדי שטור עם סימנים מתחלפים יתכנס יש להניח לכל הפחות שהאיבר הכללי שואף לאפס. גם זה לא מספיק, ונראה לכך דוגמאות בהמשך. אולם מסתבר שאם דורשים שסדרת הערכים המוחלטים של הטור היא סדרה מונוטונית אפשר להבטיח התכנסות.

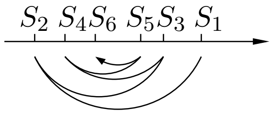
הגדרה 6.4.4 תהי $(a_n)_{n=1}^{\infty}$ סדרה אי־שלילית יורדת השואפת לאפס. טור מהצורה $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ נקרא **טור לייבניץ**.

הערות

1. הגדרנו $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ ולא $\sum_{n=1}^{\infty}(-1)^{n}a_n$ כדי שהאיבר הראשון בטור יהיה חיובי. הבחירה היא מטעמי נוחות, ואין לכך חשיבות רבה. אם היינו בוחרים באפשרות השנייה הדבר היה שקול להכפלת הטור ב־(-1).
2. כיוון שהסדרה בהגדרה יורדת ואי־שלילית, אם היא מתאפסת כל האיברים אחרי אותו מקום שווים לאפס. מקרה זה אינו מעניין עבורנו כי אז סוגיית התכנסות הטור היא טריביאלית. לכן לרוב נתעניין במקרה שהסדרה חיובית ממש.
3. שימו לב: סדרת הסכומים החלקיים (S_N) של טור לייבניץ המתאים לסדרה (a_n) איננו מוגדר על ידי $S_N = a_1 + \ldots + a_N$, כפי שאנו רגילים, אלא על ידי $S_N = a_1 - a_2 + a_3 - \ldots + (-1)^{N+1}a_N$.

לפני שנוכיח שטור לייבניץ מתכנס נוכיח את הלמה הבאה, שמבטאת במדויק את העובדה שכל מחובר בטור לייבניץ מבטל במובן מסוים את התרומה של האיבר לפניו:

למה 6.4.5 יהי $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ טור לייבניץ ותהי (S_N) סדרת הסכומים החלקיים שלו.



איור 6.4.1

1. תת־הסדרה של האיברים הזוגיים $(S_{2m})_{m=1}^{\infty}$ היא סדרה עולה, ותת־הסדרה $(S_{2m-1})_{m=1}^{\infty}$ של האיברים האי־זוגיים היא סדרה יורדת.
2. כל איבר בסדרה $(S_{2m})_{m=1}^{\infty}$ קטן מכל איבר בסדרה $(S_{2m-1})_{m=1}^{\infty}$.

הוכחה נוכיח את החלק הראשון. נשים לב ש־

$$S_{2m} - S_{2(m-1)} = a_{2m-1} - a_{2m} \geq 0$$

כי (a_n) יורדת. מכאן ש־(S_{2m}) עולה. באותו אופן, מראים ש־(S_{2m-1}) יורדת.

נפנה לחלק השני. יהיו $j,k \in \mathbb{N}$ ונבחר n טבעי גדול משניהם. אנו יודעים מהחלק הראשון שמתקיים

$$S_{2j} \leq S_{2n} \quad , \quad S_{2n-1} \leq S_{2k-1}$$

לכן אם נראה ש־ $S_{2n} \leq S_{2n-1}$ לכל n נקבל את הדרוש. ואמנם,

$$S_{2n} = S_{2n-1} + (-1)^{2n+1}a_{2n} \leq S_{2n-1}$$

כי $a_{2n} \geq 0$. ■

משפט 6.4.6 (משפט לייבניץ) טור לייבניץ $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ מתכנס למספר ממשי S המתקיים $0 \leq S \leq a_1$. יתרה מזאת, אם r_m הוא ה־m־זנב של הטור אז לכל m מתקיים $(-1)^m r_m \geq 0$ וגם $|r_m| \leq a_{m+1}$.

הוכחה תהי (S_N) סדרת הסכומים החלקיים של הטור. לפי הסעיף הראשון בלמה הקודמת $(S_{2m})_{m=1}^{\infty}$ היא סדרה עולה ולפי הסעיף השני היא חסומה, למשל על ידי S_1. לכן (S_{2m}) מתכנסת. נסמן את הגבול ב־S. באותו אופן רואים ש־ $(S_{2m-1})_{m=1}^{\infty}$ יורדת וחסומה, ולכן מתכנסת. נסמן את הגבול ב־T. אבל מתקיים

$$\begin{aligned} T - S &= \lim_{m\to\infty} S_{2m-1} - \lim_{m\to\infty} S_{2m} \\ &= \lim_{m\to\infty}(S_{2m-1} - S_{2m}) \\ &= \lim_{m\to\infty}(-(-1)^{(2m)+1}a_{2m}) \\ &= 0 \end{aligned}$$

(כי לפי ההגדרה של טור לייבניץ $a_n \to 0$). קיבלנו ש־ $S = T$. הראינו שסדרת האיברים הזוגיים של $(S_m)_{m=1}^{\infty}$ וסדרת האיברים האי־זוגיים שלה שואפות שתיהן לאותו גבול, ומכאן נובע שהסדרה (S_N) שואפת כולה ל־S (ראו תרגיל (4) בעמוד 135), כלומר, $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ מתכנס ל־S.

נוכיח כעת את הטענות הנוספות. נשים לב שכל איבר בסדרה $(S_{2m-1})_{m=1}^{\infty}$ גדול או שווה ל־S, כי זו סדרה יורדת שגבולה הוא S. לכן בפרט $S \leq S_1 = a_1$. באותו אופן, S גדול או שווה לכל איבר בסדרה $(S_{2m})_{m=1}^{\infty}$, שכן זו סדרה עולה המתכנסת ל־S, ולכן בפרט $S \geq S_2 = a_1 - a_2 \geq 0$, וקיבלנו ש־ $0 \leq S \leq a_1$.

נזכור שה־m־זנב של $\sum_{n=1}^{\infty}(-1)^{2n+1}a_n$ הוא הטור

$$r_m = \sum_{n=m+1}^{\infty}(-1)^{n+1}a_n = (-1)^m\sum_{n=1}^{\infty}(-1)^{n+1}a_{m+n}$$

הטור $\sum_{n=1}^{\infty}(-1)^{n+1}a_{m+n}$ המופיע באגף ימין הוא בעצמו טור לייבניץ (למה?) ולכן ממה שכבר הוכחנו נקבל שהוא מתכנס ומתקיים

$$0 \leq \sum_{n=1}^{\infty}(-1)^{n+1}a_{m+n} \leq a_{m+1}$$

(כי a_{m+1} הוא האיבר הראשון בטור הזנב r_m). מכאן $|r_m| \leq a_{m+1}$, וגם הטענה $(-1)^m r_m \geq 0$ ברורה. ■

דוגמאות

1. הטור $\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}$ מתכנס (ההתכנסות היא בתנאי).

2. לא ניתן לוותר על הנחת המונוטוניות של הטור. למשל ההנחה $a_n \to 0$ לבדה אינה מספיקה, כפי שמדגים המקרה

$$a_n = \begin{cases} 1/n & \text{זוגי } n \\ 0 & \text{אי־זוגי } n \end{cases}$$

(מדוע $\sum_{n=1}^{\infty}(-1)^n a_n$ מתבדר?). דוגמאות נוספות נמצאות בתרגילים בסוף הסעיף.

משפט לייבניץ נותן לא רק תנאי מספיק להתכנסות אלא גם כלי להערכת הערך של הטור, כי הוא חוסם את ערך הזנב. יהי $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ טור לייבניץ ונסמן ב־α את הערך שלו. נניח שרוצים למצוא קירוב של α עד כדי שגיאה קטנה ε, כלומר מחפשים N כך שהמרחק של הסכום החלקי S_N מ־α קטן מ־ε. לפי משפט לייבניץ מתקיים

$$|S_N - \alpha| = |r_N| \leq a_{N+1}$$

ולכן אם נבחר N כך ש־ $a_{N+1} < \varepsilon$ אזי S_N קרוב ל־α עד כדי ε.

למשל, ננסה להעריך את הגבול של $\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}$ בדיוק של $\frac{1}{10}$. כאן הסדרה a_n היא $a_n = \frac{1}{n}$ וכדי שיתקיים $a_{N+1} < \frac{1}{10}$ נוכל לבחור $N = 10$. אנו מסיקים שהגבול α של הטור שווה עד כדי עשירית לסכום

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{1}{9} - \frac{1}{10}$$

משפט לייבניץ דן בטור מהצורה $\sum a_n b_n$ שאיבריו מתקבלים כמכפלה של שתי סדרות: סדרה מונוטונית (a_n), השואפת לאפס, וסדרה (b_n) המוגדרת על ידי $b_n = (-1)^{n+1}$. מסתבר שניתן להכליל את המשפט במידה ניכרת על ידי החלפת (b_n) בסדרה כללית יותר המקיימת את התכונה הבאה:

הגדרה 6.4.7 הטור $\sum_{n=1}^{\infty} a_n$ נקרא **טור חסום** אם סדרת הסכומים החלקיים שלו היא סדרה חסומה.

דוגמאות

1. סדרת הסכומים החלקיים של הטור $\sum_{n=1}^{\infty}(-1)^{n+1}$ מקבלת רק את הערכים $0, 1$, ולכן הטור חסום.
2. עבור m קבוע, הטור $\sum_{n=1}^{\infty}(-1)^{[n/m]+1}$ הוא טור חסום (הוכיחו!).
3. כל טור מתכנס הוא טור חסום כי סדרת הסכומים החלקיים שלו מתכנסת וממילא חסומה.

להוכחת המשפט הבא נזדקק לזהות אלגברית:

למה 6.4.8 (נוסחת הסכימה של אבל)[4] יהי $m \in \mathbb{N}$ ויהיו $(a_1, \ldots, a_m), (b_1, \ldots, b_m)$ סדרות סופיות באורך m. לכל $1 \leq k \leq m$ נסמן ב־ $B_k = \sum_{i=1}^{k} b_i$ את סכום k האיברים הראשונים של $(b_i)_{i=1}^{m}$. אז

$$\sum_{i=1}^{m} a_i b_i = a_m B_m + \sum_{i=1}^{m-1} (a_i - a_{i+1}) B_i$$

הערה הלמה נראית מסובכת אך היא מסתירה זהות אלגברית פשוטה. התבוננו למשל במקרה $n = 2$: אז הלמה אומרת

$$a_1 b_1 + a_2 b_2 = a_2(b_1 + b_2) + (a_1 - a_2) b_1$$

לפני עיון בהוכחה כדאי לרשום במפורש את המקרים $n = 3, 4$.

הוכחה נסמן $B_0 = 0$. אז לכל $1 \leq i \leq m$ מתקיים $b_i = B_i - B_{i-1}$, ולכן

$$\sum_{i=1}^{m} a_i b_i = \sum_{i=1}^{m} a_i (B_i - B_{i-1}) = \sum_{i=1}^{m} a_i B_i - \sum_{i=1}^{m} a_i B_{i-1}$$

כיוון ש־ $B_0 = 0$, הסכום השני באגף הימני מתחיל למעשה באינדקס $i = 2$. על ידי שינוי משתנה הסכימה, רואים שהסכום הימני ביותר שווה ל־ $\sum_{i=1}^{m-1} a_{i+1} B_i$. לכן

$$\begin{aligned}
\sum_{i=1}^{m} a_i b_i &= \sum_{i=1}^{m} a_i B_i - \sum_{i=1}^{m-1} a_{i+1} B_i \\
&= a_m B_m + \sum_{i=1}^{m-1} (a_i B_i - a_{i+1} B_i) \\
&= a_m B_m + \sum_{i=1}^{m-1} (a_i - a_{i+1}) B_i
\end{aligned}$$

כפי שרצינו. ■

מסקנה 6.4.9 בסימוני הלמה הקודמת, אם בנוסף ידוע ש־ $(a_i)_{i=1}^{m}$ היא סדרה מונוטונית ושלכל $1 \leq k \leq m$ מתקיים $|B_k| \leq L$ אז

$$|\sum_{i=1}^{m} a_i b_i| \leq L \cdot (2|a_m| + |a_1|)$$

[4] Niels Abel, 1802-1829.

הוכחה מכך שהסדרה מונוטונית, נובע ש־ $|a_{i+1}-a_i|$ בעל סימן זהה לכל i, ולכן

$$\sum_{i=1}^{m-1}|a_{i+1}-a_i| = |\sum_{i=1}^{m-1}(a_{i+1}-a_i)| = |a_m-a_1|$$

ולכן לפי הלמה הקודמת נקבל

$$\begin{aligned}
|\sum_{i=1}^{m} a_i b_i| &\leq |a_m B_m + \sum_{i=1}^{m-1}(a_i - a_{i+1})B_i| \\
&\leq |a_m||B_m| + \sum_{i=1}^{m-1}|B_i||a_{i+1}-a_i| \\
&\leq L\cdot(|a_m| + \sum_{i=1}^{m-1}|a_{i+1}-a_i|) \\
&= L\cdot(|a_m| + |a_m - a_1|) \\
&\leq L\cdot(2|a_m| + |a_1|)
\end{aligned}$$

כמובטח. ■

המשפט הבא מכליל את משפט ההתכנסות לטורי לייבניץ.

משפט 6.4.10 (מבחן דירכלה)[5] אם $\sum_{n=1}^{\infty} b_n$ הוא טור חסום והסדרה $(a_n)_{n=1}^{\infty}$ מונוטונית ושואפת לאפס, אז הטור $\sum_{n=1}^{\infty} a_n b_n$ הוא טור מתכנס.

הוכחה תהי (T_N) סדרת הסכומים החלקיים של הטור $\sum_{n=1}^{\infty} b_n$. לפי הנתון, קיים $M>0$ עבורו $|T_N|\leq M$ לכל N. נראה שהטור $\sum_{n=1}^{\infty} a_n b_n$ מתכנס על ידי שימוש בתנאי קושי. יהי $\delta>0$. כיוון ש־ (a_n) היא סדרה יורדת המתכנסת ל־0, קיים N כך שלכל $n>N$ מתקיים $0<a_n<\delta$. נקבע $n>N$ ומספר k ונשים לב שאם נגדיר $B_k = b_{n+1}+\ldots+b_{n+k}$ אז מתקיים $B_k = T_{n+k}-T_n$ ולכן גם

$$|B_k| \leq |T_{n+k}| + |T_n| \leq 2M$$

כך נקבל בעזרת הלמה האחרונה ש־

$$|\sum_{j=n+1}^{n+k} a_j b_j| = |\sum_{i=1}^{k} a_{n+i} b_{n+i}| \leq 2M\cdot(2|a_{n+k}| + |a_{n+1}|) < 2M\cdot(2\delta+\delta) = 6M\delta$$

בהינתן $\varepsilon>0$ אם נבחר $\delta=\frac{\varepsilon}{6M}$ קיבלנו שיש N כך שלכל $n>N$ ולכל k מתקיים $|\sum_{i=n+1}^{n+k} a_i b_i| < \varepsilon$ ולכן הטור מקיים את תנאי קושי, ומתכנס. ■

[5] Lejeune Dirichlet, 1805-1859.

דוגמאות

1. משפט לייבניץ נובע ממשפט דירכלה, כי הטור $\sum(-1)^n$ הוא טור חסום.
2. הטור $\sum\frac{\{\sqrt{n+1}\}-\{\sqrt{n}\}}{n}$ מתכנס, כי $\sum(\{\sqrt{n+1}\}-\{\sqrt{n}\})$ הוא טור חסום (למה?) ו־ $\frac{1}{n}\to 0$ מונוטונית.

תנאי נוסף להתכנסות הוא התנאי הבא, שבו הדרישה מהטור $\sum b_n$ חזקה יותר, אבל הדרישה מהסדרה (a_n) חלשה יותר.

מסקנה 6.4.11 (מבחן אבל) תהי (a_n) סדרה מונוטונית וחסומה ויהי $\sum_{n=1}^{\infty} b_n$ טור מתכנס. אז הטור $\sum_{n=1}^{\infty} a_n b_n$ מתכנס.

הוכחה כיוון ש־ (a_n) מונוטונית וחסומה, היא מתכנסת למספר ממשי a. לכל N נרשום כעת

$$S_N=\sum_{n=1}^{N} a_n b_n=\sum_{n=1}^{N}(a_n-a+a)b_n=\sum_{n=1}^{N}(a_n-a)b_n+\sum_{n=1}^{N} ab_n$$

המספר a קבוע, ולכן כאשר $N\to\infty$ המחובר השני באגף ימין מתכנס לפי ההנחה. כמו־כן, המחובר השמאלי באגף ימין מתכנס לפי משפט דירכלה, כי $(a_n-a)_{n=1}^{\infty}$ היא שואפת מונוטונית לאפס, והטור $\sum_{n=1}^{\infty} b_n$ הוא טור מתכנס, ובפרט חסום. ■

לדוגמה, נוכל להסיק שלכל טור מתכנס $\sum a_n$ הטור $\sum\frac{\sqrt{n}a_n}{\sqrt{n+1}}$ מתכנסת כי הסדרה $\frac{\sqrt{n}}{\sqrt{n+1}}$ מונוטונית וחסומה. אם תנסו להוכיח זאת ישירות תגלו שאין זה קל!

תרגילים

1. אילו מהטורים הבאים מתכנסים, ואילו מתכנסים בהחלט?
   (א) $\sum\frac{(-1)^n}{\sqrt{n}}$.
   (ב) $\sum\frac{(-1)^{[\sqrt{n}]}}{\sqrt{n}}$.
   (ג) $\sum\frac{(-1)^{[\sqrt{n}]}}{n^2}$.
   (ד) $\sum\frac{(-1)^{[\sqrt{n}]}-(-1)^{[\sqrt{n+1}]}}{n^2}$.
   (ה) $\sum\frac{(-1)^n}{n^2+(-1)^{n+1}}$.
2. יהי S סכום הטור ההרמוני המתחלף, כלומר $S=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}$. הוכיחו ש־ $\frac{7}{12}<S<\frac{47}{60}$.
3. הצדיקו את התכנסות הטורים הבאים ותנו ביטוי פשוט לסכום:

$$\sum_{n=1}^{\infty}\left(\frac{1}{(2n-1)2n}+\frac{(-1)^n}{\sqrt{n}}\right)+\sum_{n=1}^{\infty}\left(\frac{1}{2n(2n+1)}+\frac{(-1)^{n+1}}{\sqrt{n}}\right)$$

4. התבוננו בטור $\dots+\frac{1}{\sqrt{2}-1}-\frac{1}{\sqrt{2}+1}+\frac{1}{\sqrt{3}-1}-\frac{1}{\sqrt{3}+1}$, והראו ש־ $S_{2N}\xrightarrow{N\to\infty}\infty$. הסיקו שדרישת המונוטוניות במשפט לייבניץ אינה מיותרת.

5. הוכיחו שאם $\sum a_n$ מתכנס בהחלט אז $\sum a_{n_k}$ מתכנס בהחלט לכל תת־סדרה (a_{n_k}) של (a_n). האם הטענה נכונה כשמחליפים התכנסות בהחלט בהתכנסות רגילה?

6. הוכיחו את משפט 6.4.3.

7. נניח ש־$\sum a_n$ מתכנס בהחלט ו־ (b_n) סדרה חסומה. הראו ש־ $\sum a_n b_n$ מתכנס בהחלט.

8. הוכיחו שאם הטורים $\sum a_n^2$ ו־ $\sum b_n^2$ מתכנסים, אז הטור $\sum a_n b_n$ מתכנס בהחלט ומתקיים

$$(\sum_{n=1}^{\infty}|a_n b_n|)^2\leq(\sum_{n=1}^{\infty}a_n^2)\cdot(\sum_{n=1}^{\infty}b_n^2)$$

(רמז : היעזרו באי־שוויון קושי־שוורץ, תרגיל (5) בעמוד 78). הסיקו שאם $\sum a_n^2$ מתכנס, אז $\sum\frac{a_n}{n}$ מתכנס.

9. הוכיחו שהטור בדוגמה (2) בעמוד 184 הוא אכן טור מתבדר.

10. הראו כיצד התכנסות של טורי לייבניץ היא מסקנה של משפט דירכלה.

11. בנו סדרה אי־שלילית (a_n) השואפת ל־0 כך שהטור $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ אינו מתכנס אפילו במובן הרחב.

12. הוכיחו שאם $\sum a_n$ מתכנס בהחלט, אז גם $\sum a_n^2$ מתכנס בהחלט. לעומת זאת, הראו שייתכן ש־ $\sum a_n$ מתכנס בעוד ש־ $\sum a_n^2$ אינו מתכנס. (*) האם ייתכן ש־ $\sum a_n$ מתכנס ו־$\sum a_n^3$ אינו מתכנס?

13. תנאי הכרחי להתכנסות של טור $\sum a_n$ היא ש־ $a_n\to 0$ ושהטור חסום. ראינו שהתנאי $a_n\to 0$ לבדו אינו מספיק. האם שני התנאים יחד גוררים התכנסות?

6.5 הכנסת סוגריים ושינוי סדר איברים

בסעיף זה נבחן מתי כלל החילוף וכלל הקיבוץ המוכרים מסכומים סופיים חלים על טורים. התשובה, בקצרה, היא שלא תמיד!

נתחיל בכלל הקיבוץ, שמבטיח שבסכום סופי $a_1+\ldots+a_n$ ניתן להכניס סוגריים בכל אופן שנרצה מבלי לשנות את התוצאה: למשל הסכום לעיל שווה לסכום $(a_1+\ldots+a_{n-1})+a_n$. שימו לב שלא שינינו את סדר האיברים.

כדי לראות שהכנסת סוגריים בטור אינסופי אינה מעשה תמים, נתבונן בדוגמה הפשוטה הבאה. יהי $a_n=(-1)^{n+1}$. אז הטור $\sum_{n=1}^{\infty}a_n=a_1+a_2+a_3+\ldots$ הוא

טור מתבדר (למשל כי האיבר הכללי אינו שואף לאפס). לעומת זאת, אם נקבץ את איברי הטור לזוגות נקבל את הטור

$$\begin{aligned}(a_1+a_2)+(a_3+a_4)+(a_5+a_6)+\ldots &= (1-1)+(1-1)+(1-1)+\ldots \\ &= 0+0+0+\ldots\end{aligned}$$

טור זה מתכנס ל־0. בחירת סוגריים אחרת נותנת את הטור

$$\begin{aligned}a_1+(a_2+a_3)+(a_4+a_5)+\ldots &= 1+(-1+1)+(-1+1)+\ldots \\ &= 1+0+0+0+\ldots\end{aligned}$$

זהו שוב טור מתכנס, אך סכומו כעת הוא 1.

נעבור לדיון פורמלי. הכנסת סוגריים לטור $\sum_{n=1}^\infty a_n$ היא הפעולה של יצירת טור חדש $\sum_{n=1}^\infty b_n$ שכל אחד מאיבריו הוא סכום של כמה איברים עוקבים מהסדרה (a_n). בדוגמה הקודמת יצרנו מהטור $\sum_{n=1}^\infty a_n = a_1+a_2+a_3+\ldots$ את הטור

$$(a_1+a_2)+(a_3+a_4)+(a_5+a_6)+\ldots$$

זהו טור חדש $\sum_{k=1}^\infty b_k$ שאיבריו הם $b_k = a_{2k-1}+a_{2k}$. באופן כללי יותר,

הגדרה 6.5.1 יהי $\sum_{n=1}^\infty a_n$ טור ו־$(n_k)_{k=1}^\infty$ סדרה עולה ממש של מספרים טבעיים עם $n_1=1$. **הטור המתקבל מ־**$\sum_{n=1}^\infty a_n$ **על ידי הכנסת סוגריים במקומות** $(n_k)_{k=1}^\infty$ הוא הטור $\sum_{k=1}^\infty b_k$ שאיבריו הם

$$b_k = a_{n_k}+\ldots+a_{n_{k+1}-1}$$

כלומר

$$\sum_{k=1}^\infty b_k = (a_{n_1}+\ldots+a_{n_2-1})+(a_{n_2}+\ldots+a_{n_3-1})+(a_{n_3}+\ldots)+\ldots$$

למשל כשכינסנו יחד כל איבר בעל אינדקס אי־זוגי עם האיבר הבא אחריו, כמו בדוגמה הראשונה בסעיף זה, השתמשנו בסדרה $n_k = 2k-1$ (וודאו שאתם מבינים למה!).

הקשר הבסיסי בין טור לטור המתקבל ממנו לאחר הכנסת סוגריים נתון בלמה הבאה:

למה 6.5.2 יהי $\sum_{n=1}^\infty a_n$ טור ויהי $\sum_{n=1}^\infty b_n$ הטור המתקבל ממנו על ידי הכנסת סוגריים במקומות (n_k). תהי (S_N) סדרת הסכומים החלקיים של $\sum_{n=1}^\infty a_n$ ו־(T_K) סדרת הסכומים החלקיים של $\sum_{n=1}^\infty b_n$. אזי $T_K = S_{n_{K+1}-1}$

הוכחה נחשב:

$$\begin{aligned} T_K &= b_1 + b_2 + \ldots + b_K \\ &= (a_1 + \ldots + a_{n_2-1}) \ + \\ &\qquad + (a_{n_2} + \ldots + a_{n_3-1}) + \ldots\ldots + (a_{n_K} + \ldots + a_{n_{K+1}-1}) \\ &= a_1 + a_2 + \ldots + a_{n_{K+1}-1} \\ &= S_{n_{K+1}-1} \end{aligned}$$

(בין השורה השנייה לשלישית השתמשנו בחוק הקיבוץ לסכומים סופיים). ■

ראינו שהכנסת סוגריים לטור מתבדר יכולה לתת טור מתכנס. ההפך אינו נכון:

משפט 6.5.3 אם $\sum_{n=1}^{\infty} a_n$ הוא טור מתכנס במובן הרחב ל־S אז כל טור המתקבל ממנו על ידי הכנסת סוגריים הוא טור מתכנס שסכומו גם־כן S.

הוכחה נסמן ב־$(S_N)_{N=1}^{\infty}$ סדרת הסכומים החלקיים של הטור $\sum_{n=1}^{\infty} a_n$, ונסמן ב־$(T_K)_{K=1}^{\infty}$ את סדרת הסכומים החלקיים של הטור המתקבל ממנו על ידי הכנסת סוגריים במקומות (n_k). אם הטור המקורי מתכנס ל־S, אז $S_N \to S$. לפי הלמה, (T_K) היא תת־סדרה של (S_N), ולכן גם $T_K \to S$. לכן הטור החדש מתכנס אף הוא ל־S. ■

יש תנאים בהם הכיוון ההפוך של המשפט נכון (ראינו שהוא אינו נכון באופן כללי). תנאים אלה דורשים שהסוגריים יוכנסו באופן "זהיר". למשל:

משפט 6.5.4 יהי $\sum_{k=1}^{\infty} b_k$ טור מתכנס ונניח שהוא התקבל על ידי הכנסת סוגריים בטור $\sum_{n=1}^{\infty} a_n$ במקומות (n_k) באופן כזה ש־ $a_{n_k}, \ldots, a_{n_{k+1}-1}$ שווי־סימן (כלומר: כולם אי־שליליים או כולם אי־חיוביים). אז הטור $\sum_{n=1}^{\infty} a_n$ מתכנס גם הוא, ו־ $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} b_n$.

הוכחה נסמן שוב ב־(S_N) את הסכומים החלקיים של $\sum a_n$ וב־(T_N) את הסכומים החלקיים של $\sum b_n$. הנחתנו היא ש־ $T_K \to T$, כאשר T סופי (ניתן להכליל זאת למקרה $T = \pm\infty$ בעזרת הוכחה דומה). עלינו להראות שגם $S_N \to T$.

יהי $\varepsilon > 0$. אנו יודעים מלמה 6.5.2 שהסדרה $(S_{n_{k+1}-1})_{k=1}^{\infty}$ מתכנסת ל־T ולכן יש K_0 כך שלכל $k > K_0$ מתקיים $|S_{n_k-1} - T| < \varepsilon$. אנו רוצים להראות שלכל i מספיק גדול מתקיים $|S_i - T| < \varepsilon$.

לכל i יש $k(i)$ יחיד כך ש־ $n_{k(i)} \leq i < n_{k(i)+1} - 1$, ומתקיים $\lim_{i\to\infty} k(i) = \infty$. כמו־כן

$$S_i = S_{n_{k(i)}-1} + \sum_{j=n_{k(i)}}^{i} a_j$$

לכל i גדול מספיק מתקיים $k(i) > K_0$ ואז

$$|S_i - T| \le |S_{n_{k(i)}-1} - T| + |\sum_{j=n_{k(i)}}^{i} a_j| < \varepsilon + |\sum_{j=n_{k(i)}}^{i} a_j|$$

לכן אם נראה שלכל i גדול מספיק מתקיים $|\sum_{j=n_{k(i)}}^{i} a_j| < \varepsilon$ נשלים את ההוכחה. מההנחה ש־ $a_{n_k}, \ldots, a_{n_{k+1}-1}$ שווי סימן נובע

$$|\sum_{j=n_{k(i)}}^{i} a_j| \le |\sum_{j=n_{k(i)}}^{n_{K(i)+1}-1} a_j| = b_{k(i)}$$

מכיוון ש־ $\lim_{k\to\infty} b_k = 0$ (למה?) גם $\lim_{i\to\infty} b_{k(i)} = 0$ (כאן השתמשנו בעובדה ש־ $k(i) \to \infty$), ולכן לכל i גדול מספיק מתקיים $|b_{k(i)}| < \varepsilon$. לכן לכל i גדול מספיק $|\sum_{j=n_{k(i)}}^{i} a_j| < \varepsilon$. מכאן ש־ $|S_i - T| < \varepsilon + \varepsilon < 2\varepsilon$, וזה גורר את מה שרצינו. ■

ישנם תנאים נוספים על הסדרה (n_k) שמבטיחים משפטים כמו האחרון: ראו למשל תרגיל (3) בסוף הסעיף.

נעבור כעת לדון בכלל החילוף. עבור סכום סופי כלל זה מבטיח ששינוי סדר האיברים בסכימה לא משפיע על התוצאה. לדוגמה, לכל $a_1, a_2, a_3 \in \mathbb{R}$ מתקיים $a_1 + a_2 + a_3 = a_1 + a_3 + a_2$. על כך שבמקרה של טורים יש בעיה תעיד הדוגמה הבאה. נתבונן בטור $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$. זהו טור לייבניץ ולכן הוא מתכנס. נסמן את סכומו ב־s. ממשפט לייבניץ אנו יודעים גם ש־ $s > 1 - \frac{1}{2} > 0$. מצד שני אם נרשה לעצמנו לשנות את סדר האיברים נקבל

$$\begin{aligned}
s &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \frac{1}{10} + \frac{1}{11} - \frac{1}{12} + \frac{1}{13} - \frac{1}{14} + \ldots \\
&= 1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \frac{1}{7} - \frac{1}{14} - \frac{1}{16} + \ldots \\
&= (1 - \frac{1}{2}) - \frac{1}{4} + (\frac{1}{3} - \frac{1}{6}) - \frac{1}{8} + (\frac{1}{5} - \frac{1}{10}) - \frac{1}{12} + (\frac{1}{7} - \frac{1}{14}) - \frac{1}{16} + \ldots \\
&= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \frac{1}{14} - \frac{1}{16} + \ldots \\
&= \frac{1}{2}(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \ldots) \\
&= \frac{1}{2}s
\end{aligned}$$

(במעבר מהשורה הראשונה לשנייה שינינו את סדר האיברים על ידי בחירת איבר חיובי, אחריו שני איברים שליליים, איבר חיובי, שניים שליליים, וכן הלאה. שימו לב שהכנסת הסוגריים בין השורה השנייה והשלישית חוקית לפי תרגיל (3) בסוף הסעיף. קיבלנו ש־ $s = \frac{1}{2}s$ ולכן $s = 0$, בסתירה לאי־שוויון $s > 0$.

ההגדרה של שינוי סדר בטור מעט יותר מסובכת מההגדרה של הכנסת סוגריים. אם $(a_n)_{n=1}^{\infty}$ היא סדרה אז שינוי סדר שלה הוא סדרה מהצורה $(a_{n_k})_{k=1}^{\infty}$ כאשר (n_k) היא סדרת אינדקסים, וכך שכל איבר של (a_n) מופיע בדיוק פעם אחת בסדרה (a_{n_k}). למשל, אם (a_1, a_2, a_3) סדרה (סופית) ו־ $(b_1, b_2, b_3) = (a_1, a_3, a_2)$ אז $(b_k)_{k=1}^{3} = (a_{n_k})_{k=1}^{3}$ עבור האינדקסים $n_1 = 1, n_2 = 3, n_3 = 2$. כפי שציינו אנו רוצים להביע את התנאי שכל איבר ב־(a_n) מופיע פעם אחת בדיוק ב־(a_{n_k}). התכונה הרלוונטית כאן היא תכונה של סדרת האינדקסים (n_k):

הגדרה 6.5.5 סדרה $(n_k)_{k=1}^{\infty}$ שאיבריה הם מספרים טבעיים חיוביים נקראת **תמורה** (permutation) של המספרים הטבעיים החיוביים אם כל מספר כזה מופיע בסדרה בדיוק פעם אחת, או במילים אחרות אם לכל $0 < i \in \mathbb{N}$ יש k יחיד כך ש־ $n_k = i$.

דוגמאות

1. ההתאמה

$$n_k = \begin{cases} k-1 & \text{זוגי } k \\ k+1 & \text{אי־זוגי } k \end{cases}$$

היא תמורה. שכן בדיקה מראה ש־n_k טבעי וחיובי אם k טבעי וחיובי. כמו־כן אם $k > 0$ זוגי אז $n_{k-1} = k$, ואילו אם $k > 0$ אי־זוגי אז $n_{k+1} = k$, ולכן לכל $i > 0$ קיים k כך ש־ $n_k = i$. קל להראות שאם $n_k = n_m$ אז $k = m$ (בדקו!).

2. אם $n_k = 2k$ אז (n_k) אינה תמורה כי אין אף מספר טבעי k כך ש־ $n_k = 1$.

3. אם $n_k = 1 + |n - 10|$ אז לכל $i > 0$ יש k כך ש־ $n_k = i$ (בדקו) אבל (n_k) אינה תמורה כי $n_9 = n_{11} = 2$, כלומר יש שני k־ים שונים כך ש־$n_k = 2$ וזה אסור לפי ההגדרה.

בעזרת מושג התמורה נוכל לנסח את רעיון שינוי הסדר של סדרה באופן הבא:

הגדרה 6.5.6 אם (a_n) סדרה נאמר שסדרה (b_k) מתקבלת ממנה על ידי **שינוי סדר** לפי התמורה (n_k) אם $b_k = a_{n_k}$. במקרה זה נאמר ש־(b_n) התקבלה מ־(a_n) בעזרת התמורה (n_k).

דוגמאות

1. נתבונן בטור שאיבריו $a_n = (-1)^{n+1}$, כלומר בטור $1-1+1-1+1-1+\ldots$. זהו טור חסום שאינו מתכנס. הטור $-1+1-1+1-1+1-\ldots$ מתקבל ממנו בעזרת התמורה

$$n_k = \begin{cases} k-1 & \text{זוגי } k \\ k+1 & \text{אי־זוגי } k \end{cases}$$

מדוגמה (1) למעלה.

2. מהטור $\dots - \frac{1}{5} + \frac{1}{4} - \frac{1}{3} + \frac{1}{2} - 1$ אפשר לקבל את $\dots - \frac{1}{5} + \frac{1}{8} - \frac{1}{6} - \frac{1}{3} + \frac{1}{4} - \frac{1}{2} - 1$ באופן הבא. נגדיר

$$n_k = \begin{cases} 2m+1 & k = 3m+1 \\ 2m & k = 3m \text{ או } k = 3m+2 \end{cases}$$

3. נתבונן שוב בטור $\dots - \frac{1}{5} + \frac{1}{4} - \frac{1}{3} + \frac{1}{2} - 1 = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$, שהוא טור לייבניץ מתכנס. נסדר אותו מחדש באופן הבא: תחילה נבחר k כך ש־ $\sum_{n=1}^{k} \frac{1}{2n-1} > 1$ (למה יש k כזה?) ונרשום את k האיברים החיוביים הראשונים מהטור, שהם $1, \frac{1}{3}, \dots, \frac{1}{2k-1}$. אז נרשום את האיבר השלילי הראשון. כעת נבחר k' כך שהסכום $\sum_{n=k+1}^{k'} \frac{1}{2n-1} > 1$ ונרשום בטור החדש את האיברים $\frac{1}{2(k+1)-1}, \frac{1}{2(k+2)-1}, \dots\dots \frac{1}{2k'-1}$ ואחריהם את האיבר השלילי הראשון בטור המקורי שלא בחרנו עדיין. נמשיך כך באינדוקציה, ונקבל את הטור

$$\underbrace{1 + \frac{1}{3}}_{>1} \; - \; \frac{1}{2} \; + \; \underbrace{\frac{1}{5} + \dots + \frac{1}{27}}_{>1} \; - \; \frac{1}{4} \; + \; \dots$$

טור זה התקבל על ידי שינוי סדר מהטור המתכנס $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$, אך הוא מתכנס ל־$\infty$ (השלימו את הפרטים).

כזכור, טור מתכנס בהחלט הוא טור $\sum a_n$ כך ש־ $\sum |a_n|$ מתכנס (הגדרה 6.4.1). מסתבר שאלו הם הטורים שניתן בבטחה לשנות בהם את סדר האיברים מבלי לשנות את תכונות ההתכנסות או את הסכום שלהן.

לפני שנוכיח את הטענה נזדקק ללמה:

למה 6.5.7 תהי (n_k) תמורה ו־$0 < M \in \mathbb{N}$. אז לכל k מספיק גדול, המספרים $n_1, n_2, \dots, n_k$ כוללים ביניהם את כל המספרים $1, 2 \dots, M$, או במילים אחרות, $\{1, \dots, M\} \subseteq \{n_1, n_2, \dots, n_k\}$.

הוכחה תהי $J = \{j : n_j \in \{1, 2, \dots, M\}\}$ קבוצת האינדקסים j כך ש־n_j שווה לאחד המספרים $1, 2, \dots, M$. לכל $i \in \{1, \dots, M\}$ יש בדיוק j אחד עם $n_j = i$ ולכן J בגודל M ובפרט סופית, וקיים לה מקסימום j_0. נובע שכל המספרים $1, \dots, M$ נמצאים בין האינדקסים $n_1, \dots, n_{j_0}$, וממילא אם $j > j_0$ אז $\{1, \dots, M\} \subseteq \{n_1, n_2, \dots, n_j\}$. ■

המשמעות של הלמה בהקשר שלנו היא שאם (b_n) מתקבלת מ־(a_n) על ידי שינוי סדר אז לכל k מספיק גדול האיברים $a_1, \dots, a_M$ מופיעים בין המחוברים של הסכום החלקי $b_1 + \dots + b_k$ של $\sum_{n=1}^{\infty} b_n$ (וודאו שאתם מבינים איך ההערה נובעת מהלמה!).

משפט 6.5.8 אם $\sum_{n=1}^{\infty} a_n$ הוא טור מתכנס בהחלט אז כל טור המתקבל ממנו על ידי שינוי סדר מתכנס גם הוא, ולאותו סכום.

הוכחה נסמן $S = \sum_{n=1}^{\infty} a_n$ ו־(S_N) סדרת הסכומים החלקיים של $\sum_{n=1}^{\infty} a_n$. תהי (n_k) תמורה, ותהי $b_k = a_{n_k}$ ו־(T_N) סדרת הסכומים החלקיים של $\sum_{k=1}^{\infty} b_k$. עלינו להראות ש־ $T_N \to S$. לשם כך מספיק שנראה ש־ $S_N - T_N \to 0$. נתבונן בהפרש הזה:

$$S_N - T_N = \sum_{n=1}^{N} a_n - \sum_{k=1}^{N} b_k = \sum_{n=1}^{N} a_n - \sum_{k=1}^{N} a_{n_k}$$

יהי $k > 0$. לפי הלמה, לכל N מספיק גדול האיברים $a_1, \ldots, a_k$ מופיעים בין האיברים $a_{n_1}, \ldots, a_{n_k}$ ולכן כל איבר a_i עבור $1 \leq i \leq k$ מופיע פעמיים באגף ימין של השוויון לעיל, פעם עם סימן חיובי ופעם עם סימן שלילי, ושתי ההופעות מבטלות זו את זו. לכן לכל N מספיק גדול מתקיים (לאחר מחיקת האיברים שהצטמצמו):

$$\begin{aligned} |S_N - T_N| &= \Big| \sum_{n\,:\,k<n\leq N} a_n - \sum_{m\,:\,1\leq m\leq N\,,\,n_m>k} a_{n_m} \Big| \\ &\leq \sum_{n\,:\,k<n\leq N} |a_n| + \sum_{m\,:\,1\leq m\leq N\,,\,n_m>k} |a_{n_m}| \\ &\leq r_k + r_k \end{aligned}$$

כאשר $r_k = \sum_{n=k+1}^{\infty} |a_n|$ הוא ה־k־זנב של הטור $\sum |a_n|$. מכיוון ש־ $r_k \to 0$ (מסקנה 6.2.5), לכל ε יש k כך ש־ $|r_k| < \frac{1}{2}\varepsilon$. עבור אותו k קיבלנו שלכל N מספיק גדול

$$|S_N - T_N| < 2r_k < \varepsilon$$

לכן $|S_N - T_N| \to 0$, כפי שרצינו להראות. ■

כפי שרמזנו, התכנסות בהחלט אינה רק תנאי מספיק לאי־רגישות של טור לשינויי סדר, אלא היא גם תנאי הכרחי. אם טור מתכנס בתנאי (כלומר, מתכנס אבל לא מתכנס בהחלט) אז חוק החילוף נכשל לחלוטין וניתן על ידי שינוי סדר לקבל טור המתכנס לכל מספר מבוקש, או טור מתבדר. לפני שנוכיח זאת נגדיר כמה סימונים חדשים.

הגדרה 6.5.9 יהי $a \in \mathbb{R}$. המספרים a^+, a^- מוגדרים על ידי

$$a^+ = \max\{a, 0\}\ ,\ a^- = \max\{-a, 0\}$$

כלומר: $a^+ = a$ אם a חיובי ואחרת שווה ל־0, ו־a^- הוא $-a$ (כלומר $|a|$) אם a שלילי ו־0 אחרת. שימו לב שלכל a מתקיים

$$\begin{aligned} a &= a^+ - a^- \\ |a| &= a^+ + a^- \end{aligned}$$

בהנתן טור $\sum a_n$, אפשר לקבל ממנו שני טורים חיוביים: $\sum a_n^+$ ו־ $\sum a_n^-$. הראשון מתקבל על ידי החלפת כל איברי הטור השליליים באפס, והשני על ידי החלפת כל איברי הטור החיוביים באפס והפיכת יתר המחוברים לחיוביים. מההערה הקודמת רואים ש־ $\sum a_n = \sum a_n^+ - \sum a_n^-$ וגם $\sum |a_n| = \sum a_n^+ + \sum a_n^-$.

למה 6.5.10 יהי $\sum a_n$ טור מתכנס בתנאי. אז $\sum a_n^+ = \sum a_n^- = \infty$, ובפרט יש בטור אינסוף איברים חיוביים ואינסוף איברים שליליים.

הוכחה מכיוון שהטורים $\sum a_n^+, \sum a_n^-$ חיוביים, הם מתכנסים במובן הרחב לגבול סופי או לאינסוף. מכיוון שהסדרה (a_n) היא הפרש הסדרות (a_n^+) ו־(a_n^-), לא ייתכן שאחד הטורים $\sum a_n^+, \sum a_n^-$ מתכנס למספר והשני לאינסוף, כי אז מאריתמטיקה של גבולות (במובן הרחב) היה נובע ש־$\sum a_n$ מתכנס ל־∞, בסתירה לנתון. כמו־כן הטור $\sum |a_n|$ מתבדר, ומכיוון שהוא סכום של הטורים $\sum a_n^+, \sum a_n^-$ לא ייתכן ששניהם מתכנסים לגבול סופי. האפשרות היחידה שנותרה היא ש־ $\sum a_n^+ = \sum a_n^- = \infty$, ומכאן שבשניהם מופיעים אינסוף איברים שונים מאפס. קיבלנו שב־$\sum a_n$ יש אינסוף איברים חיוביים ואינסוף איברים שליליים. ∎

משפט 6.5.11 (משפט רימן[6]) אם $\sum_{n=1}^{\infty} a_n$ הוא טור מתכנס בתנאי, אז לכל S (סופי או אינסופי) אפשר, על ידי שינוי סדר, לקבל מ־$\sum_{n=1}^{\infty} a_n$ טור המתכנס ל־S. כמו־כן, אפשר על ידי שינוי סדר לקבל טור שאינו מתכנס (גם במובן הרחב).

הוכחה אנו נוכיח רק את המקרה S ממשי. המקרים האחרים דומים.

באופן לא פורמלי, נסדר את האיברים של הטור באופן הבא: קודם נבחר מספיק איברים חיוביים כדי שסכומם יהיה מעט גדול מ־S, אח"כ נבחר מספיק איברים שליליים כך שהסכום הכולל עד כה יהיה מעט קטן מ־S, ושוב נבחר איברים חיוביים עד שהסכום הכולל מעט גדול מ־S, וחוזר חלילה. הסכום הזה "מתנודד סביב S", ומכיוון שהאיבר הכללי בטור שואף לאפס (כי הוא מתכנס בתנאי), בכל פעם שאנו עוברים את S איננו עוברים אותו בהרבה ולכן הטור החדש יתכנס אליו.

באופן פורמלי, עלינו לבנות תמורה (n_k) כך ש־ $\sum_{k=1}^{\infty} a_{n_k} = S$. נגדיר את n_k ברקורסיה על k באופן הבא: תחילה נקבע $n_1 = 1$. נניח כעת שהגדרנו $n_1, \ldots, n_k$ ויהי $T_k = \sum_{i=1}^{k} a_{n_i}$. נפריד לשני מקרים:

[6] (Bernhard Riemann, 1826-1866).

1. $T_k \leq S$. במקרה זה נגדיר את n_{k+1} להיות האינדקס הראשון $i \neq n_1, \ldots, n_k$ שמקיים $a_i \geq 0$.
2. $T_k > S$. במקרה זה נגדיר את n_{k+1} להיות האינדקס הראשון $i \neq n_1, \ldots, n_k$ שמקיים $a_i < 0$.

לפי הלמה הקודמת יש אינסוף מספרים חיוביים ושליליים בסדרה (a_n) ולכן בכל שלב באינדוקציה אפשר למצוא i כנדרש.

יש לוודא שסדרת האינדקסים (s_i) שהגדרנו היא תמורה. ברור מהבניה שבסדרה (n_k) אין חזרות, כלומר שאם $i \neq j$ אז $n_i \neq n_j$ (נמקו זאת!). לכן כדי להראות ש־(n_k) תמורה נותר להראות שכל מספר טבעי חיובי מופיע בה, דהיינו שלכל $i \in \mathbb{N}$ יש $j \in \mathbb{N}$ עם $n_j = i$.

נשים לב שאם עבור k מסוים מתקיים $T_k \leq S$ אז יש $k' > k$ כך ש־ $T_{k'} > S$. זו מסקנה מכך שלפי הלמה, $\sum a_n^+ = \infty$. באותו אופן, אם $T_k > S$ אז יש $k' > k$ כך ש־ $T_{k'} \leq S$. לכן במהלך ההגדרה של (n_k) כל אחד מהשלבים (1) ו־(2) מבוצע אינסוף פעמים.

יהי $i \in \mathbb{N}$, ונניח $a_i \geq 0$. אז מההגדרה ברור ש־ $n_j = i$ אמ"מ בשלב j של הגדרת הסדרה (n_k) ביצענו אפשרות (1) בפעם ה־i. לפי האמור לעיל יש i כזה, כי שלב (1) מבוצע אינסוף פעמים. באופן דומה, אם $a_i < 0$ אז יש j כך שבשלב j של הבניה מבצעים את (2) בפעם ה־i, ואז $n_j = i$.

אם כן (n_k) היא תמורה. כדי להשלים את ההוכחה עלינו להראות ש־ $T_k \to S$, כלומר, עלינו להראות שלכל $\varepsilon > 0$ מתקיים $|T_k - S| < \varepsilon$ לכל k גדול מספיק.

$a_n \to 0$ כי $\sum a_n$ מתכנס, ולכן גם $a_n^+, a_n^- \to 0$. כמו־כן $n_k \to \infty$ (למה?) ולכן $a_{n_k} \to 0$. יהי K כך שלכל $k > K$ מתקיים $|a_{n_k}| < \varepsilon$.

יהי $m > K$ אינדקס כך שבצעד m של הבנייה ביצענו את (1) ובזמן $m+1$ ביצענו את (2). יש m כזה כי כל אחת מהפעולות מבוצעת אינסוף פעמים במהלך הגדרת (n_k). כעת מכיוון שבשלב m ביצענו (1) הרי $T_m \leq S$. העובדה שבשלב $m+1$ ביצענו (2) משמעה ש־ $T_{m+1} > S$. אבל

$$S < T_{m+1} = T_m + a_{n_m} \leq S + a_{n_m} < S + \varepsilon$$

כי $m > K$ (היזכרו בהגדרת K!), וקיבלנו ש־

$$|T_{m+1} - S| < \varepsilon$$

כעת אנו טוענים שלכל $k > m$ מתקיים $|T_k - S| < \varepsilon$. ההוכחה באינדוקציה. עבור $k = m+1$ הוכחנו זאת למעלה. נניח עתה שידוע ש־ $|T_k - S| < \varepsilon$, ונראה ש־ $|T_{k+1} - S| < \varepsilon$. נניח למשל

$$S < T_k < S + \varepsilon$$

אז בשלב $k+1$ של הבנייה בחרנו את n_{k+1} כך ש־ $a_{n_{k+1}} < 0$, ולכן

$$T_{k+1} = T_k + a_{n_{k+1}} < S + \varepsilon$$

כי $-\varepsilon < a_{n_{k+1}} < \varepsilon$. באותו אופן

$$T_{k+1} = T_k + a_{n_{k+1}} > S - \varepsilon$$

ובסיכום

$$S - \varepsilon < T_{k+1} < S + \varepsilon$$

באופן דומה מראים שאם $S - \varepsilon < T_k \le S$ אז $|T_{k+1} - S| < \varepsilon$.

בסיכום, הראינו שלכל $\varepsilon > 0$ יש m כך שלכל $k > m$ מתקיים $|T_k - S| < \varepsilon$. מכאן שאכן $T_k \to S$, כפי שרצינו להוכיח. ∎

נסיים את הסעיף בהכללה של מושג הטור. לעתים מתעורר הצורך לסכום אוסף מספרים אינסופי שאינו נתון כסדרה, כלומר, אין דרך לבחור מיהו האיבר הראשון, השני וכו'. בהעדר סדר לא נוכל להגדיר את הסכומים החלקיים בצורה הרגילה, שכן הסכום החלקי ה־N הוא סכום N האיברים הראשונים, ואלה אינם מוגדרים במקרה שלפנינו. למשל, אם רוצים לסכום את המספרים הרציונליים $\mathbb{Q}$, כיצד נבחר את חמשת המספרים "הראשונים"? ההגדרה הבאה מתגברת על קושי זה על ידי שהיא עוסקת בכל הסכומים הסופיים המתקבלים מתת־קבוצות סופיות של האוסף (עבור סכום סופי הסר הסכימה אינו משנה, ולכן הוא מוגדר היטב גם בלי סדר).

הגדרה 6.5.12 תהי I קבוצה ונניח שלכל $i \in I$ נתון מספר a_i. נאמר שהטור $\sum_{i\in I} a_i$ מתכנס למספר a אם לכל $\varepsilon > 0$ קיימת תת־קבוצה סופית $J_0 \subseteq I$ כך שלכל קבוצת אינדקסים סופית $J \subseteq I$, אם $J_0 \subseteq J$ אז $|a - \sum_{j\in J} a_j| < \varepsilon$.

במקרה ש־ $I = \mathbb{N}$ אפשר להראות שהטור $\sum_{i\in I} a_i$ מתכנס אמ"מ $\sum_{n=1}^{\infty} a_n$ מתכנס בהחלט, ואז הערכים שווים. נדון בכך בתרגיל (14) למטה.

תרגילים

1. לכל זוג סדרות מהרשימה הבאה מצאו סדרה (n_k) כך ש־$\sum b_n$ מתקבל מ־$\sum a_n$ על ידי הכנסת סוגריים לפי (n_k).

(א) $(a_n) = 1,1,1,1,1,\dots$, $(b_n) = 1,2,3,4,5,\dots$.

(ב) $(a_n) = 1,-2,1,-2,1,-2,\dots$, $(b_n) = 0,-3,0,-3,0,-3,\dots$.

(ג) $(a_n) = (\frac{1}{n(n+1)})_{n=1}^{\infty}$, $(b_n) = (\frac{2}{(2n-1)(2n+1)})_{n=1}^{\infty}$.

2. מצאו טור $\sum_{n=1}^{\infty} a_n$ שאינו חסום מלעיל או מלרע ותארו הכנסת סוגריים כך שהטור המתקבל מתכנס.

3. נניח ש־ $a_n \to 0$ ושהטור $\sum b_k$ מתקבל מ־$\sum a_n$ על ידי הכנסת סוגריים במקומות (n_k). נניח גם שהסדרה $(n_{k+1} - n_k)_{k=1}^{\infty}$ חסומה. הוכיחו שאם $\sum b_k$ מתכנס אז גם $\sum a_n$ מתכנס, ולאותו סכום (גם במקרה שהסכום הוא אינסופי).

4. הוכיחו שאם $\sum b_k$ מתקבל מ־ $\sum a_n$ על ידי הכנסת סוגריים ואם $\sum a_n$ מתכנס בהחלט אז גם $\sum b_k$ מתכנס בהחלט. האם ההפך גם נכון?

5. אילו מהסדרות (n_k) הבאות הן תמורות?
   (א) $n_k = k^3$.
   (ב) $n_k = [\frac{k}{2}]$.
   (ג) $n_k = k + 1$ אם שארית החלוקה של k ב־10 קטנה מ־9, $n_k = 10[\frac{k}{10}]$ אחרת.
   (ד) $n_k = 2k$ אם k אי־זוגי ו־ $n_k = 2k - 1$ אם k זוגי.

6. תהי (n_k) תמורה. לכל k נגדיר את m_k להיות האינדקס היחיד שמקיים $n_{m_k} = k$.
   (א) הוכיחו ש־ $m_{n_k} = k$ לכל k.
   (ב) הסיקו שאם $\sum b_n$ מתקבל משינוי סדר של $\sum a_n$ אז גם $\sum a_n$ מתקבל משינוי סדר של $\sum b_n$.

7. האם ייתכן ש־$\sum a_n$ מתכנס בתנאי, אבל על ידי שינוי סדר אפשר לקבל טור מתכנס בהחלט?

8. תהי (a_n) סדרה ותהי (n_k) תמורה. הוכיחו ש־ $\lim_{n\to\infty} a_n = a$ אמ"מ $\lim_{k\to\infty} a_{n_k} = a$.

9. נאמר שתמורה (n_k) היא בעלת **תומך סופי** אם $n_k = k$ פרט למספר סופי של n־ים. הוכיחו שאם (n_k) היא תמורה בעלת תומך סופי אז $\sum a_n$ מתכנס אמ"מ $\sum a_{n_k}$ מתכנס, ואם הם מתכנסים הסכומים שווים.

10. יהי $\sum_{n=1}^{\infty} a_n$ טור מתכנס ותהי (n_k) תמורה של הטבעיים בעלת התכונה שהסדרה $(n_k - k)_{k=1}^{\infty}$ חסומה. הוכיחו שגם הטור החדש $\sum_{n=1}^{\infty} a_{\pi(n)}$ מתכנס, ולאותו סכום (בשפה ציורית יותר: אם M מספר, ואם משנים את סדר האיברים בטור באופן שאף איבר לא זז ממקומו המקורי יותר ממרחק M, אז לטור המתקבל יש אותן תכונות התכנסות כמו הטור המקורי).

11. תנו דוגמה לטור $\sum a_n$ שסכומיו החלקיים אינם חסומים מלעיל או מלרע, ותארו באופן מפורש שינוי סדר כך שהטור המתקבל מתכנס.

12. השלימו את הוכחת משפט 6.5.11 במקרה $S = \pm\infty$. כמו־כן תנו הוכחה לטענה הנוספת לגבי האפשרות למצוא תמורה עבורה הטור המתקבל אינו מתכנס אפילו במובן הרחב.

13. השתמשו בטכניקה דומה לזו שבעזרתה הוכחנו את משפט רימן כדי להוכיח שאם (a_n) היא סדרה חיובית כך ש־ $a_n \to 0$ אבל $\sum a_n = \infty$, אז לכל S ישנה סדרת סימנים (s_n) (כלומר $s_n \in \{1, -1\}$) כך שהטור $\sum s_n a_n$ מתכנס ל־S.

14. הראו שהגבול של טור במובן של הגדרה 6.5.12 הוא יחיד, כלומר אם טור $\sum_{i\in I} a_i$ מתכנס ל־s ול־t במובן של הגדרה 6.5.12 אז $s = t$.

15. תהי $(a_n)_{n=0}^{\infty}$ סדרה. הראו שהטור $\sum_{n\in\mathbb{N}} a_n$ מתכנס (במובן של הגדרה 6.5.12) אמ"מ $\sum_{n=0}^{\infty} a_n$ מתכנסת בהחלט, ובמקרה זה הם מתכנסים לאותו מספר.

16. (*) תהי I קבוצה ולכל $i \in I$ יהי $a_i \neq 0$ מספר ממשי. הראו שאם I אינה בת־מניה אז $\sum_{i\in I} a_i$ אינו מתכנס במובן של הגדרה 6.5.12 (המושג בן־מניה הוגדר בסעיף 5.12).

17. בשאלה זו נבחן את תכונות ההתכנסות של טור שהתקבל מטור אחר על ידי חלוקת איברי הטור המקורי לקבוצות אינסופיות של איברים, סכימת האיברים בכל קבוצה, וסכימת התוצאות. פעולה זו היא הכללה של שינוי סדר והכנסת סוגריים.

יהי $\sum_{n=1}^{\infty} a_n$ טור מתכנס בהחלט.

(א) הוכיחו את השוויון

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{2n} + \sum_{n=1}^{\infty} a_{2n+1}$$

(העובדה שהטורים באגף ימין מתכנסים נובעת מתרגיל (5) בעמוד 188).

(ב) (*) יהיו $(b_{1,k})_{k=1}^{\infty}, (b_{2,k})_{k=1}^{\infty}, \ldots$ סדרות ונניח שהאיברים $(b_{i,k})_{i,k=1}^{\infty}$ עוברים על האיברים (a_n) בדיוק פעם אחת (נסחו תכונה זו במדויק!) כל סדרה $(b_{n,k})_{k=1}^{\infty}$ היא תת־סדרה של $(a_n)_{n=1}^{\infty}$ ומשום כך הטורים $\sum_{k=1}^{\infty} b_{n,k}$ מתכנסים בהחלט (תרגיל (5) בעמוד 188). כעת נסמן $B_i = \sum_{k=1}^{\infty} b_{i,k}$. הוכיחו ש־$\sum_{i=1}^{\infty} B_i$ מתכנס בהחלט ושמתקיים השוויון $\sum_{n=1}^{\infty} a_n = \sum_{i=1}^{\infty} B_i$.

6.6 מכפלת טורים

נתבונן בשני סכומים סופיים $A = \sum_{n=0}^{N} a_n$ ו־ $B = \sum_{n=0}^{N} b_n$. ניתן להציג את המכפלה AB כסכום באופן הבא

$$AB = (\sum_{i=0}^{N} a_i)(\sum_{j=0}^{N} b_j) = \sum_{0\le i,j\le N} a_i b_j$$

אגף ימין הוא סכום כל הזוגות a_ib_j כש־i,j עוברים על כל הערכים $0,1,\ldots,N$. השוויון נובע מכך שכל אחד מהמחוברים a_ib_j מופיע בדיוק פעם אחת בביטוי האמצעי. באופן מפורט יותר,

$$\begin{aligned} AB &= (\sum_{i=0}^{N} a_i)(\sum_{j=0}^{N} b_j) \\ &= \sum_{i=0}^{N}(a_i \cdot \sum_{j=0}^{N} b_j) \\ &= \sum_{i=0}^{N}\sum_{j=0}^{N} a_ib_j \\ &= \sum_{0\leq i,j\leq N} a_ib_j \end{aligned}$$

כאן השתמשנו רק בחילוף, קיבוף ופילוג, ובדיון בסכומים כפולים שבעמוד 52.

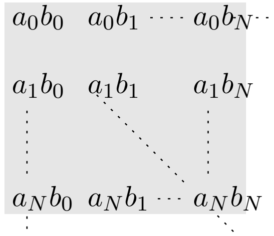
בהקבלה למקרה הסופי אפשר לצפות שכאשר כופלים יחד שני טורים מתכנסים $A=\sum_{n=0}^{\infty} a_n$ ו־ $B=\sum_{n=0}^{\infty} b_n$, המכפלה AB תהיה שווה, במובן כלשהו, לסכום האיברים a_ib_j כש־i,j עוברים על כל הערכים $0,1,2,\ldots$. אולם שלא כמו במקרה של סכומים סופיים, כאן לסדר הסכימה עלול להיות חשיבות: יש דרכים רבות לסדר את המספרים a_ib_j ולא כולם יתכנסו בהכרח למספר AB, כי, כפי שראינו בסעיף הקודם, באופן כללי שינוי סדר של איברים של טור אינו שומר על הסכום. בסעיף זה נבחן מתי וכיצד AB מתקבל מסכימת האיברים a_ib_j בצורות שונות.

בסעיף זה אנו מתחילים את הסכימה של טורים מהאינדקס 0 ולא מ־1. הסיבה היא שהניסוח של חלק מהתוצאות נוח ביותר עם המנהג הזה, אך אין לכך חשיבות רבה ואפשר לנסח גרסאות למשפטים לטורים המתחילים מכל אינדקס.

מהתוצאה שהזכרנו בתחילת הסעיף לסכומים סופיים אפשר להסיק תוצאה כללית לטורים. אם $A=\sum_{n=0}^{\infty} a_n$ הוא טור מתכנס עם סכומים חלקיים (S_N) ו־ $B=\sum_{n=0}^{\infty} b_n$ טור מתכנס עם סכומים חלקיים (T_N) אז מאריתמטיקה של גבולות מקבלים

$$\begin{aligned} AB &= \lim_{N\to\infty} S_N \cdot \lim T_N \\ &= \lim_{N\to\infty} S_N \cdot T_N \\ &= \lim_{N\to\infty} ((\sum_{n=0}^{N} a_n)(\sum_{n=0}^{N} b_n)) \\ &= \lim_{N\to\infty} \sum_{0\leq i,j\leq N} a_ib_j \end{aligned}$$

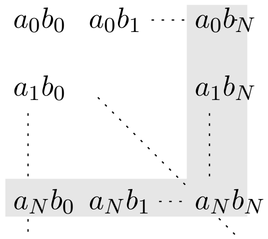
כאשר השוויון האחרון נובע מהשוויון שהזכרנו בתחילת הסעיף. אם מסדרים את האיברים a_ib_j בטבלה כמו באיור 6.6.1, רואים שהקבוצה $\{a_ib_j \,:\, 0 \le i,j \le N\}$, שאיבריה מסוכמים בשורה האחרונה לעיל, מתאימה לריבוע השמאלי העליון בטבלה, שממדיו $(N+1)\times(N+1)$. אפשר לפרק את סכום האיברים בריבוע זה לשני תת־סכומים: הראשון סכום האיברים בריבוע הקטן יותר שממדיו $N \times N$, והשני סכום האיברים הנותרים שהם האיברים בעמודה הימנית והשורה התחתונה של הטבלה, כפי שמסומן באיור 6.6.2. השורה התחתונה מורכבת מהאיברים מהצורה $a_{N,j}$ עבור $0 \le j \le N$, והעמודה הימנית מאיברים מהצורה $a_{i,N}$ עבור $0 \le i \le N$. אנו מקבלים את הפירוק הבא של הסכום:



איור 6.6.1 הסכום $\sum_{0\le i,j\le N} a_ib_j$

$$\sum_{0\le i,j\le N} a_ib_j = \sum_{0\le i,j\le N-1} a_ib_j + \sum_{\max\{i,j\}=N} a_ib_j$$

אפשר לפרק הלאה את הסכום השמאלי באגף ימין, שהוא מאותו סוג כמו הסכום שהתחלנו ממנו, ואם נמשיך כך N פעמים נקבל



איור 6.6.2 הסכום $\sum_{\max\{i,j\}=N} a_ib_j$

$$\sum_{0\le i,j\le N} a_ib_j = \sum_{k=0}^{N}(\sum_{\max\{i,j\}=k} a_ib_j)$$

ולכן

$$AB = \lim_{N\to\infty}\sum_{0\le i,j\le N} a_ib_j = \lim_{N\to\infty}\sum_{k=0}^{N}(\sum_{\max\{i,j\}=k} a_ib_j) = \sum_{k=0}^{\infty}(\sum_{\max\{i,j\}=k} a_ib_j)$$

(שימו לב שבסכום הפנימי באגף ימין יש רק מספר סופי של מחוברים). הביטוי האחרון הוא תוצאה חלקית מהסוג שרצינו: זהו טור שהתקבל על ידי סידור של a_ib_j בסדרה ואחר־כך הכנסת סוגריים.

מתי כל סידור של a_ib_j נותן טור שמתכנס ל־AB? אנו יודעים שדבר כזה יכול לקרות אם ורק אם הטור $\sum a_ib_j$ מתכנס בהחלט (התכנסות בהחלט של טור היא תכונה שאינה תלויה בסדר). מסתבר שכדי להבטיח זאת די שהטורים המקוריים יתכנסו בהחלט:

משפט 6.6.1 (משפט קושי על מכפלת טורים מתכנסים בהחלט) יהיו $A = \sum_{n=0}^{\infty} a_n$ ו־ $B = \sum_{n=0}^{\infty} b_n$ טורים מתכנסים בהחלט. אם הסדרה $(c_k)_{k=0}^{\infty}$ היא סידור כלשהו של האיברים a_ib_j $(i,j = 0,1,2,\ldots)$ אז הטור $\sum_{k=0}^{\infty} c_k$ מתכנס בהחלט וסכומו הוא AB.

הוכחה נניח ש־c_k הוא סידור של המספרים a_ib_j, כלומר, יש סדרות $(i(k))_{k=0}^{\infty}$ ו־ $(j(k))_{k=0}^{\infty}$ של אינדקסים כך ש־ $c_k = a_{i(k)}b_{j(k)}$ ולכל זוג (u,v) של אינדקסים קיים בדיוק אינדקס k אחד כך ש־ $i(k) = u$ ו־$j(k) = v$.

ראשית נראה ש־$\sum_{k=0}^\infty c_k$ מתכנס בהחלט. די שנראה שסדרת הסכומים החלקיים של $\sum_{k=0}^n |c_k|$ חסומה, כי זהו טור חיובי. נסמן $S=\sum_{k=0}^\infty |a_k|$, $T=\sum_{k=0}^\infty |b_k|$. בהינתן n, כל אחד מהאיברים $c_0,c_1,\ldots,c_n$ הוא מכפלה $c_k=a_{i(k)}b_{j(k)}$. יהי m הטבעי המקסימלי מבין $i(1),\ldots,i(n),j(1),\ldots,j(n)$. אז

$$\sum_{k=0}^n |c_k| = \sum_{k=0}^n |a_{i(k)}|\cdot|b_{j(k)}| \leq \sum_{0\leq u,v\leq m} |a_u|\cdot|b_v| = (\sum_{u=0}^m |a_u|)\cdot(\sum_{v=0}^m |b_v|)$$

כאשר האי־שוויון נובע מכך שבסכום מימינו מופיעים כל המחוברים שמופיעים משמאלו, ואולי מחוברים נוספים שכולם אי־שליליים. כיוון ש־ $\sum_{i=0}^m |a_i| \leq S$ ו־ $\sum_{j=0}^m |b_j| \leq T$, נקבל ש־ $\sum_{k=0}^n |c_k| \leq S\cdot T$. מכיוון ש־$n$ היה שרירותי, סדרת הסכומים החלקיים של $\sum_{k=0}^n |c_k|$ חסומה מלעיל (על ידי ST), ולכן הטור $\sum_{k=0}^n |c_k|$ מתכנס, כלומר $\sum_{k=0}^\infty c_k$ מתכנס בהחלט, ובפרט הוא גם מתכנס.

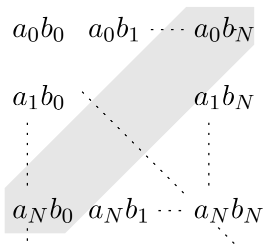
נותר להוכיח ש־ $\sum_{k=0}^\infty c_k = AB$. לאור האמור למעלה, סידור אחר של האיברים בטור $\sum_{k=0}^\infty c_n$ אינו משנה את הסכום, ולכן ניתן להניח שהאיברים מסודרים בצורה כזאת שלכל $N\geq 1$ טבעי, N^2 האיברים הראשונים $c_1,\ldots,c_{N^2}$ הם סידור של האיברים $\{a_ib_j\,:\,0\leq i,j<N\}$ (וודאו שיש סידור כזה של $\sum c_k$!). אם (R_N) היא סדרת הסכומים החלקיים של $\sum_{k=0}^\infty c_k$ מקבלים ש־

$$R_{N^2} = \sum_{0\leq i,j\leq N} a_ib_j = (\sum_{i=0}^N a_i)(\sum_{j=0}^N b_j) \to AB$$

כלומר התת־סדרה $(R_{N^2})_{N=0}^\infty$ של (R_N) מתכנסת ל־AB. אבל (R_N) מתכנסת וגבולה $\sum_{k=0}^\infty c_k$, ולכן $\sum_{k=0}^\infty c_k = AB$, כפי שרצינו להוכיח. ■

נעבור לדון בסידור אחר. עבור סכומים סופיים $A=\sum_{n=0}^N a_n$ ו־ $B=\sum_{n=0}^N b_n$ מתקיים השוויון

$$(\sum_{n=0}^N a_n)\cdot(\sum_{n=0}^N b_n) = \sum_{k=0}^{2N}(\sum_{\substack{i+j=k\\0\leq i,j\leq N}} a_ib_j)$$



איור 6.6.3 הסכום $\sum_{i+j=N} a_ib_j$

השוויון ברור כששמים לב, כרגיל, שכל ביטוי a_ib_j מופיע בדיוק פעם אחת בכל אגף. אם נסדר כמו קודם את האיברים $(a_{i,j})_{0\leq i,j\leq N}$ בטבלה ריבועית, אז לכל k הסכום $\sum_{i+j=k} a_{ij}$ מתאר את סכום האיברים הנמצאים על האלכסון ה־k בכיוון $\swarrow$ כאשר מתחילים לספור את האלכסונים מהפינה השמאלית העליונה, כמתואר באיור 6.6.3 (הוספנו את הדרישה $0\leq i,j\leq N$ כדי לא לסכום איברים הנמצאים מחוץ לריבוע הראשי $(N+1)\times(N+1)$). וודאו שאתם מבינים זאת!

נשאלת השאלה, האם נוסחה דומה תקפה לטורים, כלומר, בהינתן טורים מתכנסים $A=\sum_{n=0}^\infty a_n$ ו־ $B=\sum_{b=0}^\infty b_n$, האם מתקיים $AB=\sum_{k=0}^\infty \sum_{i+j=k} a_ib_j$? אם

שהטורים המקוריים מתכנסים בהחלט אז התשובה חיובית, שכך במקרה זה הטור $\sum_{k=0}^{\infty}\sum_{i+j=k} a_ib_j$ מתקבל על ידי שינוי סדר של הסדרה (c_k), המורכבת ממכפלות a_ib_j, ואז הכנסת סוגריים. לפי משפט קושי הטור $\sum_{k=0}^{\infty} c_k$ מתכנס בהחלט ל־AB, ולפי משפט 6.5.3 הכנסת סוגריים אינה משנה את סכומו.

המשפט הבא מבטיח ש־ $AB = \sum_{k=0}^{\infty}\sum_{i+j=k} a_ib_j$ גם כשרק אחד מהטורים המקוריים מתכנס בהחלט. כאן ובמהלך ההוכחה אנו מניחים במובלע שאינדקסים מקבלים רק ערכים אי־שליליים ואיננו מציינים זאת במפורש.

משפט 6.6.2 (משפט מרטנס)[7] יהיו $A = \sum_{n=0}^{\infty} a_n$, $B = \sum_{n=0}^{\infty} b_n$ טורים מתכנסים, ונניח ש־ $\sum a_n$ מתכנס בהחלט. נסמן $d_n = \sum_{i+j=n} a_ib_j$. אז הטור $\sum_{n=0}^{\infty} d_n$ מתכנס וסכומו הוא AB.

הוכחה לכל n נסמן

$$A_n = \sum_{k=0}^{n} a_k, \quad B_n = \sum_{k=0}^{n} b_k, \quad D_n = \sum_{k=0}^{n} d_k$$

כעת אפשר לכתוב

$$\begin{aligned} D_n &= d_0 + d_2 + \ldots + d_n \\ &= a_0b_0 + (a_0b_1 + a_1b_0) \\ &\qquad + (a_0b_2 + a_1b_1 + a_2b_0) + \ldots + (a_0b_n + \ldots + a_nb_0) \\ &= a_0B_n + a_1B_{n-1} + a_2B_{n-2} + \ldots + a_nB_0 \end{aligned}$$

השוויון האחרון התקבל על ידי קיבוץ יחד של המחוברים ש־a_0 מופיע בהם, קיבוץ המחוברים ש־a_1 מופיע בהם, וכן הלאה.

לכל $n = 0, 1, 2, \ldots$ נסמן $\beta_n = B_n - B$. קיבלנו־

$$D_n = a_0(B + \beta_n) + a_1(B + \beta_{n-1}) + \ldots + a_n(B + \beta_0)$$

על ידי פתיחת סוגריים וקיבוץ מחדש של האיברים, נקבל

$$D_n = A_n \cdot B + (a_0\beta_n + a_1\beta_{n-1} + \ldots + a_n\beta_0)$$

נסמן את סכום המחוברים בסוגריים באגף ימין ב־ $w_n = a_0\beta_n + \ldots + a_n\beta_0$. כיוון ש־ $A_n \cdot B \to A \cdot B$, ההוכחה תסתיים אם נצליח להוכיח ש־ $w_n \to 0$, כי אז $D_n \to AB$.

[7] Franz Mertens, 1840־1927.

נוכיח אם כן ש־ $w_n \to 0$. נשים לב שהסדרות (a_n) ו־(β_n) מתכנסות ולכן חסומות. יהי M חסם משותף שלהן. לכל $0 \le k < n$ מתקיים

$$\begin{aligned} |w_n| &\le |a_0\beta_n| + \ldots + |a_k\beta_{n-k}| + |a_{k+1}\beta_{n-k-1}| + \ldots + |a_n\beta_0| \\ &\le M(|\beta_n| + \ldots + |\beta_{n-k}|) + M(|a_{k+1}| + \ldots + |a_n|) \\ &\le M(|\beta_n| + \ldots + |\beta_{n-k}|) + M \cdot r_k \end{aligned}$$

כאשר $r_k = \sum_{i=k+1}^{\infty} |a_i|$ הוא ה־k־זנב של הטור $\sum_{i=1}^{\infty} |a_i|$. אנו יודעים ש־ $r_k \to 0$, כי $\sum a_n$ מתכנס בהחלט. לכן בהינתן $\varepsilon > 0$ אפשר לבחור k כך ש־ $0 \le Mr_k < \varepsilon$. אז לכל $n > k$ מתקיים

$$|w_n| \le M(|\beta_n| + \ldots + |\beta_{n-k}|) + \varepsilon$$

מצד שני מכיוון ש־k קבוע, $\lim_{n\to\infty}(|\beta_n| + \ldots + |\beta_{n-k}|) = 0$ (כי כל אחד מ־k המחוברים שואף ל־0) ולכן לכל n גדול מספיק מתקיים $M(|\beta_n|+\ldots+|\beta_{n-k}|) < \varepsilon$. לכן לכל n גדול מספיק

$$|w_n| \le \varepsilon + \varepsilon = 2\varepsilon$$

מכאן ש־ $w_n \to 0$. ■

דוגמאות

1. נתחיל בדוגמה נגדית המראה שלא ניתן לוותר על תנאי ההתכנסות בהחלט במשפטים האחרונים. יהיו $a_n = b_n = \frac{(-1)^n}{\sqrt{n+1}}$. אז $\sum a_n, \sum b_n$ מתכסים כי הם טורי לייבניץ. נגדיר

$$d_n = \sum_{i+j=n} a_i b_j = (-1)^n \sum_{i=0}^{n} \frac{1}{\sqrt{i+1}} \cdot \frac{1}{\sqrt{n-i+1}} =$$

מכיוון שלכל $0 \le i \le n$ מתקיים $\sqrt{i+1}, \sqrt{n-i+1} \le \sqrt{n+1}$ הרי ש־

$$\sqrt{i+1} \cdot \sqrt{n-i+1} \le n+1$$

ולכן

$$|d_n| = \sum_{i=1}^{n} \frac{1}{\sqrt{i}} \cdot \frac{1}{\sqrt{n-i}} \ge \sum_{i=1}^{n} \frac{1}{n+1} = \frac{n}{n+1} \to 1$$

וממילא הסדרה $\sum d_n$ אינה מתכנסת, כי $d_n \not\to 0$. זה גם גורר שאף סידור של $(a_ib_j)_{i,j\in\mathbb{N}}$ אינו מתכנס בהחלט.

2. לכל x נגדיר מספר $e(x)$ על ידי $e(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. הטור המגדיר את $e(x)$ מתכנס בהחלט (הוכיחו!). נראה כעת ש־ $e(x+y) = e(x) \cdot e(y)$. ואמנם, לפי משפט מרטנס,

$$\begin{aligned}
e(x)e(y) &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \cdot \sum_{n=0}^{\infty} \frac{y^n}{n!} \\
&= \sum_{n=0}^{\infty} \sum_{i=0}^{n} (\frac{x^i}{i!})(\frac{y^{n-i}}{(n-i)!}) \\
&= \sum_{n=0}^{\infty} \sum_{i=0}^{n} \frac{1}{n!} \cdot \frac{n!}{i!(n-i)!} x^i y^{n-i} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i=0}^{n} \frac{n!}{i!(n-i)!} x^i y^{n-i} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} (x+y)^n \\
&= e(x+y)
\end{aligned}$$

(במעבר מהשורה הראשונה לשנייה השתמשנו במשפט על שינוי סדר סכימה. במעברים האחרים ביצענו מניפולציות אלגבריות והשתמשנו בנוסחת הבינום כדי לקבל את השוויון $(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}$).

הערה בסעיף 510 נראה שלמעשה $e(x) = e^x$. שם גם הסביר כיצד הגענו לטור לטור המדובר.

תרגילים

1. הוכיחו או הפריכו את הטענות הבאות תחת ההנחה שהטורים $A = \sum_{n=0}^{\infty} a_n$ ו־ $B = \sum_{n=0}^{\infty} b_n$ מתכנסים, ותחת ההנחה שהם מתכנסים בהחלט.
   (א) $\sum_{n=0}^{\infty} (\sum_{k=0}^{\infty} a_k b_n) = AB$.
   (ב) $\sum_{n=0}^{\infty} \sum_{m=0}^{n} a_n a_{n-m} = A^2$.
   (ג) $\sum_{n=0}^{\infty} \sum_{k=0}^{n} (a_k b_{n-k} - b_k a_{n-k}) = 0$.
   (ד) $\sum_{n=0}^{\infty} \sum_{k=-n}^{\infty} a_n b_{n+k} = AB$.
2. נסחו גרסה של משפט קושי על שינוי סדר סכימה עבור מכפלת הטורים $(\sum_{n=1}^{\infty} a_n)(\sum_{n=1}^{\infty} b_n)$ (יש צורך לשנות את הניסוח כיוון שהסכימה מתחילה מ־1, במקום מ־0. אפשר לבדוק תחילה מה הנוסחה לסכומים סופיים של שלושה מחוברים).
3. הוכיחו שאם $\sum_{n=0}^{\infty} a_n$, $\sum_{n=0}^{\infty} b_n$ מתכנסים בהחלט ל־ A, B בהתאמה אז $\sum_{(i,j)\in\mathbb{N}\times\mathbb{N}} a_i b_j$ מתכנס ושווה ל־ AB במובן של הגדרה 6.5.12. הכלילו זאת למכפלה של מספר סופי שרירותי של טורים מתכנסים בהחלט.

4. לכל $x \in \mathbb{R}$ נגדיר מספרים $s(x), c(x)$ על ידי

$$\begin{aligned} s(x) &= \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n+1}}{(2n+1)!} \\ c(x) &= \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n}}{(2n)!} \end{aligned}$$

הוכיחו שלכל $x, y \in \mathbb{R}$ מתקיימים השוויונות הבאים (כדאי להיעזר במשפט הבינום):

(א) $s(x)^2 + c(x)^2 = 1$.

(ב) $s(x+y) = s(x)c(y) + s(y)c(x)$.

(ג) $c(x+y) = c(x)c(y) - s(x)s(y)$.

הערה בסעיף 510 נראה ש־ $c(x) = \cos x$ וש־ $s(x) = \sin(x)$, ונסביר את מקור הטורים האלה.

5. בשאלה זו נראה שיש אינסוף מספרים ראשוניים[8] (המספרים הראשוניים הוגדרו בתרגיל (22) בעמוד 57). נניח בשלילה שיש רק N מספרים ראשוניים שונים, ונסמן אותם ב־$p_1, p_2, \ldots, p_N$. נגדיר $a_n = \frac{1}{1-1/p_n}$ (נציין שהמספר 1 אינו נחשב ראשוני), כך ש־ $a_n = \sum_{k=0}^{\infty}(\frac{1}{p_n})^k$, ונשים לב שאלה טורים מתכנסים בהחלט. לכן לפי שאלה (3) למעלה, מתקיים

$$a_1 \cdot a_2 \cdot \ldots \cdot a_N = \sum_{i_1, i_2, \ldots, i_N \in \mathbb{N}} (p_1^{i_1} \cdot p_2^{i_2} \cdot \ldots \cdot p_N^{i_N})^{-1}$$

ובפרט הטור באגף ימין מתכנס. השתמשו בכך שכל מספר טבעי ניתן לכתיבה כמכפלה של ראשוניים כדי להראות שהסדרה ההרמונית מופיעה בין איברי הטור באגף ימין, והגיעו לסתירה.

6.7 הצגת המספרים כשברים עשרוניים אינסופיים

בסעיף זה נכליל את שיטת הייצוג העשרוני שפגשנו בסעיף 3.7 באופן שיאפשר לייצג כל מספר ממשי.

כזכור, ספרה עשרונית היא מספר שלם בין 0 ל־9, ובהינתן סדרה סופית $(a_n)_{n=1}^N$ של ספרות, הסימן $0.a_1a_2a_3\ldots a_N$ מציין את המספר הרציונלי $\sum_{n=1}^N \frac{a_n}{10^n}$ ונקרא השבר העשרוני (הסופי) שספרותיו $a_1, \ldots, a_n$. ההגדרה הבאה דומה אלא שהסכום הסופי מתחלף בטור:

[8] עובדה זו הייתה ידועה ליוונים והוכחה שלה נמצאת בכתביו של אוקלידס (Euclid of Alexandria, 325־265 לפנה"ס). ההוכחה שהצגנו כאן מאוחרת יותר, והתגלתה על ידי המתמטיקאי ליאונרד אוילר (Leonhard Euler, 1707־1783).

הגדרה 6.7.1 בהינתן סדרה אינסופית $(a_n)_{n=1}^{\infty}$ של ספרות, הסימן $0.a_1a_2a_3\dots$ הוא **הצגה העשרונית**[9] של המספר $\sum_{n=1}^{\infty}\frac{a_n}{10^n}$ ונקרא **שבר עשרוני אינסופי** בעל ספרות (a_n).

יש לוודא שהטור בהגדרה מתכנס. ואמנם, מכיוון שספרה היא מספר בין 0 ל־9 מתקיים $0 \le a_n \cdot 10^{-n} \le 10^{-(n-1)}$ לכל n, ולכן הטור החיובי $\sum_{n=1}^{\infty}\frac{a_n}{10^n}$ נשלט על־ידי הטור $\sum_{n=1}^{\infty}10^{-(n-1)}$, שהוא טור מתכנס, וממבחן ההשוואה הטור $\sum a_n \cdot 10^{-n}$ מתכנס. יתר על כן הטור מקיים

$$0 \le 0.a_1a_2a_3\ldots = \sum_{k=1}^{\infty}\frac{a_k}{10^k} \le \sum_{k=1}^{\infty}\frac{9}{10^k} = 9\cdot\sum_{k=1}^{\infty}\frac{1}{10^k} = 9\cdot\frac{1}{9} = 1$$

(השתמשנו בנוסחה לטור גאומטרי). מהאי־שוויון נובע שמספר מהצורה $0.a_1a_2\ldots$ שייך לקטע $[0,1]$. השאלה המרכזית העומדת לפנינו היא האם אפשר לייצג כל מספר בקטע $[0,1]$ כשבר עשרוני כזה. ואמנם,

משפט 6.7.2 לכל $x \in [0,1)$ יש ייצוג כשבר עשרוני אינסופי.

הוכחה נגדיר סדרת ספרות $(a_n)_{n=1}^{\infty}$ ברקורסיה באופן הבא. ראשית נחלק את הקטע $[0,1]$ לעשרה תת־קטעים באורך עשירית: $[0,\frac{1}{10}],[\frac{1}{10},\frac{2}{10}],\ldots,[\frac{9}{10},1]$. כל נקודה ב־$[0,1]$ שייכת לאחד מהקטעים האלה, ולכן יש ספרה k כך ש־ $x \in [\frac{k}{10},\frac{k+1}{10}]$. נגדיר $a_1 = k$.

בשלב הבא נתבונן בקטע $[\frac{a_1}{10},\frac{a_1+1}{10}]$ (שמכיל את x לפי הבחירה של a_1 בשלב הקודם) ונחלק אותו לעשרה קטעים באורך מאית כל אחד:

$$[\frac{a_1}{10},\frac{a_1}{10}+\frac{1}{100}] \quad , \quad [\frac{a_1}{10}+\frac{1}{100},\frac{a_1}{10}+\frac{2}{100}] \quad , \quad \ldots \quad , \quad [\frac{a_1}{10}+\frac{9}{100},\frac{a_1+1}{10}]$$

הקטעים האלה מכסים לגמרי את הקטע $[\frac{a_1}{10},\frac{a_1+1}{10}]$ ולכן קיימת ספרה k כך ש־ $x \in [\frac{a_1}{10}+\frac{k}{100},\frac{a_1}{10}+\frac{k+1}{100}]$. נקרא לספרה זו a_2.

באופן כללי, נניח שהגדרנו ספרות $a_1,\ldots,a_n$ כך ש־

$$x \in [x_n, x_n + \frac{1}{10^n}]$$

כאשר $x_n = 0.a_1a_2\ldots a_n$. נחלק את הקטע $[x_n,x_n+\frac{1}{10^n}]$, שאורכו 10^{-n}, לעשרה קטעים באורך $\frac{1}{10^{n+1}}$, דהיינו נתבונן בקטעים

$$[x_n, x_n+\frac{1}{10^{n+1}}] \quad , \quad [x_n+\frac{1}{10^{n+1}}, x_n+\frac{2}{10^{n+1}}] \quad \ldots \quad [x_n+\frac{9}{10^{n+1}}, x_n+\frac{1}{10^n}]$$

[9] לפעמים להצגה עשרונית קוראים גם **פיתוח עשרוני**.

אשר מכסים את הקטע $[x_n, x_n+\frac{1}{10^n}]$. אחד מהם בוודאי מכיל את x. נגדיר $a_n = k$ כך ש־k היא ספרה שמקיימת $x \in [x_n + \frac{k}{10^{n+1}}, x_n + \frac{k+1}{10^{n+1}}]$.

כך הגדרנו ברקורסיה סדרה $(a_n)_{n=1}^{\infty}$. נמשיך לסמן $x_n = 0.a_1a_2\ldots a_n$. מהבנייה ברור ש־ $x \in [x_n, x_n+\frac{1}{10^n}]$, כלומר

$$x_n \le x \le x_n + \frac{1}{10^n}$$

אבל x_n היא סדרה עולה וחסומה ולכן מתכנסת, וממשפט הסנדוויץ' נובע $x_n \to x$, וקיבלנו

$$x = \lim_{n\to\infty} x_n = \lim_{n\to\infty}\sum_{k=1}^{n}\frac{a_k}{10^k} = \sum_{k=1}^{\infty}\frac{a_k}{10^k} = 0.a_1a_2a_3\ldots$$

ומצאנו את הפיתוח העשרוני המבוקש ל־x. ■

עסקנו למעלה בייצוג עשרוני של מספרים בקטע $[0,1]$ אך בעזרתם ובעזרת השלמים אפשר לייצג כל מספר ממשי. נזכור שלכל מספר טבעי n יש כתיבה עשרונית $n = b_db_{d-1}\ldots b_1b_0$, כאשר b_i הן ספרות והסימון $b_db_{d-1}\ldots b_1b_0$ מייצג את המספר $\sum_{k=0}^{d} b_k10^k$ (סעיף 3.7). נוכל להשתמש בעובדה זו כדי לקבל ייצוג עשרוני של כל מספר ממשי חיובי:

הגדרה 6.7.3 עבור סדרות $(a_n)_{n=1}^{\infty}$ ו־$(b_n)_{n=0}^{d}$ של ספרות עשרוניות, המספר

$$x = \sum_{k=0}^{d} b_k10^k + \sum_{k=1}^{\infty}\frac{a_k}{10^k}$$

מסומן ב־$b_db_{d-1}\ldots b_1b_0.a_1a_2a_3\ldots$, וסימון זה נקרא **ההצגה העשרונית** של x.

לכל $x \ge 0$ מתקיים $x = [x] + \{x\}$. אנו יודעים מסעיף 3.7 למצוא ייצוג עשרוני למספר השלם $[x]$ ולפי המשפט למעלה יש ל־$\{x\}$ ייצוג עשרוני אינסופי. אנו מסיקים שיש ל־x הצגה עשרונית, כלומר לכל מספר חיובי יש פיתוח עשרוני. הוספת הסימן "$-$" מאפשר לייצג גם מספרים שליליים. נסכם זאת במשפט:

משפט 6.7.4 לכל מספר ממשי יש ייצוג עשרוני.

האם יש התאמה מלאה בין מספרים לפיתוח העשרוני שלהם? ברור שאם הספרות בפיתוח עשרוני של שני מספרים שווים בהתאמה אז המספרים שווים, כלומר, אם $a_n = b_n$ לכל n אז $0.a_1a_2\ldots = 0.b_1b_2\ldots$. ההפך אינו נכון. למשל, מתקיים $1.000\ldots = 0.999\ldots$, כי

$$0.999\ldots = \sum_{n=1}^{\infty}\frac{9}{10^n} = 9\cdot\sum_{n=1}^{\infty}10^{-n} = 9\frac{1}{10(1-1/10)} = 1$$

(השתמשנו כאן בנוסחה לטור גאומטרי). חישוב דומה מראה שאם $a_1 \dots a_k$ ספרות אז לכל ספרה $c \geq 1$ מתקיים

$$0.a_1 \dots a_k c0000 \dots = 0.a_1 \dots a_k(c-1)9999 \dots$$

(הוכיחו זאת!). שימו לב שאלה מספרים רציונליים כי אחת ההצגות שלהם היא למעשה שבר עשרוני סופי. נציין שמספרים כאלה מתאימים לקצוות של הקטעים שהגדרנו בהוכחה של משפט 6.7.4. ואמנם אם $x \in [0,1]$ ואם נבצע את תהליך ייצור הספרות של x כמו בהוכחת המשפט, אז במידה ובשלב ה־n של התהליך x הוא הקצה המשותף של שניים מהקטעים שהגדרנו באותו שלב אז יש שתי בחירות אפשריות לספרה ה־n של x, ובשלבים הבאים הספרות יהיו תמיד 0 או תמיד 9 לפי הבחירה שעשינו.

מסתבר שהמקרה שתואר לעיל הוא היחיד שבו למספר יש יותר מהצגה אחת כשבר עשרוני אינסופי:

טענה 6.7.5 אם $x = 0.a_1a_2 \dots$ ו־ $y = 0.b_1b_2 \dots$ ואם $x = y$ אז או ש־ $a_n = b_n$ לכל n, או שיש אינדקס k כך שהספרות שוות עד למקום ה־$k-1$ ומהמקום ה־k אחת הסדרות היא $c, 0, 0, 0, 0, \dots$ והשנייה היא $(c-1), 9, 9, 9, 9, \dots$ לאיזושהי ספרה $c \geq 1$. בפרט, לכל מספר ממשי יש לכל היותר שתי הצגות עשרוניות שונות.

הוכחה יהי k האינדקס כך ש־ $a_i = b_i$ לכל $i < k$ אבל $a_k \neq b_k$. נניח בלי הגבלת הכלליות ש־ $a_k > b_k$. אז השוויון $0.a_1a_2 \dots = 0.b_1b_2 \dots$ גורר

$$\frac{a_k}{10^k} + \sum_{n=k+1}^{\infty} \frac{a_n}{10^n} = \frac{b_k}{10^k} + \sum_{n=k+1}^{\infty} \frac{b_n}{10^n}$$

(למה?), ולאחר העברת אגפים וכפל ב־ $\frac{1}{10^k}$ נקבל

$$a_k - b_k = \sum_{n=1}^{\infty} \frac{b_{k+n} - a_{k+n}}{10^n}$$

אגף שמאל הוא מספר שלם כי הוא הפרש של ספרות והוא חיובי מההנחה $a_k > b_k$. מצד שני הטור $\sum_{n=1}^{\infty} \frac{b_{k+n} - a_{k+n}}{10^n}$ נשלט על ידי הטור $\sum_{n=1}^{\infty} \frac{9}{10^n} = 1$ ולכן

$$1 \leq a_k - b_k = \sum_{n=1}^{\infty} \frac{b_{k+n} - a_{k+n}}{10^n} \leq \sum_{n=1}^{\infty} \frac{9}{10^n} = 1$$

ולכן $a_k - b_k = 1$. כמו־כן מכיוון שהאגפים שווים כל האי־שוויונות הם שוויונות ולכן

$$\sum_{n=1}^{\infty} \frac{9}{10^n} = \sum_{n=1}^{\infty} \frac{b_{k+n} - a_{k+n}}{10^n}$$

אנו משאירים כתרגיל את הבדיקה שזה ייתכן רק אם $b_{k+n} - a_{k+n} = 9$ לכל $n \geq 1$.
מכיוון ש־ a_n, b_n הן ספרות נובע $b_{k+n} = 9$, $a_{n+k} = 0$ לכל $n \geq 1$, כפי שרצינו. ■

נציג עתה שיטה אלגברית יותר למציאת הפיתוח העשרוני של מספר $x \in [0,1)$. נניח שהפיתוח העשרוני של x הוא $x = 0.a_1a_2a_3\ldots$. המטרה שלנו היא לחלץ את הספרות $a_1, a_2, \ldots$. נשים לב שמכיוון שמתקיים

$$10x = 10 \cdot \sum_{k=1}^{\infty} \frac{a_k}{10^k} = a_1 + \sum_{k=2}^{\infty} \frac{a_k}{10^{k-1}} = a_1 + 0.a_2a_3a_4\ldots$$

הרי שהספרה a_1 נתונה על ידי הערך השלם של $10x$, דהיינו $a_1 = [10x]$. כמו־כן המספר $y = 0.a_2a_3\ldots$ נתון על ידי החלק השבור של $10x$, דהיינו על ידי $y = \{10x\}$. כדי לקבל את הספרה השנייה a_2 נוכל לפעול באותה דרך: מתקיים $a_2 = [10y]$, ומידע על יתר הספרות מגולם במספר $z = 0.a_3a_4\ldots = \{10y\}$. נוכל להמשיך ולקבל $a_3 = [10z]$, וכן הלאה.

ננסח מחדש את התהליך שתיארנו. בהינתן מספר $x \in [0,1)$, נגדיר ברקורסיה סדרה $(r_n)_{n=0}^{\infty}$ של מספרים ב־$[0,1[$ (זו הסדרה המתאימה לסדרה $x, y, z\ldots$ שתיארנו קודם) וסדרה של ספרות $(a_n)_{n=1}^{\infty}$. תחילה נגדיר $r_0 = x$. בהינתן r_n, נגדיר

$$\begin{aligned} a_{n+1} &= [10r_n] \\ r_{n+1} &= \{10r_n\} \end{aligned}$$

שימו לב ש־ $r_n \in [0,1)$ ולכן a_n היא אכן ספרה, ולפי מה שראינו למעלה, תהליך זה נותן את הספרות בפיתוח של x.

עובדה מעניינת היא שאפשר לאפיין את המספרים הרציונליים במונחים של הפיתוח העשרוני האינסופי שלהם.

הגדרה 6.7.6 תהי (a_n) סדרה. נאמר שהיא **מחזורית** (periodic) אם יש $m > 0$ טבעי כך ש־ $a_n = a_{n+m}$ לכל n. אם זה נכון לכל n גדול מספיק נאמר שהיא **מחזורית החל ממקום מסוים**.

למשל, הסדרה $0, 1, 0, 1, 0, 1, \ldots$ היא מחזורית עם מחזור 2. לעומתה, הסדרה $0, 0, 1, 0, 1, 0, 1, 0, \ldots$ היא מחזורית עם מחזור שתיים החל מהמקום השני, והסדרה ההרמונית $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, ,$ אינה מחזורית.

טענה 6.7.7 יהי $x \in [0,1)$. אז x רציונלי אמ"מ סדרת הספרות בהצגה העשרונית שלו מחזורית החל ממקום מסוים.

הערה אם יש למספר שני פיתוחים עשרוניים אז אנו כבר יודעים שהוא רציונלי ושני הפיתוחים בחזוריים החל ממקום מסויים. לכן בהוכחת הטענה אפשר להניח שהפיתוח ל x יחיד.

הוכחה נניח ש־ $x = 0.a_1a_2\ldots$ הוא פיתוח עשרוני של x ושהוא מחזורי, דהיינו נניח שיש N ו־ $m > 0$ כך שלכל $n > N$ מתקיים $a_n = a_{n+m}$. נגדיר $y = 10^N x$ ו־ $z = 10^{N+m}x$. אז

$$\begin{aligned} y &= u + 0.a_{N+1}a_{N+2}\ldots \\ z &= v + 0.a_{N+m+1}a_{N+m+2}\ldots \end{aligned}$$

כאשר u, v מספרים שלמים. אבל מההגדרה של N, m מתקיים $a_{N+i} = a_{N+m+i}$ לכל $i \geq 1$, כלומר

$$0.a_{N+1}a_{N+2}\ldots = 0.a_{N+m+1}a_{N+m+2}\ldots$$

ולכן

$$z - y = v - u \in \mathbb{Z}$$

מצד שני,

$$z - y = 10^{N+m}x - 10^N x = x(10^{N+m} - 10^N)$$

ולכן

$$x = \frac{v-u}{10^{N+m} - 10^N} \in \mathbb{Q}$$

כעת נוכיח את הכיוון השני. יהי $x \in [0,1)$ רציונלי ויהיו $p, q \in \mathbb{Z}$ כך ש־ $x = \frac{p}{q}$. יהיו r_n המספרים המוגדרים כמו בדיון שאחרי משפט 6.7.4, ונשים לב שהם מקיימים את יחס הרקורסיה

$$\begin{aligned} r_0 &= \frac{p}{q} \\ r_{n+1} &= 10r_n - a_{n+1} \end{aligned}$$

מאחר שהספרות a_n הן מספרים שלמים, המספרים r_n הם רציונליים ואפשר לרשום אותם בצורה

$$r_n = \frac{m_n}{q}$$

עבור מספר שלם m_n כלשהו. ואמנם, $m_0 = p$ וקל לבדוק ש־m_n מקיים את הנוסחה

$$m_{n+1} = 10m_n - qa_{n+1}$$

(וודאו זאת!).

אחת התכונות של r_n היא ש־ $r_n \in [0,1)$. נשים לב שיש בדיוק q מספרים רציונליים בקטע $[0,1]$ שהמכנה שלהם הוא q, דהיינו המספרים $0, \frac{1}{q}, \frac{2}{q}, \dots, \frac{q-1}{q}$. לכן לא ייתכן שכל ה־$r_i$־ים שונים, ובהכרח קיימים שני אינדקסים $i < j$ כך ש־ $r_i = r_j$. אבל המספר r_n קובע את r_{n+1} לחלוטין (כי $r_{n+1} = \{10r_n\}$) ונובע ש־ $r_{i+1} = r_{j+1}$, ובאינדוקציה שלכל $n \in \mathbb{N}$ מתקיים $r_{i+n} = r_{j+n}$. בפרט, עבור $k = j - i$, לכל $n > i$ יוצא ש־ $r_n = r_{i+(n-i)} = r_{j+(n-i)} = r_{n+k}$, וקיבלנו שהסדרה (r_n) מחזורית החל ממקום מסויים. מכיוון ש־ $a_n = [10r_{n-1}]$, יוצא מיד שגם (a_n) מחזורית החל ממקום מסויים, כפי שרצינו. ■

הטענה האחרונה מאפשרת לרשום במפורש מספרים אי־רציונליים חדשים. למשל, נגדיר את הסדרה

$$a_n = \begin{cases} 0 & \text{היא חזקה של 2 } n \\ 1 & \text{אחרת} \end{cases}$$

ויהי

$$x = 0.a_1a_2a_3\ldots = 0.1101000100000001000000000000000100000000\ldots$$

אז x אי־רציונלי, כי אין זה נכון שהסדרה (a_n) מחזורית החל ממקום מסוים (למשל, כי לכל $m > 0$, מתקיים $a_{2^m} = 1$ אבל $a_{2^m+m} = 0$).

כפי שהגדרנו פיתוח של מספרים שלמים בבסיסים שונים מעשר (עמוד 59) כך אפשר לעשות לגבי הפיתוח האינסופי:

הגדרה 6.7.8 יהי $b \geq 2$ מספר טבעי. ייצוג של מספר $x \in [0,1]$ **בבסיס** b היא סדרה $(a_n)_{n=1}^{\infty}$ עם $a_n \in \{0,1,2,\dots,b-1\}$ כך ש־ $x = \sum_{k=1}^{\infty} \frac{a_k}{b^k}$. אז רושמים $x = (0.a_1a_2\ldots)_b$ (ואם b מובן מההקשר רושמים פשוט $x = 0.a_1a_2\ldots$).

למשל, במדעי המחשב נוהגים לייצג מספרים בבסיס שתיים (הצגה בינרית). אז הספרות הן $0,1$. המספר חצי הוא $\frac{1}{2^1} = 0.1$ ורבע הוא $\frac{0}{2} + \frac{1}{2^2} = 0.01$. המספר שליש הוא $0.01010101\ldots$ (הוכיחו!).

כמו במקרה של הצגה עשרונית, לכל בסיס b ולכל מספר $x \in [0,1]$ קיימת הצגה של x בבסיס b. התכונות של הצגה כזאת דומות מאד למקרה $b = 10$. אנו משאירים כתרגיל את הניסוחים וההוכחות (ראו למטה).

תרגילים

1. כיצד אפשר לקבוע איזה משני מספרים גדול יותר בהינתן הכתיבה העשרונית שלהם?
2. תנו תיאור מפורש של המספרים שיש להם יותר מייצוג אחד בבסיס שתיים, וכן לגבי מספרים בבסיס שלוש.
3. האם יש שני מספרים שלמים m,n ומספר $x \in (0,1)$ כך שהפיתוח של x בבסיס m הוא מחזורי אבל הפיתוח בבסיס n אינו מחזורי החל ממקום מסוים?
4. בהינתן מספר $x = 0.a_1a_2a_3\ldots$ בכתיב עשרוני, מהו הפיתוח של x בבסיס 100?
5. יהי $\beta > 1$ מספר לא שלם. פיתוח־β (beta expansion) של מספר x הוא הצגה של המספר כטור $x = \sum_{n=0}^{\infty} \frac{a_n}{\beta^n}$, כאשר $a_n \in \{0,1,\ldots,[\beta]\}$.
   (א) לאילו מספרים יש פיתוח־β?
   (ב) (*) תנו דוגמה למספרים β ו־x כך של־x יש יותר מפיתוח־β אחד (תארו אותו, למשל, בעזרת פיתוח־β שלו!).

6.8 מכפלות אינסופיות

ההגדרה של מכפלה אינסופית מקבילה להגדרה של טור:

הגדרה 6.8.1 תהי a_n סדרה של מספרים חיוביים. תהי $(P_N)_{N=1}^{\infty}$ הסדרה המוגדרת על ידי

$$P_N = \prod_{n=1}^{N} a_n = a_1 \cdot a_2 \cdot \ldots \cdot a_N$$

הסדרה (P_N) נקראת **סדרת המכפלות החלקיות** של הסדרה (a_n), והסימן $\prod_{n=1}^{\infty} a_n$ נקרא **המכפלה האינסופית** המתאימה לסדרה (a_n). אם (P_N) מתכנסת למספר ממשי $P > 0$, נאמר שהמכפלה $\prod_{n=1}^{\infty} a_n$ **מתכנסת** ל־P, ונרשום $\prod_{n=1}^{\infty} a_n = P$. אם (P_N) אינה מתכנסת או אם $P_N \to 0$ אומרים שהמכפלה $\prod_{n=1}^{\infty} a_n$ **מתבדרת**.

אם $\sum a_n$ מתכנס אז $a_n \to 0$. הטענה המקבילה למכפלות היא

טענה 6.8.2 אם המכפלה $\prod a_n$ מתכנסת אז $a_n \to 1$.

הוכחה יהי $P = \prod_{n=1}^{\infty} a_n$. נסמן ב־$P_n$ את המכפלות החלקיות של הסדרה (a_n), ונשים לב שמתקיים $P_{n+1} = P_n \cdot a_{n+1}$, ולכן

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{P_{n+1}}{P_n} = \frac{\lim_{n\to\infty} P_{n+1}}{\lim_{n\to\infty} P_n} = \frac{P}{P} = 1$$

לפי אריתמטיקה של גבולות (מההנחה שהמכפלה מתכנסת נובע $P \neq 0$). ■

תורת המכפלות האינסופיות מקבילה לתורת הטורים, ואם תרצו תוכלו לפתח אותה בעצמכם (הדבר יהיה קל יותר לאחר שנגדיר בפרק הבא את הלוגריתם). בסעיף הנוכחי נוכיח קשר פשוט אך כללי למדי בין טורים למכפלות, ולשם כך נזדקק לשתי התוצאות הבאות, שמקבילות לגמרי למשפטים שלמדנו על טורים.

משפט 6.8.3 אם $a_n \geq 1$ לכל n אז המכפלה האינסופית $\prod_{n=1}^{\infty} a_n$ מתכנסת אמ"מ סדרת המכפלות הסופיות $P_N = \prod_{n=1}^{N} a_n$ היא סדרה חסומה מלעיל.

משפט 6.8.4 (תנאי קושי למכפלות) מכפלה אינסופית $\prod_{n=1}^{\infty} a_n$ מתכנסת אמ"מ לכל $\varepsilon > 0$ קיים N כך שלכל $n > N$ ולכל k טבעי מתקיים $|1 - \prod_{i=n}^{n+k} a_i| < \varepsilon$.

ההוכחה מושארת כתרגיל.

התוצאה העיקרית של הסעיף הזה היא משפט המקשר בין התכנסות של מכפלה אינסופית להתכנסות של טור:

משפט 6.8.5 יהי $a_n = 1 + \varepsilon_n$ כאשר $\varepsilon_n \geq 0$. אז המכפלה $\prod a_n$ מתכנסת אמ"מ הטור $\sum \varepsilon_n$ מתכנס.

הוכחה מכיוון ש־ $\varepsilon_n \geq 0$ הרי $\sum \varepsilon_n$ מתכנס אמ"מ סדרת הסכומים $S_N = \sum_{n=1}^{N} \varepsilon_n$ חסומה מלעיל, ומכיוון ש־ $1 + \varepsilon_n \geq 1$ הרי שהמכפלה $\prod(1+\varepsilon_n)$ מתכנסת אמ"מ סדרת המכפלות $P_N = \prod_{n=1}^{N}(1+\varepsilon_n)$ חסומה מלעיל.

ניעזר באי־שוויון $e^{\frac{x}{1+x}} \leq 1 + x \leq e^x$ אשר מתקיים לכל $x > -1$ (תרגיל (7) בעמוד 157). מהאי־שוויון $1 + x \leq e^x$ אנו מקבלים $1 + \varepsilon_n \leq \exp(\varepsilon_n)$ לכל n, ולכן

$$P_N = \prod_{n=1}^{N}(1+\varepsilon_n) \leq \prod_{n=1}^{N} e^{\varepsilon_n} = e^{\sum_{n=1}^{N}\varepsilon_n} = e^{S_N}$$

מכאן נובע מיד שאם (S_N) חסומה מלעיל אז (P_N) חסומה מלעיל, ולכן אם S_N מתכנסת אז P_N מתכנסת. מצד שני, מהאי־שוויון $1 + x \geq e^{\frac{x}{1+x}}$ מקבלים $1 + \varepsilon_n \geq e^{\varepsilon_n/(1+\varepsilon_n)})$ ולכן

$$P_N = \prod_{n=1}^{N}(1+\varepsilon_n) \geq \prod_{n=1}^{N} e^{\varepsilon_n/(1+\varepsilon_n)}) = e^{\sum_{n=1}^{N}\frac{\varepsilon_n}{1+\varepsilon_n}})$$

מכאן נובע שאם הסדרה (P_N) חסומה מלעיל אז הסדרה (T_N) המוגדרת על ידי $T_N = \sum_{n=1}^{N} \frac{\varepsilon_n}{1+\varepsilon_n}$ חסומה מלעיל גם כן. אבל (T_N) היא סדרת סכומים חלקיים של הטור החיובי $\sum \frac{\varepsilon_n}{1+\varepsilon_n}$, והיא חסומה אמ"מ הטור מתכנס, וזה קורה אמ"מ $\sum \varepsilon_N$ מתכנס (זהו תרגיל (6) מעמוד 179), כנדרש. ■

מסקנה 6.8.6 יהי $a_n = 1 + \varepsilon_n$ כאשר $-1 < \varepsilon_n \leq 0$. אז המכפלה $\prod a_n$ מתכנסת אמ"מ הטור $\sum \varepsilon_n$ מתכנס.

הוכחת המסקנה מושארת כתרגיל.

דוגמאות

1. $\prod_{n=1}^{\infty}(1+\frac{1}{2^n})$ מתכנס כי $\sum \frac{1}{2^n}$ מתכנס.
2. המכפלה $\prod_{n=2}^{\infty} \frac{n-1}{n}$ אינה מתכנסת, כי $\frac{n-1}{n} = 1 - \frac{1}{n}$ כי $\sum \frac{1}{n}$ מתבדר. מכיוון שסדרת המכפלות החלקיות P_N יורדת מונוטונית וחסומה מלרע על ידי 0, אנו מסיקים $P_N \to 0$, דהיינו $\prod_{n=1}^{\infty} \frac{n-1}{n}$ מתבדרת (אפשר לראות זאת גם בעזרת נימוק פשוט וישיר. איך?).

תרגילים

1. קבעו אילו מהמכפלות הבאות מתכנסות, וחשבו את גבולן:
   (א) $\prod_{n=1}^{\infty}(1+\frac{1}{n})$.
   (ב) $\prod_{n=1}^{\infty}(1-\frac{n^2-1}{n^2})$.
   (ג) $\prod_{n=1}^{\infty}(1-\frac{1}{n^2})$.
   (ד) $\prod_{n=1}^{\infty}(1-\frac{1}{n^2})^n$.
   (ה) $\prod_{n=1}^{\infty} e^{1/n}$.
2. עבור $n \geq 2$ יהי $\varepsilon_n = 1 + \frac{(-1)^n}{\sqrt{[n/2]}}$. הוכיחו שהמכפלה $\prod_{n=1}^{\infty}(1+\varepsilon_n)$ מתבדרת אף ש־ $\sum \varepsilon_n$ מתכנס. הסיקו שמשפט 6.8.5 אינו נכון באופן כללי.
3. הוכיחו את מסקנה 6.8.6 (רמז: $1-x = (1+\frac{x}{1-x})^{-1}$).
4. (*) הוכיחו את החיזוק הבא של משפט 6.8.5: אם $\sum \varepsilon_n$ טור מתכנס בהחלט (אבל לאו דווקא חיובי) אז $\prod(1+\varepsilon_n)$ מתכנס.

פרק 7

פונקציות, גבולות ורציפות

מושג הפונקציה התפתח מהצורך לתאר גדלים מספריים התלויים בגודל מספרי אחר. למשל אם חלקיק נע בקו ישר אז אפשר לתאר את מיקומו בעזרת מספר, והמיקום תלוי בזמן, שגם הוא מספר. פונקציות מתעוררות בהקשרים רבים אחרים והן אחד המושגים הבסיסיים ביותר במתמטיקה. בתחילת הפרק נגדיר פונקציות באופן כללי. נדון גם בפונקציות ממשיות, המתאימות מספרים למספרים, ובכמה דוגמאות חשובות של פונקציות כאלה.

לאחר מכן נגדיר את מושגי הגבול והרציפות של פונקציה. הגבול של פונקציה בנקודה הוא מושג הדומה למושג הגבול של סדרה, ונקדיש כמה סעיפים להגדרתו ופיתוח תכונותיו. בעזרת מושג הגבול נגדיר את מושג הרציפות של פונקציה: באופן גס אפשר לומר שפונקציה רציפה היא פונקציה שאין לה "קפיצות". נראה שפונקציות רציפות מקיימות כמה מהתכונות שאנו מצפים מפונקציות המתעוררות "בטבע".

7.1 מושג הפונקציה

הגדרה 7.1.1 פונקציה (function) f מקבוצה A לקבוצה B היא אובייקט המתאים לכל איבר $a \in A$ איבר יחיד $b \in B$ הנקרא **התמונה** (image) של a תחת פעולת f, ומסומן ב־$f(a)$ (קרי: f של a). נאמר גם ש־f **מעתיקה** את a ל־b, ונסמן זאת $f : a \mapsto b$. לפונקציה קוראים לפעמים גם **העתקה** (map).[1]

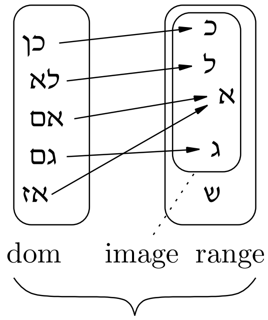
כדי לציין ש־f היא פונקציה מ־A ל־B נרשום $f : A \to B$ או $A \xrightarrow{f} B$. הקבוצה A נקראת **התחום** (domain) של f, ומסומנת $\operatorname{dom} f$. הקבוצה B נקראת **הטווח** (range) של f, ומסומנת $\operatorname{range} f$. קבוצת האיברים ב־B שהם תמונות של איברים

[1] הגדרנו את מושג הפונקציה כמושג עצמאי, אך אפשר להגדיר אותו בהתבסס על תורת הקבוצות, ראו למשל בספר [3].

ב־A, דהיינו הקבוצה $\{b \in B \mid \exists a \in A \quad b = f(a)\}$, נקראת **התמונה** (image) של f, ומסומנת $\operatorname{image} f$.[2]

יש ספרים שבהם הסימן $f(x)$ מציין את הפונקציה f ולא את התמונה של x תחת f. אז בסימון $f(x)$ האות x מפורשת כמשתנה, והסימון נועד לרמוז שאפשר להציב ערכים במקום x. אנו נקפיד להבחין בין הפונקציה f לבין התמונה $f(x)$ של x, למעט בכמה מקרים שנפרט אותם בבוא העת. כשרוצים להדגיש שהאות f מציינת פונקציה רושמים לפעמים $f(\cdot)$.

דרך אחת לתאר פונקציה היא בעזרת דיאגרמת חיצים. בתיאור זה מציירים את התחום והטווח זה מול זה, ומותחים חץ מכל איבר בתחום אל התמונה שלו.



איור 7.1.1 דיאגרמת חיצים

לדוגמה, הפונקציה f המתוארת באיור 7.1.1 מתאימה לכל מילה בקבוצת מילים מסוימת את האות הראשונה שלה. התחום הוא קבוצת המילים בצד שמאל של האיור, דהיינו

$$\operatorname{dom} f = \{\text{ כן,לא,אם,גם,אז }\}$$

והטווח הוא קבוצת האותיות בצד ימין של האיור, ש־

$$\operatorname{range} f = \{\text{ כ,ל,א,ג,ש }\}$$

יש איברים בטווח שאינם בתמונה, כמו האיבר "ש" (הוא אינו בתמונה כי אף חץ לא מגיע אליו). התמונה היא

$$\operatorname{image} f = \{\text{ כ,ל,א,ג }\}$$

ההגדרה של פונקציה אוסרת על יציאת שני חיצים מאותה נקודה בתחום, כי לכל איבר בתחום מתאים איבר יחיד בטווח, וזה אכן לא קורה בדוגמה למעלה. לעומת זאת, מותר שיהיו איברים בתמונה שנכנס אליהם יותר מחץ אחד. למשל, ל־"א" נכנסים שני חיצים: היא התמונה של "אם" וגם של "אז".

את שתי ההערות האחרונות ננסח באופן כללי כך: בהינתן פונקציה $f : A \to B$, ייתכן שיש $b \in B$ שאינו תמונה של אף איבר בתחום, כלומר $b \neq f(a)$ לכל $a \in A$, וייתכן גם שיש $b \in B$ שהוא תמונה של שני איברים בתחום, כלומר ייתכן שיש $a', a'' \in A$ עם $a' \neq a''$ אבל $f(a') = f(a'') = b$.

דוגמאות

1. תהי $f : \{\text{ בני אדם }\} \to \{\text{ בני אדם }\}$ הפונקציה שמתאימה לכל אדם את אמו. תחומה היא קבוצת כל בני האדם, הטווח שלה גם הוא קבוצת כל בני האדם, והתמונה היא קבוצת האמהות.

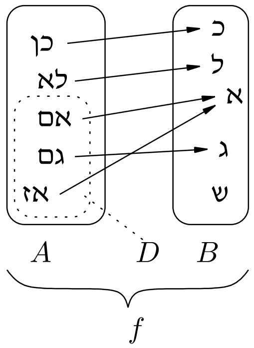
[2] התחום $\operatorname{dom} f$ נקרא לפעמים גם תחום ההגדרה של f, ואומרים ש־f מוגדרת ב־x אם $x \in \operatorname{dom} f$.

2. יהי g הכלל שמתאים לכל אדם את בנו הבכור. אז g אינו פונקציה, כי מהניסוח עולה שהתחום אמור להיות קבוצת בני האדם, אבל g אינו מוגדר לכל בן אדם אלא רק לאותם אנשים שיש להם בן, וקיימים בני אדם שאין להם ילדים.

3. לכל קבוצה A תהי $\mathrm{id}_A : A \to A$ הפונקציה המוגדרת על ידי $\mathrm{id}_A(x) = x$ לכל $x \in A$. אז id_A נקראת **פונקציית הזהות** (identity function) על A, כי היא מתאימה לכל איבר את עצמו.

4. פגשנו פונקציות בכמה מקומות בעבר. סדרה היא פונקציה $a : \mathbb{N}^+ \to \mathbb{R}$, אלא שהתמונה של n מסומנת ב־a_n במקום ב־$a(n)$. גם פעולות החיבור והכפל ב־$\mathbb{R}$ הן פונקציות: התחום שלהן הוא קבוצת הזוגות של מספרים ממשיים, וטווח (וגם התמונה) הוא קבוצת הממשיים.

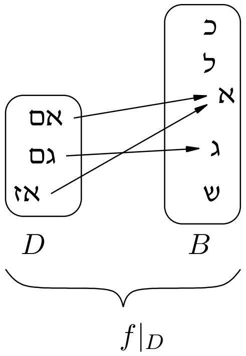
הגדרה 7.1.2 שתי פונקציות f, g הן שוות אם הן מתארות את אותה התאמה, כלומר אם הן מוגדרות באותן נקודות ($\mathrm{dom}\, f = \mathrm{dom}\, g$) ולכל איבר x בתחום המשותף מתקיים $f(x) = g(x)$.



אפשר באופן מלאכותי להקטין את התחום של פונקציה נתונה:

הגדרה 7.1.3 תהי $f : A \to B$ פונקציה ו־ $D \subseteq A$ תת־קבוצה. אז $f|_D$ מציין את הפונקציה מ־D ל־B המוגדרת על ידי $f|_D(d) = f(d)$ עבור $d \in D$. הפונקציה $f|_D$ נקראת **הצמצום** (restriction) של f ל־D, ונקראת גם f **מצומצם ל־**D.[3]

שימו לב שאם $f : A \to B$ ואם $D \subseteq A$ תת־קבוצה ממש של התחום (כלומר $D \neq A$) אז הפונקציות f ו־$f|_D$ אינן שוות, כי יש להן תחומים שונים.



איור 7.1.2 הצמצום של פונקציה

7.2 פונקציות ממשיות

בספר זה נתעניין בעיקר בפונקציות המתאימות מספרים למספרים. פונקציות כאלה מופיעות באופן טבעי במדעים המדויקים: למשל אפשר לדבר על הפונקציה f שהערך $f(t)$ שלה שווה למרחק של כדור הארץ מהשמש בזמן t. מרבית הפיזיקה הקלסית מבוססת מבחינה מתמטית על ניתוח של פונקציות כאלה והכללות שלהן.

הגדרה 7.2.1 פונקציה ממשית היא פונקציה שתחומה ותמונתה הם תת־קבוצות של המספרים הממשיים.[4]

[3] לעתים רושמים $f|_D$ גם כאשר f אינה מוגדרת בכל נקודה ב־D. אז הכוונה היא הצמצום של f לקבוצה $D \cap \mathrm{dom}\, f$.

[4] ישנם ספרים בהם מגדירים פונקציה ממשית להיות כל פונקציה עם ערכים ממשיים, ללא הגבלה על התחום.

פורמלית, ייתכנו פונקציות ממשיות שהתחום שלהן הוא קבוצה מסובכת של מספרים (למשל, פונקציה שתחומה היא קבוצת המספרים הרציונליים). אנו לא נתעניין במקרים כאלה: **התחומים של כל הפונקציות הממשיות שנפגוש יהיו קטעים או איחודים של קטעים**.

לעתים קרובות מגדירים פונקציה ממשית בעזרת נוסחה, כלומר כלל חשבוני או אחר המתאר כיצד לחשב את הערך $f(x)$ של הפונקציה בנקודה x. במקרים אלה נתכוון תמיד שהפונקציה ממשית גם אם לא ציינו זאת במפורש. כשנרצה להגדיר כך פונקציה נרשום "$f(x) = \ldots$" כאשר במקום שלוש הנקודות מופיע הכלל לחישוב של $f(x)$ מ־x.[5]

לדוגמה, הנוסחה $f(x) = x^2$ מגדירה את הפונקציה הממשית המתאימה לכל מספר ממשי את הריבוע שלו (כמובן, הנוסחה $f(y) = y^2$ מגדירה את אותה הפונקציה בדיוק).

לפעמים מגדירים פונקציה בעזרת נוסחה שאינה מוגדרת לכל מספר. במקרים אלה התחום הוא קבוצת המספרים שבה הנוסחה מוגדרת. למשל אם $f(x) = \frac{1}{x}$ אז התחום הוא $\mathbb{R} \setminus \{0\}$.

הערות

1. ניתן לכתוב נוסחאות שונות עבור אותה פונקציה. למשל יהיו $f, g : \mathbb{R} \to \mathbb{R}$ הפונקציות $f(x) = x$ ו־ $g(x) = 1 \cdot x$. אף שהנוסחאות המגדירות את f ו־g שונות, שתי הנוסחאות מתארות את אותה פונקציה, כי $f(x) = g(x)$ לכל $x \in \mathbb{R}$. בדוגמה זו רואים מיד שהפונקציות שוות, אך יש דוגמאות מסובכות יותר. למשל, נתבונן בפונקציות $f, g : (-1,1) \to \mathbb{R}$ המוגדרות על ידי

$$f(x) = \sqrt{1+x} \qquad , \qquad g(x) = 1 + \sum_{k=1}^{\infty}\left(\prod_{m=1}^{k} \frac{3-2m}{2m}\right)x^k$$

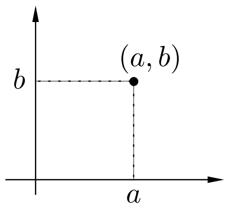
   למראית עין הפונקציות שונות לחלוטין, אך האמת היא שהפונקציות שוות (נוכיח זאת בפרק מאוחר יותר). דוגמה זו ממחישה שלפעמים אם לא נדע מראש ששתי נוסחאות שוות לא נוכל לנחש זאת, וגם אם ננחש, לא נוכל בהכרח להוכיח את השוויון. אין דרך כללית לקבוע אם שתי נוסחאות מגדירות את אותה הפונקציה.

2. לא כל הפונקציה ממשית ניתנת לתיאור בעזרת נוסחה. למעשה, "רוב" הפונקציות אינן ניתנות לתיאור כזה.

אפשר לתאר פונקציות ממשיות בצורה גאומטרית. נעזר בתיאור של המישור בעזרת קואורדינטות קרטזיות (Cartesian coordinates), שבו כל נקודה במישור מתאימה

[5] בחלק מהספרים מוצאים הגדרות מהסוג $y = x^2$. כאן הכוונה היא ש־y היא פונקציה של x, והיה מוטב לציין את התלות במפורש ולכתוב $y(x) = x^2$.

לזוג סדור של מספרים ממשיים. כזכור מפרק 2, קבוצת הזוגות הסדורים של מספרים ממשיים מסומנת ב־ $\mathbb{R}\times\mathbb{R}$, או בקיצור ב־ $\mathbb{R}^2$.

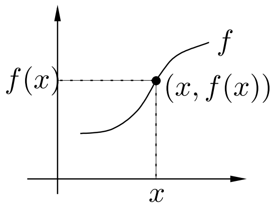


איור 7.2.1 נקודה במישור

כדי לזהות את $\mathbb{R}^2$ עם המישור נקבע במישור שני ישרים ניצבים, שאחד מהם אופקי ונקרא **ציר ה־x** והשני אנכי ונקרא **ציר ה־y**. נחשוב על הצירים כעל שני עותקים של ציר המספרים שמגמתם ימינה ולמעלה בהתאמה, ונניח שנקודת החיתוך בין הצירים היא נקודת האפס בשניהם. לזוג הסדור $(a,b)\in\mathbb{R}\times\mathbb{R}$ נתאים את הנקודה במישור שמרחקה מציר ה־x הוא b, ושמרחקה מציר ה־y הוא a. משיקולים גאומטריים יש נקודה יחידה כזו, ולכל נקודה P במישור יש זוג יחיד $(a,b)\in\mathbb{R}^2$ המזוהה עמו.

הגדרה 7.2.2 אם $f:A\to B$ פונקציה אז **הגרף** (graph) שלה היא קבוצת הזוגות הסדורים ב־$A\times B$ מהצורה $(x,f(x))$ עבור $x\in\operatorname{dom}f$, כלומר הקבוצה

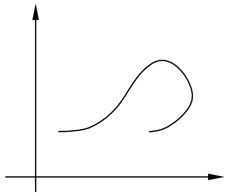
$$\operatorname{graph}f=\{(x,y)\,:\,x\in A\,,\,y=f(x)\}\subseteq A\times B$$



איור 7.2.2 גרף של פונקציה

הגרף של פונקציה ממשית הוא תת־קבוצה של $\mathbb{R}^2$, ואם נזהה את $\mathbb{R}^2$ עם המישור נוכל לחשוב על הגרף של f כעל קבוצה של נקודות במישור, או כעל "צורה", המורכבת מכל הנקודות $(x,f(x))$ עבור $x\in\operatorname{dom}f$, כפי שמתואר באיור 7.2.2.

שימו לב שישר אנכי ℓ העובר דרך הנקודה $(x,0)$ שעל ציר ה־x חותך את הגרף של פונקציה f אמ"מ $x\in\operatorname{dom}f$, ואז החיתוך יתרחש בנקודה $(x,f(x))$ בלבד. בפרט, העובדה ש־f היא פונקציה גורר שכל קו מאונך ℓ במישור חותך את הגרף בנקודה אחת לכל היותר. דוגמה של צורה במישור שאינה פונקציה מתוארת באיור 7.2.3. באופן דומה, ישר מאוזן במישור העובר דרך הנקודה $(0,y)$ שעל ציר ה־y חותך את גרף הפונקציה f אמ"מ y בתמונה של f, והחיתוך יתרחש בנקודה (x,y) אם ורק אם $y=f(x)$. ייתכן שיש מספר x־ים כאלה.

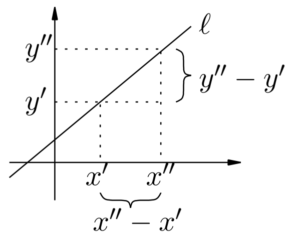

איור 7.2.3 צורה שאינה גרף של פונקציה

דוגמאות

1. יהי c מספר. הפונקציה $f:\mathbb{R}\to\mathbb{R}$ המוגדרת על ידי $f(x)=c$ לכל $x\in\mathbb{R}$ נקראת **הפונקציה הקבועה** (עם ערך c). הגרף שלה מורכב כולו מנקודות שהקואורדינטה השנייה שלהן היא c, כלומר, נקודות שמרחקן מציר ה־x הוא $|c|$ ונמצאות מעל ציר ה־x אם $c>0$ ומתחת לציר ה־x אם $c<0$ (אם $c=0$ מדובר בציר ה־x עצמו). לכן הגרף של f הוא קו ישר אופקי.

   כדי לומר בקצרה ש־f היא פונקציה קבועה עם ערך c רושמים $f\equiv c$, ואומרים ש־f שווה **זהותית** ל־c.

2. באופן כללי יותר יהיו $a,b\in\mathbb{R}$ ותהי $\ell:\mathbb{R}\to\mathbb{R}$ הפונקציה $\ell(x)=ax+b$. אז ℓ מתארת קו **ישר** אשר חותך את ציר ה־y בנקודה $\ell(0)=b$. אומרים שלישר יש **שיפוע** a. השיפוע מתאר את מידת השינוי של הקואורדינטה



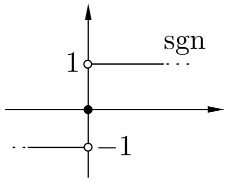
איור 7.2.4 קו ישר

השנייה של נקודה על הגרף כשמשנים את הקואורדינטה הראשונה. ואמנם, לכל $x',x''\in\mathbb{R}$ מתקיים

$$\ell(x')-\ell(x'')=(ax'+b)-(ax''+b)=a(x'-x'')$$

כלומר השינוי בקואורדינטה השנייה (ערך הפונקציה) גדול פי a מהשינוי בקואורדינטה הראשונה. מכאן שבהינתן פונקציה $f:\mathbb{R}\to\mathbb{R}$, אם ידוע ש־$f$ הוא ישר אז אפשר לשחזר את הנוסחה של f מהערכים של f כי השיפוע ניתן לחישוב מערכי הפונקציה, למשל על ידי הנוסחה $a=f(1)-f(0)$, ו־$b=f(0)$.

שימו לב שגאומטרית יש במישור גם ישרים המקבילים לציר ה־y. אלו אינם גרפים של פונקציות.



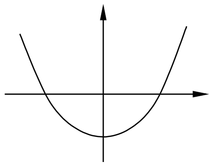
איור 7.2.5 פונקציית הסימן

3. לפעמים נגדיר פונקציה על ידי חלוקה של ההגדרה למקרים, כלומר חלוקת התחום לכמה תת־קבוצות ומתן הגדרה שונה בכל אחת מהן. למשל, **פונקציית הסימן** (sign function) היא הפונקציה $\operatorname{sgn}:\mathbb{R}\to\mathbb{R}$ המתאימה לכל מספר $x\neq 0$ את המספר ±1 או 0 לפי הסימן של x:

$$\operatorname{sgn}x=\begin{cases}-1 & x<0\\ 0 & x=0\\ 1 & x>0\end{cases}$$

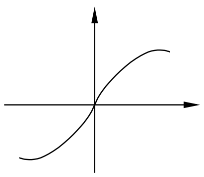
(מקובל לסמן $\operatorname{sgn}x$ במקום $\operatorname{sgn}(x)$ כשהדבר לא גורם בלבול). שימו לב ש־$x\cdot\operatorname{sgn}x=|x|$ לכל $x\in\mathbb{R}$.

הגרף של sgn מתואר באיור 7.2.5. שימו לב לשימוש בעיגולים המלאים והריקים כדי להדגיש שבנקודה 0 הפונקציה מקבלת ערך 0, ולא 1 או -1. נשתמש בסימונים דומים בהמשך בכל מקום שיש צורך להדגיש איזה מבין כמה ערכים מקבלת פונקציה בנקודה: נקודה ריקה על הגרף פירושה שהיא אינה שייכת לגרף.


איור 7.2.6 פונקציה זוגית

רבים מהמושגים הקשורים לפונקציות ממשיות מקורם ברעיונות גאומטריים, ובעתיד כדאי תמיד לשאול מה המשמעות הגאומטרית של הגדרה או טענה. למשל, יש כמה סוגי סימטריה שפונקציות ממשיות עשויות לקיים ושיש להן פירוש גאומטרי פשוט. הנה כמה מהן.

הגדרה 7.2.3 פונקציה ממשית f נקראת **זוגית** (even) אם לכל $x\in\operatorname{dom}f$ גם $-x\in\operatorname{dom}f$, ומתקיים $f(-x)=f(x)$. הפונקציה f נקראת **אי־זוגית** (odd) אם לכל $x\in\operatorname{dom}f$ גם $-x\in\operatorname{dom}f$ ומתקיים $f(-x)=-f(x)$.

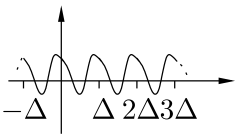

איור 7.2.7 פונקציה אי־זוגית

מבחינה גאומטרית, f זוגית אם החלק של הגרף שלה הנמצא משמאל לציר ה־y הוא השיקוף דרך ציר ה־y של החלק בגרף הנמצא בצד ימין של ציר ה־y. באופן דומה, f היא אי־זוגית אם החלק של הגרף שלה הנמצא משמאל לציר ה־y מתקבל מהחלק הימני של הגרף על־ידי שיקופו דרך ציר ה־y ואחר־כך דרך ציר ה־x.

דוגמאות

1. פונקציה מהצורה x^n עבור n טבעי היא זוגיות או אי־זוגיות בדיוק כאשר n זוגי או אי־זוגי, בהתאמה (למה?).
2. הפונקציה $h(x) = x + 1$ אינה זוגית או אי־זוגית, כי $h(1) = 2$ ואילו $h(-1) = 0$.

סוג נוסף של סימטריה הוא מחזוריות:



איור 7.2.8 פונקציה מחזורית

הגדרה 7.2.4 תהי f פונקציה ממשית. מספר $\Delta > 0$ נקרא **מחזור** (period) של f אם מתקיים ש־ $x \in \operatorname{dom} f$ אמ"מ $x + \Delta \in \operatorname{dom} f$ ולכל x כזה מתקיים

$$f(x + \Delta) = f(x)$$

אם לפונקציה יש מחזור אז היא נקראת **מחזורית** (periodic).

אם f מחזורית עם מחזור Δ, הגרף של f מורכב מחזרות תקופתיות של החלק מהגרף מעל הקטע $[0, \Delta)$ (או של כל קטע אחר באורך Δ). אם Δ הוא מחזור של f אז גם 2Δ כזה, כי לכל $x \in \mathbb{R}$

$$f(x + 2\Delta) = f((x + \Delta) + \Delta) = f(x + \Delta) = f(x)$$

ובאופן כללי $n \cdot \Delta$ הוא מחזור לכל $n > 0$ טבעי (הוכיחו!). אם יש מחזור קטן ביותר (כלומר: לקבוצת המחזורים יש איבר מינימלי) אז הוא נקרא **המחזור** של f (בהא הידיעה). לא תמיד יש מחזור קטן ביותר: למשל פונקציה קבועה היא מחזורית וכל מספר חיובי הוא מחזור שלה.

דוגמאות

1. תהי $f(x) = \{x\}$ כאשר $\{x\}$ הוא החלק השבור של x. אז f מחזורית עם מחזור 1, כי לכל x מתקיים $\{x + 1\} = \{x\}$. זהו גם המחזור הקטן ביותר (הוכיחו זאת!).
2. הפונקציה $f(x) = x$ אינה מחזורית, כי לכל $\Delta > 0$ מתקיים $f(0 + \Delta) = \Delta$ ולכן $f(0 + \Delta) \neq f(0)$ וממילא Δ אינו מחזור של f.

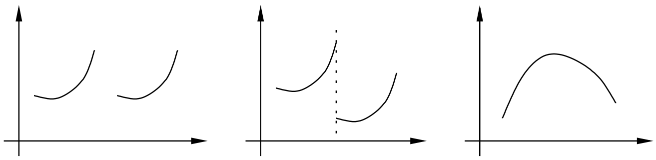
תרגילים

1. מהם תחומי ההגדרה של הפונקציות הממשיות הבאות?
   (א) $f(x) = \frac{1}{\sqrt{x} - 1}$.
   (ב) $f(x) = \frac{1}{x^2 - 4} + \frac{1}{x - 4}$.

(ג) $f(x) = \lim_{n\to\infty} x^n$.

(ד) $f(x) = \sum_{k=1}^{\infty} x^k$.

2. בתרשימים הבאים, סמנו על ציר ה־x וציר ה־y את התחום ואת התמונה של הפונקציות.



איור 7.2.9

3. לכל אחת מהפונקציות הבאות, קבעו האם היא זוגית, אי־זוגית, או מחזורית (ייתכנו צירופים שונים של התכונות וגם ייתכן שאף אחת אינה מתקיימת).

(א) $f \equiv 0$.

(ב) $f(x) = x^4 + 1$.

(ג) $f(x) = x^3 - 2x$.

(ד) $f(x) = x^4 + x^2$.

(ה) $f(x) = \sqrt{x}$.

(ו) $f(x) = [x]$.

(ז) $f(x) = \{x\}$.

(ח) $f(x) = |x| + x$.

(ט) sgn.

4. תהי $f : \mathbb{R} \to \mathbb{R}$ ונסמן ב־f^2 את הפונקציה $f^2(x) = f(x) \cdot f(x)$.

(א) נניח ש־ f^2 אי־זוגית. הוכיחו ש־f קבועה.

(ב) נניח ש־f^2 זוגית. האם בהכרח f זוגית או אי־זוגית?

5. יהיו f, g פונקציות עם מחזורים n, m בהתאמה, $n, m \in \mathbb{N}$. הוכיחו שהפונקציה $h(x) = f(x) + g(x)$ מחזורית, ומצאו את המחזור שלה. (*) מצאו דוגמה לפונקציות מחזוריות f, g כך ש־h אינה מחזורית.

7.3 הפונקציות האלמנטריות, חלק א'

מתוך שפע בלתי נתפש של פונקציות ממשיות אנו מסוגלים לתאר רק קומץ בצורה מפורשת. בסעיף זה נכיר כמה משפחות של פונקציות שתיאורן הוא אלגברי או גאומטרי. בסעיף 7.12 נכיר שתי משפחות נוספות. המשפחות האלה מהוות את אבני הבניין של מחלקת הפונקציות האלמנטריות. הפונקציות במחלקה זו מוכרות לכם בוודאי מבית הספר. הן חשובות מבחינה תאורטית ומופיעות בתפקידים חשובים גם ובמדעים אחרים.

פולינומים

הגדרה 7.3.1 פולינום (polynomial) הוא פונקציה $p:\mathbb{R}\to\mathbb{R}$ מהצורה

$$p(x)=\sum_{k=0}^{d}a_kx^k=a_0+a_1x+a_2x^2+\ldots+a_dx^d$$

כאשר $a_0,\ldots,a_d$ הם מספרים ממשיים הנקראים **מקדמי הפולינום** (coefficients). המקדם a_0 נקרא **המקדם החפשי** (free term) של p. ההצגה $p(x)=\sum_{k=0}^{d}a_kx^k$ של p נקראת **קנונית** אם $a_d\neq 0$, ואז a_d נקרא **המקדם הראשי** של p, אומרים ש־d הוא **המעלה** (degree) של p, ורושמים $\deg=d$. הפולינום $p\equiv 0$ נקרא **פולינום האפס**.[6]

למשל, הפונקציות $p(x)=x+1$, $p(x)=x^2-x+1$ ו־ $p(x)=x^5-2^2+8$ הן פולינומים. גם $f(x)=(x+1)(x-1)$ היא פולינום, כפי שניתן לראות ע"י פתיחת הסוגריים. באופן כללי, הפולינומים הם אותן פונקציות ממשיות שערכן בנקודה $x\in\mathbb{R}$ ניתנת לחישוב בעזרת פעולות חיבור וכפל של x עם עצמו ועם מספרים קבועים.

המעלה של פולינום הוא החזקה המקסימלית של x אשר מופיעה בה. אך שימו לב שהמעלה של הפולינום $p(x)=0x^3+x^2-1$ הוא 2 ולא 3, כי המקדם של x^3 הוא 0 (והפונקציה היא למעשה $p(x)=x^2-1$). אבל בהנתן הצגה של פולינום, אפשר למחוק חזקות גבוהות בעלות מקדם 0, עד להחזקה הגבוהה ביותר המופיעה איננו אפס, וכך לקבל כתיבה קנונית של אותה פונקציה. יתר על כן המשפט הבא אומר שאם שני פולינומים שווים כפונקציות אז יש להם אותה כתיבה קנונית.

משפט 7.3.2 לפולינום p יש כתיבה קנונית יחידה, כלומר יש רק דרך אחת לבחור d ומקדמים $a_0,\ldots,a_d$ כך ש־ $a_d\neq 0$ ומתקיים $p(x)=\sum_{k=0}^{d}a_kx^k$.

אפשר למצוא הוכחה למשפט בתרגילים בסוף הסעיף (ראו תרגיל (4) למטה).

פונקציות רציונליות

הגדרה 7.3.3 פונקציה רציונלית (rational function) היא פונקציה מהצורה $\frac{p}{q}$ כאשר p,q פולינומים ו־ $q\not\equiv 0$.[7]

למשל, $r(x)=\frac{x}{1+x}$ ו־ $r(x)=\frac{x+1}{x^2+1}$ הן פונקציות רציונליות. כפי שפולינומים הן הפונקציות שערכן ניתן לחישוב בעזרת חיבור וכפל, הפונקציות הרציונליות הן

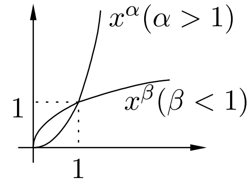
[6] פולינום מהצורה $p(x)=a\cdot x^k$ נקרא **מונום** (*monomial*). פולינום הוא סכום של כמה מונומים, ומכאן שמו.

[7] הפונקציות הרציונליות נקראות כך משום שהן מנות של פולינומים, כפי שמספר רציונלי הוא מנה של מספרים שלמים. אין הכוונה שהיא מקבלת רק ערכים רציונליים.

הפונקציות הממשיות שערכן בנקודה ניתן לחישוב בעזרת חיבור, כפל וחילוק. שימו לב שכל פולינום הוא בפרט פונקציה רציונלית, כי אפשר לרשום פולינום p כ־$\frac{p}{1}$, ו־1 הוא פולינום.

התחום של פונקציה רציונלית $\frac{p}{q}$ הוא קבוצת הנקודות x עבורן $q(x) \neq 0$. יש לכל היותר מספר סופי של נקודות כאלה ומספר ה־x־ים עם $q(x) = 0$ אינו עולה על $\deg q$. ראו תרגיל (5) בסוף הסעיף.

פונקציות חזקה

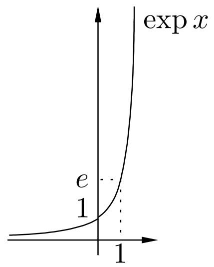


איור 7.3.1 פונקציית חזקה

הגדרה 7.3.4 יהי $\alpha \in \mathbb{R}$. **פונקציית חזקה**[8] (power function) עם מעריך α היא הפונקציה $f(x) = x^\alpha$.

כאשר α מספר טבעי הגדרה זו נותנת פולינום (בעצם, מונום). תחום ההגדרה של פונקציית חזקה תלוי בערך α: כאשר $\alpha \in \mathbb{N}$ התחום הוא $\mathbb{R}$. כאשר α שלם שלילי התחום ייחשב ל־ $\mathbb{R} \setminus \{0\}$. ביתר המקרים (דהיינו כש־α אינו שלם) התחום מוגדר להיות $[0, \infty)$.

פונקציות מעריכיות



איור 7.3.2 פונקציית האקספוננט

הגדרה 7.3.5 יהי $a > 0$. הפונקציה $\exp_a : \mathbb{R} \to (0, \infty)$ הנתונה על ידי הנוסחה $\exp_a(x) = a^x$ נקראת **הפונקציה המעריכית** (exponential function) עם **בסיס** a. עבור $a = e$ מסמנים אותה בקיצור $\exp$ (כאן e הוא המספר שהוגדר בסעיף 5.11).

מקובל לרשום $\exp_a x$ במקום $\exp_a(x)$ כאשר הדבר לא יוצר דו־משמעות.

לפונקציות המעריכיות התנהגות שונה מאד מפונקציות החזקה, אף ששתיהן מוגדרות בעזרת פעולת החזקה. נראה לכך דוגמאות בהמשך.

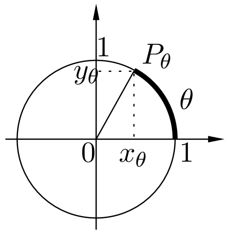
פונקציות טריגונומטריות

נסיים עם דיון במשפחת הפונקציות הטריגונומטריות. ההגדרה הקלאסית שלהן משתמשת במושגים מהגאומטריה של המישור. יהי S מעגל היחידה במישור, כלומר אוסף הנקודות $(x, y) \in \mathbb{R}^2$ שמרחקן מהראשית הוא 1. לפי משפט פיתגורס,

$$S = \{(x, y) \in \mathbb{R}^2 \,:\, x^2 + y^2 = 1\}$$

נסמן ב־π את מחצית ההיקף של S.

[8] ישנם ספרים בהם המונח "פונקציית חזקה" משמש לציון פונקציות מעריכיות.



איור 7.3.3 הנקודה המתאימה לקשת P_θ באורך θ על מעגל היחידה

לכל מספר ממשי θ תהי $P_\theta = (x_\theta, y_\theta)$ הנקודה על המעגל S אשר מתקבלת באופן הבא: מתחילים מהנקודה $(1,0)$ שבה המעגל חותך את הקרן החיובית של ציר ה־x, ונעים מרחק $|\theta|$ על גבי המעגל, כאשר תנועה בכיוון החיובי היא נגד כיוון השעון ובכיוון השלילי היא עם כיוון השעון. כך לכל מספר ממשי θ התאמנו נקודה (x_θ, y_θ). הפונקציות הטריגונומטריות הן **סינוס** (sine): $\sin : \mathbb{R} \to \mathbb{R}$, **וקוסינוס** (cosine): $\cos : \mathbb{R} \to \mathbb{R}$, מוגדרות על ידי

$$\cos\theta = x_\theta \qquad , \qquad \sin\theta = y_\theta$$

מקובל לרשום $\sin x$ במקום $\sin(x)$, וכותבים $\sin^2 x$ במקום $(\sin x)^2$ כשאין סכנת בלבול. כך גם לגבי $\cos$.

ההגדרה למעלה בעייתית כיוון שהיא מבוססת על מושגים גאומטריים שלא הגדרנו. בעיה נוספת היא שההגדרה אינה נותנת שום כלי כדי לחשב את המספר π או את הערכים של $\sin, \cos$ למעט בכמה נקודות מיוחדות. ובכל זאת הפונקציות הטריגונומטריות חשובות מכדי שנוותר עליהן. לכן נחרוג באופן חד פעמי ממנהגנו ונקבל את קיומן, יחד עם כמה תכונות שלהן, ללא הוכחה.

הנחה : קיים מספר $\pi > 0$ ופונקציות $\sin, \cos : \mathbb{R} \to [-1, 1]$ המקיימות את התכונות הבאות:

1. לכל $x \in \mathbb{R}$ מתקיים $\cos^2 x + \sin^2 x = 1$.
2. $\sin$ פונקציה אי־זוגית ו־$\cos$ פונקציה זוגית.
3. לכל $x \in \mathbb{R}$ מתקיים $\sin(x+y) = \sin x \cos y + \cos x \sin y$.
4. לכל $x \in \mathbb{R}$ מתקיים $\cos(x+y) = \cos x \cos y - \sin x \sin y$.
5. $\sin 0 = \cos\frac{\pi}{2} = 0$, $\cos 0 = \sin\frac{\pi}{2} = 1$, ו־ $0 < \sin x, \cos x < 1$ לכל $x \in (0, \frac{\pi}{2})$.
6. לכל $x \in \mathbb{R}$ מתקיים

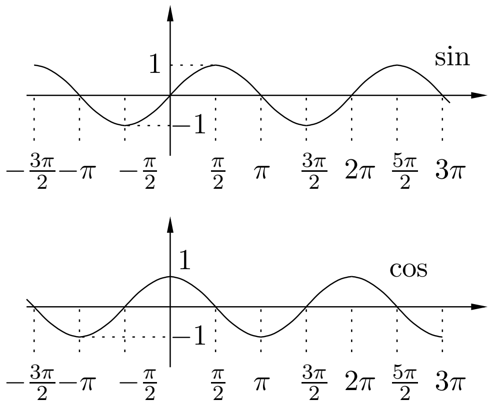
$$\begin{array}{ccccc} \cos(x+\frac{\pi}{2}) = -\sin x & , & \cos(x+\pi) = -\cos x & , & \cos(x+2\pi) = \cos x \\ \sin(x+\frac{\pi}{2}) = \cos x & , & \sin(x+\pi) = -\sin x & , & \sin(x+2\pi) = \sin x \end{array}$$

ו־2π הוא המחזור המינימלי של $\sin, \cos$.

7. לכל $x \in (0, \frac{\pi}{2})$ מתקיים $\sin x \le x \le \frac{\sin x}{\cos x}$.

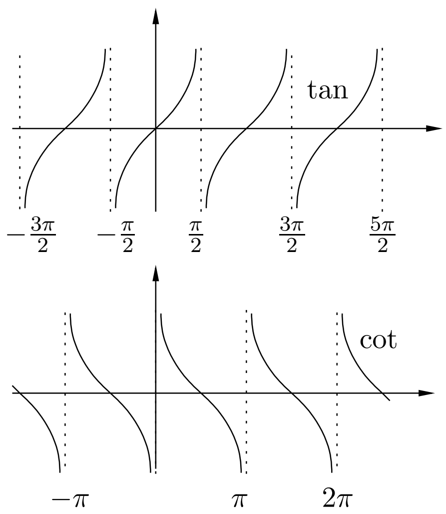
במהלך הפרקים הבאים ננתח את הפונקציות הטריגונומטריות באופן יסודי על סמך תכונות אלה. בעזרת הכלים שנפתח נמצא תיאור מפורש שלהן ושל π, ולבסוף אף נמצא דרכים לחשב אותן בכל רמת דיוק רצויה, ואז גם יהיו לנו הכלים להגדיר פונקציות המקיימות את התכונות למעלה ללא צורך בהנחות כלשהן (ראו תרגיל (5) בעמוד 533). תהליך זה דומה להתפתחות ההיסטורית של ההבנה של הפונקציות הללו: תחילה התבססו מתמטיקאים על תיאור גאומטרי, ורק בשלב מאוחר יחסית מצאו תיאור אנליטי ומדויק שלהן.

הגרפים של הפונקציות $\sin, \cos$ מופיעות באיור 7.3.4. בהמשך נראה שההנחות שלנו על הפונקציות אכן גוררות שהגרפים שלהן נראים כך.



איור 7.3.4 הגרפים של הפונקציות $\sin$ ו־$\cos$

שתי פונקציות טריגונומטריות נוספות הן המנה $\tan x = \frac{\sin x}{\cos x}$, הנקראת **טנגנס**, והמנה $\cot x = \frac{\cos x}{\sin x}$, הנקראת **קוטנגנס**. המספר $\tan\theta$ הוא השיפוע של הישר העובר דרך הראשית והנקודה $P_\theta = (\cos\theta, \sin\theta)$ שעל מעגל היחידה (ראו איור 7.3.3 בעמוד 227), ואילו $\cot$ הוא אחד חלקי השיפוע הזה. הפונקציות $\tan, \cot$ מוגדרות היכן שהמכנה בהגדרה שלהן לא מתאפס: לפי תרגיל (9) למטה פירוש הדבר ש־$\tan$ מוגדרת בכל מקום פרט לנקודות $\pi k + \frac{\pi}{2}$ עבור k שלם, ו־$\cot$ מוגדרת בכל מקום למעט נקודות מהצורה πk עבור k שלם. מהתכונות של $\sin, \cos$ אפשר להסיק תכונות דומות של הפונקציות $\tan, \cot$. נתייחס לכך בתרגילים למטה. הגרפים של $\tan, \cot$ נראים כך:



איור 7.3.5 הפונקציות $\tan$ ו־$\cot$

תרגילים

1. בהינתן זוג נקודות $(x',y'),(x'',y'')$ במישור עם $x'\neq x''$, מצאו נוסחה לישר (היחיד!) העובר דרכן. הסיקו שישר נקבע על ידי כל שתי נקודות על הגרף שלו.

2. כתבו את הפולינומים הבאים בכתיב קנוני ומצאו את מעלת הפולינום.

   (א) $3x+1+2x+x(4x^2-x)-4x^3$.

   (ב) $(x-2)(x^2-2)(x^3+1)-x^6$.

3. הוכיחו את הטענות הבאות:

   (א) אם p,q פולינומים אז הפונקציה $r(x)=p(x)+q(x)$ היא פולינום ממעלה לכל היותר $\max\{\deg p,\deg q\}$, וייתכן שהמעלה קטנה יותר.

   (ב) אם p,q פולינומים אז הפונקציה $s(x)=p(x)\cdot q(x)$ היא פולינום ממעלה $\deg p+\deg q$, אלא אם כן $p=0$ או $q=0$ ואז גם $s\equiv 0$.

4. יהי $p(x)=\sum_{n=0}^{d}a_nx^n$ ונניח ש־ $a_d\neq 0$. הראו ש־ $p\not\equiv 0$. הסיקו שהכתיבה הקנונית של פולינום היא יחידה, כלומר, שאם מתקיים גם $p(x)=\sum_{n=0}^{d}b_nx^n$ אז $b_i=a_i$ לכל $0\leq i\leq d$ (לשם כך שימו לב ש־ $p(x)-p(x)=0$ לכל x; כתבו את ההפרש כפולינום בעזרת המקדמים a_i,b_i).

5. משפט החלוקה עם שארית למספרים טבעיים קובע שאם $m<n$ טבעיים אז יש $d\in\mathbb{N}$ ו־ $r\in\{0,\ldots,m-1\}$ יחידים כך ש־ $n=dm+r$ (הוכחנו זאת בתרגיל (1) בעמוד 48). בשאלות הבאות נראה שלפולינומים יש תכונה דומה.

   הגדרה יהיו p,q פולינומים. נאמר ש־q **מחלק** את p, ונרשום $q|p$, אם קיים פולינום s כך ש־ $p(x)=s(x)\cdot q(x)$ לכל x.

   (א) הוכיחו שלמספרים $\alpha\neq\beta$, אם נגדיר $p(x)=x-\alpha$ ו־ $q(x)=x-\beta$, אז p אינו מחלק את q, ולהפך.

   (ב) יהיו $p(x)=\sum_{k=0}^{n}a_kx^k$ ו־ $q(x)=\sum_{k=0}^{m}b_kx^k$ פולינומים עם $m\leq n$ ו־ $a_n,b_m\neq 0$. הוכיחו שקיימים פולינומים d,r עם $\deg r<m$ כך ש־ $p(x)=q(x)d(x)+r(x)$ לכל x. (רמז: הוכיחו באינדוקציה על n שלכל $m\leq n$ הטענה מתקיימת. כדי לבצע את שלב האינדוקציה, שימו לב שקיים קבוע c כך שההפרש $p(x)-cx^{n-m}\cdot q(x)$ הוא פולינום ממעלה קטנה מ־n). שימו לב ש־ $\deg r<m$ מבטיח ש־ q אינו מחלק את r.

   (ג) אם p פולינום ואם $p(\alpha)=0$ אז α נקרא **שורש** (root) של p. הוכיחו ש־α שורש של p אמ"מ $x-\alpha$ מחלק את p, כלומר אמ"מ יש פולינום q כך ש־ $p(x)=(x-\alpha)q(x)$. כיוון אחד הוא קל; בשביל הכיוון השני היעזרו בסעיף הקודם, והראו שאם $p(x)=(x-\alpha)d(x)+r(x)$ היא החלוקה עם שארית של p בפולינום $x-\alpha$, אז $r\equiv 0$ (הציבו $x=\alpha$).

   (ד) הסיקו שלפולינום p ממעלה n יש לכל היותר n שורשים. (רמז: הראו שאם $\alpha_1,\ldots,\alpha_k$ הם שורשים של הפולינום p אז יש פולינום q שמקיים

$p(x) = (x-\alpha_1)\cdot\ldots\cdot(x-\alpha_k)q(x)$. השוו את מעלות הפולינומים בשני האגפים והסיקו ש־q קבוע).

6. תהי $f:[0,1]\to\mathbb{R}$ פונקציה. מתי קיימת פונקציה מחזורית $g:\mathbb{R}\to\mathbb{R}$ עם מחזור 1 כך ש־ $g|_{[0,1]} = f|_{[0,1]}$?

7. הראו שתכונה (6) של הפונקציות הטריגונומטריות, כפי שהוצגו בעמוד 227, נובעת מהתכונות האחרות. (רמז: השתמשנו בנוסחאות עבור $\sin(x+y)$ ו־ $\cos(x+y)$ ובערכים של הפונקציות ב־0 ו־$\frac{\pi}{2}$).

8. הוכיחו את השוויונות הבאים:

$$\begin{aligned}
\cos x &= \sin(\frac{\pi}{2}-x)\\
\sin x &= \cos(\frac{\pi}{2}-x)\\
\sin 2x &= 2\sin x\cos x\\
\cos 2x &= 2\cos^2 x - 1\\
&= 1-2\sin^2 x\\
\sin x+\sin y &= 2\cdot\sin(\frac{x+y}{2})\cdot\cos(\frac{x-y}{2})\\
\cos x+\cos y &= 2\cdot\cos(\frac{x+y}{2})\cdot\cos(\frac{x-y}{2})
\end{aligned}$$

9. הראו ש־$\sin$ מתאפסת רק בנקודות πk עבור k שלם, וש־$\cos$ מתאפסת רק בנקודות $\pi k+\frac{\pi}{2}$ עבור k שלם.

10. הראו ש־$\tan$ היא פונקציה אי־זוגית ומחזורית עם מחזור π, ושלכל מספר שלם k היא מתאפסת בנקודות πk, חיובית בנקודות $\pi k < x < \pi k + \frac{\pi}{2}$ ושלילית בנקודות $\pi k+\frac{\pi}{2} < x < \pi k$. נסחו והוכיחו תכונות דומות עבור $\cot$.

11. הוכיחו את השוויונות

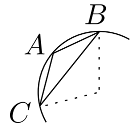
$$\begin{aligned}
\sin x &= \frac{2\tan(x/2)}{1+\tan^2(x/2)}\\
\cos x &= \frac{1-\tan^2(x/2)}{1+\tan^2(x/2)}
\end{aligned}$$

עבור $x\in[0,\frac{\pi}{2}]$ השתמשו בנוסחאות ל־$\cos 2x$ ו־$\sin 2x$ על מנת לבטא את $\cos\frac{x}{2}$ בעזרת $\cos x$ ואת $\sin\frac{x}{2}$ בעזרת $\sin x$.

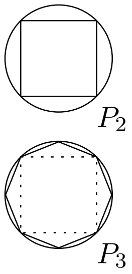
12. בשאלה זו נמצא ביטוי למספר π, המבוסס על שיקולים גיאומטריים.

(א) יהי BC מיתר על מעגל היחידה ויהי s אורך הקשת הקצרה בין B ו־C. הראו ש־ $BC = 2\sin\frac{s}{2}$.

(ב) יהיו B,C כמו קודם ותהי A נקודת האמצע של הקשת בין B ל־C; ראו איור 7.3.6. הוכיחו ש־ $AB+AC = \sqrt{2-\sqrt{4-BC^2}}$. היעזרו בנוסחה ל־$\sin\frac{x}{2}$ שמצאתם בשאלה הקודמת.



איור 7.3.6



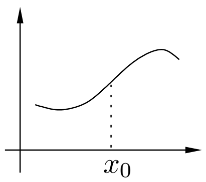
איור 7.3.7

(ג) נבנה ברקורסיה סדרת מצולעים P_n בעלי 2^n צלעות, שקדקדיהם על מעגל היחידה. P_2 הוא ריבוע, ובהנתן P_n נקבל את P_{n+1} על ידי כך שנחליף כל צלע BC של P_n בשתי צלעות AB, AC שהתקבלו כמו בסעיף הקודם; ראו איור 7.3.7 . הוכיחו שההיקף של P_n הוא

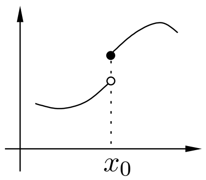
$$\ell_n = 2^n\sqrt{2-\sqrt{2+\sqrt{2+\sqrt{2+\ldots\sqrt{2}}}}}$$

כאשר מספר סימני השורש הוא $n-1$. הוכיחו שהגבול $\lim_{n\to\infty}\ell_n$ קיים, ונמקו באופן לא פורמלי מדוע אפשר לפרש אותו בתור האורך 2π של המעגל.

7.4 הגבול של פונקציה בנקודה



איור 7.4.1 פונקציה עם גבול ב־x_0



איור 7.4.2 פונקציה ללא גבול ב־x_0

נניח שאנו צופים בתנועה של חלקיק, ובזמן מסויים t_0 האור כבה ואיננו רואים באותו רגע את החלקיק. למרות החשכה הרגעית נוכל לנחש בוודאות גבוהה את מיקום החלקיק על בסיס הזמנים שבהם כן צפינו במיקומו. הסיבה היא שאם $f(t)$ מתאר את מיקום החלקיק בזמן t אז סביר להניח שמיקום החלקיק בזמן t_0 קרוב למיקום שלו בזמנים t אשר קרובים ל־t_0. במילים אחרות, כאשר t קרוב ל־t_0 סביר ש־$f(t)$ קרוב ל־$f(t_0)$. אמירה זו משקפת את העובדה שלפונקציה f יש גבול כאשר t שואף ל־t_0. מושג זה דומה במידה רבה למושג הגבול של סדרה a_n, שהוא המספר אליו מתקרבים האיברים a_n כאשר n שואף לאינסוף.

ניסיון הגדרה ראשון: מספר L הוא הגבול של פונקציה ממשית f בנקודה x_0 אם $f(x)$ קרוב ל־L כאשר x קרוב ל־x_0.

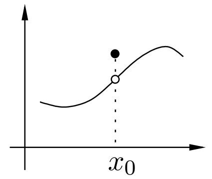
איורים 7.4.2,7.4.1 מתארים שתי פונקציות, שלאחת מהן יש גבול ב־x_0 ולשנייה אין.

נעבור לדיון פורמלי יותר. תחילה עלינו לפרט למה הכוונה כשאומרים ש־$f(x)$ קרוב ל־L, וכשאומרים ש־x קרוב ל־x_0. בהקבלה להגדרת הגבול של סדרה, נאמר שמספר הוא קרוב ל־L אם הוא נמצא בסביבה קטנה של L. גם את התנאי "לכל x קרוב ל־x_0 מתקיים..." אפשר לפרש בלשון של סביבות, דהיינו בתור התכונה "יש סביבה של x_0 שלכל x בה מתקיים...". ליתר דיוק, אם $P(x)$ תכונה התלויה במספר x, נאמר ש־P **מתקיימת בסביבה** של x_0 אם יש סביבה V של x_0 כך שלכל $x \in V$ מתקיים $P(x)$.

ניסיון הגדרה שני: מספר L הוא הגבול של פונקציה ממשית f בנקודה x_0 אם לכל סביבה U של L, קטנה ככל שתהיה, יש סביבה של x_0 שבה מתקיים $f(x) \in U$.

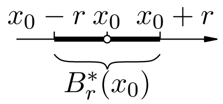
הגדרה זו קרובה להגדרה הסופית, אך נותרו שתי נקודות עדינות הדורשות טיפול. הראשונה היא שבתכונה $f(x) \in U$ מסתתרת הטענה ש־f מוגדרת ב־x. כדי לא לציין זאת במפורש בכל פעם שהשאלה תתעורר, נאמץ את המוסכמה הבאה:

מוסכמה : אם f פונקציה ממשית ו־x מספר ממשי אז בכל פעם שנאמר ש־$f(x)$ מקיימת תכונה מסוימת נתכוון בנוסף ש־f מוגדת ב־x.



איור 7.4.3 ערך הפונקציה בנקודה אינו משפיע על קיום הגבול

בעיה נוספת היא שההגדרה למעלה מחייבת את f להיות מוגדרת בנקודה x_0 עצמה. ואמנם, מכיוון ש־x_0 שייכת לכל סביבה שלה, התנאי "קיימת סביבה V של x_0 כך שלכל $x \in V$ מתקיים $f(x) \in U$" גוררת ש־f מוגדרת ב־x_0 (השתמשנו כאן במוסכמה האחרונה) ויתר על כן ש־ $f(x_0) \in U$. זו תוצאת לוואי לא רצויה, ויש מקרים שבהם נרצה לדבר על הגבול של פונקציה שאינה מוגדרות ב־x_0, או שנרצה להתעלם מהערך שלה ב־x_0. באיור 7.4.3 מופיע פונקציה בעלת גבול ב־x_0 למרות שיש לה קפיצה בנקודה x_0 עצמה. כדי להמנע מתוצאת לוואי זו, נעזר בהגדרה הבאה.



איור 7.4.4 סביבה מנוקבת

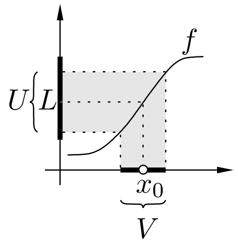
הגדרה 7.4.1 יהי $x \in \mathbb{R}$ ו־ $r > 0$. **הסביבה המנוקבת** של x ברדיוס $r > 0$ היא הקבוצה $B_r^*(x) = B_r(x) \setminus \{x\}$, כלומר, זו הסביבה ברדיוס r של x שממנה הוצאה הנקודה x עצמה.

לפעמים לצורך הדגשה נאמר על סביבה שהיא סביבה **מלאה** כדי להדגיש שאינה מנוקבת. שימו לב שסביבה מנוקבת לעולם אינה קבוצה ריקה. בדיקה קלה מראה שחיתוך של שתי סביבות מנוקבות של נקודה x_0 הוא גם־כן סביבה מנוקבת של x_0.

כמו קודם, אם $P(x)$ תכונה התלויה במספר x אז נאמר ש־P מתקיימת בסביבה מנוקבת של x_0 אם יש סביבה מנוקבת של x_0 כך שלכל x בה מתקיים $P(x)$.

הדיון עד כה מוביל להגדרה הבאה:

הגדרה 7.4.2 תהי f פונקציה ממשית ו־ $x_0 \in \mathbb{R}$. מספר ממשי L נקרא **הגבול** (limit) של f ב־x_0 אם לכל סביבה U של L, קטנה ככל שתהיה, קיימת סביבה מנוקבת של x_0 שמתקיים בה $f(x) \in U$. במקרה שקיים L כזה אומרים ש־f **מתכנסת** (ל־L) כאשר x שואף ל־x_0, ומסמנים $L = \lim_{x \to x_0} f(x)$, או $f(x) \xrightarrow{x \to x_0} L$.



איור 7.4.5 לכל x בסביבה המנוקבת V של x_0 מתקיים $f(x) \in Y$

באופן מפורש יותר, $\lim_{x \to x_0} f(x) = L$ אמ"מ לכל סביבה U של L יש סביבה מנוקבת V של x_0 כך שלכל $x \in V$ מתקיים $f(x) \in U$. אם נרשום במלואה את ההגדרה של סביבה ונתרגם את התנאי לסימנים, נקבל

$$\lim_{x \to x_0} f(x) = L \quad \Longleftrightarrow$$
$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \; (0 < |x - x_0| < \delta \implies |f(x) - L| < \varepsilon)$$

תיאור זה נקרא לפעמים **תיאור הגבול בלשון** ε־δ. שימו לב שהאי־שוויון $|x - x_0| > 0$ היא דרך עקיפה לומר $x \neq x_0$. וודאו שאתם מבינים כיצד הפרמטרים ε, δ מתאימים לבחירה של סביבות!

איור 7.4.5 מתאר סביבה U של L וסביבה מנוקבת V של x_0 המתאימה לה. כפי שרואים, לכל נקודה $x \in V$, הערך $f(x)$ (דהיינו הגובה של הנקודה על הגרף מעל

x) נמצא ב־U. הסביבה המנוקבת V אינה יחידה, ואפשר היה לבחור סביבה מעט גדולה יותר, או כל סביבה קטנה יותר, כך שתתקיים אותה התכונה. יתר על כן, מהאיור רואים שככל שנקטין את הסביבה U של L אפשר להקטין את V בהתאמה, כך שעדיין יתקיים $f(x) \in U$ לכל $x \in V$. מסיבה זו האיור מתאר פונקציה שגבולה ב־x_0 הוא L.

הערות

1. מההגדרה נובע שאם ל־f יש גבול ב־x_0 אז קיימת סביבה מנוקבת של x_0 שבה f מוגדרת בכל נקודה. במילים אחרות, תנאי הכרחי לקיום הגבול של f ב־x_0 הוא ש־f תהיה מוגדרת בסביבה מנוקבת של x_0. אין כאן הפתעה: אילולא f הייתה מוגדרת קרוב ל־x_0 לא היה טעם לדבר על גבול שם. למשל במקרה של הפונקציה $f(x) = \sqrt{x}$ אין טעם לדבר על הגבול של f ב־-1 כי f אינה מוגדרת באף נקודה קרובה ל־-1.
2. הסביבה המנוקבת V של x_0 המתאימה ל־U בהגדרת הגבול אינה יחידה. כל סביבה מנוקבת קטנה יותר מתאימה ל־U גם־כן.
3. התפקיד של האות x בסימונים מהסוג $\lim_{x\to x_0} f(x)$ הוא כמשתנה סרק, כמו האינדקס n בגבול $\lim_{n\to\infty} a_n$ של סדרה. כמובן שמתקיים

$$\lim_{x\to x_0} f(x) = \lim_{y\to x_0} f(y)$$

ומאידך, באופן כללי

$$\lim_{x\to x_0} f(x) \neq \lim_{x\to y_0} f(x) \neq \lim_{y\to y_0} f(x)$$

כמו במקרה של גבול של סדרה, עלינו לבדוק שהגבול הוא יחיד כדי להצדיק את השימוש בהא הידיעה:

טענה 7.4.3 אם $\lim_{x\to x_0} f(x) = L$ וגם $\lim_{x\to x_0} f(x) = M$ אז $L = M$.

הוכחה נניח בשלילה $L \neq M$. אז קיימות סביבות U_L, U_M של L, M בהתאמה כך ש־ $U_L \cap U_M = \emptyset$ (טענה 5.2.4). לפי הגדרת הגבול קיימות סביבות מנוקבות V_M, V_L של x_0 כך ש־

$$\begin{aligned} x \in V_L \quad &\Longrightarrow \quad f(x) \in U_L \\ x \in V_M \quad &\Longrightarrow \quad f(x) \in U_M \end{aligned}$$

(ובפרט f מוגדרת ב־V_L, V_M). תהי $V = V_L \cap V_M$. אז גם V היא סביבה מנוקבת לא ריקה של x_0. יהי $x \in V$. אז $x \in V_L$ ולכן $f(x) \in U_L$. מצד שני גם $x \in V_M$, ולכן $f(x) \in U_M$. קיבלנו $f(x) \in U_L \cap U_M$, בסתירה לכך ש־U_M, U_L זרות. ■

כדאי להשוות את ההוכחה האחרונה עם הוכחת יחידות הגבול של סדרה (משפט 5.2.5). שימו לב במיוחד לתפקיד של הטענה שחיתוך של סביבות מנוקבות היא סביבה מנוקבת, והשוו אותה עם למה 5.2.6 והתפקיד שהיא משחקת בהוכחת משפט 5.2.5.

דוגמאות

1. יהי c מספר. נתבונן בפונקציה הקבועה $f \equiv c$ ונראה ש־ $\lim_{x\to x_0} f(x) = c$ בכל נקודה x_0. ואמנם, תהי U סביבה של c. לכל $x \in R$ מתקיים $f(x) = c$, כלומר $f(x) \in U$. בפרט בהינתן x_0 אפשר לבחור סביבה מנוקבת שרירותית V של x_0, כמו למשל $V = B_1^*(x_0)$, ואז f מוגדרת בכל $x \in V$ ומקיימת $f(x) \in U$.

2. נראה שלפונקציה $f(x) = 2x$ יש גבול ב־ $x_0 = 1$. תחילה עלינו לנחש מה הערך של הגבול. ברור למדי שכאשר x קרוב ל־1 אז $2x$ קרוב ל־2, ולכן ננחש שהגבול הוא 2. אפשר גם לחשב את הערך של הפונקציה בכמה נקודות קרובות ל־1. אם ננהג כך נקבל

$$f(1.1) = 2.2 \quad , \quad f(0.99) = 1.98 \quad , \quad f(1.001) = 2.002 \quad , \quad \ldots$$

זו עדות נוספת לכך שכאשר x שואף ל־1 הגבול של f הוא 2.

יהי $\varepsilon > 0$ ותהי $U = B_\varepsilon(2)$ הסביבה של 2 ברדיוס ε. עלינו להראות שקיימת סביבה מנוקבת $V = B_\delta^*(1)$ של 1 כך שלכל $x \in V$ שבו f מוגדרת מתקיים $f(x) \in U$. מאחר ש־f מוגדרת בכל נקודה התנאי הראשון מתקיים תמיד. אנו מחפשים לכן $\delta > 0$ כך שלכל x,

$$0 < |x-1| < \delta \quad \Longrightarrow \quad |f(x) - 2| < \varepsilon$$

נשים לב שלכל δ, אם $|x-1| < \delta$ אז $|2x-2| < 2\delta$, ולכן

$$|x-1| < \delta \quad \Longrightarrow \quad |f(x) - 2| < 2\delta$$

(כי $f(x) = 2x$). לכן אם נבחר $\delta = \varepsilon/2$ מתקיים

$$0 < |x-1| < \delta \quad \Longrightarrow \quad |f(x) - 2| < \varepsilon$$

וקיבלנו את הדרוש.

באופן כללי יותר, אם a, b מספרים אז לפונקציה $f(x) = ax + b$ יש גבול בכל נקודה ולכל x_0 מתקיים $\lim_{x\to x_0} f(x) = ax_0 + b$. ההוכחה מושארת כתרגיל.

3. הנה דוגמה מעט יותר מורכבת. תהי $g(x) = x^2 + 1$ ונחשב את הגבול ב־1. סביר להניח שכאשר x קרוב ל־1 אז $x^2 + 1$ קרוב ל־2. לכן ננסה להוכיח ש־ $\lim_{x\to 1} g(x) = 2$.

יהי $\varepsilon > 0$. אנו מחפשים מספר חיובי δ כך שלכל מספר x מתקיים

$$0 < |x-1| < \delta \quad \Longrightarrow \quad |g(x) - 2| < \varepsilon$$

(כמו בדוגמה הקודמת, g מוגדרת בכל מקום ולכן אין צורך להוסיף את התנאי "g מוגדרת ב־x" באופן מפורש). נשים לב שמתקיים

$$g(x) - 2 = x^2 - 1 = (x-1)(x+1)$$

לכן לכל $\delta > 0$ מתקיים

$$|x-1| < \delta \quad \Longrightarrow \quad |g(x) - 2| = |x-1| \cdot |x+1| < \delta|x+1|$$

מצד שני אם $|x-1| < \delta$ אז $|x+1| < 2 + \delta$ (הוכיחו זאת!), ולכן

$$|x-1| < \delta \quad \Longrightarrow \quad |g(x) - 2| < \delta(2+\delta)$$

ולכן אם נמצא δ כך ש־ $\delta(2+\delta) \leq \varepsilon$, סיימנו.

ואכן יש δ כזה: למשל, נבחר $\delta = \min\{1, \frac{\varepsilon}{4}\}$. אז $\delta \leq 1$ ולכן $\delta(\delta+2) \leq \delta \cdot 3$, ומצד שני $\delta \leq \frac{\varepsilon}{4}$ ולכן $\delta(\delta+2) \leq 3\delta < \varepsilon$, כפי שרצינו.

הערה הסביבה המנוקבת $B^*_\delta(1)$ שמצאנו כאן אינה הסביבה הגדולה ביותר עם התכונה המבוקשת. במילים אחרות, ה־δ שמצאנו אינו ה־δ הגדול ביותר האפשרי. אין בכך פגם. כאשר מוכיחים קיום של גבול ב־x_0 אין צורך למצוא את הסביבה המנוקבת הגדולה ביותר של x_0 שמתאימה לסביבה הנתונה של הגבול. די למצוא סביבה מתאימה כלשהי.

4. נגדיר $\widehat{g} : \mathbb{R} \to \mathbb{R}$ על ידי

$$\widehat{g}(x) = \begin{cases} x^2 + 1 & x \neq 1 \\ 0 & x = 1 \end{cases}$$

קל לוודא של־$\widehat{g}$ יש אותו גבול ב־1 כמו ל־g. ההוכחה זהה להוכחה במקרה של g (יש רק לשים לב שהערך של g ב־1 לא השפיע באמת על מהלך ההוכחה).

הגבול של פונקציה בנקודה אינו תמיד קיים. סיבה טריביאלית לכך יכולה להיות שהפונקציה אינה מוגדרת בסביבה מנוקבת של הנקודה, אך זו אינה הסיבה היחידה.

טענה 7.4.4 תהי f פונקציה המוגדרת בסביבה מנוקבת של x_0 ויהי L מספר. אז התנאים הבאים שקולים:

1. L אינו הגבול של f ב־x_0.
2. קיימת סביבה U של L כך לכל סביבה מנוקבת V של x_0 קיים $x \in V$ שבו f מוגדרת ומקיימת $f(x) \notin U$.

הוכחה נניח ש־(1) לא מתקיים, כלומר ש־ $\lim_{x\to x_0} f(x) = L$. אז לכל סביבה U של L יש סביבה מנוקבת V של x_0 כך ש־f מוגדרת בכל $x \in V$ ומקיימת $f(x) \in U$, כלומר, (2) אינו מתקיים. מכאן ש־(2) גורר את (1).

להפך, נניח ש־(1) מתקיים. אז מהגדרת הגבול קיימת סביבה U של L כך שלכל סביבה מנוקבת V של x_0, לא נכון ש־f מוגדרת בכל $x \in V$ וש־ $f(x) \in U$. נראה ש־U היא הסביבה המבוקשת, כלומר ש־(2) מתקיים עבור ה־U הזו. ואמנם, תהי V סביבה מנוקבת של x_0. עלינו להראות שיש $x \in V$ עם $f(x) \notin U$. לפי הנתון f מוגדרת בסביבה מנוקבת של x_0, ולכן יש סביבה מנוקבת $W \subseteq V$ שבה f מוגדרת (לא מובטח ש־f מוגדרת בכל V, ולכן יש צורך לעבור לתת־סביבה). כיוון ש־f מוגדרת בכל W, לפי בחירת U יש $x \in W$ עם $f(x) \notin U$. מכיוון ש־ $W \subseteq V$ הרי $x \in V$, וזו הנקודה המבוקשת. ■

עבור f המוגדרת בסביבה מנוקבת של x_0 אפשר לרשום את הטענה האחרונה גם כך:

$$\lim_{x\to x_0} f(x) \neq L \quad\iff\quad \exists\varepsilon>0\ \forall\delta>0\ \exists x\in B^*_\delta(x_0)\ |f(x)-L|\geq\varepsilon$$

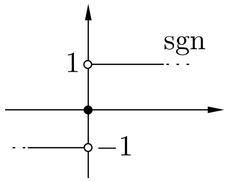
ולכן ל־f אין גבול ב־x_0 אם התנאי הזה מתקיים לכל מספר L, דהיינו:

$$\text{ל־ } f \text{ אין גבול ב־ } x_0 \quad\iff\quad \forall L\ \exists\varepsilon>0\ \forall\delta>0\ \exists x\in B^*_\delta(x_0)\ |f(x)-L|\geq\varepsilon$$

דוגמאות

1. תהי $f(x)=\frac{1}{x}$. אז $\operatorname{dom} f = (-\infty,0)\cup(0,\infty)$, והפונקציה מוגדרת בסביבה מנוקבת של 0. נראה של־f אין גבול ב־0. יהי $L\in\mathbb{R}$ ויהי $\varepsilon=1$. נראה שאין סביבה מנוקבת של 0 שלכל x בה מתקיים $f(x)\in B_1(L)$. ואמנם, בכל סביבה מנוקבת V של 0 יש מספר x המקיים $0<x<\frac{1}{L+1}$, ועבור x זה מתקיים $f(x)=L+1\notin B_1(L)$.
2. נתבונן בפונקציית הסימן

$$\operatorname{sgn}(x)=\begin{cases}-1 & x<0\\ 0 & x=0\\ 1 & x>1\end{cases}$$

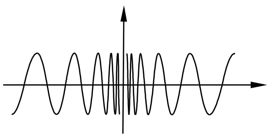


איור 7.4.6 פונקציית הסימן

לגרף של הפונקציה יש קפיצה בנקודה 0, ופירוש הדבר הוא שכש־x מתקרב ל־0 הוא יכול לקבל כמה ערכים רחוקים זה מזה. לכן נצפה שאין ל־sgn גבול ב־0. לשם כך עלינו להראות שלכל מספר $L \in \mathbb{R}$, המספר L אינו גבול של sgn ב־0.

ואמנם, יהי $L \in \mathbb{R}$. נבחר סביבה U של L כך שלפחות אחד המספרים 1 או -1 אינו שייך ל־U. למשל, $U = B_{1/2}(L)$ היא סביבה כזו. בכל סביבה מנוקבת V של 0 יש מספרים חיוביים ומספרים שליליים, ולכן בכל סביבה מנוקבת V של 0 יש נקודות בהן הפונקציה מקבלת ערכים 1 וערכים -1 (אם $V = B_r^*(0)$ אז $\pm r/2$ הן נקודות כאלה). לכן בכל סביבה של 0 יש נקודות x בהן $\operatorname{sgn}(x) \notin U$, וזה מראה ש־$L$ איננו גבול של sgn ב־0.

3. תהי $s : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ נתונה על ידי $s(x) = \sin \frac{1}{x}$. הגרף של s מתואר באיור 7.4.7. הסיבה שהגרף של s נראה כך היא, באופן גס, שכאשר x מתקרב ל־0 מהכיוון החיובי $1/x$ מקבל את כל הערכים הממשיים החיוביים, וכיוון ש־$\sin$ מחזורית זה אומר ש־ $\sin \frac{1}{x}$ מתנודדת בין 1 ל־-1 אינסוף פעמים. הסבר דומה נכון לגבי החלק של הגרף משמאל לאפס.


איור 7.4.7 הפונקציה $\sin(1/x)$

מהאיור ברור שלפונקציה זו אין גבול ב־0. כדי להוכיח זאת נקבע $L \in \mathbb{R}$ ונראה ש־L אינו הגבול של s ב־0. נבחר סביבה U של L כך שלפחות אחד המספרים $1, -1$ לא שייך ל־U. בכל סביבה מנוקבת V של 0 יש נקודות בהן s מקבלת את הערכים 1 ו־-1, למשל נקודות מהצורה $\frac{1}{\pi k + \pi/2}$ עבור k־ים זוגיים ואי־זוגיים בהתאמה (למה כל סביבה מנוקבת של 0 מכילה נקודות כאלה?). לכן אין סביבה מנוקבת של 0 שבה הערכים של f מוכלים ב־U, ולכן L אינה הגבול של s ב־0.

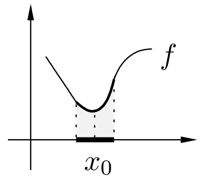
4. **פונקציית דירכלה** היא הפונקציה $D : \mathbb{R} \to \mathbb{R}$ המוגדרת על ידי

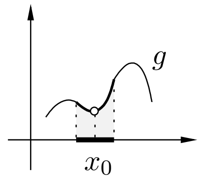
$$D(x) = \begin{cases} 0 & x \notin \mathbb{Q} \\ 1 & x \in \mathbb{Q} \end{cases}$$

הגרף של D מורכב מנקודות הנמצאות על שני ישרים: ציר ה־x והישר המאוזן בגובה 1 מעל ציר ה־x. בכל אחד מהישרים הגרף מפוזר כמו אבקה ואינו מכיל אף קטע רצוף, אך כל קטע באחד הישרים מכיל נקודות מהגרף. מטבע הדברים קשה לצייר גרף כזה!

נראה של־D אין גבול באף נקודה. יהי $L \in \mathbb{R}$ ונתבונן בסביבה $U = B_{1/3}(L)$ של L. אחת מהאפשרויות $0 \notin U$ או $1 \notin U$ מתקיימת. לכל $x_0 \in \mathbb{R}$, בכל סביבה מנוקבת של x_0 יש מספרים רציונליים ומספרים אי־רציונליים (משפטים 4.3.5 ו־4.5.3). לכן בכל סביבה מנוקבת של x_0 הפונקציה D מקבלת את שני הערכים $0, 1$. לכן בכל סביבה מנוקבת של x_0 יש x כך ש־$D(x) \notin U$, ומכאן ש־L אינה הגבול של D ב־x_0. מכיוון ש־L, x_0 היו מספרים שרירותיים אנו מסיקים שאין ל־D גבול באף נקודה.

בפרק על סדרות ראינו שהגבול של סדרה אינו מושפע משינויים סופיים בסדרה, או ליתר דיוק, שאם האיברים של שתי סדרות שווים החל ממקום מסוים אז התכונות הגבוליות של הסדרות זהות. על כך אמרנו שהגבול של סדרה הוא תכונה אסימפטוטית. התכונה המקבילה בהקשר הנוכחי היא שהגבול של פונקציה f בנקודה x_0 תלויה רק בהתנהגות הפונקציה בסביבות קטנות של x_0, או במילים אחרות, לכל סביבה מנוקבת V של x_0, קטנה ככל שתהיה, שינוי של הערכים של f מחוץ ל־V לא משפיע על קיום הגבול או על ערכו. משום כך אומרים שהגבול של פונקציה בנקודה היא תכונה **אינפיניטסימלית** (infinitesimal).





איור 7.4.8 פונקציות המסכימות בקטע סביב x_0

משפט 7.4.5 אם f, g פונקציות ממשיות ואם יש סביבה מנוקבת של x_0 שבה הן מסכימות, אז הגבול של f ב־x_0 קיים אמ"מ הגבול של g ב־x_0 קיים, ואם הגבולות קיימים אז הם שווים.

הוכחה נראה שאם $\lim_{x\to x_0} f(x) = L$ אז $\lim_{x\to x_0} g(x) = L$. הכיוון ההפוך נובע על ידי החלפת התפקידים של f, g.

תהי W סביבה מנוקבת של x_0 שבה שתי הפונקציות מוגדרות ושוות. תהי U סביבה של L. מכיוון ש־ $\lim_{x\to x_0} f(x) = L$, יש סביבה מנוקבת V_f של x_0 כך שלכל $x \in V$ מתקיים $f(x) \in U$. תהי $V_g = V_f \cap W$, שהיא גם סביבה מנוקבת של x_0. הפונקציה g מוגדרת בכל נקודה $x \in V_g$ כי $V_g \subseteq W$, ואם $x \in V_g$ אז $x \in V_f$ ולכן

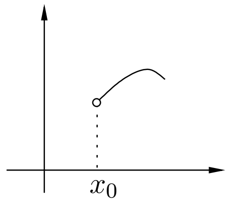
$$g(x) = f(x) \in U$$

בסיכום, הראינו שלכל סביבה U של L יש סביבה מנוקבת V_g של x_0 כך שאם $x \in V_g$ אז g מוגדרת ב־x ו־ $g(x) \in U$, ולכן $\lim_{x\to x_0} g(x) = L$. ■

העובדה שהגבול הוא תכונה מקומית באה לידי ביטוי בכך שקיום הגבול של f בנקודה x_0 אינו מבטיח קיום גבול בנקודות סמוכות ל־x. למשל, תהי

$$f(x) = \begin{cases} x & x \in \mathbb{Q} \\ -x & x \notin \mathbb{Q} \end{cases}$$

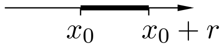
אנו משאירים כתרגיל את הבדיקה ש־ $\lim_{x\to 0} f(x) = 0$ אבל של־f אין גבול באף נקודה אחרת (ההוכחה דומה להוכחה שלפונקציית דירכלה אין גבול באף נקודה. ראו דוגמה (4) בעמוד 237).



איור 7.4.9 פונקציה עם גבול מימין ב־x_0

נסיים את הסעיף בווריאציה על הגדרת הגבול. יש מקרים שבהם פונקציה אינה מוגדרת בסביבה מנוקבת של נקודה אבל עדיין היינו רוצים לדבר על הגבול שלה שם. זהו למשל המקרה באיור 7.4.9, שם f אינה מוגדרת משמאל ל־x_0 אבל כש־x מתקרב ל־x_0 מצד ימין, $f(x)$ מתקרב ל־L.

כדי לתאר קרבה לנקודה מאחד הצדדים נזדקק למושג הבא:



איור 7.4.10 סביבה ימנית של x_0

הגדרה 7.4.6 יהי $x_0 \in \mathbb{R}$ ו־ $r > 0$. **הסביבה ימנית** (המלאה) ברדיוס r של x_0 היא הקטע $[x_0, x_0+r)$, **והסביבה הימנית המנוקבת** ברדיוס r של x_0 היא הקטע (x_0, x_0+r). באופן דומה, **הסביבה השמאלית** (המלאה) ברדיוס r של x_0 היא הקטע $(x_0-r, x_0]$, **והסביבה השמאלית המנוקבת** ברדיוס r של x_0 היא הקטע (x_0-r, x_0).

הגדרה 7.4.7 תהי f פונקציה ממשית ו־ $x_0 \in \mathbb{R}$. מספר ממשי L נקרא **הגבול מימין** (right limit) של f ב־x_0 (או: הגבול של f כש־x שואף ל־x_0 מימין) אם לכל סביבה U של L, קטנה ככל שתהיה, קיימת סביבה ימנית מנוקבת V^+ של x_0 כך שלכל $x \in V^+$ מתקיים $f(x) \in U$ (ובפרט, f מוגדרת בכל נקודה ב־V^+). במקרה זה מסמנים $\lim_{x\to x_0+} f(x) = L$.

באופן דומה, מספר ממשי L נקרא **הגבול משמאל** (left limit) של f ב־x_0 (או: הגבול של f כש־x שואף ל־x_0 משמאל) אם לכל סביבה U של L, קטנה ככל שתהיה, קיימת סביבה שמאלית מנוקבת V^- של x_0 כך שלכל $x \in V^-$ מתקיים $f(x) \in U$ (ובפרט, f מוגדרת בכל נקודה ב־V^-). במקרה זה מסמנים $\lim_{x\to x_0-} f(x) = L$.[9]

כשנרצה להדגיש שמדובר בגבול במובן הרגיל ולא בגבול חד־צדדי נאמר גבול **דו־צדדי**.

לגבול החד־צדדי יש תכונות דומות לגבול הדו־צדדי. למשל, תנאי הכרחי לקיום של גבול חד צדדי הוא שהפונקציה תהייה מוגדרת בסביבה חד־צדדית מתאימה של הנקודה. הגבולות החד־צדדיים יחידים. אם שתי פונקציות מסכימות בסביבה ימנית של x_0 אז הגבול מימין קיים עבור אחת מהן אמ"מ הוא קיים עבור השנייה, ואם הגבולות קיימים הם שווים. תכונה דומה מתקיימת לגבולות משמאל. ההוכחה של תכונות אלה דומה למקרה הדו־צדדי, ומושארת כתרגיל.

דוגמאות

1. לפונקציית הסימן יש גבולות מימין ומשמאל ב־0, והם שווים בהתאמה ל־1 ול־ -1. נוכיח למשל ש־ $\lim_{x\to 0+} \operatorname{sgn} x = 1$: אם U סביבה של 1 אז $(0,1)$ היא סביבה ימנית מנוקבת של x_0, ולכל $x \in (0,1)$ מתקיים $f(x) = 1 \in U$.

2. נראה שלפונקציה $f(x) = \sqrt{x}$ יש גבול מימין ב־0 והוא שווה ל־0. שימו לב שפונקציה זו אינה מוגדרת כלל משמאל ל־0. אנו ננחש שהגבול שווה לערך של הפונקציה באפס, דהיינו לערך $f(0) = 0$. אפשר להשתכנע על ידי חישוב כמה ערכים.

   נשים לב שלכל $\delta > 0$, אם $0 < x < \delta$ אז f מוגדרת ב־x ו־ $0 < f(x) < \sqrt{\delta}$. לכן בהינתן סביבה $U = B_\varepsilon(0)$ של 0, אם נבחר $\delta = \varepsilon^2$ אז לכל $0 < x < \delta$

[9] ישנם ספרים שבהם הגבול מימין מסומן ב־ $\lim_{x\searrow x_0} f(x)$ או ב־$f(x+)$, והגבול משמאל מסומן ב־ $\lim_{x\nearrow x_0} f(x)$ או ב־$f(x-)$.

מתקיים $\varepsilon = \sqrt{\delta} < f(x) < 0$. לכן עבור הסביבה הימנית המנוקבת $V = (0, \delta)$ של 0 מתקיים

$$x \in V \quad \Longrightarrow \quad f(x) \in U$$

ומכאן ש־ $\lim_{x \to 0+} f(x) = 0$, כנדרש.

מהדוגמה הראשונה אנו רואים שאי קיום הגבול אינו פוסל את קיום הגבולות החד־צדדיים בנקודה, או במילים אחרות קיום הגבולות החד־צדדיים אינו מבטיח שהגבול הדו־צדדי קיים. בכל זאת ניתן לאפיין את קיום הגבול במונחים של גבולות חד־צדדיים:

טענה 7.4.8 תהי f פונקציה. אז הגבול הדו־צדדי של f ב־x_0 קיים אמ"מ הגבולות החד־צדדיים של f ב־x_0 קיימים ושווים זה לזה. במקרה זה ערך הגבול שווה לערך המשותף של הגבולות החד־צדדיים.

הוכחה נניח שהגבולות החד־צדדיים קיימים ושווים לערך משותף L. תהי U סביבה של L. יש סביבה ימנית מנוקבת V^+ וסביבה שמאלית מנוקבת V^- של x_0 כך ש־f מוגדרת בהן ומתקיים

$$\begin{aligned} x \in V^+ \quad &\Longrightarrow \quad f(x) \in U \\ x \in V^- \quad &\Longrightarrow \quad f(x) \in U \end{aligned}$$

קל לבדוק מההגדרות שהקבוצה $V^- \cup V^+$ מכילה סביבה מנוקבת של x_0, שנסמנה W (הוכיחו זאת!). אז לכל $x \in W$ מתקיים אחד מהתנאים $x \in V^-$ או $x \in V^+$ ולכן f מוגדרת ב־x ומתקיים $f(x) \in U$. לכן $\lim_{x \to x_0} f(x) = L$.

הכיוון השני מידי. בהנחה ש־ $\lim_{x \to x_0} f(x) = L$ נראה למשל שהגבול מימין ב־x_0 קיים ושווה ל־L (המקרה של הגבול משמאל דומה). אם U סביבה של L אז יש $\delta > 0$ כך שהסביבה מנוקבת $V = B^*_\delta(x_0)$ של x_0 מקיימת שאם $x \in V$ אז $f(x) \in U$. נגדיר $V^+ = (x_0, x_0 + \delta)$, שהיא סביבה ימנית מנוקבת של x_0. מכיוון ש־ $V^+ \subseteq V$ הרי לכל $x \in V^+$ מתקיים $f(x) \in U$. מכאן ש־ $\lim_{x \to x_0+} f(x) = L$. באופן דומה הסביבה השמאלית $V^- = (x_0 - \delta, x_0)$ מוכיחה שהגבול משמאל קיים ושווה ל־L. ■

תרגילים

1. הראו שאם U, V סביבות מנוקבות של נקודה $x \in \mathbb{R}$ אז גם $U \cap V$ ו־$U \cup V$ הן סביבות מנוקבות של x. כמו־כן הראו שאם V^+ היא סביבה ימנית מנוקבת של x_0 ו־V^- היא סביבה שמאלית מנוקבת של x_0 אז $V^+ \cup V^-$ מכילה סביבה מנוקבת של x_0.

2. הוכיחו את השוויונות הבאים:

(א) $\lim_{x\to 0} x = 0$. (ב) $\lim_{x\to 0}(x^2+1) = 1$. (ג) $\lim_{x\to 0} 2^x = 1$.

3. לכל אחת מהפונקציות הבאות, מצאו באילו נקודות יש לה גבול וחשבו אותו:

(א) $-x$. (ב) $ax+b$ (a, b קבועים). (ג) $|x|$.
(ד) $\sqrt{x}$. (ה) x^3. (ו) $\frac{x-1}{x+1}$.
(ז) $\frac{x^2-1}{x+1}$. (ח) $\frac{x^2-2}{x+1}$. (ט) $x \cdot [x]$.
(י) $\frac{[x]}{x}$. (יא) $x \cdot [\frac{1}{x}]$. (יב) $x \cdot \{x\}$.
(יג) $\frac{\{x\}}{x}$. (יד) $x\{\frac{1}{x}\}$.

4. לאילו ערכים של a, b יש לפונקציה

$$f(x) = \begin{cases} ax & x \le 1 \\ x^2 + b & x > 1 \end{cases}$$

גבול ב־1?

5. לאילו מהפונקציות הבאות יש גבול ב־0? לאילו יש גבולות חד־צדדיים?

(א) $1/x$.
(ב) $\{1/x\}$.
(ג) $[1/x]$.
(ד) $e^{1/x}$.
(ה) $e^{-1/|x|}$.

6. תהי $\{a_n\}_{n=1}^{\infty}$ סדרה ותהי $f : (-1, 0) \cup (0, 1) \to \mathbb{R}$ מוגדרת על ידי

$$f(x) = a_{[1/|x|]}$$

הראו ש־ $\lim_{x\to 0} f(x)$ קיים אמ"מ $\lim_{n\to\infty} a_n$ קיים, ואם הם קיימים אז הם שווים.

7. נגדיר $f : \mathbb{R} \to \mathbb{R}$ על ידי

$$f(x) = \begin{cases} 1 & x = \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots \\ 0 & \text{אחרת} \end{cases}$$

באיזה נקודות יש ל־f גבול?

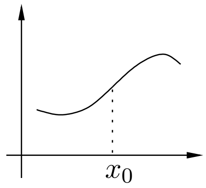
8. שאלה זו מכלילה את השאלה הקודמת. עבור קבוצה $A \subseteq \mathbb{R}$ נגדיר את **הפונקציה המציינת** (indicator function) של A להיות הפונקציה המסומנת 1_A ומוגדרת לכל $x \in \mathbb{R}$ על ידי

$$1_A(x) = \begin{cases} 1 & x \in A \\ 0 & \text{אחרת} \end{cases}$$

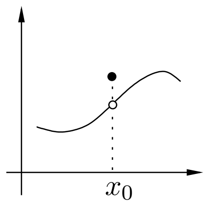
הוכיחו של־1_A יש גבול בנקודה x_0 אמ"מ A מכילה סביבה מנוקבת של x_0 או $A \setminus \mathbb{R}$ מכילה סביבה מנוקבת של x_0. מתי יש לה גבולות מימין או משמאל?

9. הוכיחו את יחידות הגבולות החד־צדדיים. הוכיחו גם שהגבול החד־צדדי הוא תכונה "אינפיניטסימלית חד־צדדית": הראו שאם שתי פונקציות מסכימות בסביבה ימנית מנוקבת של x_0 אז לאחת יש גבול מימין ב־x_0 אמ"מ גם לשנייה יש, ואם הגבולות קיימים אז הם שווים. נסחו את הטענה גם למקרה של גבול השמאלי.

7.5 רציפות בנקודה



איור 7.5.1 פונקציה רציפה ב־x_0



איור 7.5.2 פונקציה שאינה רציפה ב־x_0, למרות שיש לה שם גבול

נדמיין חלקיק שנע בקו ישר, ותהי s הפונקציה כך ש־$s(t)$ מתארת את מיקומו בזמן t. משיקולים פיזיקליים ברור של־s יש גבול בכל נקודה, אך יותר מכך: אם החלקיק הולך ומתקרב לנקודה L כאשר t שואף ל־t_0, אז בזמן t_0 הוא מגיע ממש ל־L. באופן מדויק,

הגדרה 7.5.1 פונקציה ממשית f נקראת **רציפה** (continuous) בנקודה x_0 אם היא מוגדרת ב־x_0, יש לה גבול ב־x_0, ומתקיים $\lim_{x\to x_0} f(x) = f(x_0)$. אומרים אז ש־$x_0$ היא **נקודת רציפות** של f.

מבין שתי הפונקציות המתוארות באיורים 7.5.1,7.5.2, הראשונה רציפה ב־x_0 והשנייה לא.

שימו לב שאם f רציפה ב־x_0 אז היא בפרט מוגדרת שם, וקיום הגבול של f ב־x_0 גורר ש־f מוגדרת בסביבה מנוקבת של x_0. מכאן שתנאי הכרחי לרציפות ב־x_0 היא שהפונקציה תהיה מוגדרת בסביבה מלאה של x_0.

דוגמאות

1. תהי $g(x) = x^2+1$. ראינו בסעיף הקודם ש־

$$\lim_{x\to 1} g(x) = 2 = g(1)$$

ולכן g רציפה ב־1.

2. נגדיר $\widehat{g}:\mathbb{R}\to\mathbb{R}$ על ידי

$$\widehat{g}(x) = \begin{cases} x^2+1 & x\neq 1 \\ 0 & x=1 \end{cases}$$

אז $g,\widehat{g}$ מסכימות בסביבה מנוקבת של 1 ולכן

$$\lim_{x\to 1}\widehat{g}(x) = \lim_{x\to 1} g(x) = 2 \neq 0 = g(1)$$

לכן $\widehat{g}$ אינה רציפה ב־1.

3. לפונקציה sgn אין גבול ב־0 ולכן ממילא אינה רציפה שם.

כמו מושג הגבול, גם רציפות היא תכונה אינפיניטסימלית של פונקציה, אלא שהיא תלויה גם בערך של f ב־x_0.

טענה 7.5.2 יהיו f,g פונקציות ממשיות. אם קיימת סביבה מלאה V של x_0 כך ש־$f|_V = g|_V$, אז f רציפה ב־x_0 אמ"מ g רציפה ב־x_0.

הוכחה די להראות שאם f רציפה ב־x_0 גם g רציפה שם (הטענה ההפוכה זהה למעט שינוי התפקידים של f ו־g). מההנחה ש־f,g מסכימות בסביבה של x_0 נובע בפרט שהן מסכימות בסביבה מנוקבת ולכן

$$\lim_{x\to x_0} g(x) = \lim_{x\to x_0} f(x)$$

מכיוון ש־f רציפה ב־x_0 מתקיים

$$\lim_{x\to x_0} f(x) = f(x_0)$$

לבסוף, מכיוון ש־f,g מסכימות בסביבה מלאה של x_0 מתקיים בפרט

$$f(x_0) = g(x_0)$$

וחיבור שלושת השוויוניות נותן $\lim_{x\to x_0} g(x) = g(x_0)$, דהיינו g רציפה ב־x_0. ■

אפשר לאפיין רציפות מבלי להזכיר במפורש גבולות. האפיונים מעט פשוטים יותר מהאפיונים של הגבול כי אין בהם צורך להבחין בין הנקודה x_0 לנקודות אחרות בקרבתה.

משפט 7.5.3 תהי f פונקציה ממשית ו־ $x_0 \in \mathbb{R}$. התנאים הבאים שקולים:

1. f רציפה ב־x_0.
2. f מוגדרת בסביבה מלאה של x_0 ולכל סביבה U של $f(x_0)$ קיימת סביבה (מלאה) V של x_0 כך שאם $x \in V$ אז $f(x) \in U$ (ובפרט f מוגדרת ב־x).
3. f מוגדרת בסביבה מלאה של x_0 ולכל $\varepsilon > 0$ קיים $\delta > 0$ כך שאם $|x-x_0| < \delta$ אז $|f(x) - f(x_0)| < \varepsilon$ (ובפרט f מוגדרת ב־x).

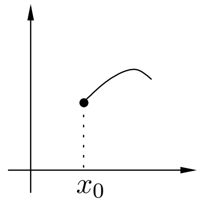
הערה לפני שניתן את ההוכחה, שימו לב להבדלים בין סעיפים (2),(3) של המשפט לבין הגדרת הגבול 7.4.2 ואפיון הגבול בלשון $\varepsilon-\delta$. בסעיף (2), V היא סביבה מלאה ולא סביבה מנוקבת, בסעיף (3) התנאי על x הוא $|x-x_0| < \delta$ ולא $0 < |x-x_0| < \delta$.

הוכחה השקילות בין סעיפים (2) ו־(3) היא מידית מההגדרה של סביבה. ולכן די שנראה שסעיפים (1),(2) שקולים.

נניח את (2). רציפות ב־x_0 פירושה $\lim_{x\to x_0} f(x) = f(x_0)$, ולכן עלינו להראות שאם U סביבה של $f(x_0)$ אז שיש סביבה מנוקבת של x_0 שבכל נקודה x בה מקיימת $f(x) \in U$. מההנחה קיימת סביבה מלאה V עם תכונה זו, ולכן התכונה בוודאי מתקיימת בסביבה המנוקבת $W = V \setminus \{x_0\}$, כנדרש.

להפך, נניח את (1). עלינו להראות שלכל סביבה U של $f(x_0)$ יש סביבה מלאה של x_0 שבכל נקודה x בה מקיימת $f(x) \in U$. לפי ההנחה $\lim_{x\to x_0} f(x) = f(x_0)$, ולכן יש סביבה מנוקבת V של x_0 כך שלכל $x \in V$ מתקיים $f(x) \in U$. מאחר שממילא $f(x_0) \in U$ (כי U היא סביבה של $f(x_0)$), הרי שלכל נקודה x בסביבה המלאה $W = V \cup \{x_0\}$ של x_0 מתקיים $f(x) \in U$ (ובפרט f מוגדרת ב־x), והראינו ש־(2) מתקיים. ■

נעבור לדון בגרסה החד־צדדית של רציפות:

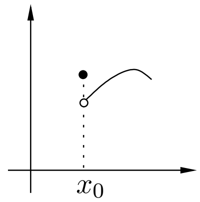


איור 7.5.3 פונקציה רציפה מימין ב־x_0

הגדרה 7.5.4 תהי f פונקציה ממשית. נאמר ש־f **רציפה מימין** (right continuous) ב־x_0 אם f מוגדרת ב־x_0 ומתקיים $\lim_{x\to x_0+} f(x) = f(x_0)$ (ובפרט הגבול קיים). באותו אופן, נאמר ש־f **רציפה משמאל** (left continuous) ב־x_0 אם f מוגדרת ב־x_0, ל־f יש גבול משמאל ב־x_0, ומתקיים $\lim_{x\to x_0-} f(x) = f(x_0)$. אם פונקציה רציפה מימין או משמאל אומרים שהיא רציפה **חד־צדדית**.

דוגמאות לרציפות ואי־רציפות חד־צדדית מופיעים באיורים 7.5.3,7.5.4.

טענה 7.5.5 פונקציה ממשית f רציפה בנקודה x_0 אמ"מ היא רציפה ב־x_0 מימין ומשמאל.



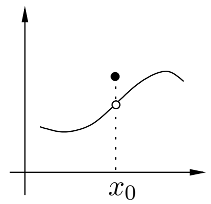
איור 7.5.4 פונקציה שאינה רציפה מימין ב־x_0, למרות שיש לה שם גבול מימין

הטענה נובעת בקלות מטענה 7.4.8, והוכחתה מושארת כתרגיל.

לרציפות מימין ומשמאל יש אפיונים ישירים כמו במשפט 7.5.3. כמו־כן זו תכונה אינפיניטסימלית, כלומר היא מקיימת משפט המקביל למשפט 7.4.5. אנו משאירים את מלאכת הניסוח וההוכחה של טענות אלה כתרגיל.

עבור פונקציות המוגדרות בסביבה מנוקבת של x_0 אפשר למיין את סוגי האי־רציפות האפשריים לשלושה סוגים. אנו נתעניין בפונקציות המוגדרות בסביבה דו־צדדית של x_0, אם כי אפשר לתת הגדרות דומות גם במקרה החד־צדדי.

הגדרה 7.5.6 תהי f פונקציה ממשית המוגדרת בסביבה מנוקבת של x_0, ונניח ש־f אינה רציפה ב־x_0. אז x_0 נקראת נקודת אי־רציפות **סליקה** של f אם ל־f קיים גבול ב־x_0.



איור 7.5.5 אי־רציפות סליקה

הסיבה שאי־רציפות כזו נקראת "סליקה" היא שניתן לסלק אותה על־ידי שינוי הערך של f בנקודה x_0 בלבד. אם נסמן $L = \lim_{x\to x_0} f(x)$, ונגדיר $\hat{f} : \operatorname{dom} f \cup \{x_0\} \to \mathbb{R}$ על־ידי

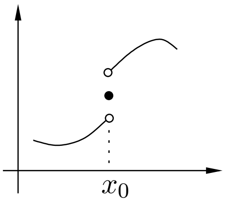
$$\hat{f}(x) = \begin{cases} f(x) & x_0 \neq x \in \operatorname{dom} f \\ L & x = x_0 \end{cases}$$

אז $f, \hat{f}$ נבדלות זו מזו רק בערך שלהן ב־x_0 וממילא הן מסכימות בסביבה מנוקבת של x_0, ולכן

$$\lim_{x\to x_0} \hat{f}(x) = \lim_{x\to x_0} f(x) = L = \hat{f}(x_0)$$

כלומר: $\hat{f}$ רציפה ב־x_0.

קל לייצר דוגמאות לאי־רציפות סליקה: אם f רציפה ב־x_0 פשוט נשנה את הערך שלה ב־x_0 ונקבל אי־רציפות סליקה שם.



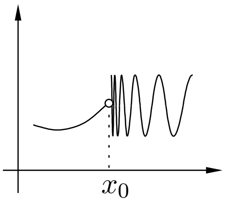
איור 7.5.6 אי־רציפות מסוג ראשון

הגדרה 7.5.7 תהי f פונקציה ממשית המוגדרת בסביבה מנוקבת של x_0, ונניח ש־f אינה רציפה ב־x_0. אז x_0 נקראת נקודת אי־רציפות **מסוג ראשון** אם הגבולות החד־צדדיים של f ב־x_0 קיימים, אבל אינם שווים.

דוגמאות

1. לפונקציה sgn יש אי־רציפות מסוג ראשון בנקודה 0, כי $\lim_{x\to 0-} \operatorname{sgn} = -1$ ו־ $\lim_{x\to 0+} \operatorname{sgn} = 1$.
2. לפונקציה $f(x) = \frac{1}{1+e^{1/x}}$ יש אי־רציפות מסוג ראשון ב־0 (הוכיחו זאת! או ראו דוגמה (1) מעמוד 250 להלן).

אי אפשר לתקן אי רציפות מסוג ראשון על־ידי שינוי ערך הפונקציה בנקודה x_0, שכן שינוי הערך בנקודה אינו משנה את העובדה שהגבול בנקודה לא קיים. אבל על ידי שינוי ערך הפונקציה בנקודה ניתן להפוך אותה לרציפה מימין או רציפה משמאל (איך?), אם כי לא בו זמנית.



איור 7.5.7 אי־רציפות מסוג שני

הגדרה 7.5.8 תהי f פונקציה ממשית המוגדרת בסביבה מנוקבת של x_0, ונניח ש־f אינה רציפה ב־x_0. אז x_0 נקראת נקודת אי־רציפות **מסוג שני**, אם לפחות אחד מהגבולות החד־צדדיים של f ב־x_0 אינו קיים.

למשל, ראינו בסעיף על גבולות חד־צדדיים שלפונקציות $\frac{1}{x}$ ו־ $\sin\frac{1}{x}$ אין גבולות חד־צדדיים ב־0, ולכן האי־רציפות שלהן ב־0 הוא מסוג שני.

כפי שראינו בעמוד 238, קיום הגבול בנקודה אינו גורר קיום גבול בנקודות סמוכות. באופן דומה רציפות בנקודה אינה גוררת רציפות בנקודות סמוכות. נסיים את הסעיף בדוגמה קיצונית עוד יותר שבה קיימות זו לצד זו נקודות רציפות ונקודות אי־רציפות רבות. בדוגמה להלן, בין כל שתי נקודות רציפות יש נקודת אי רציפות, ובין כל שתי נקודות אי־רציפות יש נקודת רציפות.

פונקציית רימן היא הפונקציה $R : \mathbb{R} \to \mathbb{R}$ הנתונה על ידי

$$R(x) = \begin{cases} \frac{1}{q} & x = \frac{r}{q} \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

לצורך ההגדרה הזו נסכים שכשמציגים מספר רציונלי כמנה $\frac{r}{q}$ של מספרים שלמים השבר הוא שבר מצומצם והמכנה חיובי. לכל מספר רציונלי יש הצגה יחידה כזו (ראו עמוד 45), ולכן ההגדרה של R חד־משמעית.

למה 7.5.9 לכל $x_0 \in \mathbb{R}$ ולכל n טבעי, קיימת סביבה מנוקבת U של x_0 כך שלכל נקודה רציונלית $\frac{r}{q} \in U$ מתקיים $q > n$

הוכחה יהי $n \in \mathbb{N}$. נראה תחילה שבכל קטע מהצורה $(k-1, k+1)$ (k שלם) יש מספר סופי של מספרים רציונליים עם מכנה קטן מ־n. לכל $q > 0$ טבעי יש לכל היותר $2q-1$ שברים בין $k-1$ ל־ $k+1$ שהמכנה שלהם הוא q, שכן הם נמנים בין המספרים

$$\frac{kq-q+1}{q}, \frac{kq-q+2}{q}, \ldots, \frac{kq}{q}, \ldots \frac{kq+q-2}{q}, \frac{kq+q-1}{q}$$

(ייתכן שלחלק מהמספרים האלה יש גם הצגה עם מכנה קטן יותר מ־q). לכן מספר השברים בקטע $(k-1, k+1)$ בעלי מכנה שאינו עולה על n חסום על ידי

$$\sum_{q=1}^{n}(2q-1) \leq \sum_{q=1}^{n}(2n-1) = n(2n+1)$$

וממילא יש מספר סופי של שברים כאלה.

יהי $x_0 \in \mathbb{R}$ ויהי $k = [x_0]$ הערך השלם של x_0. מתקיים $x_0 \in (k-1, k+1)$. בקטע זה יש מספר סופי של שברים עם מכנה קטן או שווה ל־n, ולכן יש סביבה מנוקבת של x_0 שלא מכילה אף אחד מהם. באופן מפורט יותר, נגדיר

$$r = \min\{|x_0 - \frac{r}{q}| \,:\, x_0 \neq \frac{r}{q} \in (k-1, k+1) \,,\, q \leq n\}$$

(המינימום קיים כי הקבוצה מימין סופית, והיא חיובית כי הקבוצה מכילה רק מספרים חיוביים). אז בסביבה המנוקבת $B_r^*(x_0)$ אין מספרים רציונליים עם מכנה קטן או שווה ל־q. ■

נראה כעת שלכל $x_0 \in \mathbb{R}$ מתקיים $\lim_{x\to x_0} R(x) = 0$. יהי $x_0 \in \mathbb{R}$, יהי $\varepsilon > 0$, ויהי n טבעי כך ש־ $\frac{1}{n} < \varepsilon$. לפי הלמה יש סביבה מנוקבת V של x_0 כך שכל מספר רציונלי $\frac{r}{q} \in V$ מקיים $q > n$. עבור $x \in V$, אם x רציונלי אפשר לרשום $x = \frac{r}{q}$ ואז

$$|R(x) - 0| = |R(\frac{r}{q})| = \frac{1}{q} < \frac{1}{n} < \varepsilon$$

אחרת x אי־רציונלי, ואז ממילא

$$|R(x) - 0| = |0 - 0| = 0 < \varepsilon$$

קיבלנו שלכל $x \in V$ מתקיים $|R(x) - 0| < \varepsilon$ ולכן $\lim_{x\to x_0} R(x) = 0$, כפי שרצינו.

לבסוף, מכיוון ש־ $R(x) = 0$ אמ"מ $x \in \mathbb{R} \setminus \mathbb{Q}$, אנו מסיקים ש־$R$ רציפה בדיוק בנקודות האי־רציונליות.

תרגילים

1. מצאו את נקודות הרציפות והרציפות החד־צדדית של הפונקציות מתרגילים (3) ו־(5) בסוף סעיף 7.4, וסווגו את נקודות האי־רציפות שלהן.
2. נניח ש־f רציפה בנקודה x_0. האם ניתן לשנות את הערך שלה ב־x_0 כך שתישאר רציפה?
3. תהי $f:(-1,1)\to\mathbb{R}$ פונקציה רציפה ב־0. האם ייתכן ש־f שלילית ב־$(-1,0)$ וחיובית ב־$(0,1)$? האם ייתכן שיש $t>0$ כך ש־$f(x)<-t$ עבור $x<0$ ו־$f(x)>t$ עבור $x>0$?
4. תהי $f:\mathbb{R}\to\mathbb{R}$ פונקציה רציפה בכל נקודה.
   (א) נניח ש־$f(r)=0$ לכל $r\in\mathbb{Q}$. הוכיחו כי $f\equiv 0$.
   (ב) באופן כללי יותר, הראו שאם $f,g:\mathbb{R}\to\mathbb{R}$ רציפות בכל נקודה ואם $f(r)=g(r)$ לכל $r\in\mathbb{Q}$ אז $f=g$.
5. תהי $f:\mathbb{R}\to\mathbb{R}$ פונקציה אי־זוגית. הוכיחו ש־f רציפה ב־0 אמ"מ מתקיים $\lim_{x\to 0+} f(x)=0$.
6. הוכיחו מהמשפטים בסעיף זה שהפונקציה

$$f(x)=\begin{cases}\frac{1}{2}x & x\le\frac{1}{2}\\ x^2 & x>\frac{1}{2}\end{cases}$$

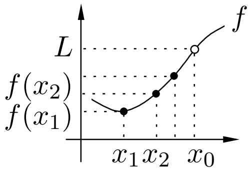
   רציפה ב־$\frac{1}{2}$ (אם תרצו, הוכיחו זאת גם מההגדרות).
7. נסחו והוכיחו גרסה של משפט 7.5.3 לרציפות חד־צדדית. נסחו והוכיחו גם משפט על כך שרציפות חד־צדדית היא תכונה אינפיניטסימלית חד־צדדית.
8. עבור $n=1,2,3,\ldots$ יהיו x_n נקודות רציפות של פונקציה $f:(a,b)\to\mathbb{R}$. נניח ש־$x_n\to x_0\in(a,b)$. האם x_0 בהכרח נקודת רציפות של f?
9. תהי $A=\{q\sqrt{2}\,:\,q\in\mathbb{Q}\}$ ותהי $f=1_A$ הפונקציה המציינת של A (כלומר $f(x)=1$ אם $x\in A$ ו־$f(x)=0$ אם $x\notin A$). הוכיחו ש־f אינה רציפה באף נקודה.
10. יהיו $f,g:\mathbb{R}\to\mathbb{R}$ הפונקציות

$$f(x)=\begin{cases}k & x=\frac{k}{2^n}\,,\,k,n\in\mathbb{N}\\ 0 & \text{אחרת}\end{cases}\qquad,\qquad g(x)=\begin{cases}\frac{1}{n} & x=\frac{k}{2^n}\,,\,k,n\in\mathbb{N}\\ 0 & \text{אחרת}\end{cases}$$

   (השבר $\frac{k}{2^n}$ מצומצם). מצאו את נקודות הרציפות של f,g.

7.6 אפיון היינה ותנאי קושי

האינטואיציה מאחורי מושג הגבול היתה ש־L הוא הגבול של f ב־x_0 אם הערכים $f(x)$ קרובים ל־L כאשר x קרוב ל־x_0. לכן נצפה שלכל סדרת מספרים (x_n) המתכנסת ל־x_0 הסדרה $f(x_n)$ תתכנס ל־L. המשפט הבא מראה שאמנם כך הדבר, ויתר על כן שאפשר לאפיין בצורה זו את הגבול:



איור 7.6.1

משפט 7.6.1 (אפיון היינה[10] לגבול ורציפות) תהי f פונקציה ממשית.

1. $\lim_{x\to x_0} f(x) = L$ אמ"מ f מוגדרת בסביבה מנוקבת של x_0 ולכל סדרת נקודות (x_n) המקיימת $x_0 \neq x_n \to x_0$ מתקיים $f(x_n) \to L$.
2. f רציפה ב־x_0 אמ"מ f מוגדרת בסביבה מלאה של x_0 ולכל סדרת נקודות (x_n) המקיימת $x_n \to x_0$ מתקיים $f(x_n) \to L$.

הערה איננו מניחים ש־f מוגדרת בכל הנקודות x_n, אבל ההנחות מבטיחות ש־f מוגדרת ב־x_n החל ממקום מסוים.

הוכחה נוכיח רק את הסעיף הראשון. הסעיף השני נובע ממנו בקלות (איך?).

נניח ש־ $\lim_{x\to x_0} f(x) = L$. תהי (x_n) סדרה כך ש־ $x_0 \neq x_n \to x_0$. עלינו להראות ש־ $f(x_n) \to L$. תהי U סביבה של L. אנו רוצים להראות שלכל n גדול מספיק מתקיים $f(x_n) \in U$. מה שאנו יודעים הוא שיש סביבה מנוקבת $V = B^*_\delta(x_0)$ של x_0 כך ש־f מוגדרת בכל V ומתקיים $f(x) \in U$ לכל $x \in V$. תהי W הסביבה המלאה סביב x_0 עם אותו רדיוס כמו V, כלומר $W = B_\delta(x_0)$. מכיוון ש־ $x_n \to x_0$ אנו יודעים שלכל n מספיק גדול מתקיים $x_n \in W$, וכיוון ש־ $x_n \neq x_0$ לכל n נובע ש־ $x_n \in V$ החל ממקום מסוים. לכן החל ממקום מסוים $f(x_n) \in U$.

בכיוון השני, נניח בשלילה ש־ $\lim_{x\to x_0} f(x) \neq L$ ונראה שקיימת סדרה (x_n) כך ש־ $x_0 \neq x_n \to x_0$ אבל $f(x_n) \not\to L$. לפי ההנחה ש־ $\lim_{x\to x_0} f(x) \neq L$, יש סביבה U של L כך שלכל סביבה מנוקבת V של x_0, יש $x \in V$ עם $f(x) \notin U$.

עבור $n \in \mathbb{N}$ נגדיר $V_n = B^*_{1/n}(x_0)$, ולפי האמור בפסקה הקודמת קיימת נקודה $x_n \in V_n$ שבה f מוגדרת ושמתקיים $f(x_n) \notin U$. מכאן שהסדרה $(f(x_n))_{n=1}^\infty$ אינה מתכנסת ל־L. מצד שני לפי בחירת ה־x_n־ים מתקיים $0 < |x_n - x_0| < 1/n$ ולכן מכלל הסנדוויץ' $|x_n - x_0| \to 0$, כלומר $x_n \to x_0$. הסדרה (x_n) היא הסדרה המבוקשת. ■

אפשר לרשום את המסקנה של אפיון היינה גם כך: f רציפה ב־x_0 אמ"מ לכל סדרה $x_n \to x_0$ מתקיים

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$$

[10] Heinrich Heine, 1821־1881.

משום כך אומרים לפעמים שכשהגבול קיים, פעולת הגבול מתחלפת עם הפעולה של הפונקציה.

אפיון היינה מאפשר להשתמש בשיטות שפיתחנו עבור סדרות כדי לחשב גבולות של פונקציות:

דוגמאות

1. תהי $g(x) = x^2 + 1$, כמו בדוגמה (3) מעמוד 235. נראה שוב שהגבול של g ב־1 הוא 2, הפעם בעזרת אפיון היינה. g מוגדרת בכל הישר ובפרט בסביבה מנוקבת של 1. לכן עלינו להראות שלכל סדרה (x_n) עם $1 \neq x_n \to 1$ מתקיים $g(x_n) \to 2$. ואמנם מאריתמטיקה של גבולות
$$\lim_{n\to\infty} g(x_n) = \lim_{n\to\infty}(x_n^2 + 1) = (\lim_{n\to\infty} x_n)^2 + 1 = 1^2 + 1 = 2$$
כנדרש.

2. יהי $\alpha \in \mathbb{R}$ ותהי $f : (0, \infty) \to \mathbb{R}$ פונקציית החזקה $f(x) = x^\alpha$. לכל מספר $x_0 > 0$ הפונקציה f מוגדרת בסביבה של x_0 ולפי כללי האריתמטיקה של החזקה, לכל סדרה x_n המתכנסת ל־x_0 מתקיים
$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_n^\alpha = x_0^\alpha = f(x_0)$$
(משפט 5.10.8). לכן f רציפה בכל נקודה $x_0 > 0$, ובפרט יש לה גבול שם. היא גם רציפה מימין ב־0 כאשר $\alpha \geq 0$ (הוכיחו!).

3. יהי $a > 0$ ותהי $f : \mathbb{R} \to \mathbb{R}$ הפונקציה $f(x) = a^x$. אז f מוגדרת בסביבה מלאה של כל נקודה x_0, ולפי כללי האריתמטיקה של החזקה, לכל סדרה x_n המתכנסת ל־x_0 מתקיים
$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} a^{x_n} = a^{x_0} = f(x_0)$$
(למה 5.10.5). לכן f רציפה בכל נקודה בתחום הגדרתה, ובפרט יש לה גבול בכל נקודה.

4. תהי sgn פונקציית הסימן ונראה שוב שאין לה גבול ב־0. לשם כך מספיק למצוא סדרה $0 \neq x_n \to 0$ כך שהסדרה $(\operatorname{sgn}(x_n))_{n=1}^\infty$ מתבדרת. נבחר את הסדרה כך ש־x_n תהיה לסירוגין חיובית ושלילית, למשל $x_n = \frac{(-1)^n}{n}$. אז (x_n) מקיימת את הדרישות באפיון היינה אבל הסדרה $\operatorname{sgn}(x_n) = (-1)^n$ מתבדרת ולכן $\lim_{n\to\infty} \operatorname{sgn}(x_n)$ אינו קיים.

גם לגבולות ולרציפות חד־צדדית יש אפיונים בעזרת סדרות. ננסח למשל את התנאי לגבולות משמאל:

משפט 7.6.2 (אפיון היינה לגבול משמאל) תהי f פונקציה ממשית. אז מתקיים $\lim_{x\to x_0-} f(x) = L$ אמ"מ f מוגדרת בסביבה שמאלית מנוקבת של x_0 ולכל סדרת נקודות (x_n) המקיימת $x_n < x_0$ ו־ $x_n \to x_0$ מתקיים $f(x_n) \to L$.

ניסוח שאר האפיונים והוכחתם מושארים כתרגיל.

דוגמאות

1. תהי $f(x) = \frac{1}{1+e^{1/x}}$. אם $0 < x_n \to 0$ אז $\frac{1}{x_n} \to \infty$, ולכן $e^{1/x_n} \to \infty$. מכאן ש־ $f(x_n) = (1+e^{1/x_n})^{-1} \to 0$. מאפיון היינה לגבול מימין נובע ש־ $\lim_{x\to 0+} f(x) = 0$. מצד שני, אם $0 > x_n \to 0$ אז $\frac{1}{x_n} \to -\infty$, לכן $e^{1/x_n} \to 0$ לכן $f(x_n) = (1+e^{1/x_n})^{-1} \to 1$, ומכאן $\lim_{x\to 0-} f(x) = 1$. לכן לפונקציה f אין גבול ב־0, כי הגבולות החד־צדדיים שלה שונים שם.

2. תהי $f : (0,\infty) \to \mathbb{R}$ נתונה על ידי $f(x) = \frac{1}{x}$. ל־f אין גבול מימין כי $f(\frac{1}{n}) = n \to \infty$ בעוד ש־$(\frac{1}{n})_{n=1}^{\infty}$ היא סדרה חיובית השואפת לאפס, ולפי אפיון היינה אם לפונקציה היה גבול L אז היה מתקיים $\lim_{n\to\infty} f(\frac{1}{n}) = L$.

3. תהי $g : \mathbb{R}\setminus\{0\} \to \mathbb{R}$ נתונה על ידי $g(x) = \sin\frac{1}{x}$ ונראה שאין ל־g גבול מימין ב־0. שכן g מקבלת ערך 1 בנקודות $\frac{1}{2\pi k+\pi/2}$ וערך -1 בנקודות $\frac{1}{2\pi k+3\pi/2}$ (k שלם). נקודות אלה חיוביות ונותנות שתי סדרות $(x_k),(x'_k)$ כך ש־ $g(x_k) \to 1$ ו־ $g(x'_k) \to -1$. לכן ל־g אין גבול מימין ב־0.

בעזרת אפיון היינה אפשר לנסח תנאים נוספים להתכנסות אשר דומים לתנאי הפנימיים שמצאנו לסדרות. הלמה הבאה למשל דומה לטענה שסדרה מתכנסת אמ"מ כל תת־סדרה שלה מתכנסת:

למה 7.6.3 תהי f פונקציה ממשית ו־ $x_0 \in \mathbb{R}$. אז $\lim_{x\to x_0} f(x)$ קיים אמ"מ f מוגדרת בסביבה מנוקבת של x_0 ולכל סדרה $(x_n)_{n=1}^{\infty}$ המקיימת $x_0 \neq x_n \to x_0$, הסדרה $(f(x_n))_{n=1}^{\infty}$ מתכנסת.

הוכחה אם הגבול $\lim_{x\to x_0} f(x)$ קיים ושווה למספר L אז לפי אפיון היינה, לכל סדרה (x_n) עם $x_0 \neq x_n \to x_0$ מתקיים $\lim_{n\to\infty} f(x_n) = L$ ובפרט הסדרה $(f(x_n))$ מתכנסת.

להפך, נניח שלכל סדרה (x_n) עם $x_0 \neq x_n \to x_0$ הסדרה $(f(x_n))$ מתכנסת. נבחר סדרה (x_n) כזאת ויהי $L = \lim_{n\to\infty} f(x_n)$. אנו טוענים שהגבול של f ב־x_0 הוא L. לשם כך די להוכיח שלכל סדרה (y_n) עם $x_0 \neq y_n \to x_0$ מתקיים $f(y_n) \to L$ (זה אפיון היינה). ואמנם, אילו הייתה סדרה (y_n) כזאת עם $f(y_n) \not\to L$ אז היה אפשר להגדיר סדרה (z_n) על ידי

$$(z_n)_{n=1}^{\infty} = x_1, y_2, x_2, y_2, x_3, y_3, \ldots$$

והיה מתקיים $x_0 \neq z_n \to x_0$ אף על פי ש־$(f(z_n))$ היא סדרה מתבדרת (למה?), וזאת בסתירה להנחה. ■

משפט 7.6.4 (תנאי קושי לגבול של פונקציה בנקודה) תהי f פונקציה ממשית ו־$x_0 \in \mathbb{R}$. אז הגבול $\lim_{x\to x_0} f(x)$ קיים אמ"מ לכל $\varepsilon > 0$ קיימת סביבה מנוקבת V של x_0 כך שלכל $x, x' \in V$ מתקיים $|f(x) - f(x')| < \varepsilon$.

הוכחה אם הגבול של f ב־x_0 קיים ושווה ל־L אז יש סביבה מנוקבת V של x_0 כך שלכל $x', x'' \in V$ מתקיים

$$\begin{aligned} |f(x') - L| &< \varepsilon \\ |f(x'') - L| &< \varepsilon \end{aligned}$$

ומכאן שלכל x', x'' כאלה

$$|f(x') - f(x'')| \le |f(x') - L| + |L - f(x'')| < 2\varepsilon$$

וזה נותן את הדרוש.

בכיוון ההפוך, נניח ש־ $x_0 \neq x_n \to x_0$. לפי הלמה הקודמת די שנראה ש־ $(f(x_n))$ מתכנסת. יהי $\varepsilon > 0$, ותהי V סביבה מנוקבת של x_0 כך שלכל $x', x'' \in V$ שבהם f מוגדרת מתקיים $|f(x') - f(x'')| < \varepsilon$ (קיומה של הסביבה מובטח מההנחה). החל ממקום מסוים $x_n \in V$, כלומר יש N כך שלכל $m, n > N$ מתקיים $x_n, x_m \in V$, ולכן לכל $m, n > N$ מתקיים $|f(x_m) - f(x_n)| < \varepsilon$. קיבלנו ש־$(f(x_n))_{n=1}^{\infty}$ מקיימת תנאי קושי לסדרות, ולכן מתכנסת, כנדרש. ■

קיים גם אפיון פנימי מקביל לרציפות:

משפט 7.6.5 (תנאי קושי לרציפות של פונקציה בנקודה) תהי f פונקציה ממשית. אז f רציפה ב־x_0 אמ"מ לכל $\varepsilon > 0$ קיימת סביבה מלאה V של x_0 כך שלכל $x', x'' \in V$ מתקיים $|f(x') - f(x'')| < \varepsilon$.

ההוכחה דומה להוכחה של משפט 7.6.4 (למעשה היא מעט קלה יותר) ומושארת כתרגיל.

תרגילים

1. פתרו את תרגילים (3),(5) בסוף סעיף 7.4 בעזרת אפיון היינה.
2. תהי $f(x) = \sin\frac{1}{x}$ אם $x \neq 0$ ו־ $f(0) = 0$. יהי $x_n = \frac{1}{\pi n}$. אז $0 \neq x_n \to 0$ ומתקיים $f(x_n) = \sin \pi n = 0 \to 0$. נסיק מכאן ש־ $\lim_{x\to 0} f(x) = 0$. אבל ראינו קודם של־f אין גבול ב־0. היכן השגיאה?

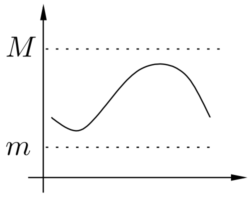
3. הוכיחו את סעיף (2) של משפט 7.6.1. הוכיחו את משפט 7.6.2, ונסחו והוכיחו גרסה של משפט זה לרציפות חד־צדדית.
4. הוכיחו ש־ $\lim_{x\to x_0+} f(x) = L$ אמ"מ לכל סדרה עולה ממש (a_n) המתכנסת ל־x_0 מתקיים $\lim_{n\to\infty} f(a_n) = L$.
5. תהי f פונקציה המוגדרת וחסומה בסביבה מנוקבת של x_0. נגדיר

$$\begin{aligned}\limsup_{x\to x_0} f(x) &= \sup\{\limsup_{n\to\infty} f(x_n) \,:\, x_0 \neq x_n \to x_0\}\\ \liminf_{x\to x_0} f(x) &= \inf\{\liminf_{n\to\infty} f(x_n) \,:\, x_0 \neq x_n \to x_0\}\end{aligned}$$

הוכיחו שהחסם העליון והתחתון בהגדרות קיים והראו ש־$\lim_{x\to x_0} f(x)$ קיים אמ"מ $\limsup_{x\to x_0} f(x) = \liminf_{x\to x_0} f(x)$, ובמקרה זה כל השלושה שווים.

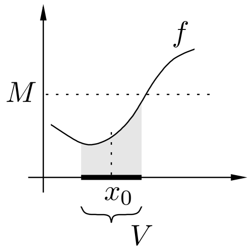
7.7 אי־שוויונות ואריתמטיקה של גבולות

בסעיף זה נוכיח משפטי אריתמטיקה ואי־שוויונות לגבולות של פונקציות. התוצאות מקבילות למשפטים שהוכחנו על גבולות של סדרות.



איור 7.7.1 פונקציה החסומה מלעיל על־ידי M ומלרע על־ידי m

הגדרה 7.7.1 תהי f פונקציה ממשית ו־M מספר ממשי. נאמר ש־M **חסם מלעיל** (upper bound) של f, ונסמן $f \leq M$, אם $f(x) \leq M$ לכל $x \in \operatorname{dom} f$. אם לפונקציה קיים חסם מלעיל אז אומרים שהפונקציה **חסומה מלעיל** (bounded above). באותו אופן M הוא **חסם מלרע** (lower bound) של f אם $f(x) \geq M$ לכל $x \in \operatorname{dom} f$, ובמקרה זה נסמן $M \leq f$. אם יש ל־f חסם מלרע אומרים שהפונקציה **חסומה מלרע** (bounded below). מספר M נקרא **חסם** של f אם $-M \leq f(x) \leq M$ (כלומר, $|f(x)| \leq M$) לכל $x \in \operatorname{dom} f$. אם יש ל־$f$ חסם אומרים ש־f **חסומה** (bounded). את היחסים $f > M$ ו־ $f < M$ מגדירים בצורה דומה.



איור 7.7.2 פונקציה החסומה מלעיל על־ידי M בסביבה V של x_0

מבחינה גאומטרית, $f \leq M$ אם הגרף של f נמצא כולו מתחת לישר המאוזן בגובה M (ואולי נוגעים בו). פונקציה היא חסומה אם הגרף מוכל כולו בפס שבין שני ישרים מאוזנים. איור 7.7.2 מדגים יחסים אלה.

כרגיל, נאמר שאי־שוויון $f \leq M$ מתקיים בקבוצה מסוימת D אם $f|_D \leq M$, דהיינו אם $f(x) \leq M$ לכל $x \in D$, ונאמר שהיא מתקיימת בסביבה מלאה או מנוקבת של x_0 אם יש סביבה מתאימה של x_0 שבה זה מתקיים. כך גם לגבי תכונות אחרות של הפונקציה.

דוגמאות

1. הפונקציה הקבועה $f \equiv c$ חסומה על ידי $|c|$.

2. הפונקציה $f(x) = x$ חסומה בקטע $[0,1]$ כי שם מתקיים $0 \leq f(x) \leq 1$, אבל היא אינה חסומה ב־$\mathbb{R}$ כי התמונה שלה היא כל $\mathbb{R}$.
3. הפונקציה $f : (0,\infty) \to R$ המוגדרת על ידי $f(x) = \frac{1}{x}$ אינה חסומה מלעיל, כי לכל $M > 0$ מתקיים $\frac{1}{2M} \in (0,\infty)$ ו־ $f(\frac{1}{2M}) = 2M > M$, אבל היא חסומה מלרע כי $f(x) \geq 0$ לכל $x \in (0,\infty)$.
4. הפונקציה $g(x) = \frac{1}{x}$ אינה חסומה מלעיל או מלרע (הוכיחו!).

נשוב לדיון בתכונות של הגבול. אם לסדרה יש גבול הרי היא חסומה. התכונה המקבילה של פונקציות היא חסימות מקומית:

טענה 7.7.2 אם $\lim_{x\to x_0} f(x) = L$ ואם A,B הם מספרים כך ש־ $A < L < B$ אז יש סביבה מנוקבת של x_0 שבה מתקיים $A < f(x) < B$. אם בנוסף f רציפה ב־x_0 אז אפשר לבחור את V להיות סביבה מלאה.

הוכחה תהי U סביבה של L המקיימת $U \subseteq (A,B)$, כמו למשל $U = B_\varepsilon(L)$ עבור $\varepsilon = \min\{B-L, L-A\}$. מקיום הגבול קיימת סביבה מנוקבת V של x_0 כך שלכל $x \in V$ מתקיים $f(x) \in U$, וזה גורר $A < f(x) < B$ לכל $x \in V$, כפי שרצינו.

במקרה ש־f רציפה $L = f(x_0)$, והסביבה $V \cup \{x_0\}$ מקיימת את הדרוש. ■

מסקנה 7.7.3 אם ל־f יש גבול ב־x_0 ו־f מוגדרת ב־x_0 אז f חסומה בסביבה מלאה של x_0.

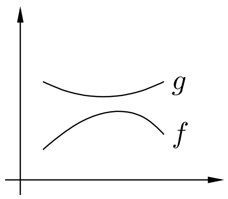
הוכחה לפי המשפט יש $A < B$ וסביבה מנוקבת V של x_0 שבה $A < f < B$. אז $W = V \cup \{x_0\}$ היא סביבה מלאה של x_0 ולכל $x \in W$ מתקיים

$$\min\{A, f(x_0)\} \leq f(x) \leq \max\{B, f(x_0)\}$$

ובפרט f חסומה ב־W. ■

מסקנה 7.7.4 אם $\lim_{x\to x_0} f(x) > 0$ אז יש $A > 0$ כך ש־ $f > A$ בסביבה מנוקבת של x_0. בפרט אם f רציפה ב־x_0 ואם $f(x_0) > 0$ יש $A > 0$ כך ש־ $f > A$ בסביבה מלאה V של x_0.

עד כאן הסקנו מסקנות מקיום הגבול. נפנה לשיטות להערכת הגבול על ידי השוואה עם פונקציות אחרות:



איור 7.7.3 אי־שוויון בין פונקציות: g גדולה מ־f

הגדרה 7.7.5 יהיו f,g פונקציות ממשיות המוגדרות בקבוצה $D \subseteq \mathbb{R}$. נאמר ש־g גדולה מ־f ב־D, או בקיצור: $f < g$ ב־D, אם $f(x) < g(x)$ לכל $x \in D$. אי־שוויון חלש בין פונקציות מוגדר באופן דומה.

מבחינה גאומטרית, אם $f < g$ אז הגרף של g נמצא כולו מעל הגרף של f, כמו באיור 7.7.3.

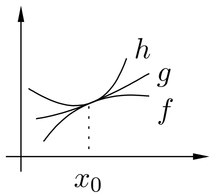
טענה 7.7.6 יהיו f, g פונקציות ממשיות ונניח ש־ $f \le g$ בסביבה מנוקבת של x_0. אם ל־f, g יש גבולות ב־x_0 אז $\lim_{x\to x_0} f(x) \le \lim_{x\to x_0} g(x)$.

הוכחה תהי V סביבה מנוקבת של x_0 כך ש־ $f \le g$ ב־V. בפרט, נובע מכך ש־ f, g מוגדרת בכל $x \in V$. תהי (x_n) סדרה כך $x_0 \ne x_n \to x_0$. אז החל ממקום מסוים $x_n \in V$ ולכן $f(x_n) \le g(x_n)$. מכאן

$$\lim_{x\to x_0} f(x) = \lim_{n\to\infty} f(x_n) \le \lim_{n\to\infty} g(x_n) = \lim_{x\to x_0} g(x)$$

השוויונות הקיצוניים נובעים מאפיון היינה לגבול, והאי־שוויון האמצעי מטענה 5.3.4 על גבולות של סדרות. ■

שימו לב שאי־שוויון חזק $f < g$ בסביבה מנוקבת של x_0 אינו מספיק כדי להבטיח ש־ $\lim_{x\to x_0} f(x) < \lim_{x\to x_0} g(x)$, כפי שאי־שוויון חזק $a_n < b_n$ בין האיברים של סדרות אינו מבטיח $\lim a_n < \lim b_n$. דוגמה קונקרטית בהקשר הנוכחי הן הפונקציות $f(x) = |x|$ ו־ $g(x) = 0$, המקיימת $f > g$ בסביבה מנוקבת של 0, אבל הגבול של שתי הפונקציות ב־0 הוא 0.



איור 7.7.4

משפט 7.7.7 (משפט הסנדוויץ') יהיו f, g, h פונקציות ממשיות ונניח ש־ $f \le g \le h$ בסביבה מנוקבת של x_0 (ובפרט שיש סביבה מנוקבת בה כל שלוש הפונקציות מוגדרות). אם ל־f, h יש גבולות ב־x_0 ואם

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} h(x) = L$$

אז גם ל־g יש גבול ב־x_0 והוא שווה ל־L.

הערה כמו במשפט המקביל לסדרות, החשיבות של המשפט היא בכך שהוא מבטיח את קיום הגבול של g ב־x_0. אילו ידענו שהגבול קיים אז השוויון בין הגבולות היה נובע מהטענה הקודמת.

הוכחה תהי V סביבה מנוקבת של x_0 שבה $f \le g \le h$. תהי (x_n) סדרה כך ש־ $x_0 \ne x_n \to x_0$. החל ממקום מסוים $x_n \in V$ ולכן מהאי־שוויון הנתון החל ממקום מסוים בסדרה מתקיים

$$f(x_n) \le g(x_n) \le h(x_n)$$

מאחר ו־

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} h(x) = L$$

נובע מאפיון היינה ש־

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} h(x_n) = L$$

ולכן מהאי־שוויון $f(x_n) \le g(x_n) \le h(x_n)$ וממשפט הסנדוויץ' לסדרות נסיק ש־ $\lim_{n\to\infty} g(x_n) = L$. מכיוון ש־$(x_n)$ סדרה שרירותית כנ"ל, נובע מאפיון היינה ש־ $\lim_{x\to x_0} g(x) = L$, כפי שרצינו. ■

מסקנה 7.7.8 יהיו f, g, h פונקציות ממשיות ונניח ש־ $f \leq g \leq h$ בסביבה מלאה של x_0. אם f, h רציפות ב־x_0 ואם $f(x_0) = h(x_0)$ אז g רציפה ב־x_0.

הוכחה ממשפט הסנדוויץ' מתקיים $\lim_{x \to x_0} g(x) = f(x_0) = h(x_0)$ ולכן רציפות g ב־x_0 ינבע מהשוויון $f(x_0) = g(x_0)$. אבל נתון ש־ $f(x_0) \leq g(x_0) \leq h(x_0)$ וגם $f(x_0) = h(x_0)$, כלומר $f(x_0) \leq g(x_0) \leq f(x_0)$, ולכן יש שוויון, כפי שרצינו. ■

הטענה הבאה מקבילה לטענה על סדרות לפיה אם $a_n \to 0$ ו־(b_n) חסומה אז $a_n b_n \to 0$. ההוכחה דומה להוכחות הקודמות, ומושארת כתרגיל.

טענה 7.7.9 יהיו $f, g : D \to \mathbb{R}$ פונקציות, ונניח ש־g חסומה בסביבה מנוקבת של x_0, וש־ $\lim_{x \to x_0} f(x) = 0$. אז $\lim_{x \to x_0} (f \cdot g)(x) = 0$.

דוגמאות

1. תהי $f(x) = x \cdot \sin e^x$ ונראה ש־ $\lim_{x \to 0} f(x) = 0$. שכן $-1 \leq \sin e^x \leq 1$ לכל x ולכן $\sin e^x$ היא פונקציה חסומה, ואילו $\lim_{x \to 0} x = 0$. לכן מטענה 7.7.9 מתקיים $\lim_{x \to 0} x \cdot \sin e^x = 0$.

2. תהי $f(x) = x + e^{-x} x^2$ ונראה ש־ $\lim_{x \to 0} f(x) = 0$. שכן לכל $-1 < x < 1$ מתקיים

$$x < x + e^{-x} x^2 < (1+e)|x|$$

(למה?). קל לוודא שהפונקציות x ו־$(1+e)|x|$ שואפות ל־0 כאשר x שואף לאפס, ולכן מכלל הסנדוויץ' מקבלים את הטענה.

3. נראה שהפונקציה $\sin$ רציפה ב־0. שכן לכל $x \in (0, \frac{\pi}{2})$ מתקיים $0 < \sin x < x$ ומאי־זוגיות של $\sin$ נובע שלכל $x \in (-\frac{\pi}{2}, 0)$ מתקיים $-x < \sin x < 0$. אנו מסיקים שלכל $x \in B^*_{\pi/2}(0)$ מתקיים $-x < \sin x < x$ ומכלל הסנדוויץ', $\lim_{x \to 0} \sin x = 0 = \sin 0$.

נעבור לפעולות החשבוניות. כפי שהגדרנו פעולות אריתמטיות על סדרות, ניתן להגדיר פעולות אריתמטיות על פונקציות. יהיו f, g פונקציות ממשיות בעלות תחום משותף D. נוכל להגדיר פונקציה חדשה $s : D \to \mathbb{R}$ שערכה בנקודה $x \in D$ הוא $s(x) = f(x) + g(x)$. פונקציה זו נקראת **הסכום** של f, g ומסומנת בקיצור $f + g$. באותו אופן ניתן להגדיר פונקציית **הפרש** $f - g$, פונקציית **מכפלה** $f \cdot g$, פונקציית **מנה** $\frac{f}{g}$, פונקציית **חזקה** f^g, ופונקציית **הערך המוחלט** $|f|$. עבור $c \in \mathbb{R}$ אפשר להגדיר פונקציות cf, c^f, f^c, וכן הלאה. חלק מהפונקציות האלה מוגדרות רק על תת־קבוצה של התחום המשותף D. למשל הפונקציה $\frac{f}{g}$ מוגדרת רק באותן נקודות $x \in D$ שבהן $g(x) \neq 0$.

הערה יש הבדל בין הביטוי $f(x)+g(x)$ לבין הביטוי $(f+g)(x)$, אף שהם מציינים את אותו המספר. הראשון הוא סכום של שני המספרים $f(x)$ ו־$g(x)$, והשני הוא הערך של הפונקציה $f+g$ בנקודה x.

הפעולות בין פונקציות מקיימות תכונות דומות לפעולות בין מספרים. הנה למשל ההוכחה לכך שהפונקציות $f+g$ ו־ $f+g$ שוות: לכל x בתחום המשותף שלהן מתקיים

$$(f+g)(x) = f(x)+g(x) = g(x)+f(x) = (g+f)(x)$$

(בשוויון האמצעי השתמשנו באקסיומת החילוף של המספרים הממשיים). קיבלנו איפוא ש־$(f+g)(x) = (g+f)(x)$ לכל x בתחום, $f+g = g+f$.

טענה 7.7.10 תהי f פונקציה ממשית. התנאים הבאים שקולים:

1. $\lim_{x\to x_0} f(x) = L$
2. $\lim_{x\to x_0}(f-L)(x) = 0$
3. $\lim_{x\to x_0} |f-L|(x) = 0$

טענה זו מקבילה לטענה 5.4.1 על סדרות. גם ההוכחה דומה, ומושארת כתרגיל.

משפט 7.7.11 (כללי תחשיב לגבולות) יהיו f,g פונקציות ממשיות עם תחום משותף ויהי $c\in\mathbb{R}$. נניח ש־ $\lim_{x\to x_0} f(x) = L$ ו־ $\lim_{x\to x_0} g(x) = M$. אז לפונקציות $|f|, cf, f+g, f\cdot g$ יש גבולות ב־x_0 ומתקיים

$$\begin{aligned}
\lim_{x\to x_0} |f|(x) &= |L| \\
\lim_{x\to x_0} (cf)(x) &= cL \\
\lim_{x\to x_0} (f+g)(x) &= L+M \\
\lim_{x\to x_0} (f\cdot g)(x) &= L\cdot M
\end{aligned}$$

אם בנוסף $M\neq 0$ אז לפונקציה $\frac{f}{g}$ יש גבול ב־x_0 ומתקיים

$$\lim_{x\to x_0} \frac{f}{g}(x) = \frac{L}{M}$$

הערה הפונקציה $\frac{f}{g}$ אינה בהכרח מוגדרת בכל התחום של f,g אלא רק בקבוצת הנקודות בהן g אינה מתאפסת. אבל לפי מסקנה 7.7.4, ההנחה $\lim_{x\to x_0} g(x)\neq 0$ גוררת שיש סביבה מנוקבת של x_0 שבה g אינה מתאפסת, ולכן גם $\frac{f}{g}$ מוגדרת בסביבה מנוקבת של x_0.

הוכחה ההוכחה מבוססת על משפטי האריתמטיקה של גבולות לסדרות ועל אפיון היינה.

נוכיח לדוגמה את הטענה על פונקציית הסכום. די שנראה שלכל סדרה (x_n), אם $x_0 \neq x_n \to x_0$ אז $(f+g)(x_n) \to L+M$. אבל אם (x_n) סדרה כזו אז $f(x_n) \to L$ ו־ $g(x_n) \to M$, ומהגדרת $f+g$ מתקיים

$$(f+g)(x_n) = f(x_n) + g(x_n) \to L + M$$

(השוויון נובע מכלל הסכום לסדרות), כנדרש. ■

כרגיל, המשפט מניח את קיום הגבולות של f ו־g. אין זה נכון, למשל, שאם $\lim_{x \to x_0}(f \cdot g)(x)$ קיים אז הגבולות $\lim_{x \to x_0} f(x)$ ו־$\lim_{x \to x_0} g(x)$ קיימים. למשל, אנו יודעים כבר שלפונקציה sgn אין גבול ב־0, אבל $\mathrm{sgn}^2 = \mathrm{sgn} \cdot \mathrm{sgn} = 1$ בסביבה מנוקבת של 0 ולכן יש לה גבול שם.

מסקנה 7.7.12 יהיו $f, g : D \to \mathbb{R}$ פונקציות ממשיות ו־$c \in \mathbb{R}$. אם f, g רציפות ב־x_0 אז $|f|, cf, f+g, f \cdot g$ רציפות ב־x_0 ואם $g(x_0) \neq 0$ אז גם $\frac{f}{g}$ רציפה ב־x_0.

הוכחה נוכיח למשל את כלל המנה. נתון ש־ $g(x_0) \neq 0$ ולפי הנחת הרציפות, $\lim_{x \to x_0} g(x) = g(x_0) \neq 0$. לכן מאריתמטיקה של גבולות ומהרציפות של f, g ב־x_0,

$$\lim_{x \to x_0} \frac{f}{g}(x) = \frac{\lim_{x \to x_0} f(x)}{\lim_{x \to x_0} g(x)} = \frac{f(x_0)}{g(x_0)} = \frac{f}{g}(x_0)$$

ולכן $\frac{f}{g}$ רציפה ב־x_0. ■

דוגמאות

1. מהמשפט אנו מסיקים, למשל, שלפונקציה $x^2 + \sqrt{x}$ יש גבול בכל תחום הגדרתה, כי זה נכון לכל אחד מהמחוברים.
2. פולינום רציף בכל נקודה. ואמנם, פונקציית הזהות $i(x) = x$ רציפה בכל הישר, וחזקה שלמה שלה $i^k(x) = x^k$ רציפה גם־כן, כי זו מכפלה חוזרת של פונקציות רציפות. כך לכל a קבוע הפונקציה ax^k רציפה. כעת על ידי הפעלה חוזרת של הכלל שסכום של פונקציות רציפות הוא פונקציה רציפה מקבלים שכל פונקציה מהצורה $p(x) = \sum_{k=0}^{d} a_k x^k$, כלומר כל פולינום רציף בכל נקודה בישר.
3. לפעמים ניתן להשתמש במשפט האריתמטיקה של גבולות כדי להוכיח שאין גבול. נראה למשל שלפונקציה $f(x) = \mathrm{sgn}(x) + 2x^3 - 2x + 1$ אין גבול ב־0. נסמן $p(x) = 2x^3 - 2x + 1$. אז ל־p יש גבול ב־0 (למה?). אילו היה קיים

ל־f גבול ב־0, אז לפי משפט האריתמטיקה גם ל־ $f-p$ היה קיים גבול ב־0. אבל $f-p=\operatorname{sgn}$ ואנו יודעים של־sgn אין גבול ב־0, ולכן לא ייתכן של־f יש גבול שם.

אפשר להוכיח גם משפט אריתמטיקה לפונקציות מהצורה f^α עבור $a>0$. משפט זה גם ינבע משיקולים שנציג בסעיף הבא.

תרגילים

1. אילו מהפונקציות הבאות חסומות בתחום הגדרתן?
   (א) x.
   (ב) $\frac{1}{1+x}$.
   (ג) $\frac{1}{1+x^2}$.
   (ד) e^{x^2}.
   (ה) e^{-x^2}.
2. יהיו f,g פונקציות עם תחום משותף. הוכיחו או הפריכו:
   (א) סכום ומכפלה של פונקציות חסומות הן פונקציות חסומות.
   (ב) אם f חסומה אז $\frac{1}{f}$ חסומה.
   (ג) אם f,g חסומות, $\sup f=A$ ו־ $\sup g=B$ אז $\sup(f+g)=A+B$ (כאן $\sup f=\sup\{f(x)\,:\,x\in\operatorname{dom} f\}$ וכו').
3. מצאו דוגמה לפונקציה $f:\mathbb{R}\to\mathbb{R}$ שאינה חסומה באף קטע (נסו לשנות את ההגדרה של פונקציית רימן מעמוד 245).
4. יהיו f,g פונקציות עם תחום משותף. הוכיחו או הפריכו:
   (א) אם f חסומה בסביבה של x_0 אז $\lim_{x\to x_0} f$ קיים.
   (ב) אם $\lim_{x\to x_0} f$ קיים ו־g חסומה ב־x_0 אז $\lim_{x\to x_0} fg$ קיים.
   (ג) אם $\lim_{x\to x_0} f=M$ אז יש סביבה מנוקבת של x_0 שבה $f\le M$.
   (ד) אם g אינה חסומה בסביבה של x_0 ואם f חסומה שם אז $\lim_{x\to x_0} fg$ אינו קיים.
5. חשבו את הגבולות הבאים:
   (א) $\lim_{x\to 0}\frac{x^2}{x+1}$.
   (ב) $\lim_{x\to 0} x\cdot e^{-(1+x^2)}$.
   (ג) $\lim_{x\to 0} x\cdot e^{1+x^2}$.
   (ד) $\lim_{x\to 0} 2^{x/(1+|x|)}$.
6. תהי f פונקציה חיובית. אילו מהגבולות הבאים קיימים, ומה ערכם, תחת כל אחת מההנחות (i) f מוגדרת וחסומה בסביבה מנוקבת של 0, (ii) ל־f יש גבול ב־0, (iii) f רציפה ב־0:

(א) $\lim_{x\to 0} xf(x)$.

(ב) $\lim_{x\to 0} \frac{f(x)}{x+1}$.

(ג) $\lim_{x\to 0} \frac{1}{f(x)}$.

(ד) $\lim_{x\to 0} \frac{1}{1+f(x)^2}$.

(ה) $\lim_{x\to 0} \frac{x}{f(x)}$.

7. הוכיחו שהגבול $\lim_{x\to 0}(e^x \operatorname{sgn} x)$ אינו קיים.

8. תהי f פונקציה המוגדרת בסביבה מנוקבת של 0 ונניח של־f אין גבול ב־0. אילו מהגבולות הבאים קיימים? האם התשובה משתנה אם מניחים ש־f חסומה בסביבה מנוקבת של 0?

(א) $\lim_{x\to 0}(x + f(x))$.

(ב) $\lim_{x\to 0} xf(x)$.

(ג) $\lim_{x\to 0}(x+1)f(x)$.

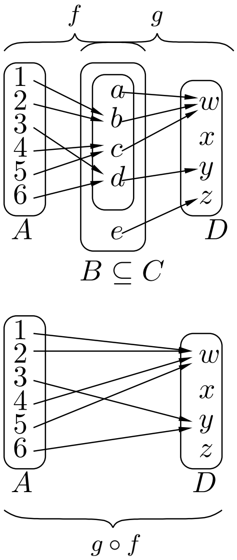
(ד) $\lim_{x\to 0} \frac{x}{f(x)}$.

9. הוכיחו ש־$\cos$ רציפה ב־0 (היעזרו בזהויות הטריגונומטריות בסעיף 7.3).

7.8 פעולת ההרכבה

בהינתן שתי פונקציות f, g, אפשר לקבל פונקציה חדשה על ידי הפעלת הפונקציות f, g בזו אחר זו. ליתר דיוק:

הגדרה 7.8.1 יהיו $f : A \to B$, $g : C \to D$ פונקציות ונניח ש־ $\operatorname{image} f \subseteq C$. **ההרכבה** (*composition*) של g עם f היא הפונקציה מ־A ל־D המתאימה לאיבר $a \in A$ את האיבר $g(f(a))$. פונקציה זו מסומנת על ידי $g \circ f$, ונקראת g **מורכב על** f.



איור 7.8.1 ההרכבה של פונקציות

ההנחה $\operatorname{image} f \subseteq C$ בהגדרה מבטיחה שלכל $a \in A$ האיבר $f(a)$ שייך לתחום של g ולכן $g(f(a))$ אכן מוגדר. מספיק כמובן להניח $B \subseteq C$.

מדיאגרמות החיצים של f ושל g אפשר לקבל דיאגרמה של $g \circ f$ על ידי "שרשור" החיצים, כלומר: חיבור ראשי החיצים בדיאגרמה של f עם זנבות החיצים בדיאגרמה של g. כך מקבלים חיצים "ארוכים" מהתחום של f לטווח של g, אשר מתארים את $g \circ f$. ראו איור 7.8.1.

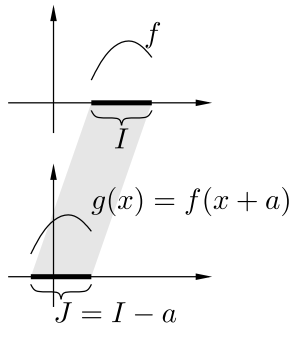
אפשר להרכיב יותר משתי פונקציות אחת אחרי השנייה. אם f, g, h פונקציות כך ש־ $\operatorname{image} f \subseteq \operatorname{dom} g$ ו־ $\operatorname{image} g \subseteq \operatorname{dom} h$ אז אפשר ליצור את הפונקציה $h \circ g \circ f : A \to D$ המתאימה ל־ $a \in A$ את האיבר $h(g(f(a)))$. הפונקציה $h \circ g \circ f$ שווה הן ל־$(h \circ g) \circ f$ והן ל־$h \circ (g \circ f)$. באופן דומה אפשר להגדיר הרכבה של כל מספר סופי של פונקציות.

דוגמאות

1. תהי f הפונקציה המתאימה לכל אדם את אמו, ותהי g הפונקציה המתאימה לכל אדם את אביו. אז $g \circ f$ היא הפונקציה המתאימה לכל אדם את סבא שלו מצד אמו.
2. הרכבה היא פעולה מוכרת מאד כשמדובר בפונקציות ממשיות, שם היא נקראת לפעמים "הצבה". אם $w(x) = \sqrt{x+1}$ אז $w = v \circ u$ כאשר $u(x) = x+1$ ו־ $v(y) = \sqrt{y}$.
3. בסעיף הקודם תיארנו כמה פונקציות מורכבות. אם f פונקציה ממשית אז $|f|$ היא ההרכבה $g \circ f$ כאשר $g(y) = |y|$. אם c מספר אז הפונקציה f^c היא ההרכבה $h \circ f$ כאשר $h(y) = y^c$.
4. תהי $(a_n)_{n=1}^{\infty}$ סדרה ותהי $(n_k)_{k=1}^{\infty}$ סדרה עולה ממש של אינדקסים. נתבונן בתת־סדרה $(a_{n_k})_{k=1}^{\infty}$. אם נחשוב על (a_n) כפונקציה $a : \mathbb{N}^+ \to \mathbb{R}$ ועל (n_k) כפונקציה $n : \mathbb{N}^+ \to \mathbb{N}$ אז התת־סדרה (a_{n_k}) היא הפונקציה המורכבת $a \circ n : \mathbb{N}^+ \to \mathbb{R}$.

שימו לב לסדר הכתיבה של פעולת ההרכבה: $g \circ f$ מפעיל קודם את f ואחר־כך את g. באופן כללי לא נכון ש־ $f \circ g = g \circ f$. ייתכן שאחת ההרכבות מוגדרת והשנייה לא. למשל אם f כמו בדוגמה (1) ו־v כמו בדוגמה (2) אז $v \circ f$ אינה מוגדרת, כי התמונה של f מוכלת בקבוצת האנשים וזו אינה מוכלת בתחום של v, שהיא קבוצת המספרים האי־שליליים. גם $f \circ v$ אינה מוגדרת.

גם במקרה ששתי ההרכבות $g \circ f$, $f \circ g$ מוגדרות ייתכן שאינן שוות ביניהן. למשל בדוגמה (1) למעלה, $g \circ f$ מתאים לכל אדם את סבא שלו מצד אימא אבל $f \circ g$ מתאים לכל אדם את סבתא שלו מצד אבא.



איור 7.8.2 הזזה של פונקציה

דוגמה חשובה של הרכבה של פונקציות הן פעולות ה"הזזה" וה"מתיחה" של התחום של פונקציה ממשית. תהי $f : I \to \mathbb{R}$ כאשר $I \subseteq \mathbb{R}$. בהינתן מספר a אפשר להגדיר פונקציה g על ידי

$$g(x) = f(x+a)$$

כלומר $g = f \circ \ell$ כאשר $\ell(x) = x + a$. תחום ההגדרה של g הוא קבוצת כל ה־x־ים כך ש־ $x + a \in I$, כלומר

$$\operatorname{dom} g = \{y - a \,:\, y \in I\} = I - a$$

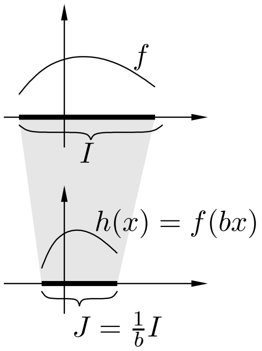
זו הקבוצה המתקבלת מהקבוצה I על ידי הזזתה מרחק a על ציר המספרים. כאשר $a > 0$ ההזזה היא לכיוון שמאל, וכאשר $a < 0$ ההזזה היא ימינה.

נשים לב ש־f מקבלת ערך y בנקודה x אמ"מ g מקבלת את הערך y בנקודה $x - a$, כי

$$g(x-a) = f((x-a)+a) = f(x)$$

במילים אחרות, (x,y) שייך לגרף של f אמ"מ $(x-a,y)$ שייך לגרף של g. מכאן שמבחינה גאומטרית, הגרף של g מתקבל מהגרף של f על ידי הזזה: אם $a>0$ מזיזים את הגרף שמאלה במרחק a, ואם $a<0$ ההזזה היא לכיוון ימין במרחק $|a|$.

הקשר בין f ל־g מתואר באיור 7.8.2. אפשר לקרוא את האיור כמו דיאגרמת חיצים, מלמטה למעלה, כפי שעשינו בדיון על הרכבת פונקציות כלליות. כאן התחומים I,J של f,g מתוארים כקטעים מודגשים. כדי לתאר את הפעולה של $f\circ\ell$ על נקודה $x\in J$, קודם מעתיקים (מזיזים) את x ל־$x+a$ ומקבלים נקודה בקטע I, ואז מעתיקים את $x+a$ לנקודה על הגרף שגובהו הוא $f(x+a)$. בסך הכול, x מועתק ל־ $f(x+a)$ תחת פעולת $g=f\circ\ell$.



איור 7.8.3 מתיחה של פונקציה

באותו אופן, בהינתן מספר $b\neq 0$ אפשר להגדיר פונקציה ממשית h על ידי

$$h(x)=f(bx)$$

ואז f מוגדרת ב־x אמ"מ $bx\in I$, כלומר

$$\operatorname{dom} h=\{\frac{y}{b}\ :\ y\in I\}=\frac{1}{b}\cdot I$$

זו הקבוצה המתקבלת מ־I על ידי "כיווץ" או "מתיחה" פי b. משיקולים דומים לקודמים אנו רואים שאם $b>1$ הגרף של h מתקבל מכיווץ הגרף של f בכיוון מאוזן לקבוצה $J=\frac{1}{b}I$. אם I קטע אז גם J קטע שאורכו $\frac{1}{|b|}$ מהאורך של I. שימו לב שבמקרה זה גם המיקום של המרכז של הקטע I משתנה. מצב זה מתואר באיור 7.8.3. עבור $b>1$ התחום "מתכווץ", עבור $0<b<1$ התחום דווקא "מתנפח", וכאשר b שלילי נוסף לכיווץ או למתיחה יש גם שיקוף דרך ציר ה־y. אנו משאירים כתרגיל בחינה מדויקת של מקרים אלו.

נשוב לענייננו העיקרי שהוא חישוב גבולות. נתבונן למשל בפונקציה $e^{\sqrt{x}}$. זו הרכבה של $g(y)=e^y$ על $f(x)=\sqrt{x}$. כאשר x קרוב ל־2 אנו יודעים ש־ $f(x)=\sqrt{x}$ קרוב ל־$\sqrt{2}$. אנו יודעים שכש־y קרוב ל־$\sqrt{2}$ הערך $g(y)$ קרוב ל־$e^{\sqrt{2}}$. לכן נצפה שכאשר x קרוב ל־2 המספר $g(f(x))=e^{\sqrt{x}}$ יהיה קרוב ל־$e^{\sqrt{2}}$.

באופן כללי יותר, יהיו f,g פונקציות ממשיות כך ש־ $g\circ f$ מוגדרת, ונניח ש־ $\lim_{x\to x_0}f(x)=y_0$ ו־ $\lim_{y\to y_0}g(y)=L$. האם זה גורר שיש ל $g\circ f$ גבול ב־x_0? באופן כללי התשובה שלילית. נתבונן למשל בפונקציות

$$f=\begin{cases} x\sin\frac{1}{x} & x\neq 0\\ 0 & x=0\end{cases}\qquad,\qquad g(y)=\begin{cases} 0 & y\neq 0\\ 1 & y=0\end{cases}$$

אז $\lim_{x\to x_0}f(x)=0$ ו־ $\lim_{y\to 0}g(y)=0$. אבל בכל סביבה של 0 הפונקציה f מקבלת גם את הערך 0 וגם ערכים שונים מ־0, ולכן $g\circ f$ מקבלת את הערכים $0,1$ בכל סביבה של 0, וממילא אין לה גבול שם.

בדוגמה האחרונה הבעיה נובעת מכך שהגבול של g ב־0 אינו תלוי בערך של g ב־0, ולעומת זאת f מתאפסת בכל סביבה של 0 ולכן הערך של g ב־0 כן משפיע על

תכונות הגבול של $g \circ f$ ב־0. המשפט הבא מראה שלמעט הבעיה הזו הגבול של פונקציה מורכבת מתנהג כמצופה:

משפט 7.8.2 יהיו f, g פונקציות ממשיות ונניח ש־ $g \circ f$ מוגדר. נניח גם שמתקיים $\lim_{x\to x_0} f(x) = y_0$ ו־ $\lim_{y\to y_0} g(y) = L$. אם מתקיים אחד מהתנאים הבאים:

1. g רציפה.
2. יש סביבה מנוקבת של x_0 שבה f אינה מקבלת את הערך y_0.

אז $\lim_{x\to x_0}(g \circ f)(x) = L$.

הוכחה עלינו להראות ש־ $g(f(x_n)) \to L$ לכל סדרה (x_n) המקיימת $x_0 \neq x_n \to x_0$. בהינתן סדרה (x_n) כזו נסמן $y_n = f(x_n)$. מכיוון ש־ $\lim_{x\to x_0} f(x) = y_0$ הרי שלפי אפיון היינה לגבול, $y_n \to y_0$. עם זאת, איננו יכולים להסיק ש־ $y_n \neq y_0$.

במקרה (1), g רציפה ב־y_0. אז לפי אפיון היינה לרציפות ומהעובדה $y_n \to y_0$, נובע ש־ $g(y_n) \to g(y_0) = L$, דהיינו $g(f(x_n)) \to L$ כפי שרצינו.

במקרה (2), תהי V סביבה מנוקבת של x_0 שבה f לא מקבלת את הערך y_0. אז החל ממקום מסוים $x_n \in V$ ולכן החל ממקום מסוים $y_n = f(x_n) \neq y_0$. לכן לפי אפיון היינה לגבולות מתקיים $\lim_{n\to\infty} g(y_n) = L$, ושוב קיבלנו $g(f(x_n)) \to L$, כפי שרצינו. ∎

מסקנה 7.8.3 יהיו f, g פונקציות ממשיות. אם f רציפה ב־x_0 ואם g רציפה ב־ $y_0 = f(x_0)$ אז $g \circ f$ רציפה ב־ x_0.

הוכחה לפי ההנחה, תנאי המשפט הקודם מתקיימים, ולכן

$$\lim_{x\to x_0} g \circ f(x) = \lim_{y\to y_0} g(y) = g(y_0) = g(f(x_0)) = g \circ f(x_0)$$

(השוויון השמאלי נובע מהמשפט הקודם, השוויון הבא מרציפות של g ב־y_0, השוויון השלישי מהגדרת y_0 והאחרון מהגדרת ההרכבה). ∎

המשפט והמסקנה תקפים בגרסאות מתאימות גם כאשר הגבול או הרציפות של f ב־x_0 הוא חד־צדדי. כמו־כן אפשר לנסח גרסאות למקרה שהגבול של g ב־y_0 הוא חד־צדדי, בתנאי שתמונת f מוכלת בסביבה (מנוקבת) של y_0 מאותו צד. ההוכחה של מקרים אלה דומה.

דוגמאות

1. הפונקציה $f(x) = e^{\sqrt{x}}$ רציפה בכל נקודה $x > 0$, כי f היא הרכבה של הפונקציה $y \mapsto e^y$ על הפונקציה $x \mapsto \sqrt{x}$. והשנייה רציפה בכל נקודה. למעשה היא רציפה מימין גם ב־0; כדי להסיק זאת מתבססים על אחת הגרסאות החד־צדדיות של המשפט.

2. תהי f פונקציה ממשית. יהי $a \in \mathbb{R}$ ותהי $\ell(x) = x + a$. אז לכל $x \neq x_0$ מתקיים $\ell(x) \neq x_0 + a$ ובפרט זה נכון בסביבה מנוקבת של x_0. לכן אם $\lim_{y \to x_0 + a} f(y)$ קיים אנו מסיקים ש־

$$\begin{aligned} \lim_{x \to x_0} f(x+a) &= \lim_{x \to x_0} f \circ \ell(x) \\ &= \lim_{y \to \ell(x_0)} f(y) \\ &= \lim_{y \to x_0 + a} f(y) \end{aligned}$$

ניתן לכתוב זאת גם כך:

$$\lim_{y \to y_0} f(y) = \lim_{x \to y_0 - a} f(x+a)$$

אפשר לפרש את הנוסחה הזאת כ"החלפת משתנה": במקום y הצבנו $x + a$ (מתחת לסימן הגבול היינו צריכים לקבל "$x + a \to y_0$", אבל העברנו אגף וקיבלנו $x \to y_0 - a$).

מבחינה גאומטרית הגרף של $f(x+a)$ הוא ההזזה של הגרף של f, ולכן תוצאה זו צפויה: פירושה שהגבול של הפונקציה המוזזת בנקודה המוזזת שווה לגבול של הפונקציה המקורית בנקודה המקורית.

שיקולים דומים מראים שאם $b \neq 0$ אז

$$\lim_{y \to y_0} f(y) = \lim_{x \to y_0 / b} f(bx)$$

הוכיחו זאת!

מקרה חשוב במיוחד של הדוגמה האחרונה היא המסקנה הבאה:

מסקנה 7.8.4 תהי f פונקציה ממשית. אז הגבול $\lim_{x \to x_0} f(x)$ קיים אמ"מ הגבול $\lim_{h \to 0} f(x_0 + h)$ קיים, ואם הם קיימים אז הם שווים.

נוכיח לדוגמה ש־$\sin$ רציפה בכל נקודה בישר. ראינו כבר ש־$\sin$ רציפה ב־0 (דוגמה (3) מעמוד 255). מכאן נובע ש־$\cos$ רציפה ב־0, כי

$$\lim_{x \to 0} \cos x = \lim_{x \to 0} (1 - 2\sin^2(\frac{x}{2})) = 1 - 2(\lim_{x \to 0} \sin(\frac{x}{2}))^2 = 1 - 0 = 1 = \cos 0 = 1$$

השתמשנו כאן בנוסחה $\cos x = 1 - 2\sin^2(\frac{x}{2})$, באריתמטיקה של גבולות, ובעובדה ש־ $\lim_{x \to 0} \sin \frac{x}{2} = 0$ (למה זה נכון?).

נראה כעת ש־$\sin$ רציפה בכל נקודה $x_0 \in \mathbb{R}$. לפי נוסחת הסכום של $\sin$ מתקיים

$$\sin(x_0 + h) = \sin x_0 \cos h + \sin h \cos x_0$$

לכן

$$\begin{aligned}\lim_{x\to x_0}\sin x &= \lim_{h\to 0}\sin(x_0+h)\\ &= \lim_{h\to 0}(\sin x_0\cos h+\sin h\cos x_0)\\ &= \lim_{h\to 0}\sin x_0\cdot\lim_{h\to 0}\cos h+\lim_{h\to 0}\sin h\cdot\lim_{h\to 0}\cos x_0)\end{aligned}$$

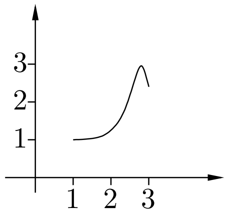
היות ו־$\sin x_0, \cos x_0$ הם קבועים (ביחס למשתנה h) הרי ש־ $\lim_{h\to 0}\sin x_0=\sin x_0$ ו־ $\lim_{h\to 0}\cos x_0=\cos x_0$. מצד שני, כפי שראינו מתקיים ש־ $\lim_{h\to 0}\sin h=0$ ו־ $\lim_{h\to 0}\cos h=1$ ולכן קיבלנו

$$\lim_{x\to x_0}\sin x=\sin x_0\cdot 1+0\cdot\cos x_0=\sin x_0$$

באופן דומה אפשר להראות ש־$\cos$ רציפה בכל נקודה, תוך שימוש בנוסחה ל־ $\cos(x+h)=\cos x\cos h-\sin x\sin h$. רציפות $\cos, \sin$ גוררת את רציפות הפונקציות $\tan, \cot$ בכל נקודה בתחום הגדרתן. הפרטים מושארים כתרגיל.

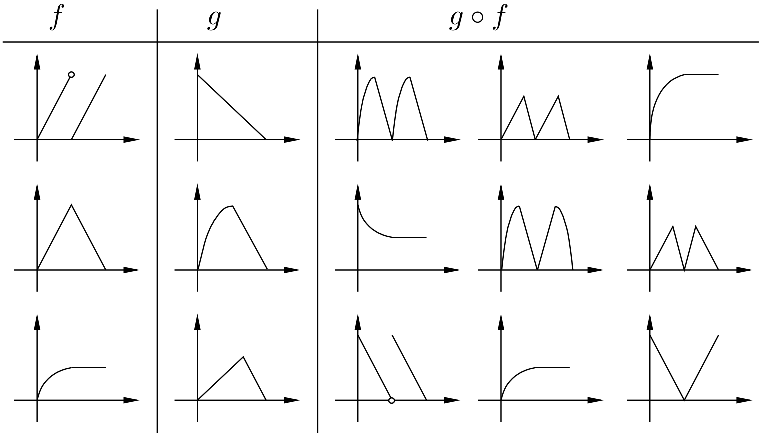
תרגילים

1. תנו דוגמה לפונקציות ממשיות f, g כך ש־$g\circ f$ מוגדר אבל $f\circ g$ אינו מוגדר.
2. האם בהכרח $\mathrm{image}\, g\circ f=\mathrm{image}\, g$?
3. תהי f הפונקציה שהגרף שלה מופיע באיור 7.8.4. ציירו גרף מקורב עבור כל אחת מהפונקציות $f(x+1)$, $f(x-2)$, $f(2x)$, $f(\frac{1}{2}x)$, $f(-2x)$, $f(-\frac{1}{2}x)$, $f(|x|)$.



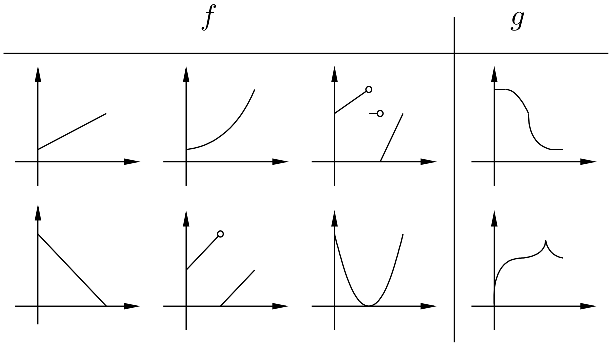
איור 7.8.4

4. יהיו $f, g:\mathbb{R}\to\mathbb{R}$. הוכיחו או הפריכו:
   (א) אם g זוגית ו־f זוגית או אי זוגית אז $g\circ f$ זוגית.
   (ב) אם f, g שתיהן אי־זוגיות אז $f\cdot g$ אי־זוגית.
   (ג) אם f, g שתיהן אי־זוגיות אז $g\circ f$ אי־זוגית.
   (ד) אם f זוגית אז $g\circ f$ זוגית לכל g
   (ה) אם f אי־זוגית אז $f+1$ אי־זוגית.
5. נניח ש־f פונקציה חסומה ו־$g\circ f$ מוגדרת. האם $g\circ f$ חסומה
6. הגרפים להלן מתארים פונקציות. לכל פונקציה f בעמודה הראשונה ולכל פונקציה g בעמודה השנייה, מצאו גרף בעמודה השלישית המתארת את $g\circ f$.



איור 7.8.5

7. לכל פונקציה f בעמודה הראשונה ולכל פונקציה g בעמודה השנייה, ציירו גרף מקורב של $g \circ f$.



איור 7.8.6

8. הוכיחו את מסקנה 7.8.3 ישירות מהגדרת הגבול.

9. חשבו את הגבולות הבאים:

(א) $\lim_{x\to 0}\sin(\sin(x))$.

(ב) $\lim_{x\to 0} e^{\sin x}$.

(ג) $\lim_{x\to 0}\left(\sin x\cdot\sin\frac{1}{x}\right)$.

(ד) $\lim_{x\to 2}\left(x^2+\sqrt{x}\right)$.

(ה) $\lim_{x\to 3} x^{\frac{3}{2}}$.

(ו) $\lim_{x\to 1+} e^{\cos x}$.

(ז) $\lim_{x\to 0}\cos(x\cdot\sin\frac{1}{x})$.

(ח) $\lim_{x\to 0+}\sqrt{\sin x}$.

(ט) $\lim_{x\to 0-}\operatorname{sgn}(e^x\cdot\sin x)$.

(י) $\lim_{x\to 2}\left(e^x\cos x\right)$.

(יא) $\lim_{x\to 0}\left(\sin^3 x+\cos^3 x\right)$.

(יב) $\lim_{x\to 1}\frac{e^x}{\cos x}$ (הניחו כי $\cos 1\neq 0$).

10. תהי f פונקציה המוגדרת בסביבה מנוקבת של 0. קבעו אילו מהגבולות הבאים קיימים תחת כל אחת מההנחות (i) f חסומה בסביבה מנוקבת של 0, (ii) ל־f יש גבול ב־0, (iii) f רציפה ב־0.

(א) $\lim_{x\to 0} f(x^2)$.

(ב) $\lim_{x\to 0}\sin f(x)$

(ג) $\lim_{x\to 0} e^{f(x)}$.

(ד) $\lim_{x\to 0} f(f(x))$.

(ה) $\lim_{x\to 0} f(x\sin\frac{1}{x})$.

(ו) $\lim_{x\to 0} f(\sin\frac{1}{x})$.

11. מתי (כלומר, תחת אילו הנחות על f) הפונקציה $g(x)=f([x])$ רציפה בכל הישר?

12. מצאו דוגמה לפונקציות ממשיות f,g כך של־f אין גבול ב־x_0 אבל ל־$g\circ f$ יש.

13. מצאו דוגמה לפונקציות ממשיות f,g כך ש־ $\lim_{x\to x_0} f(x)=y_0$, ל־g אין גבול ב־y_0, אבל ל־$g\circ f$ יש גבול ב־x_0.

14. הוכיחו ש־$\cos$ רציפה בכל נקודה.

15. באילו נקודות יש לפונקציות הבאות גבול?

(א) $\sqrt{|\sin x|}$.

(ב) $\operatorname{sgn}(\sin x+2)$.

(ג) $\cos(\operatorname{sgn} x)$.

(ד) $\sin(\operatorname{sgn} x)$ (הניחו $\sin 1\neq 0$).

(ה) $\operatorname{sgn}(\sin(\sin x))$.

7.9 פונקציות רציפות בקטע סגור

לפונקציות שמקורן בטבע או באלגברה יש כל מיני תכונות מיוחדות. רציפות היא תכונה אחת כזו, אך יש אחרות. ניסיוננו מלמד, למשל, שאם חלקיק נמצא בזמנים שונים בשתי נקודות שונות אז במהלך התנועה שלו הוא עובר דרך כל נקודות הביניים. תכונה נוספת של חלקיק היא שבזמן סופי הוא יכול להתרחק ממקומו ההתחלתי רק מרחק סופי, ושבין כל שני זמנים נתונים יש רגע שבו המרחק הזה גדול ביותר. כל אחת מהתכונות האלה ניתנת לניסוח כתכונה של פונקציות ממשיות.

ואולם ההגדרה המופשטת של פונקציות אינה מחייבת פונקציות ממשיות להתנהג כך. מתעוררת השאלה מה הקשר בין התכונות השונות האלה. בסעיף זה נראה ושתכונת הרציפות מספיקה כדי להבטיח את קיומן של התכונות האחרות שהזכרנו (בניסוחן המתמטי, כמובן).

הגדרה 7.9.1 תהי f פונקציה ממשית. נאמר ש־f **רציפה בקטע** I אם f רציפה בכל נקודה פנימית של I, ובמידה שאחת מנקודות הקצה שייכת לקטע, הפונקציה רציפה חד־צדדית שם (כלומר רציפה משמאל בקצה הימני אם הוא שייך ל־I, ורציפה מימין בקצה השמאלי אם הוא שייך ל־I).

בפרט, רציפות של f בקטע סגור $[a,b]$ פירושה רציפות בכל נקודה ב־(a,b), רציפות משמאל ב־b ורציפות מימין ב־a.

הערות

1. לפעמים מתארים פונקציה רציפה בקטע כפונקציה שאת הגרף שלה אפשר לצייר מבלי להרים את העיפרון מהנייר. תיאור זה מביע בצורה טובה למדי את הרעיון אך יש להיזהר שלא ליחס לה משמעות עמוקה מדי, שכן קיימות פונקציות רציפות שלא נוכל לצייר בפועל.
2. אם f רציפה בקטע I ואם $J \subseteq I$ גם־כן קטע אז f רציפה ב־J (הוכיחו!)
3. התכונה שפונקציה f רציפה בקטע $[a,b]$ היא בעצם תכונה של הצמצום $f|_{[a,b]}$. שימו לב שייתכן ש־f רציפה ב־$[a,b]$ ובכל זאת f אינה רציפה בנקודות a,b. למשל נתבונן בפונקציה

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

כאן f אינה רציפה ב־0 אבל f רציפה בקטע $[0,1]$ (קל לוודא ש־f רציפה מימין ב־0).

כדי לציין שהאיברים של סדרה שייכים לקבוצת מספרים מסוימת, נאמץ את הסימון הבא:

סימון 7.9.2 אם D קבוצה ו־(x_n) סדרה, נסמן $(x_n) \subseteq D$ אם $x_n \in D$ לכל n.

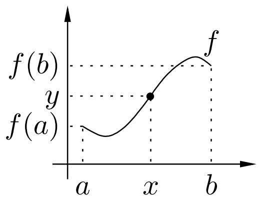
הלמה הטכנית העיקרית שנשתמש בה כאן היא הלמה הבאה:

למה 7.9.3 תהי f פונקציה רציפה ב־$[a,b]$. אם $(x_n) \subseteq [a,b]$ ואם $x_n \to x_0$ אז $x_0 \in [a,b]$ ומתקיים $f(x_n) \to f(x_0)$.

הערה למעשה זה איפיון, כלומר אם כל סדרה כמו בלמה מקיימת את המסקנה אז f רציפה ב־$[a,b]$. אנו משאירים כיוון זה כתרגיל.

הוכחה העובדה ש־ $x_0 \in [a,b]$ נובעת מכך ש־ $a \le x_n \le b$ לכל n ולכן גם נקודת הגבול מקיימת $a \le x_0 \le b$. לגבי הטענה השנייה, נשתמש באיפיון היינה לרציפות: אם x_0 נקודה פנימית של $[a,b]$ הדבר נובע מכך ש־f רציפה ב־x_0, ואם x_0 נקודת קצה אז הדבר נובע מכך ש־f רציפה חד־צדדית ב־x_0 (מהכיוון המתבקש) ו־$(x_n) \subseteq [a,b]$. ■

כאשר חלקיק נמצא בזמן מסוים בנקודה x_1 ובזמן אחר בנקודה x_2 נצפה שבזמני ביניים החלקיק יעבור בכל הנקודות שבין x_1 ו־x_2. בשפה מתמטית יותר,



איור 7.9.1 תכונת ערך הביניים

הגדרה 7.9.4 תהי $f : [a,b] \to \mathbb{R}$. נאמר ש־f מקיימת את **תכונת ערך הביניים** בקטע אם לכל y בין $f(a)$ ל־$f(b)$ קיים $x \in [a,b]$ עם $f(x) = y$.[11]

לא כל פונקציה מקיימת את תכונת ערך הביניים:

דוגמאות

1. תהי $f(x) = x^2$ ויהיו $0 \le a < b$. אז הערכים של f ב־a,b הם a^2, b^2 בהתאמה, ולכל מספר y ביניהם (כלומר לכל $a^2 < y < b^2$) יש $x \in [a,b]$ עם $f(x) = y$ (המספר הוא כמובן $x = \sqrt{y}$).
2. נתבונן בפונקציה sgn ויהיו $a = -1, b = 1$. אז $1/2$ נמצא בין $f(-1) = -1$ לבין $f(1) = 1$, אבל אין אף נקודה x בין a ל־b שבה $f(x) = \frac{1}{2}$.

מסתבר שרציפות היא תנאי מספיק כדי להבטיח את תכונת ערך הביניים:

משפט 7.9.5 (משפט ערך הביניים) תהי $f : [a,b] \to \mathbb{R}$ רציפה ויהי y מספר בין $f(a)$ ל־$f(b)$. אז קיים $x \in [a,b]$ המקיים $f(x) = y$.

הוכחה בלי הגבלת הכלליות נניח ש־ $f(a) \le f(b)$ (המקרה $f(b) < f(a)$ מוכח בצורה דומה). יהי y מספר כך ש־ $f(a) < y < f(b)$ (אם $y = f(a)$ או $y = f(b)$ אין מה להוכיח). הרעיון מאחורי ההוכחה הוא לבחור את x_0 להיות המספר המקסימלי בקטע $[a,b]$ שמקיים $f(x_0) \le y$ (כמובן שיש להוכיח שיש מספר מקסימלי כזה), ואז להראות שהוא לא יכול לקיים $f(x_0) < y$.

נסמן

$$A = \{x \in [a,b] \,:\, f(x) \le y\}$$

אז A אינה ריקה כי היא מכילה את a, והיא חסומה מלעיל על־ידי b. לכן יש לה חסם עליון, ונוכל להגדיר

$$x_0 = \sup A$$

[11] לפעמים אומרים שפונקציה מקיימת את תכונת ערך הביניים בקטע I אם היא מקיימת את התכונה בהגדרה לכל תת־קטע $[a,b] \subseteq I$, כלומר, אם לכל $a < b$ ב־I ולכל y בין $f(a), f(b)$ קיים $x \in [a,b]$ כך ש־ $f(x) = y$.

אז מתקיים $x_0 \in [a,b]$. יתר ההוכחה מוקדש להוכחת השוויון $f(x_0) = y$.

נבחר סדרה $A \supseteq (x_n)$ כך ש־ $x_n \to x_0$ (למה יש סדרה כזאת?). כל ה־x_n־ים שייכים ל־A ולכן מהגדרת A מתקיים $f(x_n) \leq y$. מכיוון ש־f רציפה ב־$[a,b]$, ובפרט ב־x_0, הרי ש־

$$f(x_0) = \lim_{n\to\infty} f(x_n) \leq y$$

אם נראה ש־ $f(x_0) \geq y$ נקבל את השוויון המבוקש $f(x_0) = y$. נניח בשלילה ש־ $f(x_0) < y$. מכך נובעות שתי מסקנות:

1. כיוון ש־ $f(b) \geq y$, מהנחת השלילה נובע ש־ $x_0 \neq b$, ולכן $x_0 < b$. מכאן ש־f רציפה מימין ב־x_0.
2. מהאי־שוויון $f(x_0) < y$ והרציפות מימין של f ב־x_0 נובע ממסקנה 7.7.4 שיש סביבה ימנית V^+ של x_0 כך שלכל $x \in V^+$ שבה מתקיים $f < y$.

האבחנה האחרונה גוררת שיש בקטע $[a,b]$ מספר $x_1 > x_0$ שמקיים $f(x_1) < y$. אבל אז $x_1 \in A$, בסתירה לכך ש־ $x_0 = \sup A$. ■

הערות

1. יש פונקציות המקיימות את תכונת ערך הביניים ואינן רציפות. שאלות בכיוון זה תמצאו בסוף הסעיף.
2. הנקודה x_0 שאת קיומה מבטיח משפט ערך הביניים לא חייבת להיות יחידה. למשל אם $f : \mathbb{R} \to \mathbb{R}$ פונקציה קבועה עם ערך c ואם $a < b$ אז כל ערך y בין $f(a)$ ל־$f(b)$ מתקבל באינסוף נקודות (כי יש רק ערך אחד y כזה, דהיינו c, והוא מתקבל בכל נקודה בקטע $[a,b]$).
3. המשפט אינו נותן כלים כדי למצוא את הנקודה שאת קיומה הוא מבטיח. באופן כללי, אין דרך למצוא אותה ממש, אבל יש מגוון שיטות המוצאות קירובים לנקודה. בתרגיל (15) למטה נציג שיטה פשוטה לכך, ובסעיף 8.9 נפתח שיטה מתוחכמת יותר.

מסקנה 7.9.6 אם $I \subseteq \mathbb{R}$ קטע ואם $f : I \to \mathbb{R}$ רציפה אז התמונה של f היא קטע.

הוכחה נסמן ב־J את התמונה של f. לפי משפט ערך הביניים, J מכילה, יחד עם כל שתי נקודות ב־J, את כל הנקודות שביניהן, ולכן J היא קטע (ראו שאלה (8) בעמוד 67). ■

דוגמאות

1. הנה הוכחה חדשה לקיומם של שורשים ריבועיים. יהי $a > 0$ ונתבונן בפונקציה $f(x) = x^2$, שהיא פונקציה רציפה בקטע $[0, a+1]$. מתקיים $f(0) = 0 < a$ ו־ $f(a+1) = a^2 + 2a + 1 > a$, ולכן ממשפט ערך הביניים יש $x \in [0, a+1]$ כך ש־ $f(x) = x^2 = a$, דהיינו, יש ל־a שורש ריבועי חיובי.

2. נראה שלמשוואה $x = \cos x$ יש פתרון, כלומר שהפונקציה $f(x) = \cos x - x$ מתאפסת. נשים לב ש־

$$f(0) = \cos 0 - 0 = 1 > 0$$

ואילו

$$f(\frac{\pi}{2}) = \cos\frac{\pi}{2} - \frac{\pi}{2} = 0 - \frac{\pi}{2} < 0$$

כמו־כן f רציפה ב־$[0, \frac{\pi}{2}]$ ולכן ממשפט ערך הביניים f מתאפסת בקטע $(0, \frac{\pi}{2})$. שימו לב שאף שהוכחנו שיש פתרון למשוואה הזו איננו יודעים לחשב אותו!

3. בדוגמה זו ניתן נימוק לטענה הפיזיקלית הבאה: בכל רגע נתון יש שתי נקודות מנוגדות על קו המשווה של כדור הארץ שבהן הטמפרטורה זהה.

תחילה יש לתאר את הבעיה בשפה מתמטית. יהי c ההיקף של כדור הארץ בקו המשווה, ונקבע נקודה P על קו המשווה. נזהה מספר x עם הנקודה על קו המשווה המתקבלת מתזוזה של x קילומטרים מערבה מ־P (אם x שלילי התנועה היא מזרחה). יהי $T(x)$ הטמפרטורה ב־x. הפונקציה T מקיימת:

(א) T רציפה. זו הנחה שבאה מההיכרות שלנו עם העולם הפיזיקלי, שבו טמפרטורות אינן משתנות בפתאומיות מנקודה לנקודה.

(ב) T מחזורית ו־c מחזור שלה, דהיינו $T(x+c) = T(x)$ לכל $x \in \mathbb{R}$. עובדה זו נובעת מכך שאם יוצאים מנקודה Q על קו המשווה ונעים מערבה מרחק c אנו חוזרים ל־Q.

הנקודה על קו המשווה המנוגדת ל־x היא הנקודה $x + \frac{c}{2}$. לכן כדי להוכיח את הטענה די לנו למצוא נקודה $x \in \mathbb{R}$ כך ש־ $T(x) = T(x+\frac{c}{2})$, או לחלופין $T(x) - T(x+\frac{c}{2}) = 0$. נסמן $f(x) = T(x) - T(x+\frac{c}{2})$. די שנראה שיש נקודה שבה f מתאפסת.

נשים לב שלכל x מתקיים

$$f(x+\frac{c}{2}) = T(x+\frac{c}{2}) - T(x+c) = T(x+\frac{c}{2}) - T(x) = -f(x)$$

(השתמשנו כאן במחזוריות של T) לכן אם $f(0) \neq 0$ אז ל־$f(\frac{c}{2})$ יש סימן הפוך מ־$f(0)$. אבל f רציפה, ולכן ממשפט ערך הביניים נובע שיש $x \in (0, \frac{c}{2})$ כך ש־ $f(x) = 0$, כנדרש.

הנה שימוש חשוב נוסף של משפט ערך הביניים:

טענה 7.9.7 לכל פולינום ממעלה אי־זוגית יש שורש.

הוכחה יהי $p(x) = \sum_{k=0}^{n} a_k x^k$ פולינום בכתיב קנוני עם n אי־זוגי. נניח $a_n > 0$ ונראה של־p יש שורש. במקרה ש־$a_n < 0$ אפשר לקבל את התוצאה מהתבוננות ב־$-p$ שהיא פולינום ממעלה אי־זוגית ועם מקדם ראשי חיובי. מתקיים

$$p(x) = x^n(a_n + \sum_{k=0}^{n-1} \frac{a_k}{x^{n-k}})$$

קל לבדוק שיש M כך שאם $|x| > M$ אז $a_n + \sum_{k=1}^{n-1} \frac{a_k}{x^{n-k}} > 0$ (הוכיחו זאת בפירוט!). מכיוון ש־n אי־זוגי, פירוש הדבר ש־$p(x)$ חיובי אם $x > M$ ושלילי אם $x < -M$. בפרט נובע ש־p מקבלת ערכים חיוביים ושליליים. אבל p רציפה בכל הישר, ולכן ממשפט ערך הביניים נובע ש־p מתאפסת. ■

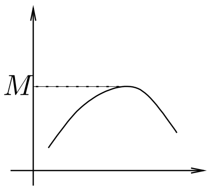
הערה לפולינום ממעלה זוגית אין בהכרח שורשים, כמו במקרה של $p(x) = x^2 + 1$.

נעבור לשאלה מסוג אחר: מתי פונקציה ממשית היא חסומה? התשובה היא שלא תמיד.

דוגמאות

1. פונקציית הזהות $i(x) = x$ המוגדרת על $\mathbb{R}$ היא רציפה אך אינה חסומה מלעיל או מלרע, כי היא מקבלת כל ערך ממשי.
2. גם בקטעים חסומים פונקציה ממשית (ואפילו רציפה) אינה חייבת להיות חסומה: למשל הפונקציה $f(x) = \frac{1}{x}$ בקטע $(0,1)$ אינה חסומה מלעיל.

מסתבר שעבור פונקציות רציפות בקטע סגור המצב טוב יותר. נזדקק לכמה הגדרות חדשות־ישנות:



איור 7.9.2 M הוא החסם העליון של הפונקציה

הגדרה 7.9.8 אם f פונקציה ממשית חסומה מלעיל אז **החסם העליון** (upper bound) של f הוא המספר

$$\sup f = \sup\{f(x) \,:\, x \in \operatorname{dom} f\}$$

החסם העליון של f בתת־קבוצה $D \subseteq \operatorname{dom} f$ (דהיינו החסם העליון של $f|_D$) מסומן ב־$\sup_{x\in D} f(x)$. באופן דומה מגדירים את החסם התחתון של פונקציה חסומה מלרע.

למה 7.9.9 תהי f פונקציה ממשית (לאו דווקא רציפה). נניח ש־f חסומה מלעיל ונסמן $M = \sup f$. אז יש סדרה $(x_n) \subseteq \operatorname{dom} f$ כך ש־ $f(x_n) \to M$. אם f אינה חסומה מלעיל אז יש סדרה $(x_n) \subseteq \operatorname{dom} f$ כך ש־ $f(x_n) \to \infty$.

הוכחה נוכיח את המקרה שבו f חסומה. המקרה האחר דומה. לכל $n \in \mathbb{N}$, אפשר לבחור x_n כך ש־ $f(x_n) > M - 1/n$, כי אחרת $M - 1/n$ היה חסם מלעיל של f בסתירה להגדרת M. כמו־כן מהגדרת M נובע ממילא $f(x_n) \le M$, וקיבלנו

$$M - \frac{1}{n} \le f(x_n) \le M$$

לכן ממשפט הסנדוויץ' לסדרות מתקיים $\lim_{n\to\infty} f(x_n) = M$. ■

משפט 7.9.10 (משפט החסימות של ויירשטראס) פונקציה רציפה $f : [a,b] \to \mathbb{R}$ היא חסומה.

הוכחה נראה ש־f חסומה מלעיל. ניתן להוכיח באופן דומה ש־f חסומה מלרע, או להסיק זאת על ידי התבוננות ב־$-f$.

נניח בשלילה ש־f אינה חסומה מלעיל. לפי הלמה יש סדרה $(x_n) \subseteq [a,b]$ כך ש־ $f(x_n) \to \infty$. לפי משפט בולצאנו־ויירשטראס 5.7.8 יש תת־סדרה $(x_{n_k})_{k=1}^{\infty}$ המתכנסת לנקודה $x_0 \in [a,b]$. מרציפות מתקיים

$$\lim_{k\to\infty} f(x_{n_k}) = f(x_0)$$

מצד שני, מכיוון ש־ $(f(x_{n_k}))$ היא תת־סדרה של $(f(x_n))$ מתקיים

$$\lim_{k\to\infty} f(x_{n_k}) = \infty$$

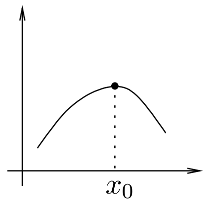
וזו סתירה. לכן f חסומה, כנדרש. ■

ובכן, מצאנו תנאי שמבטיח חסימות. נעבור לשאלה האם לכל פונקציה יש ערך מקסימלי. ליתר דיוק,

הגדרה 7.9.11 אם f פונקציה ממשית אז **המקסימום** שלה הוא הערך

$$\max f = \max\{f(x) \,:\, x \in \operatorname{dom} f\}$$

אם הוא קיים. המקסימום של f מסומן $\max f$. אם $x \in \operatorname{dom} f$ היא נקודה המקיימת $f(x) = \max f$ אז x נקראת **נקודת מקסימום** של f. אומרים אז גם שהמקסימום מתקבל ב־x. המקסימום של f בתת־קבוצה $D \subseteq \operatorname{dom} f$ (כלומר המקסימום של $f|_D$) מסומן ב־$\max_{x\in D} f(x)$. מינימום ונקודת מינימום של פונקציה מוגדרים בצורה דומה.



איור 7.9.3 x_0 היא נקודת מקסימום של הפונקציה

אף שהמקסימום של פונקציה הוא יחיד, ייתכנו נקודות מקסימום רבות. למשל, פונקציה קבועה מקבלת מקסימום בכל נקודה.

גם כאשר פונקציה חסומה מלעיל המקסימום לא תמיד קיים:

דוגמאות

1. תהי $f:[0,1]\to[0,1]$ מוגדרת על ידי $f(x)=x$ עבור $x\in[0,1)$ ו־ $f(1)=0$. אז f חסומה בקטע הסגור $[0,1]$ ו־ $\sup f=1$ אבל הערך 1 לא מתקבל ולכן לפונקציה אין מקסימום.
2. באופן כללי לא די להניח רציפות כדי להבטיח שלפונקציה יש מקסימום, גם אם הפונקציה חסומה. למשל תהי $i(x)=x$ פונקציית הזהות בקטע $(0,1)$. כאן שוב החסם העליון הוא 1 אך אינו מתקבל באף נקודה וממילא אין לפונקציה מקסימום.

בקטעים סגורים המצב טוב יותר:

משפט 7.9.12 (משפט המקסימום של ויירשטראס) פונקציה רציפה $f:[a,b]\to\mathbb{R}$ מקבלת מקסימום ומינימום בקטע.

הוכחה נראה ש־f מקבלת מקסימום. ההוכחה שהיא מקבלת מינימום דומה.

לפי המשפט הקודם, f חסומה מלעיל ב־$[a,b]$ ולכן נוכל להגדיר את המספר

$$M=\sup\{f(x)\,:\,x\in[a,b]\}$$

עלינו להראות שיש $x_0\in[a,b]$ עם $f(x_0)=M$.

לפי למה 7.9.9 קיימת סדרה $(x_n)_{n=1}^\infty\subseteq[a,b]$ כך ש־ $f(x_n)\to M$. לכן לפי משפט בולצאנו־ויירשטראס קיימת תת־סדרה $(x_{n_k})_{k=1}^\infty$ המתכנסת לנקודה $x_0\in[a,b]$. כיוון ש־f רציפה ב־x_0 מתקיים

$$f(x_0)=\lim_{k\to\infty}f(x_{n_k})=\lim_{n\to\infty}f(x_n)=M$$

ולכן x_0 נקודת מקסימום של f, כנדרש. ■

המשפט מבטיח את קיומו של מקסימום אך אינו נותן לנו כלים למצוא את הערך של המקסימום או את הנקודות בהן הוא מתקבל. נשוב לשאלה זו בפרק 8.

תרגילים

1. הראו שאם g פונקציה בקטע I שאינה קבועה ואם היא מקבלת מספר סופי של ערכים אז יש לה נקודת אי רציפות בקטע.
2. הראו שאם ל־f יש ב־x_0 אי־רציפות סליקה או מסוג ראשון אז יש קטע שבו היא אינה מקיימת את תכונת ערך הביניים.
3. מצאו דוגמה לפונקציה חסומה f המוגדרת בסביבה U של 0, כך של־f אין גבול ב־0 אבל היא מקיימת את תכונת ערך הביניים ב־U.

4. תהי $f:\mathbb{R}\to R$ פונקציה רציפה ומחזורית, כלומר יש מספר $\Delta>0$ כך ש־$f(x+\Delta)=f(x)$ לכל x. הראו שאם f אינה קבועה אז ל־f מחזור מינימלי, כלומר לקבוצת ה־Δ־ות עם התכונה האמורה יש מינימום.

5. תהי f פונקציה רציפה בקטע $[a,b]$. נניח ש־f מתאפסת לפחות פעם אחת ב־$[a,b]$, ונגדיר את קבוצת האפסים שלה:

$$A=\{x\in I\;:\;f(x)=0\}$$

הראו של־A יש איבר קטן ביותר.

6. יהי I קטע ו־ $f:I\to\mathbb{R}$ פונקציה רציפה.
   (א) אם I קטע סגור האם התמונה של f היא קטע סגור?
   (ב) אם I קטע פתוח האם התמונה של f היא קטע פתוח?

7. בסעיף 2.2 הזכרנו את משפט נקודת השבת של בראואר. ניסחנו אותו כך: בכל דרך שמערבבים דלי מלא נוזל, חלק מהנוזל חוזר למקומו המקורי. משפט זה מתייחס לפעולה במרחב תלת־ממדי, ולא נוכל להוכיחו בספר זה, אך תוכלו להוכיח את הגרסה החד־ממדית הבאה שלו: הראו שאם $f:[0,1]\to[0,1]$ רציפה אז יש נקודה שלא זזה, כלומר יש $x\in[0,1]$ כך ש־ $f(x)=x$.

8. אילו מהפונקציות הבאות חסומות מלעיל ו\או מלרע?
   (א) $\frac{1}{x^2}$.
   (ב) $e^{-|x|}$.
   (ג) x^2-2^x.
   (ד) $(1+\cos(x^2))\sin x$.

9. הוכיחו או הפריכו:
   (א) קיימת פונקציה ממשית על קטע סגור $[a,b]$ שאינה חסומה.
   (ב) קיימת פונקציה ממשית חסומה על קטע סגור $[a,b]$ שאינה מקבלת מקסימום.
   (ג) קיימת פונקציה רציפה בקטע סגור שאינה קבועה אך יש לה אינסוף נקודות מקסימום.
   (ד) קיימת פונקציה רציפה ולא קבועה בקטע סגור עם אינסוף נקודות מקסימום שבין כל שתיים מהן יש נקודת מינימום.

10. היכן בהוכחת משפט החסימות של ויירשטראס השתמשנו בכך שהקטע סגור?

11. הוכיחו או הפריכו:
   (א) פונקציה רציפה בקטע חסום היא חסומה.
   (ב) פונקציה רציפה וחסומה בקטע חסום מקבלת מקסימום.
   (ג) פונקציה חסומה המוגדרת על כל הישר מקבלת מקסימום.

(ד) פונקציה רציפה ומחזורית המוגדרת על כל הישר מקבלת מקסימום.

(ה) פונקציה מחזורית המוגדרת על כל הישר ורציפה מימין בכל נקודה מקבלת מקסימום.

12. הוכיחו שהתמונה של הפונקציות $\sin, \cos$ היא הקטע $[-1,1]$ ושהתמונה של הפונקציות $\tan, \text{cotan}$ היא $\mathbb{R}$.

13. תהי f פונקציה בקטע I. נאמר ש־f **פונקציית ליפשיץ** (Lipschitz function) ב־I אם קיים קבוע $M > 0$ כך שלכל זוג נקודות $x', x'' \in I$ מתקיים $|f(x') - f(x'')| \le M|x' - x''|$. אז M נקרא **קבוע ליפשיץ**[12] (Lipschitz constant) של f. הראו שפונקציה ליפשיצית היא רציפה ב־I.

14. בשאלה זו נוכיח את **עקרון ההעתקה המכווצת** (contraction mapping principle). תהי $f : I \to I$ כאשר I קטע סגור, ונניח שיש קבוע $0 < \lambda < 1$ כך ש־ $|f(x) - f(y)| < \lambda|x - y|$ לכל $x, y \in I$. פונקציה כזו נקראת **פונקציה מכווצת** (contraction) (חשבו מדוע!). לפי השאלה הקודמת פונקציה כזו היא ליפשיצית ולכן רציפה.

תהי $x_1 \in I$ נקודה שרירותית ונגדיר ברקורסיה $x_{n+1} = f(x_n)$.

(א) הראו שהסדרה (x_n) מתכנסת. לשם כך הוכיחו שיש קבוע C כך ש־ $|x_{n+1} - x_n| < C\lambda^n$ והסיקו ש־(x_n) סדרת קושי.

(ב) תהי $x_0 = \lim x_n$. הראו ש־x_0 היא נקודת שבת, כלומר ש־ $f(x_0) = x_0$.

(ג) הראו ש־x_0 היא נקודת השבת היחידה, כלומר, שאם $f(y) = y$ אז $y = x_0$.

15. בשאלה זו ניתן הוכחה חלופית למשפט ערך הביניים, הנקראת **שיטת החצייה**. תהי $f : [a,b] \to \mathbb{R}$ פונקציה רציפה, ונניח ש־ $f(a) < y < f(b)$. נבנה סדרת קטעים $I_n = [a_n, b_n]$ ברקורסיה כך $f(a_n) \le y \le f(b_n)$ לכל n. לשם כך נגדיר $I_1 = [a,b]$, ובהנתן $I_n = [a_n, b_n]$ נגדיר את c_n להיות נקודת האמצע של הקטע, $c_n = \frac{1}{2}(a_n + b_n)$, ונבחין בין שני מקרים: אם $f(c_n) \le y$ נגדיר $I_{n+1} = [c_n, b_n]$, ואחרת נגדיר $I_{n+1} = [a_n, c_n]$.

(א) הראו שקיימת נקודה (יחידה) $x_0 \in [a,b]$ כך ש־ $x_0 = \lim a_n = \lim b_n$.

(ב) הראו ש־ $f(x_0) = y$ (התבוננו בגבולות $\lim_{n\to\infty} f(a_n)$, $\lim_{n\to\infty} f(b_n)$ והשתמשו ברציפות של f ב־x_0).

(ג) הראו ש־ $|a_n - x_0| \le (b-a)2^{-n+1}$.

(ד) השתמשו בשיטה זו כדי למצוא קירוב למספר $\sqrt{2}$ בדיוק של $\frac{1}{16}$.

16. (*) יהי I קטע חסום ותהי $f : I \to I$ פונקציה רציפה. נגדיר $I_1 = I$ ובהינתן I_n יהי I_{n+1} התמונה של I_N תחת f, דהיינו $I_{n+1} = \{f(x) : x \in I_n\}$.

(א) הוכיחו ש־I_n הם קטעים וש־ $I_{n+1} \subseteq I_n$.

[12] Rudolph Lipschitz, 1832-1903.

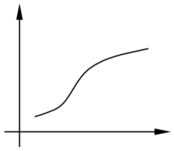
(ב) תהי $J = \cap_{n=1}^{\infty} I_n$ קבוצת הנקודות השייכות לכל הקטעים. הראו ש־ $f(x) \in J$ לכל $x \in J$.

(ג) יהי $x \in I$ ונגדיר סדרה ברקורסיה על ידי $x_1 = x$ ו־ $x_{n+1} = f(x_n)$, כך ש־ $(x_n) = x, f(x), f(f(x)), \ldots$. הראו שכל גבול חלקי של (x_n) שייך ל־J.

17. (*) ראינו שקיום גבול אינו גורר רציפות בנקודה וגם לא בנקודות סמוכות. בשאלה זו נראה שאם הגבול קיים בכל נקודה אז חייבות להיות נקודות רציפות רבות. נניח ש־ $f : [0,1] \to \mathbb{R}$ פונקציה ושלכל $x_0 \in [0,1]$ קיים הגבול $\lim_{x \to x_0} f(x)$ (אם x_0 נקודת קצה הגבול קיים במובן החד־צדדי). הוכיחו שקבוצת נקודות האי־רציפות של f היא בת־מניה (מושג זה הוגדר בסעיף 5.12) והסיקו בעזרת טענה 5.12.6 שלפונקציה כזאת בהכרח קיימות הרבה נקודות רציפות. (רמז: הראו שלכל n יש מספר סופי של נקודות x ב־$[0,1]$ כך ש־$|\lim_{y \to x} f(y) - f(x)| > \frac{1}{n}$, שימו לב ש־

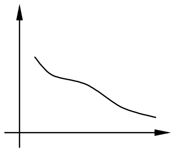
$$\left\{ \begin{array}{c} \text{נקודות אי־רציפות} \\ \text{של } f \text{ ב־ } [0,1] \end{array} \right\} = \bigcup_{n=1}^{\infty} \{x \in [0,1] \ : \ |\lim_{y \to x} f(y) - f(x)| > \frac{1}{n}\}$$

והיעזרו בלמה 5.12.2).

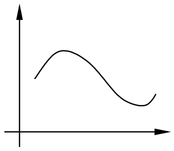

איור 7.10.1 פונקציה עולה

7.10 פונקציות מונוטוניות

הגדרה 7.10.1 פונקציה ממשית f נקראת **עולה** (increasing) אם לכל $x < y$ שבהם היא מוגדרת מתקיים $f(x) \leq f(y)$, ונקראת **יורדת** (decreasing) אם לכל $x < y$ שבהם היא מוגדרת מתקיים $f(x) \geq f(y)$. אם האי־שוויונות מתקיימים בגרסה החזקה אז f נקראת **עולה ממש** (strictly increasing) או **יורדת ממש** (strictly decreasing), בהתאמה. פונקציה נקראת **מונוטונית** monotone אם היא עולה או יורדת, ונקראת **מונוטונית ממש** (strictly monotone) אם היא עולה ממש או יורדת ממש.


איור 7.10.2 פונקציה יורדת

כשנרצה להדגיש שפונקציה היא מונוטונית אך לא מונוטונית ממש נאמר שהיא מונוטונית **חלש**.


איור 7.10.3 פונקציה שאינה עולה ואינה יורדת

מבחינה גאומטרית, f עולה אם הגרף שלה עולה משמאל לימין (כמו $\nearrow$), ו־f היא יורדת אם הגרף יורד משמאל לימין (כמו $\searrow$). לפונקציה מונוטונית יכולים להיות גם קטעים מאוזנים בגרף, אבל אז היא אינה מונוטונית ממש. ראו דוגמאות בתרשימים ממול.

דוגמאות

1. יהי $a > 0$. אז הפונקציה $f(x) = ax$ עולה ממש, כי אם $x' < x''$ אז $f(x') = ax' < ax'' = f(x'')$. אם $a < 0$ הפונקציה יורדת ממש.
2. אם $f(x) = x^2$ אז f אינה עולה כי $-1 < 0$ אבל $f(-1) = 1 > 0 = f(0)$. היא גם אינה יורדת כי $0 < 1$ אבל $f(0) = 0 < 1 = f(1)$.
3. תהי $f : [0, \infty) \to \mathbb{R}$ הפונקציה $f(x) = x^2$ (היא מוגדרת על ידי אותה נוסחה כמו בדוגמה הקודמת אך התחום קטן יותר). אז f עולה ממש, כי אם $x' < x''$ ושניהם אי־שליליים אז $f(x') = (x')^2 < (x'')^2 = f(x'')$.
4. נראה שהפונקציה $\sin$ עולה ממש בקטע $[-\frac{\pi}{2}, \frac{\pi}{2}]$ ויורדת ממש בקטע $[\frac{\pi}{2}, \frac{3\pi}{2}]$, ושהפונקציה $\cos$ יורדת ממש בקטע $[0, \pi]$ ועולה ממש בקטע $[\pi, 2\pi]$ (בעזרת תכונות המחזוריות של $\sin, \cos$ אפשר להסיק שהן מונוטוניות גם בקטעים נוספים).

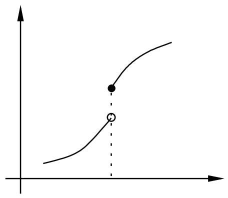
   לפי תרגיל (8) בעמוד 230 לכל x, y מתקיים

$$\sin x - \sin y = \sin x + \sin(-y) = 2 \sin \frac{x-y}{2} \cos \frac{x+y}{2}$$

   אם $-\frac{\pi}{2} \le x < y \le \frac{\pi}{2}$ אז מתקיים $-\frac{\pi}{2} < \frac{x-y}{2} < 0$ ולכן $\sin \frac{x-y}{2} < 0$, ומתקיים גם $-\frac{\pi}{2} \le \frac{x+y}{2} \le \frac{\pi}{2}$ ולכן $\cos \frac{x+y}{2} \ge 0$. מכאן ומהשוויון לעיל אנו מסיקים שאם $-\frac{\pi}{2} \le x < y \le \frac{\pi}{2}$ אז $\sin x - \sin y < 0$, דהיינו: $\sin$ עולה ממש ב־$[-\frac{\pi}{2}, \frac{\pi}{2}]$.

   הטענה ש־$\cos$ יורדת ממש בקטע $[0, \pi]$ נובעת מהזהות $\cos x = \sin(\frac{\pi}{2} - x)$ שהוכחה גם־כן בתרגיל (8) בעמוד 230.

הגבול של פונקציה מונוטונית בנקודה אינו תמיד קיים, כפי שרואים בדוגמה שבאיור 7.10.4. מצד שני לפונקציה זו יש גבולות חד־צדדיים ב־x_0. מצב דומה קיים באופן כללי. המשפט הבא מקביל למשפט על גבולות של סדרות מונוטוניות:



איור 7.10.4 אי־רציפות של פונקציה מונוטונית

משפט 7.10.2 תהי f פונקציה ממשית עולה וחסומה מלעיל. אם f מוגדרת בסביבה שמאלית מנוקבת של x_0 אז הגבול משמאל של f ב־x_0 קיים ושווה ל־

$$\lim_{x \to x_0-} f(x) = \sup_{x < x_0} f(x)$$

באופן דומה, אם f עולה, חסומה מלרע ומוגדרת בסביבה ימנית מנוקבת של x_0 אז הגבול שלה מימין ב־x_0 קיים ושווה ל־

$$\lim_{x \to x_0+} f(x) = \inf_{x > x_0} f(x)$$

הוכחה נוכיח את הטענה לגבי הגבול משמאל ב־x_0, ונשאיר לכם את המקרה השני. נגדיר

$$L = \sup_{x<x_0} f(x)$$

כאשר החסם העליון קיים לפי הנתון. אנו רוצים להראות ש־ $\lim_{x\to x_0-} f(x) = L$. יהי $\varepsilon > 0$ ויהי $U = B_\varepsilon(L)$. די שנמצא סביבה שמאלית מנוקבת W של x_0 כך ש־f מוגדרת בכל $x \in W$ ומקיימת שם $f(x) \in U$.

מהגדרת החסם העליון, קיים $x_1 < x_0$ שבו f מוגדרת ומקיימת $f(x_1) > L - \varepsilon$. נשים לב שאם $x \in (x_1, x_0)$ ואם f מוגדרת ב־x אז ממונוטוניות $f(x) \geq f(x_1)$ ולכן

$$L - \varepsilon \leq f(x) \leq L$$

(האי־שוויון הימני מהגדרת L). לכן אם $x \in (x_1, x_0)$ ו־f מוגדרת ב־x אז $f(x) \in U$. נותר רק לציין שעל אף ש־f אינה מוגדרת בהכרח בכל נקודה בקטע (x_1, x_0), הרי נתון שהיא מוגדרת בסביבה שמאלית כלשהי של x_0, ולכן יש סביבה כזו W של x_0 המוכלת בקטע (x_1, x_0). זו הסביבה המבוקשת. ∎

יש גרסה של הטענה לפונקציות יורדות. ניסוחה והוכחתה מושארים כתרגיל.

מסקנה 7.10.3 אם f מוגדרת בסביבה מנוקבת של x_0 ועולה שם אז הגבולות החד־צדדיים ב־x_0 קיימים ומתקיים $\lim_{x\to x_0-} f(x) \leq \lim_{x\to x_0+} f(x)$.

הוכחה נגדיר

$$A = \{f(x) \,:\, x < x_0\} \qquad , \qquad B = \{f(x) \,:\, x > x_0\}$$

A, B אינן ריקות כי f מוגדרת בסביבה מנוקבת של x_0. לכל $a \in A\,,\, b \in B$ מתקיים $a = f(x)\,,\, b = f(y)$ עבור מספרים x, y שמקיימים $x < x_0 < y$, ולכן ממונוטוניות $f(x) \leq f(y)$, כלומר $a \leq b$. ממשפט 4.4.5 נובע ש־A חסומה מלעיל ו־B חסומה מלרע ושמתקיים $\sup A \leq \inf B$, ולפי המשפט הקודם,

$$\lim_{x\to x_0-} f(x) = \sup A \leq \inf B = \lim_{x\to x_0+} f(x)$$

כנדרש. ∎

במקרה ש־f מוגדרת גם ב־x_0 נוכל להסיק מעט יותר:

מסקנה 7.10.4 תהי f פונקציה עולה המוגדרת בסביבה שמאלית של x_0 וגם ב־x_0. אז הגבול משמאל של f ב־x_0 קיים ו־ $\lim_{x\to x_0-} f(x) \leq f(x_0)$. באופן דומה, אם f מוגדרת בסביבה ימנית של x_0 וגם ב־x_0 אז הגבול מימין של f ב־x_0 קיים ו־ $\lim_{x\to x_0+} f(x) \geq f(x_0)$.

הוכחה אם f מוגדרת ב־x_0 אז לכל $x < x_0$ שבה f מוגדרת מתקיים $f(x) \le f(x_0)$, ולכן $f(x_0)$ הוא חסם מלעיל של הקבוצה $\{f(x) : x < x_0\}$. כיוון שהגבול של f משמאל ב־x_0 הוא החסם העליון של קבוצה זו, המסקנה נובעת. ■

אם f עולה ומוגדרת בסביבה מלאה של x_0 אז המסקנה נותנת את האי־שוויונות

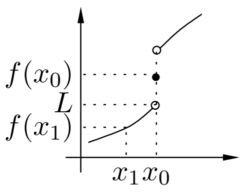
$$\lim_{x\to x_0-} f(x) \le f(x_0) \le \lim_{x\to x_0+} f(x)$$

מכאן שאם הגבולות החד־צדדיים שווים (כלומר אם הגבול הדו־צדדי ב־x_0 קיים) אז הערך המשותף שלהם הוא $f(x_0)$, ואז f רציפה ב־x_0. מאידך הגבולות החד־צדדיים קיימים (כי הפונקציה מונוטונית). אנו מסיקים שאם f מונוטונית ומוגדרת בסביבת x_0 ואם היא אינה רציפה ב־x_0 אז האי־רציפות היא ממין ראשון, ומתקיים $\lim_{x\to x_0-} f(x) < \lim_{x\to x_0+} f(x)$.

מסקנה 7.10.5 יהי $I \subseteq \mathbb{R}$ קטע ותהי $f : I \to \mathbb{R}$ פונקציה עולה. אז f רציפה ב־I אמ"מ התמונה של f היא קטע.

הוכחה ראינו כבר שאם f רציפה אז תמונתה קטע (מסקנה 7.9.6). לכן די שנראה שאם הפונקציה אינה רציפה אז תמונתה אינה קטע. תהי $x_0 \in I$ נקודת אי־רציפות של f. נניח לדוגמה ש־ $\lim_{x\to x_0-} f(x) \ne f(x_0)$ (בפרט אנו מניחים שהגבול משמאל קיים, ולכן x_0 אינו הקצה השמאלי של הקטע I). אם זה לא המצב אז בהכרח $\lim_{x\to x_0+} f(x) \ne f(x_0)$, וההוכחה במקרה זה דומה.

נסמן $L = \lim_{x\to x_0-} f(x)$, ונזכור שלפי משפט 7.10.2 מתקיים $L \le f(x_0)$. מאחר ש־ $L \ne f(x_0)$ הרי $L < f(x_0)$.



איור 7.10.5

יהי $x \in I$. אם $x < x_0$ אז $f(x) \le L$ ואם $x \ge x_0$ אז ממונוטוניות $f(x) \ge f(x_0)$. לכן אין אף $x \in I$ כך ש־ $L < f(x) < f(x_0)$.

מצד שני x_0 איננו הקצה השמאלי של I ולכן יש $x_1 \in I$ קטן יותר מ־x_0, ומתקיים $L \ge f(x_1)$ (למה?). קיבלנו אפוא ששתי הנקודות $f(x_1), f(x_0)$ שייכות לתמונה של f אבל שהתמונה אינה מכילה אף מספר בקטע $(L, f(x_0))$ וממילא אינה מכילה את כל הנקודות בין $f(x_1)$ ל־$f(x_0)$. מכאן שהתמונה איננה קטע, כפי שרצינו. ■

קיימות פונקציות מונוטוניות עם הרבה נקודות אי־רציפות. תהי $(r_n)_{n=1}^{\infty}$ סדרה של מספרים שונים בקטע $[0,1]$. נגדיר $f : [0,1] \to \mathbb{R}$ על ידי

$$f(x) = \sum_{n: r_n < x} 2^{-n}$$

f מוגדרת בכל נקודה כי בטור המגדיר את $f(x)$ משתתפים חלק מאיברי הטור $\sum_{n=1}^{\infty} 2^{-n}$, שהוא טור מתכנס. f עולה חלש כי לכל $x < y$ מתקיים

$$f(y) - f(x) = \sum_{n\,:\,x \le r_n < y} 2^{-n} \ge 0$$

(כי בסכום משתתפים מספרים אי־שליליים). יתר על כן, אם $\{r_n : n \in \mathbb{N}\}$ קבוצה צפופה אז f מונוטונית ממש כי אז יש n כך ש־ $x \le r_n < y$ ולכן הסכום המגדיר את ההפרש $f(y) - f(x)$ מכיל לפחות מחובר אחד חיובי, ומכאן $f(y) - f(x) > 0$.

אם x_0 היא אחת מהנקודות בסדרה אז יש n כך ש־ $x_0 = r_n$, ואז לכל $x > x_0$,

$$f(x) - f(x_0) \ge \sum_{k\,:\,x_0 \le r_k < x_0+h} 2^{-k} \ge 2^{-n}$$

לכן הגבול

$$\lim_{x \to x_0+} f(x) \ge f(x_0) + 2^{-n} > f(x_0)$$

(הגבול קיים ממונוטוניות). מכאן ש־f אינה רציפה מימין ב־x_0, ומכיוון שהיא עולה אנו מסיקים שאין לה גבול דו־צדדי שם. קיבלנו אפוא ש־f אינה רציפה באף אחת מהנקודות r_n ולכן ל־f יש אינסוף נקודות אי־רציפות. בפרט, אם בוחרים את (r_n) כך שתהיה צפופה ב־$[0,1]$ אז קבוצת נקודות האי־רציפות צפופה.

אפשר להראות עוד ש־f למעלה רציפה משמאל בכל נקודה ושאם $x \notin \{r_n : n \in \mathbb{N}\}$ אז f רציפה ב־x (ראו תרגיל (6) למטה).

תרגילים

1. האם פונקציה יכולה להיות גם עולה ממש וגם יורדת ממש? מה אפשר לומר על פונקציה שהיא גם עולה וגם יורדת?
2. הוכיחו שסכום של פונקציות עלולת הוא עולה. מה לגבי מכפלה? ומה לגבי סכום של פונקציה עולה ופונקציה יורדת?
3. תהי f פונקציה רציפה בקטע I ונניח שלכל $r, s \in \mathbb{Q}$ אם $r < s$ אז $f(r) < f(s)$. הוכיחו ש־f עולה ממש.
4. תהי f פונקציה רציפה ב־$[a,b]$ ומונוטונית ב־(a,b). הוכיחו שהיא מונוטונית ב־$[a,b]$. האם הטענה נכונה אם מחליפים את מונוטוניות במונוטוניות ממש?
5. השלימו את ההוכחה של משפט 7.10.2 ונסחו גרסה של המשפט לפונקציות יורדות.
6. תהי f כמו בדוגמה בסוף הסעיף, כלומר $f : [0,1] \to \mathbb{R}$ נתונה על ידי

$$f(x) = \sum_{n\,:\,r_n < x} 2^{-n}$$

עבור סדרה של מספרים שונים $(r_n)_{n=1}^{\infty}$ ב־$[0,1]$. הראו ש־f רציפה ב־x אמ"מ $x \notin \{r_n : n \in \mathbb{N}\}$, שבנקודה r_n יש ל־f אי־רציפות מסוג ראשון, וש־f רציפה משמאל בכל נקודה.

7. תהי $f:\mathbb{R}\to\mathbb{R}$ פונקציה עולה ורציפה. תהי $x_0\in\mathbb{R}$ ונגדיר ברקורסיה $x_{n+1}=f(x_n)$, כך שמקבלים את הסדרה $x_0,f(x_0),f(f(x_0)),\ldots$.
   (א) הוכיחו שהסדרה (x_n) מונוטונית. מתי היא עולה ומתי היא יורדת?
   (ב) הראו שאם הסדרה (x_n) חסומה אז יש $x\in\mathbb{R}$ עם $f(x)=x$, ואחרת $x_n\to\infty$ או $x_n\to-\infty$.
   (ג) האם התשובות משתנות אם f אינה רציפה?
8. (*) תהי $f:\mathbb{R}\to\mathbb{R}$ פונקציה עולה. הוכיחו של־f יש לכל היותר מספר בן־מניה של נקודות אי־רציפות (רמז: לכל נקודת אי־רציפות x_0 נגדיר קטע $I_{x_0}=(\lim_{x\to x_0-}f(x),\lim_{x\to x_0+}f(x))$. הראו שאם x_0,x_1 נקודות אי־רציפות אז $I_{x_0}\cap I_{x_1}=\emptyset$. היעזרו בתרגיל (5) בעמוד 161).
9. (*) השתמשו בשאלה הקודמת כדי לתת הוכחה חדשה לכך שהפונקציה e^x רציפה בכל נקודה (רמז: הראו שיש לה נקודות רציפות והיעזרו בתכונות האלגבריות של פעולת החזקה כדי להראות שאם יש לה נקודת רציפות אחת אז היא רציפה בכל נקודה).
10. תהי f פונקציה ממשית המוגדרת בסביבה של x_0. תהי

$$\omega_f(x_0,r)=\sup\{|f(x)-f(y)|\,:\,x,y\in B_r(x_0)\}$$

הראו ש־ $0\ge\omega_f(x_0,r)$ ושאם $s<r$ אז $\omega_f(x_0,r)\le\omega_f(x_0,s)$. הסיקו שהגבול

$$\omega_f(x_0)=\lim_{r\to0+}\omega_f(x_0,r)$$

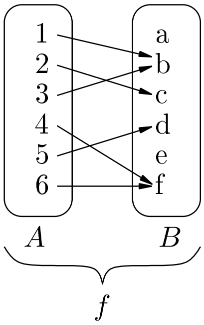
קיים. המספר $\omega_f(x_0)$ נקרא **התנודה** (*oscilation*) של f ב־x_0. הוכיחו ש־f רציפה ב־x_0 אמ"מ $\omega_f(x_0)=0$.

7.11 פונקציות הפוכות

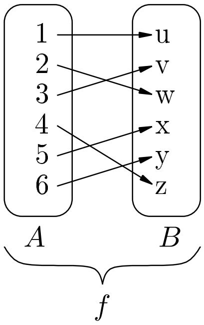
נעבור לדון בפעולה נוספת על פונקציות, הפעולה של יצירת פונקציה הפוכה. בהינתן פונקציה $f:A\to B$ אפשר לנסות ליצור פונקציה המתאימה לכל $b\in B$ איבר $a\in A$ כך ש־ $f(a)=b$. איבר a כזה, גם אם איננו יחיד, נקרא **מקור** (perimage) של b.

למשל אם $f(x)=x^2$ אז למספר חיובי a המספר $\sqrt{a}$ הוא מקור של a כי $f(\sqrt{a})=a$, וגם $-\sqrt{a}$ הוא מקור של a. מאידך, אם a שלילי אז אין לו מקור. כפי שהדוגמה הזו מדגימה, בהינתן פונקציה $f:A\to B$ ואיבר $b\in B$ אנו עשויים להיתקל בשתי בעיות בבואנו למצוא את "האיבר" a כך ש־ $f(a)=b$: ראשית אם $b\notin\operatorname{image}f$ אז לא קיים איבר a כזה, ושנית, ייתכן שיש יותר מאיבר a אחד המקיים $f(a)=b$, ובאופן כללי אין דרך להעדיף אחד על פני האחרים.

בסעיף הנוכחי נצמצם את הדיון לפונקציות עבורן בעיות אלה אינן קיימות:



איור 7.11.1 פונקציה שאינה חד־חד־ערכית ואינה על



איור 7.11.2 פונקציה חד־חד־ערכית ועל

הגדרה 7.11.1 תהי $f : A \to B$.

1. f נקראת **חד־חד־ערכית** (one-to-one או injective), או בקיצור חח"ע, אם לאיברים שונים בתחום יש תמונות שונות, דהיינו, אם לכל $a', a'' \in A$ כך ש־ $a' \neq a''$ מתקיים $f(a') \neq f(a'')$.
2. f היא **על** (onto או surjective) אם לכל איבר $b \in B$ קיים לפחות איבר אחד $a \in A$ כך ש־ $f(a) = b$.

מבחינת דיאגרמות החיצים, f חח"ע אם לכל איבר בטווח יש **לכל היותר** חץ אחד שנכנס אליו, ו־f על אם לכל איבר בטווח יש **לפחות** חץ אחד שנכנס אליו. למשל הפונקציה באיור 7.11.1 אינה על, כי לאותיות a ו־e לא נכנס אף חץ, והיא גם אינה חח"ע כי לאותיות b ו־f נכנסים יותר מחץ אחד. לעומת זאת הפונקציה f המופיעה באיור 7.11.2 היא חח"ע ועל.

פונקציה ממשית f היא חח"ע אם כל ישר מאוזן חותך את הגרף לכל היותר בנקודה אחת. באופן דומה, פונקציה ממשית היא על אם לכל y בטווח הישר המאוזן בגובה y (דהיינו הישר המקביל לציר ה־x שעובר דרך הנקודה $(0, y)$ שעל ציר ה־y) חותך את גרף הפונקציה.

דוגמאות

1. הפונקציה $f : \mathbb{R} \to \mathbb{R}$ הנתונה על ידי $f(x) = x^2$ אינה על כי $-1 \neq x^2$ לכל x, ואינה חח"ע כי למספרים $-1, 1$ יש אותה תמונה תחת f.
2. הפונקציה $f(x) = ax + b$ היא חח"ע ועל אם $a \neq 0$. שכן אם $x' \neq x''$ אז $f(x') - f(x'') = a(x' - x'') \neq 0$ ולכן היא חח"ע, ובהינתן $y \in \mathbb{R}$ מתקיים $f(\frac{y-b}{a}) = y$, דהיינו y בתמונה של f.
3. תהי $f : \mathbb{R} \to \mathbb{R}$ נתונה על ידי

$$f(x) = \begin{cases} x - 1 & x \leq 0 \\ x + 1 & x > 0 \end{cases}$$

אז f חח"ע (הוכיחו!) אבל איננה על כי 0 אינו בתמונה.

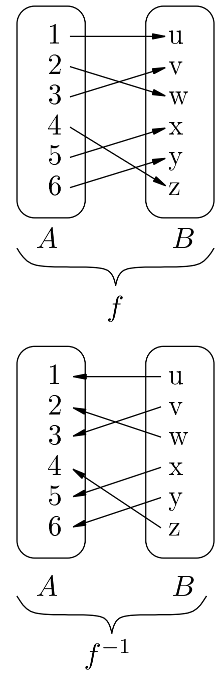
4. תהי $f : \mathbb{R} \to \mathbb{R}$ נתונה על ידי $f(x) = \tan x$ כאשר $\tan x$ מוגדר ו־$f(x) = 0$ היכן ש־$\tan$ אינו מוגדר. אז f על כי התמונה של $\tan$ הוא $\mathbb{R}$, אך אינה חח"ע כי כל מספר מתקבל אינסוף פעמים (שהרי $\tan$ מחזורית).

אם $f : A \to B$ חח"ע ועל אז העובדה ש־f היא על מבטיחה שלכל $b \in B$ יש לפחות $a \in A$ אחד כך ש־ $f(a) = b$ והחד־חד־ערכיות מבטיחה שיש לכל היותר a אחד כזה, ולכן בהינתן b בטווח התנאי $f(a) = b$ מגדיר את a ביחידות. מכאן שההגדרה הבאה אכן מגדירה פונקציה.

הגדרה 7.11.2 תהי $f : A \to B$ פונקציה חד־חד־ערכית ועל. הפונקציה f^{-1} היא הפונקציה מ־B ל־A המוגדרת על ידי

$$f^{-1}(b) = a \qquad \Longleftrightarrow \qquad b = f(a)$$

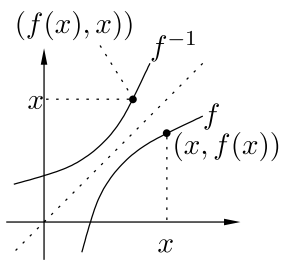
f^{-1} נקראת **הפונקציה ההפוכה** (inverse function) ל־f. אם ל־f קיימת פונקציה הפוכה (כלומר אם היא חח"ע ועל) אז f נקראת **פונקציה הפיכה** (invertible).[13]



איור 7.11.3 פונקציה והפונקציה ההפוכה לה

הפיכות היא תכונה התלויה בטווח B של הפונקציה והיה מוטב לומר ש־f הפיכה ביחס ל־B. שימו לב שאם $f : A \to B$ חח"ע ואם $C = \operatorname{image} f$ אז f הפיכה ביחס ל־C (כי הפונקציה $f : A \to C$ היא חח"ע ועל) אף שייתכן ש־f אינה הפיכה ביחס ל־B כי היא אולי אינה על. עם זאת מקובל לומר פשוט ש־f הפיכה, ואנו ננהג כך גם־כן. בפרט אם f חח"ע אבל לאו דווקא על, אז f^{-1} מציין את הפונקציה ההפוכה של $f : \operatorname{dom} f \to \operatorname{image} f$.

אם מתבוננים בדיאגרמת חיצים של f, הפונקציה ההפוכה מתקבלת על ידי הפיכת הכיוון של כל החיצים. התמונה הופכת לתחום, התחום הופך לטווח (וגם לתמונה), ויש חץ מ־b ל־a אמ"מ היה קודם חץ מ־a ל־b. הדרישה ש־f תהיה חח"ע מבטיחה שאין בתחום של f^{-1} (שהוא הטווח של f) נקודות שיוצא מהן יותר מחץ אחד, והדרישה ש־f על מבטיחה שמכל נקודה בתחום של f^{-1} יוצא לפחות חץ אחד, ולכן הדיאגרמה המתקבלת מתארת פונקציה. ראו איור 7.11.3.



איור 7.11.4 גרף של פונקציה ושל הפונקציה ההפוכה לה

אפשר גם לפרש בצורה גאומטרית את פעולת ההיפוך של פונקציה ממשית f. נשים לב שאם (x,y) נקודה בגרף של f אז $y = f(x)$, כלומר $f^{-1}(y) = x$, ולכן (y,x) נמצא בגרף של f^{-1}. הנקודה (y,x) היא הנקודה שמתקבלת מ־(x,y) על ידי שיקוף דרך הישר ℓ העובר דרך ראשית הצירים ושיפועו 1. מכאן שהגרף של f^{-1} מתקבל משיקוף של הגרף של f דרך הישר ℓ. תהליך זה מתואר באיור 7.11.4.

דוגמאות

1. אם f היא הפונקציה שממירה טמפרטורות בצלזיוס לטמפרטורות בפרנהייט, אז f^{-1} ממירה טמפרטורות בפרנהייט לטמפרטורות בצלזיוס.
2. תהי $f : [0,\infty) \to [0,\infty)$ הפונקציה המתאימה למספר את ריבועו, כלומר $f(x) = x^2$. הפונקציה ההפוכה ל־f היא פונקציית השורש $g(y) = \sqrt{y}$.

התכונות הבסיסיות של הפונקציה ההפוכה מסוכמות בטענה הבאה:

טענה 7.11.3 תהי $f : A \to B$ חח"ע ועל. אז

1. לכל $a \in A$ מתקיים $(f^{-1} \circ f)(a) = a$ ולכל $b \in B$ מתקיים $(f \circ f^{-1})(b) = b$.

[13] לפעמים משתמשים בסימון $f^{-1}(y)$ גם כאשר f לא הפיכה. אז $f^{-1}(y)$ מציין את קבוצת המקורות של y, דהיינו $f^{-1}(y) = \{x \in \operatorname{dom} f \,:\, f(x) = y\}$. לא נעשה שימוש בסימון זה.

2. f^{-1} היא חח"ע ועל מ־B ל־A, היא הפיכה, ומתקיים $(f^{-1})^{-1} = f$.

הוכחה בחלק הראשון השוויונות $f^{-1}(f(a)) = a$ ו־ $f(f^{-1}(b)) = b$ נובעים מידית מההגדרות (וודאו שזה ברור לכם!).

לגבי החלק השני, f^{-1} היא פונקציה מ־B ל־A ולכן כדי לראות ש־f^{-1} על, די לשים לב שלכל $a \in A$ מתקיים $f^{-1}(f(a)) = a$ ולכן a בתמונה של f^{-1}. כדי לראות ש־f^{-1} חח"ע, יהיו $b', b'' \in B$ ויהיו $a' = f^{-1}(b')$ ו־ $a'' = f^{-1}(b'')$. אילו היה מתקיים $a' = a''$ אז היה מתקיים $b' = f(a') = f(a'') = b''$ בסתירה להנחה.

קיבלנו ש־ f^{-1} חח"ע ועל ולכן הפיכה. נסמן $g = f^{-1}$, כך ש־ $g : B \to A$. עלינו להראות ש־ $f = g^{-1}$. הפונקציה g^{-1} היא מ־A ל־B, ומהגדרת הפונקציה ההפוכה לכל $a \in A$ מתקיים $g^{-1}(a) = b$ אמ"מ $g(b) = a$. אבל $g(b) = f^{-1}(b)$ ולכן $g(b) = a$ אמ"מ $f(a) = b$. קיבלנו שלכל $a \in A$ מתקיים $g^{-1}(a) = b$ אמ"מ $f(a) = b$, ולכן $f = g^{-1} = (f^{-1})^{-1}$ כנדרש. ■

כדי לבדוק שפונקציה ממשית היא חח"ע בדרך כלל בודקים אם היא מונוטונית.

טענה 7.11.4 תהי f פונקציה מונוטונית ממש. אז f היא חח"ע, ויתר על כן גם f^{-1} היא מונוטונית ממש, ויש לה אותה מגמה כמו f (כלומר, עולה אם f עולה ויורדת אם f יורדת).

הוכחה תהי f פונקציה ממשית מונוטונית ממש. נניח למשל שהיא עולה. לכל שני איברים $x' \neq x''$ בתחום של f, אחד מהם בהכרח קטן מהשני. נניח בלי הגבלת הכלליות ש־ $x' < x''$. מכך ש־f עולה ממש נסיק $f(x') < f(x'')$, ולכן $f(x') \neq f(x'')$.

כעת יהיו y', y'' איברים בתמונה של f. יש איברים x', x'' בתחום כך ש־$y' = f(x')$ ו־ $y'' = f(x'')$ נניח בה"כ ש־ $y' < y''$, ונראה ש־ $x' < x''$. ואמנם, לא ייתכן $x' = x''$ כי זה היה גורר $y' = f(x') = f(x'') = y''$. גם לא ייתכן $x' > x''$ כי אז העובדה ש־f עולה ממש גוררת $y' = f(x') > f(x'') = y''$. לכן האפשרות היחידה היא ש־ $x' < x''$. ■

משבנינו את כל היסודות נוכל לתת הוכחה קצרה למשפט החשוב הבא:

משפט 7.11.5 יהי $I \subseteq \mathbb{R}$ קטע ותהי $f : I \to \mathbb{R}$ פונקציה מונוטונית ממש ורציפה. אז התמונה של f היא קטע J ו־ $f^{-1} : J \to I$ רציפה.

הוכחה העובדה שהתמונה J של f היא קטע נובעת ממשפט ערך הביניים והוכחה במסקנה 7.9.6. התחום של f^{-1} הוא הקטע J, וכיוון ש־f מונוטונית, לפי טענה 7.11.4 גם f^{-1} מונוטונית. התמונה של f^{-1} אינה אלא הקטע I. לכן f^{-1} רציפה לפי טענה 7.10.5. ■

למשל, תהי $f : [0, \infty) \to \mathbb{R}$ הפונקציה $f(x) = x^2$. ברור ש־0 היא נקודת המינימום היחידה של f וש־f אינה חסומה מלעיל. מכיוון ש־f רציפה נובע שהתמונה שלה

היא הקטע $[0,\infty)$. מהמשפט האחרון נובע שהפונקציה ההפוכה ל־f רציפה על $[0,\infty)$. אבל $f^{-1}(y)=\sqrt{y}$, וקיבלנו הוכחה חדשה לרציפות של פונקציית השורש.

לפעמים נוח לנסח את המשפט מבלי להזכיר במפורש מונוטוניות:

משפט 7.11.6 יהי $I\subseteq\mathbb{R}$ קטע ותהי $f:I\to\mathbb{R}$ פונקציה רציפה והפיכה. אז התמונה של f היא קטע J ו־ $f^{-1}:J\to I$ רציפה.

ההוכחה מבוססת על משפט 7.11.5 והעובדה שאם f פונקציה רציפה והפיכה בקטע אז היא מונוטונית ממש בקטע. הוכחת טענה זו והמשפט מושארים כתרגיל (ראו שאלה 5 למטה).

תרגילים

1. אילו מהפונקציות הבאות על ואילו חח"ע?

   (א) $f(x)=x^3$.

   (ב) $f(x)=x^3$ אם $x\le 0$ ו־ $f(x)=x^2-1$ אם $x>0$.

   (ג) $f(x)=x$ אם $x<0$ או $x>1$ ו־ $f(x)=1-x$ אז $x\in[0,1]$.

   (ד) $f(x)=x\sin x$.

   (ה) $f(x)=\{\arctan x\}$ בקטע $[0,\infty)$.

   (ו) $\tan(x^2)$.

   (ז) $\tan(\sin x)$ (הניחו $\frac{\pi}{2}>1$).

2. הוכיחו או הפריכו:

   (א) אם $f:A\to B$ היא על ייתכן שיש $C\subseteq A$ כך ש־ $f|_C:C\to B$ איננה על.

   (ב) אם $f:A\to B$ היא חח"ע ייתכן שיש $C\subseteq A$ כך ש־ $f|_C:C\to B$ איננה חח"ע.

   (ג) אם $f:A\to B$ היא על ואם $g:B\to C$ היא על אז $g\circ f$ היא על.

   (ד) אם $f:A\to B$ היא חח"ע ואם $g:B\to C$ היא חח"ע אז $g\circ f$ היא חח"ע.

3. הראו שאם $f:A\to B$ ו־ $g:B\to C$ הן חח"ע ועל אז $(g\circ f)^{-1}=f^{-1}\circ g^{-1}$ (שימו לב ש־ $g\circ f:A\to C$ חח"ע ועל ולכן ההפכי של $g\circ f$ קיים).

4. הראו שאם $f:[a,b]\to\mathbb{R}$ מקבלת כל ערך בתמונה שלה בדיוק פעמיים אז f אינה רציפה.

5. הוכיחו שאם פונקציה רציפה בקטע היא חח"ע אז היא מונוטונית ממש. הסיקו שאם f רציפה והפיכה בקטע I אז f^{-1} רציפה.

6. תהי $f:[-1,0)\cup(0,1]\to\mathbb{R}$ חח"ע ונניח ש־f רציפה בכל נקודה בתחום שלה (חד־צדדית ב־±1). נסמן $E=\operatorname{image} f$. האם $f^{-1}:E\to(-1,0]\cup(0,1]$ רציפה?

7. תהי $f:[a,b]\to\mathbb{R}$ רציפה ועולה ממש. בשאלה זו ניתן הוכחה אחרת לרציפות של f^{-1}. הראו שאם $y_0=f(x_0)$ ואם $y_n\to y_0$ אז $f^{-1}(y_n)\to x_0$. לשם כך הראו שלא ייתכן שיש ל־$f^{-1}(y_n)$ תת־סדרה המתכנסת למספר אחר בקטע $[a,b]$ (השוו זאת עם ההוכחה של סעיף (3) במשפט 5.4.9).

8. (*) מצאו דוגמה לפונקציה $f:[-1,1]\to[-1,1]$ שהיא חח"ע ועל ורציפה ב־0, אבל כך ש־f^{-1} אינה רציפה ב־0.

7.12 הפונקציות האלמנטריות, חלק ב'

בסעיף זה נגדיר את הפונקציות ההפוכות לפונקציות המעריכיות והטריגונומטריות, ונשלים את הגדרת משפחת הפונקציות האלמנטריות.

בסעיף 4.6 הגדרנו את פונקציית החזקה $\exp_a(x)=a^x$ עבור $a>0$. פונקציה זו מוגדרת על כל הישר, ועבור $a\neq1$ תכונות החזקה גם מכתיבות שזו פונקציה מונוטונית ממש, ולכן חח"ע (היא עולה ממש עבור $a>1$ ויורדת ממש עבור $a<1$. עבור $a=1$ זוהי הפונקציה הקבועה 1). כמו־כן ראינו ש־$\exp_a$ רציפה בכל הישר.

טענה 7.12.1 לכל $1\neq a>0$ התמונה של $\exp_a$ היא $(0,\infty)$.

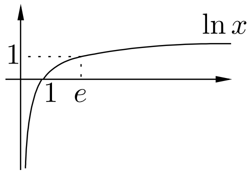
הוכחה נניח למשל ש־ $a>1$. מתכונות החזקה אנו יודעים ש־ $\exp_a(x)>0$ לכל x, כלומר התמונה של $\exp_a$ מוכלת ב־$(0,\infty)$. יהי $y\in(0,\infty)$, ונרצה להראות שיש x עם $a^x=y$. מתכונת ערך הביניים ורציפות הפונקציה $\exp_a$ די שנראה שיש x',x'' כך ש־ $a^{x'}<y<a^{x''}$. אבל זה נובע מכך ש־ $\lim_{n\to\infty}a^n=\infty$ ו־ $\lim_{n\to\infty}a^{-n}=0$ (אלו תכונות של גבולות של סדרות גאומטריות שהוכחנו בפרק 5). ■

מכיוון ש־$\exp_a$ מונוטונית ממש ורציפה יש לה פונקציה הפוכה:

הגדרה 7.12.2 יהי $1\neq a>0$. הפונקציה $\log_a:(0,\infty)\to\mathbb{R}$ היא הפונקציה ההפוכה לפונקציה המעריכית $\exp_a$ ונקראת **הלוגריתם** (logarithm) עם בסיס a. הלוגריתם בבסיס e (דהיינו הפונקציה ההפוכה ל־e^x) מסומן בקיצור $\ln$, ונקרא **הלוגריתם הטבעי** (מקור השם $\ln$ הוא מהלטינית: logarithmus naturalis).

אם כן, $\log_a(y)$ הוא המספר (היחיד) x שעבורו מתקיים $a^x=y$. במילים אחרות, $\log_a y$ הוא מספר שעונה על השאלה "a בחזקת מה שווה ל־y?". כאשר הדבר לא יוצר דו־משמעות מקובל לרשום $\log_a y$ במקום $\log_a(y)$.

התכונות של הלוגריתם נגזרות מתכונות הפונקציה המעריכית. ראינו בסעיף הקודם שהגרף של פונקציה הפוכה הוא השיקוף דרך הישר $y=x$ של הגרף של הפונקציה



איור 7.12.1 הלוגריתם הטבעי

המקורית. מכאן שהגרף שך $\ln$ הוא השיקוף של הגרף של $\exp$, ומופיע באיור 7.12.1. באופן פורמלי יותר, מתכונות הפונקציה ההפוכה שנוסחו בטענה 7.11.3 אנו מסיקים שלכל $a \neq 1$, לכל $x \in \mathbb{R}$ ולכל $y > 0$ מתקיים

$$\log_a a^x = x \qquad , \qquad a^{\log_a y} = y$$

מכיוון ש־ $a^0 = 1$ ו־ $a^1 = a$ אנו מקבלים גם ש־

$$\log_a 1 = 0 \qquad , \qquad \log_a a = 1$$

אם $a > 1$ אז $\exp_a$ עולה ממש ולכן גם $\log_a$ עולה ממש (טענה 7.11.4). לכן לכל $1 > y > 0$ מתקיים $\log_a y < 0$, ולכל $y > 1$ מתקיים $\log_a y > 0$. באופן דומה, עבור $1 > a > 0$ הפונקציה $\log_a$ יורדת ממש והאי־שוויונות האלה מתהפכים. ממשפט 7.11.5 אנו יודעים ש־$\log_a$ רציפה בכל תחומה.

ללוגריתם תכונות אלגבריות המשקפות את התכונות האלגבריות של פעולת ההעלאה בחזקה:

טענה 7.12.3 יהי $1 \neq a > 0$. אז:

1. לכל $y', y'' > 0$, מתקיים $\log_a(y' \cdot y'') = \log_a y' + \log_a y''$.
2. לכל $y > 0$ ולכל r ממשי, $\log_a y^r = r \cdot \log_a y$.

הוכחה כדי לראות את (1), נשים לב שלפי כללי החזקה מתקיים

$$a^{\log_a y' + \log_a y''} = a^{\log_a y'} \cdot a^{\log_a y''} = y' \cdot y''$$

(השוויון הראשון נובע מתכונות החזקה, והשני מהגדרת $\log_a$). לכן $\log_a(y'y'')$ שווה ל־ $\log_a y' + \log_a y''$.

את (2) מוכיחים באופן דומה: לכל $y > 0$ ול־r ממשי מתקיים

$$a^{r \cdot \log_a y} = \left(a^{\log_a y}\right)^r = y^r$$

ולכן $\log_a y^r = r \cdot \log_a y$. ■

קל להמיר לוגריתם בבסיס אחד לבסיס אחר:

טענה 7.12.4 לכל $1 \neq a, b > 0$, מתקיים $\log_a y = (\log_a b) \log_b y$.

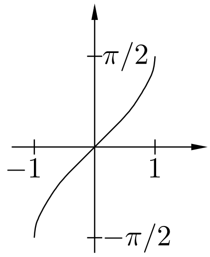
הוכחה די שנראה שלכל $y > 0$ מתקיים $a^{(\log_a b) \log_b (y)} = y$. ואכן,

$$a^{(\log_a b) \log_b (y)} = \left(a^{(\log_a b)}\right)^{\log_b y} = b^{\log_b y} = y$$

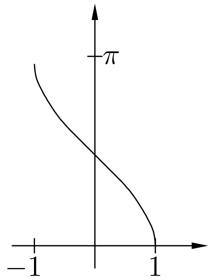
כפי שרצינו. ■

מעניין לדעת שהתכונה האלגברית $\ln(xy) = \ln x + \ln y$ הייתה המניע להמצאת הלוגריתם, בשנת 1614, על ידי ג'ון נפייר.[14] כפי שכל ילד יודע קל יותר לחבר מאשר לכפול: על מנת לחבר שני מספרים בני n ספרות עשרוניות כל אחת דרושים כ־$3n$ פעולות חשבוניות של חיבור ספרות, בעוד שדרושים n^2 פעולות כאלה על מנת להכפיל את המספרים. מכאן שיש עניין רב בהמרת בעיות כפל בבעיות חיבור. הלוגריתם מאפשר המרה כזו, כי אם נדע לחשב בקלות לוגריתמים וחזקות אז חישוב המספר $x\cdot y$ שקול לחישוב המספר $\ln(x\cdot y)$, שהוא שווה ל־ $\ln x+\ln y$. כך כדי לחשב מכפלה של שני מספרים, די לחשב סכום של הלוגריתמים של אותם המספרים. כדי להפעיל שיטה זו יש לדעת לחשב חזקות ולוגריתמים, ולכן עם המצאת הלוגריתם החלו לפרסם רשימות (או טבלאות) של ערכים של הלוגריתם, ברמות דיוק הולכות וגדלות. השימוש בטבלאות כאלה ושיטות דומות כדי לחשב מכפלות נמשך עד להמצאת המחשב, באמצע המאה העשרים.

נעבור לדון בפונקציות הטריגונומטריות ההפוכות. כפי שראינו בדוגמה (4) מעמוד 277 הפונקציות $\cos : [0,\pi] \to [-1,1]$ ו־ $\sin : [-\frac{\pi}{2},\frac{\pi}{2}] \to [-1,1]$ מונוטוניות ממש ולכן חח"ע, ומכיוון שהן רציפות ומקבלות את הערכים ± 1 בקטעים הנתונים הן גם על. לכן קיימות להן פונקציות הפוכות בתחומים אלה.



איור 7.12.2 הפונקציה $\arcsin$



איור 7.12.3 הפונקציה $\arccos$

הגדרה 7.12.5 הפונקציה ההפוכה ל־$\sin$ בתחום $[-\frac{\pi}{2},\frac{\pi}{2}]$ נקראת **ארקסינוס** ומסומנת $\arcsin$. הפונקציה ההפוכה ל־$\cos$ בתחום $[0,\pi]$ נקראת **ארקוסינוס** ומסומנת $\arccos$.

מההגדרה נובע ש־ $\arccos : [-1,1] \to [0,\pi]$ ו־ $\arcsin : [-1,1] \to [-\frac{\pi}{2},\frac{\pi}{2}]$. לפי העיקרון שהגרף של פונקציה הפוכה מתקבל מגרף הפונקציה המקורית על ידי שיקוף סביב הישר $y=x$, נובע שהגרפים של $\arccos$ ו־$\arcsin$ נראים כמו התרשימים מצד שמאל. שימו לב ש־$\arcsin$ עולה ו־$\arccos$ יורדת לפי טענה 7.11.4 והן רציפות לפי משפט 7.11.5.

כדי לחשב ביטויים מהצורה $\cos(\arcsin y)$ או $\arccos(\sin\theta))$ ניעזר באבחנות הבאות. מכיוון ש־ $\cos^2\theta + \sin^2\theta = 1$ לכל θ זה נכון גם עבור $\theta = \arcsin y$. אם נציב זאת נקבל

$$1 = \cos^2\arcsin y + \sin^2\arcsin y = \cos^2\arcsin y + y^2$$

ולכן

$$(\cos(\arcsin y))^2 = 1 - y^2$$

לכל y המספר $\arcsin y$ שייך לקטע $[-\frac{\pi}{2},\frac{\pi}{2}]$ שבו $\cos$ חיובי. לכן $\cos(\arcsin y) > 0$, ומהשוויון לעיל נובע ש־

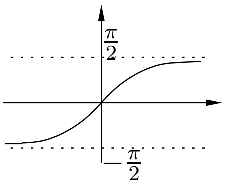
$$\cos(\arcsin y) = \sqrt{1-y^2}$$

[14] John Napier, 1550-1617

לגבי הביטוי $\arccos(\sin\theta)$, נשים לב ש־ $\sin\theta = \cos(\frac{\pi}{2} - \theta)$ (תרגיל (8) בעמוד 230) ואם $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ אז $\frac{\pi}{2} - \theta \in [0, \pi]$ ולכן

$$\arccos(\sin\theta) = \arccos(\cos(\frac{\pi}{2} - \theta)) = \frac{\pi}{2} - \theta$$

בנוסף לפונקציות ההפוכות ל־$\sin$ ו־$\cos$, ניתן להגדיר פונקציות הפוכות ל־$\tan$ ו־$\cot$an. אנו משאירים כתרגיל את הבדיקה ש־$\tan$ עולה ממש בקטע $(-\frac{\pi}{2}, \frac{\pi}{2})$ ותמונתה היא $\mathbb{R}$, וש־$\cot$ מונוטונית יורדת ממש בקטע $(0, \pi)$ ותמונתה $\mathbb{R}$.

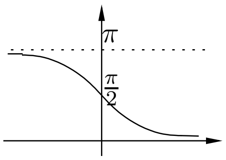


איור 7.12.4 הפונקציה arctan

הגדרה 7.12.6 הפונקציות ההפוכה ל־$\tan$ בקטע $(-\frac{\pi}{2}, \frac{\pi}{2})$ נקרא **ארקטנגנס** ומסומן arctan. הפונקציה ההפוכה ל־$\cot$an בקטע $(0, \pi)$ נקרא **ארקוטנגנס** ומסומן arccotan.

הגרפים של arctan ו־arccotan מופיעים באיורים 7.12.4 ו־7.12.5. אנו משאירים כתרגיל את תיאור התכונות של פונקציות אלה.

נסיים את הסעיף עם ההגדרה שבכותרת.



איור 7.12.5 הפונקציה arccotan

הגדרה 7.12.7 נאמר שפונקציה היא **פונקציה אלמנטרית** (elementary function) אם היא פונקציה קבועה, פולינום, פונקציה רציונלית, פונקציה מעריכית, לוגריתם, פונקציה טריגונומטרית או פונקציה טריגונומטרית הפוכה; או שהיא מתקבלת מפונקציות אלה על ידי מספר סופי של פעולות אלגבריות ופעולות הרכבה.

במילים אחרות, הפונקציות האלמנטריות הן כל הפונקציות שאנו יכולים "לרשום אותן בנוסחה סופית" בעזרת הפונקציות המצוינות בהגדרה.

משפט 7.12.8 הפונקציות האלמנטריות רציפות בכל קטע שבו הן מוגדרות.

הוכחה ראינו שהפולינומים, פונקציות החזקה, הפונקציות המעריכיות והלוגריתמים, והפונקציות הטריגונומטריות והפונקציות הטריגונומטריות ההפוכות כולן רציפות בכל קטע בו הן מוגדרות. כמו־כן ראינו שהרכבה של פונקציות רציפות בקטע היא פונקציה רציפה בקטע. ושהפעולות האריתמטיות שומרות על רציפות בקטע. כיוון שכל פונקציה אלמנטרית מתקבלת בעזרת פעולות אלו מפונקציות רציפות הרי שכל פונקציה אלמנטרית היא רציפה בכל קטע שבה היא מוגדרת. $\blacksquare$

כיוון שיש פונקציות ממשיות המוגדרות אך לא רציפות בקטע, אנו מסיקים:

מסקנה 7.12.9 ישנן פונקציות ממשיות שאינן אלמנטריות.

למשל, פונקציית הסימן אינה אלמנטרית. חשבו רגע על המשמעות של קביעה זו: זו מסקנה מאד לא טריביאלית. הרי יש אינסוף פונקציות אלמנטריות, ולכן אי

אפשר לבדוק אותן אחת אחת כדי לוודא שאף אחת מהן אינה שווה לפונקציית הסימן. במקום זאת, הוכחנו שפונקציית הסימן אינה אלמנטרית על ידי כך שמצאנו תכונה המשותפת לפונקציות האלמנטריות (רציפות), והשתמשנו בעובדה שלפונקציית הסימן אין תכונה זו.

נציין שהפונקציות האלמנטריות הן מחלקה קטנה למדי ויש פונקציות חשובות רבות שאינן כאלה. הפעולות שבעזרתן אפשר ליצור פונקציות אלמנטריות מוגבלות מדי ובהמשך נראה שיש דרכים טבעיות נוספות להציג פונקציות שמאפשרות לתאר גם פונקציות שאינן אלמנטריות.

תרגילים

1. הוכיחו כי $\sin(\arccos y) = \sqrt{1-y^2}$, וש־ $\arcsin(\cos x) = \frac{\pi}{2} - x$.
2. מצאו נוסחה עבור $\arcsin(\cos\theta)$.
3. מצאו נוסחה עבור $\tan(\arcsin y)$.

7.13 גבולות במובן הרחב וגבולות באינסוף

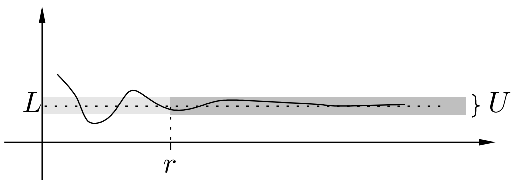
בסעיף זה נרחיב את מושג הגבול בכמה אופנים. הרעיונות כבר מוכרים ברובם ולכן לא נתעכב על פרטים מיותרים אלא נביא תוצאות והוכחות מדגמיות. יתר התכונות וההגדרות מופיעות בתרגילים בסוף הסעיף.

ההכללה הראשונה של מושג הגבול עוסקת בערך $f(x)$ של פונקציה f כאשר x גדל ללא סוף, כלומר כאשר "x שואף ל־∞". כאן יש דמיון רב למושג הגבול של סדרה, וגם לגבול חד־צדדי של פונקציה בנקודה, שכן אפשר להתקרב לאינסוף "רק משמאל". באופן דומה אפשר להגדיר גבול של פונקציה ב־$-\infty$. אנו ניתן את שתי ההגדרות, אך בהמשך נתייחס למקרה של גבול באינסוף ונשאיר את הגרסה המקבילה כתרגיל.

הגדרה 7.13.1 תהי f פונקציה ממשית ו־L מספר. נאמר ש־L הוא הגבול של f כאשר x שואף לאינסוף, ונרשום $\lim_{x\to\infty} f(x) = L$, אם לכל סביבה U של L קיים מספר r כך שלכל $x > r$ מתקיים $f(x) \in U$ (ובפרט f מוגדרת ב־x).[15]

נאמר שהגבול של f כש־x שואף ל־$-\infty$ הוא L, ונרשום $\lim_{x\to-\infty} f(x) = L$, אם לכל סביבה U של L קיים מספר r כך שלכל $x < r$ מתקיים $f(x) \in U$.

[15] אם נגדיר סביבה ("מנוקבת") של אינסוף להיות קרן מהצורה (r,∞) אז ההגדרה של גבול באינסוף זהה בצורתה להגדרת הגבול בנקודה: $\lim_{n\to\infty} f(x) = L$ אמ"מ לכל סביבה U של L יש סביבה V של אינסוף כך שאם $x \in V$ אז $f(x) \in U$.



איור 7.13.1 פונקציה שגבולה ב־∞ הוא L. לכל $x > r$ מתקיים $f(x) \in U$

בסימנים,

$$\lim_{x\to\infty} f(x) = L \qquad \Longleftrightarrow \qquad \forall \varepsilon > 0 \quad \exists r \quad \forall x > r \quad |f(x) - L| < \varepsilon$$

כרגיל, על מנת שהגבול באינסוף יהיה קיים הפונקציה חייבת להיות מוגדרת בקרן (r,∞) כלשהי.

כמו במקרה של גבול בנקודה, הגבול באינסוף הוא יחיד אם הוא קיים. הגבול באינסוף הוא תכונה אסימפטוטית במובן זה שאם f,g מסכימות בקרן (כלומר אם יש קרן (r,∞) כך שלכל $x \in (r,\infty)$ שתי הפונקציות מוגדרות ומתקיים $f(x) = g(x)$) אז הגבול של אחת באינסוף קיים אמ"מ הגבול של השנייה קיים באינסוף ואז הגבולות שווים.

משפט 7.13.2 (אפיון היינה לגבולות באינסוף) $\lim_{x\to\infty} f(x) = L$ אמ"מ f מוגדרת בקרן (r,∞) כלשהי, ולכל סדרה (x_n) עם $x_n \to \infty$ מתקיים $\lim_{n\to\infty} f(x_n) = L$.

הוכחת אפיון היינה דומה מאד למקרה של גבול בנקודה ומושארת כתרגיל.

דוגמאות

1. נראה ש־ $\lim_{x\to\infty} \frac{1}{x} = 0$. שכן עבור $\varepsilon > 0$ נתון, יהי $r = \frac{1}{\varepsilon}$. אם $x > r$ אז $\left|\frac{1}{x} - 0\right| = \frac{1}{x} < \frac{1}{r} = \varepsilon$ ולכן הבחירה r מתאימה. לחלופין אפשר להשתמש באפיון היינה ובאריתמטיקה של גבולות של סדרות: לכל סדרה $x_n \to \infty$ מתקיים $\frac{1}{x_n} \to 0$, ולכן $\lim_{x\to\infty} \frac{1}{x} = 0$.
2. נראה שלפונקציה $\sin x$ אין גבול באינסוף. ניעזר באפיון היינה. מתקיים $\pi k + \frac{\pi}{2} \to \infty$ כאשר $k \to \infty$, אבל מצד שני, $\lim_{k\to\infty} \sin(\pi k + \frac{\pi}{2}) = (-1)^k$, ולכן הסדרה $(\sin(\pi k + \frac{\pi}{2}))_{k=1}^{\infty}$ אינה מתכנסת.

הגבול באינסוף מקיים תכונות חסימות ואריתמטיקה הדומות לתכונות של הגבול בנקודה. אם לפונקציה יש גבול באינסוף אז יש קרן שבה היא חסומה. אם ל־f,g יש גבולות L,M באינסוף אז גם ל־ $f+g, f\cdot g$ יש גבולות באינסוף והם שווים ל־ $L+M, L\cdot M$ בהתאמה, ואם $M \neq 0$ אז הגבול של $\frac{f}{g}$ באינסוף קיים ושווה ל־$\frac{L}{M}$.

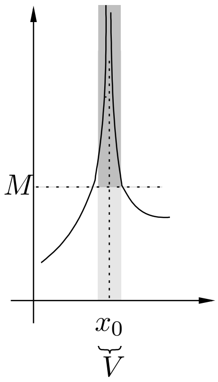
אם f,g,h מקיימות בקרן כלשהי את האי־שוויון $f \leq g \leq h$ ואם הגבול של f,h באינסוף קיים ושווה ל־L אז גם ל־g יש גבול באינסוף והוא שווה ל־L. וכן הלאה.

גם התורה של פונקציות מונוטוניות תקפה לגבולות ב־$\pm\infty$ עם השינויים המתבקשים. למשל אם f פונקציה עולה וחסומה ומוגדרת בקרן אז הגבול באינסוף קיים ו־$\lim_{x\to\infty} f(x) = \sup_{x\in\mathbb{R}} f(x)$.

את כל הטענות מוכיחים באותה צורה כמו הטענות המקבילות לגבולות בנקודה. אנו משאירים את הפרטים כתרגיל.

נעבור להכללה נוספת של הגבול אשר מקבילה לגבול במובן הרחב של סדרה, כלומר היא עוסקת במקרה שפונקציה גדלה בלי סוף. אפשר להגדיר גבול כזה בנקודה באופן דו־צדדי וחד־צדדי, ואפשר גם לתת הגדרה של גבול במובן הרחב של פונקציה כש־x שואף ל־$\pm\infty$. הנה הגדרה טיפוסית:

הגדרה 7.13.3 תהי f פונקציה ממשית. נאמר שהגבול של f ב־x_0 הוא אינסוף, ונרשום $\lim_{x\to x_0} f(x) = \infty$, אם לכל M קיימת סביבה מנוקבת V של x_0 כך שלכל $x \in V$ מתקיים $f(x) > M$ (ובפרט f מוגדרת ב־x).



איור 7.13.2 פונקציה שגבולה ב־x_0 הוא ∞ לכל $x \in V$ מתקיים $f(x) \in U$

באופן דומה, נאמר שהגבול של f ב־x_0 הוא $-\infty$ אם לכל מספר M קיימת סביבה מנוקבת של x_0 כך שלכל $x \in V$ מתקיים $f(x) < M$.

בסימנים,

$$\lim_{x\to x_0} f(x) = \infty \quad\iff\quad \forall M\ \exists\delta>0\ \forall x\ (0<|x-x_0|<\delta \implies f(x)>M)$$

אנו משאירים כתרגיל את ההגדרה של גבולות אינסופיים חד־צדדיים ושל גבולות אינסופיים כאשר x שואף ל־$\pm\infty$, ואת ההכללות של התכונות הבאות למקרים אלה.

כמו במקרה של סדרות, נשתמש במילה "גבול" לציין רק גבול סופי (ממשי). גבול שהוא סופי או אינסופי מכונה **גבול במובן הרחב**.

הגבול במובן הרחב בנקודה הוא יחיד והוא תכונה אינפיניטסימלית, דהיינו אם f,g מסכימות בסביבה מנוקבת של נקודה x_0 אז אחת מתכנסת במובן הרחב אמ"מ השנייה מתכנסת במובן הרחב, והגבולות שווים. ברור מההגדרה גם שאם $\lim_{x\to x_0} f(x) = \infty$ אז יש סביבה מנוקבת של x_0 שבה f חסומה מלרע, ומאידך היא אינה חסומה מלעיל באף סביבה מנוקבת של x_0.

משפט 7.13.4 (אפיון היינה לגבולות אינסופיים) $\lim_{x\to x_0} f(x) = \infty$ אמ"מ f מוגדרת בסביבה מנוקבת של x_0 ולכל סדרה (x_n) המקיימת $x_0 \neq x_n \to x_0$ מתקיים $\lim_{n\to\infty} f(x_n) = \infty$.

השקילות בין שני התנאים מוכחת כמו במקרה הסופי, עם השינויים המתבקשים.

דוגמאות

1. תהי $f(x)=\frac{1}{x^2}$ ונראה ש־f שואפת ל־∞ ב־0. יהי M כלשהו. אז יש $n>1$ טבעי כך ש־ $n>M$, ולכן $n^2>n>M$. אם נבחר $\delta=\frac{1}{n}$ ואם $0<|x-0|<\delta$ אז $f(x)=\frac{1}{x^2}>n^2>M$. לכן $\lim_{x\to 0} f(x)=\infty$.
2. ייתכן שלפונקציה יש גבולות חד־צדדיים אינסופיים אבל שונים. למשל תהי $g(x)=\frac{1}{x}$. הגבול מימין ב־0 של g הוא ∞ בעוד שהגבול משמאל הוא $-\infty$.

   נוכיח למשל ש־ $\lim_{x\to 0-} g(x)=-\infty$ (נסחו תחילה את ההגדרה של גבול כזה!). יהי M מספר כלשהו. אז יש $n\neq 0$ טבעי כך ש־ $-n<M$. אם נבחר $\delta=\frac{1}{n}$ נקבל שאם x בקטע $(-\delta,0)$ אז $x>-\frac{1}{n}$ ולכן $g(x)=\frac{1}{x}<-n<M$. ההוכחה ש־ $\lim_{x\to 0} g(x)=\infty$ דומה.
3. נראה ש־ $\lim_{x\to\infty}\sqrt{x}=\infty$. שכן יהי M כלשהו. אם $x>|M|^2$ אז $\sqrt{x}>|M|\geq M$, ולכן הבחירה $r=|M|^2$ מתאימה.
4. נוכיח ש־

$$\lim_{x\to\infty} x^\alpha=\begin{cases}\infty & \alpha>0\\ 1 & \alpha=0\\ 0 & \alpha<0\end{cases}$$

   ואמנם מאריתמטיקה של גבולות לסדרות, אם $x_n\to\infty$ אז $x_n^\alpha\to\infty$ אם $\alpha>0$, $x_n^\alpha\to 1$ אם $\alpha=0$ ו־ $x_n^a\to 0$ אם $\alpha<0$. לכן הטענה נובעת מאפיון היינה (לא קשה להוכיח זאת גם ישירות מתכונות החזקה!).
5. עבור n שלם מתקיים

$$\lim_{x\to-\infty} x^n=\begin{cases}\infty & n \text{ זוגי}\\ -\infty & n \text{ אי־זוגי}\end{cases}$$

   (הוכיחו!).

כללי האריתמטיקה וההשוואה לגבולות במובן הרחב של פונקציה דומים לכללים לגבולות במובן הרחב של סדרות. הנה דוגמה המעידה על הכלל:

טענה 7.13.5 יהיו f,g פונקציות ממשיות המוגדרות בסביבה מנוקבת של x_0 ונניח ש־ $\lim_{x\to x_0} f(x)=\infty$.

1. אם $g\geq f$ בסביבה מנוקבת של x_0 אז $\lim_{x\to x_0} g(x)=\infty$.
2. אם g חסומה מלרע בסביבה מנוקבת של x_0 אז $\lim_{x\to x_0}(f+g)(x)=\infty$.
3. אם g חסומה מלרע על ידי מספר חיובי בסביבה מנוקבת כלשהי של x_0 אז $\lim_{x\to x_0}(f\cdot g)(x)=\infty$. אם g חסומה מלעיל על ידי מספר שלילי בסביבה מנוקבת כלשהי של x_0 אז $\lim_{x\to x_0}(f\cdot g)(x)=-\infty$

השוו את הטענה הבאה עם טענה 5.5.9 מהפרק על סדרות:

טענה 7.13.6 תהי f פונקציה ממשית.

1. אם $\lim_{x\to x_0} f(x) = \infty$ אז $\frac{1}{f}$ מוגדרת בסביבה מנוקבת של x_0 ומתקיים $\lim_{x\to x_0} \frac{1}{f}(x) = 0$.
2. אם $\lim_{x\to x_0} f(x) = 0$ ואם יש סביבה מנוקבת של x_0 שבה f חיובית אז $\frac{1}{f}$ מוגדרת בסביבה מנוקבת של x_0 ומתקיים $\lim_{x\to x_0} \frac{1}{f}(x) = \infty$.

ישנה גרסה של המשפט על גבולות של פונקציות מורכבות במקרה של גבול באינסוף וגבולות במובן הרחב:

טענה 7.13.7 יהיו f, g פונקציות ממשיות. נניח שמתקיים $\lim_{x\to x_0} f(x) = y_0$ ו־$\lim_{y\to y_0} g(y) = M$ כאשר x_0, y_0, M יכולים להיות מספרים ממשיים או $\pm\infty$. אם יש סביבה מנוקבת של x_0 שבה f אינה מקבלת את הערך y_0 (כאשר $x_0 = \pm\infty$ מפרשים סביבה מנוקבת כקרן) אז $\lim_{x\to x_0} g\circ f(x) = M$.

ההוכחה דומה למקרה של גבול רגיל (משפט 7.8.2) ומושארת כתרגיל.

לגבי פונקציות מונוטוניות, נוכל כעת לתת גרסה של משפט 7.10.2 במקרה של פונקציות לא חסומות:

טענה 7.13.8 תהי f פונקציה עולה המוגדרת בסביבה שמאלית V של x_0. אם f אינה חסומה מלעיל ב־V אז $\lim_{x\to x_0-} f(x) = \infty$.

הוכחה יהי $M > 0$. לפי ההנחה יש $x_1 \in V$ כך ש־ $f(x_1) > M$. ממונוטוניות לכל x בסביבה השמאלית (x_1, x_0) של x_0 מתקיים $f(x) \geq f(x_1) > M$. לכן $\lim_{x\to x_0-} f(x) = \infty$, כנדרש. ■

תרגילים

1. הוכיחו:
   (א) $\lim_{x\to\infty} \frac{x^2+1}{x^3+1} = 0$.
   (ב) $\lim_{x\to\infty} \frac{\sin x}{x} = 0$.
   (ג) $\lim_{x\to\infty} \sin x$ לא קיים.
   (ד) $\lim_{x\to 0+} \frac{1}{x} = \infty$.
   (ה) $\lim_{x\to 4} \frac{1}{x(x-4)^2} = \infty$.
2. חישבו את הגבולות הבאים או הראו שאינם קיימים:
   (א) $\lim_{x\to 0} \frac{1}{\sqrt{x^2+1}-1}$.

(ב) $\lim_{x\to\infty} e^{-x}$.

(ג) $\lim_{x\to\infty} 2^x$.

(ד) $\lim_{x\to 1} \tan \frac{\pi x}{2}$.

(ה) $\lim_{x\to 0+} \ln x$.

(ו) $\lim_{x\to\infty} \arctan x$.

(ז) $\lim_{x\to\infty} \frac{\cos x}{\ln x}$.

(ח) $\lim_{x\to 0} \frac{1}{1-e^x}$.

3. תהי $f:(a,b)\to\mathbb{R}$ מונוטונית יורדת שאינה חסומה מלרע. הוכיחו שמתקיים $\lim_{x\to b-} f(x) = -\infty$.

4. לאילו מהפונקציות הבאות קיים גבול ב־0 מימין? למה הוא שווה? הוכיחו!

(א) $f(x) = \begin{cases} \frac{1}{\sin(1/x)} & x \neq \frac{1}{\pi k} \\ 0 & x = \frac{1}{\pi k} \end{cases}$ (k שלם).

(ב) $e^{1/x}$.

(ג) $\frac{\sin(1/x)}{x^2}$.

5. תהי $\{a_n\}_{n=0}^{\infty}$ סדרה ממשית. נגדיר פונקציה $A:[0,\infty)\to\mathbb{R}$ על ידי

$$A(x) = a_{[x]}$$

הוכיחו ש־ $\lim_{n\to\infty} a_n = L$ אמ"מ $\lim_{x\to\infty} A(x) = L$.

6. הראו ש־ $\lim_{x\to\infty} f(x) = L$ אמ"מ $\lim_{x\to 0+} f(\frac{1}{x}) = L$. השתמשנו באפיון זה כדי להסיק מתכונות הגבול החד־צדדי את התכונות של הגבול באינסוף שהופיעו בעמוד 291.

7. הוכיחו את טענות 7.13.5, 7.13.6 ו־7.13.7.

8. הוכיחו ש־ $\lim_{x\to x_0} f(x) = \infty$ אמ"מ $\lim_{x\to x_0} (-f)(x) = -\infty$

9. הראו שאם מתקיים $\lim_{x\to x_0} f(x) = \infty$ וגם $\lim_{x\to x_0} g(x) = -\infty$ אז מתקיים $\lim_{x\to x_0} (f\cdot g)(x) = -\infty$.

10. הוכיחו או הפריכו:

(א) $\lim_{x\to\infty} f(x) = L$ אמ"מ $\lim_{x\to -\infty} f(-x) = L$.

(ב) $\lim_{x\to\infty} f(x) = L$ אמ"מ $\lim_{x\to 0+} f(\frac{1}{x}) = L$.

(ג) $\lim_{x\to -\infty} f(x) = L$ אמ"מ $\lim_{x\to -\infty} f(e^x) = L$.

11. נסחו והוכיחו תנאי קושי לקיום גבול (במובן הצר) באינסוף.

12. יהי p פולינום. חשבו את הגבול $\lim_{x\to -\infty} p(x)$ כפונקציה של המקדמים והמעלה של p.

13. הוכיחו שאם $\lim_{x\to\infty} f(x) = \infty$ ו־ $\lim_{x\to -\infty} f(x) = -\infty$ ואם f רציפה ב־$\mathbb{R}$ אז f מקבלת כל ערך ממשי.

14. הראו שאם $f:\mathbb{R}\to\mathbb{R}$ רציפה ואם הגבולות של f ב־$\pm\infty$ קיימים אז f חסומה ב־$\mathbb{R}$.
15. תהי $f:\mathbb{R}\to\mathbb{R}$ פונקציה רציפה ונניח ש־ $\lim_{x\to-\infty}f(x)=\infty$ וגם $\lim_{x\to\infty}f(x)=-\infty$. הראו ש־$f$ מקבלת מינימום ב־$\mathbb{R}$.
16. נניח ש־f עולה ממש ו־ $\lim_{x\to\infty}f(x)=\infty$. מה ניתן לומר על f^{-1}?

7.14 רציפות במידה שווה

תהי f פונקציה רציפה בקטע I, ויהי $\varepsilon>0$. לכל נקודה $x_0\in I$ יש התכונה שלכל x קרוב מספיק ל־x_0, המרחק של $f(x)$ מ־$f(x_0)$ קטן מ־ε. ואולם מידת הקרבה הדרושה על מנת שזה יתקיים עשויה להשתנות בהתאם לנקודת הבסיס x_0. כאשר מידת הקרבה הדרושה היא אחידה על פני כל הנקודות x_0 בקטע, וכאשר זה נכון לכל ε, אומרים שהפונקציה רציפה במידה שווה. באופן מדויק,

הגדרה 7.14.1 יהי I קטע ותהי f פונקציה המוגדרת ב־I. נאמר ש־f **רציפה במידה שווה** (uniformly continuous) ב־I, או בקיצור רבמ"ש, אם לכל $\varepsilon>0$ קיים $\delta>0$ כך שלכל $x',x''\in I$ המקיימים $|x'-x''|<\delta$ מתקיים $|f(x')-f(x'')|<\varepsilon$.

שימו לב שרציפות במידה שווה היא תכונה המתייחסת תמיד לקטע מסוים.

כדאי להשוות את הגדרת הרציפות והרציפות במ"ש:

$$f \text{ רציפה} \iff \begin{cases} \forall\varepsilon>0 \quad \boxed{\forall x_0\in I \quad \exists\delta>0} \quad \forall x\in I \\ (|x-x_0|<\delta \implies |f(x)-f(x_0)|<\varepsilon \end{cases}$$

$$f \text{ רבמ"ש} \iff \begin{cases} \forall\varepsilon>0 \quad \boxed{\exists\delta>0 \quad \forall x_0\in I} \quad \forall x\in I \\ (|x-x_0|<\delta \implies |f(x)-f(x_0)|<\varepsilon \end{cases}$$

(כאן החלפנו את הסימנים x',x'' בסימנים x_0,x. כמובן שאין לכך משמעות). ההבדל, המסומן למעלה במלבן, וטמון בסדר הכמתים בלבד, ומשיקולים לוגיים נובע שהתנאי השני גורר את הראשון: בהינתן ε התנאי השני מבטיח שקיים δ שמתאים בו־זמנית לכל ה־x_0־ים וה־x־ים ב־I, וממילא בהינתן x_0 אפשר לבחור את אותו δ והוא יתאים לכל ה־x־ים ב־I. נובע מכך שאם f רבמ"ש ב־I אז היא רציפה ב־I. וודאו שזה ברור לכם!

דוגמאות

1. יהי $a\in\mathbb{R}$ קבוע ותהי $f(x)=ax$. יהיו $\varepsilon>0$ ו־ $\delta=\frac{\varepsilon}{a}$. אם $|x'-x''|<\delta$ אז לכל $x',x''\in\mathbb{R}$ מתקיים

$$|f(x')-f(x'')|=|ax'-ax''|=a|x'-x''|<\varepsilon$$

לכן f רציפה במ"ש ב־$\mathbb{R}$.

2. תהי $p(x) = x^2$ ונראה ש־p אינה רבמ"ש בכל הישר. עלינו להראות שקיים $\varepsilon > 0$ שעבורו אף בחירה של אינה טובה. נראה זאת ל־$\varepsilon = 1$ (אפשר להראות זאת גם לכל ε אחר). יהי $\delta > 0$ ונראה שיש מספרים x', x'' המקיימים $|x' - x''| < \delta$ אבל $|p(x') - p(x'')| > 1$. די להראות שקיים מספר x כך ש־ $|p(x+\frac{\delta}{2}) - p(x)| > 1$ (כלומר, הטענה תתקיים לבחירה $x' = x\ ,\ x'' = x + \frac{\delta}{2}$). חישוב קצר מראה ש־

$$p(x+\frac{\delta}{2}) - p(x) = x\delta + \frac{\delta^2}{4}$$

ולכן אם נבחר $x = \frac{1}{\delta}$ מתקיים $|p(x+\frac{\delta}{2}) - p(x)| = 1 + \frac{\delta^2}{4} > 1 = \varepsilon$, כפי שרצינו.

3. נתבונן בפונקציה הרציפה $f(x) = \frac{1}{x}$ בקטע $(0,1)$, ונראה שאינה רציפה במידה שווה שם. נקבע $\varepsilon = 1$, ויהי $\delta > 0$. די למצוא שני מספרים $x', x'' \in (0,\delta)$ כך ש־ $|f(x') - f(x'')| > 1$, כי התנאי $x', x'' \in (0,\delta)$ מבטיח $|x' - x''| < \delta$. אבל בוודאי שקיימים מספרים כאלה, כי $\lim_{x\to 0+} f(x) = \infty$ ולכן לכל $x' \in (0,\delta)$ שנבחר אפשר למצוא $x'' \in (0,\delta)$ עם $f(x'') > f(x') + 1$.

4. יהי $0 < r < 1$ ונראה שהפונקציה $f(x) = \frac{1}{x}$ רבמ"ש בקטע $[r,1]$. ואמנם נשים לב שלכל $x', x'' \in [r,1]$ מתקיים

$$\begin{aligned} |f(x') - f(x'')| &= |\frac{1}{x'} - \frac{1}{x''}| \\ &= \frac{|x'' - x'|}{|x'x''|} \\ &\leq \frac{1}{r^2}|x'' - x'| \end{aligned}$$

לכן בהינתן $\varepsilon > 0$, אם $x', x'' \in [r,1]$ ו־ $|x' - x''| < r^2\varepsilon$ אז מתקיים $|f(x') - f(x'')| < \varepsilon$, כלומר הבחירה $\delta = r^2\varepsilon$ מתאימה.

הדוגמאות למעלה מראות שרציפות לא גוררת באופן כללי רציפות במידה שווה. אולם יש מקרה חשוב שבו יש גרירה כזאת:

משפט 7.14.2 (משפט קנטור) פונקציה רציפה בקטע סגור רציפה בו במידה שווה.

הוכחה תהי $f : [a,b] \to \mathbb{R}$ רציפה ונניח בשלילה שאיננה רבמ"ש. אז קיים $\varepsilon > 0$ כך שלכל $\delta > 0$ יש נקודות $x', x'' \in [a,b]$ כך ש־ $|x' - x''| < \delta$ אבל $|f(x') - f(x'')| \geq \varepsilon$.

בפרט, לכל n קיימות נקודות $u_n, w_n \in [a,b]$ כך שמתקיים $|u_n - w_n| < \frac{1}{n}$, אבל $|f(u_n) - f(w_n)| \geq \varepsilon$ (לכל n בחרנו $\delta = \frac{1}{n}$ והסתמכנו על הפסקה הקודמת).

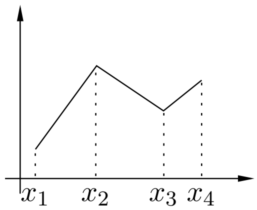
לפי משפט בולצאנו־ויירשטראס, קיימת תת־סדרה $(u_{n_k})_{k=1}^{\infty}$ של (u_n) המתכנסת לנקודה $x_0 \in [a,b]$. מכיוון ש־ $|u_n - w_n| < \frac{1}{n}$ לכל n ו־ $u_{n_k} \to x_0$ הרי ש־

$w_{n_k} \to x_0$ (למה?). מכיוון ש־f רציפה הרי שהסדרות $(f(u_{n_k}))_{k=1}^{\infty}$ ו־$(f(w_{n_k}))_{k=1}^{\infty}$ מתכנסות שתיהן ל־$f(x_0)$, ולכן

$$|f(u_{n_k}) - f(w_{n_k})| \to 0$$

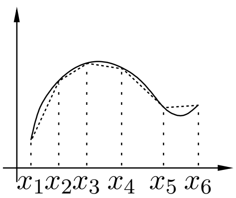
אבל זה לא ייתכן, כי $|f(u_{n_k}) - f(w_{n_k})| \geq \varepsilon$ לכל k. ■

התנאי במשפט 7.14.2 הוא מספיק אבל לא הכרחי: ישנן פונקציות רבמ"ש על קטע לא סגור. למשל ראינו למעלה ש־ $f(x) = x$ רבמ"ש על כל הישר.



איור 7.14.1 קו פוליגוני

השימושים העיקריים של תכונת הרציפות במ"ש ידונו בהמשך, אך הנה שימוש נאה של המושג. תהי $f : [a,b] \to \mathbb{R}$ פונקציה. טבעי לשאול באילו תנאים אפשר לקרב את הגרף של f על־ידי פוליגון. ליתר דיוק, נאמר שפונקציה $p : [a,b] \to \mathbb{R}$ מתארת **פוליגון** (polygon) אם אפשר לפרק את $[a,b]$ למספר סופי של תת־קטעים $[x_0,x_1],[x_1,x_2],\ldots,[x_{n-1},x_n]$ שאיחודם הקטע $[a,b]$ כולו ושנפגשים רק בקצוות, וכך שבכל קטע $[x_{i-1},x_i]$ הפונקציה p מתארת קו ישר.[16] במילים אחרות p הוא פוליגון אם יש נקודות $a = x_0 \leq x_1 \leq \ldots \leq x_{n-1} \leq x_n = b$ וקבועים α_i, β_i כך שלכל $x \in [x_{i-1},x_i]$ מתקיים $p(x) = \alpha_i x + \beta_i$. שימו לב ש־$p$ רציפה: שכן מההגדרה נובע ש־p רציפה מימין ומשמאל בכל נקודה $x \in (a,b)$ ולכן היא רציפה ב־x, ורציפה חד־צדדית ב־a,b.



איור 7.14.2 קירוב פוליגוני לפונקציה ביחס לנקודות $x_1, x_2, \ldots, x_6$

אם כן, השאלה שאנו שואלים היא עבור אילו פונקציות $f : [a,b] \to \mathbb{R}$ אפשר, עבור כל $\varepsilon > 0$, למצוא פונקציה פוליגונית p כך שלכל $x \in [a,b]$ מתקיים $|f(x) - p(x)| < \varepsilon$. תשובה חלקית נתונה במשפט הבא:

משפט 7.14.3 אם f רציפה בקטע סגור $[a,b]$ אז לכל $\varepsilon > 0$ יש פונקציה פוליגונית p כך שלכל $x \in [a,b]$ מתקיים $|f(x) - p(x)| < \varepsilon$.

הוכחה תהי f רציפה ב־$[a,b]$ ויהי $\varepsilon > 0$. לפי משפט קנטור, f רבמ"ש ב־$[a,b]$ ולכן יש $\delta > 0$ כך שאם $u,v \in [a,b]$ מקיימים $|u - v| \leq \delta$ אז $|f(u) - f(v)| < \varepsilon$.

יהיו $a = x_1 < x_2 < \ldots < x_n = b$ נקודות ב־$[a,b]$ כך ש־ $|x_{i+1} - x_i| \leq \delta$, למשל אפשר לבחור

$$x_0 = a\ ,\ x_1 = a + \delta\ ,\ x_2 = a + 2\delta\ ,\ldots,\ x_{n-1} = a + n\delta\ ,\ x_n = b$$

כאשר n הוא המספר הטבעי היחיד המקיים $a + n\delta < b \leq a + (n+1)\delta$.

נגדיר את הפונקציה הפוליגונית p כך שבקטע $[x_i, x_{i+1}]$ היא מתארת את הישר שמחבר בין הנקודות $(x_i, f(x_i))$, $(x_{i+1}, f(x_{i+1}))$ שעל הגרף של f. באופן מפורש, עבור $x \in [x_i, x_{i+1}]$ נגדיר

$$h(x) = f(x_i) + \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}(x - x_i)$$

[16] פונקציה המתארת פוליגון נקראת לפעמים פונקציה **לינארית למקוטעין**.

עלינו להראות שלכל $x \in [x_{i-1}, x_i]$ מתקיים $|f(x) - p(x)| < \varepsilon$. נניח בלי הגבלת הכלליות ש־ $f(x_{i-1}) \leq f(x_i)$. מהנוסחה עבור p ברור ש־ $p(x_i) = f(x_i)$ ו־ $p(x_{i+1}) = f(x_{i+1})$, ולכן די להוכיח את הטענה עבור $x \in (x_{i-1}, x_i)$. נשים לב שמהנוסחה עבור p מתקיים שלכל $x \in (x_{i-1}, x_i)$ הערך של $p(x)$ נמצא בין $f(x_{i-1})$ ל־$f(x_i)$ (בדקו!), כלומר

$$f(x_{i-1}) \leq p(x) \leq f(x_i)$$

מצד שני לכל $x \in (x_{i-1}, x_i)$ מתקיים $|x - x_{i-1}| < \delta$ וגם $|x - x_i| < \delta$ (כי $x_i - x_{i-1} \leq \delta$) ולכן

$$|f(x) - f(x_{i-1})| < \varepsilon \qquad , \qquad |f(x) - f(x_i)| < \varepsilon$$

מכאן נובע

$$f(x_{i-1}) > f(x) - \varepsilon \qquad , \qquad f(x_i) < f(x) + \varepsilon$$

ובשילוב עם האי־שוויון $f(x_{i-1}) \leq p(x) \leq f(x_i)$ מקבלים

$$f(x) - \varepsilon < p(x) < f(x) + \varepsilon$$

דהיינו $|f(x) - p(x)| < \varepsilon$, כפי שרצינו להראות. ■

מסתבר שגם הכיוון ההפוך של המשפט נכון: אם אפשר לקרב באופן טוב כרצוננו את הגרף של f על ידי פוליגונים אז f רציפה. עובדה זו תוכח במועד מאוחר יותר (סעיף 10.3).

תרגילים

1. בדקו מההגדרה אילו מהפונקציות הבאות רציפות במ"ש בתחומים הנתונים.
   (א) x^3 ב־$[0,1]$ וב־$\mathbb{R}$
   (ב) e^x ב־$[0,1]$ וב־$\mathbb{R}$.
   (ג) x^2 ב־$[0,2]$.
   (ד) $\sin\frac{1}{x}$ ב־$(0,1)$.
   (ה) $\sqrt{x}$ ב־ $[0,\infty)$.
2. תהי $f : \mathbb{R} \to \mathbb{R}$ פונקציה מחזורית ורציפה. הראו ש־f רבמ"ש בכל הישר.
3. תהי $f : \mathbb{R} \to \mathbb{R}$ פונקציה רציפה ונניח שהיא קבועה בקטעים $(-\infty, a)$ ו־(b, ∞). הוכיחו שהיא רציפה במ"ש ב־$\mathbb{R}$.
4. הוכיחו או הפריכו את הטענות הבאות:

(א) אם f רבמ"ש בקטע I ואם $J \subseteq I$ תת־קטע אז f רבמ"ש על J.

(ב) אם $f : \mathbb{R} \to \mathbb{R}$ רבמ"ש בכל קטע $[n, n+1]$ עבור $n \in \mathbb{N}$ אז f רבמ"ש ב־$\mathbb{R}$.

(ג) אם f, g רבמ"ש בקטע I אז גם $f+g$ רבמ"ש ב־I.

(ד) אם f, g רבמ"ש בקטע I אז גם $f \cdot g$ רבמ"ש ב־I.

(ה) אם f, g רבמ"ש בקטע I ואם I סגורה אז גם $f \cdot g$ רבמ"ש ב־I.

5. תהי f פונקציה רציפה בקטע (a, b). הראו ש־f רבמ"ש ב־(a, b) אמ"מ קיימים לה גבולות חד־צדדיים (במובן הצר) ב־a, b.

6. תהי $f : [a, \infty) \to \mathbb{R}$ רציפה. הראו שאם $\lim_{x\to\infty} f(x) = L$ אז f רבמ"ש על $[a, \infty)$. האם ההפך נכון?

7. נניח ש־f רבמ"ש ב־$[a, b]$ ושבהגדרת הרציפות במ"ש למספר $\varepsilon = 1$ מתאים המספר δ. תנו הערכה לגודל

$$\max_{x\in[a,b]} f(x) - \min_{x\in[a,b]} f(x)$$

(על ההערכה להיות תלויה ב־δ).

8. תרגיל זה מתייחס לתנאי ליפשיץ שהוגדר בתרגיל (13) בעמוד 275.

(א) תהי f פונקציה בקטע I בעלת קבוע ליפשיץ M. הראו ש־f רבמ"ש ב־I.

(ב) הראו על ידי מציאת קבוע ליפשיץ שהפונקציות הבאות רבמ"ש בקטעים הנתונים:

i. $\sin x$ בכל הישר.

ii. x^2 בקטע $[0, a)$ עבור $a > 0$ (האם x^2 ליפשיצית בכל $\mathbb{R}$?).

iii. xe^{-x} רבמ"ש על $(0, \infty)$.

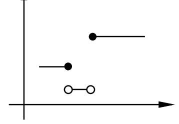
iv. $x \sin \frac{1}{x}$ ב־ $(0, \infty)$.

(ג) האם פונקציה רציפה במ"ש היא בהכרח ליפשיצית?

9. נאמר שפונקציה g היא **פונקציית מדרגה** (step function) אם אפשר להציג את I כאיחוד של מספר סופי של תת־קטעים שבהם g קבועה.

(א) הראו שפונקציית מדרגה בקטע היא רציפה אמ"מ היא קבועה.

(ב) הראו שאם f פונקציה רבמ"ש על קטע I אז לכל $\varepsilon > 0$ יש פונקציית מדרגה $g : I \to R$ כך ש־ $|f(x) - g(x)| < \varepsilon$ לכל $x \in I$.



איור 7.14.3 פונקציית מדרגה

פרק 8

הנגזרת

אם פונקציה f רציפה בנקודה x_0 אז כאשר x קרוב ל־x_0 הערכים $f(x)$ ו־$f(x_0)$ קרובים. תיאור זה הוא תיאור איכותי, ולא כמותי: הוא אינו נותן מידע על האופן שבו המרחק בין $f(x)$ ל־$f(x_0)$ תלוי במרחק בין x ל־x_0. מסיבה זו המסקנות שהסקנו בפרק הקודם היו בעלות אופי תיאורטי, ולא יישומי: יכולנו למשל להסיק קיום של מינימום ומקסימום של פונקציה אך לא מצאנו שיטה למצוא היכן הם מתקבלים. כך גם לגבי משפט ערך הביניים.

בפרק זה נגדיר את הנגזרת, המודדת את קצב ההתקרבות של $f(x)$ ל־$f(x_0)$ כפונקציה של המרחק בין x ל־x_0. אנו נפתח שיטות שונות לניתוח של פונקציות שיש להן נגזרת (בין השאר נמצא פתרונות לבעיות שהוזכרו בפסקה הקודמת). נראה גם מגוון שימושים אחרים של הנגזרת, בין השאר לחישוב גבולות (שיטת לופיטל), הוכחת אי־שוויונות (בעזרת אי־שוויון ינסן לפונקציות קמורות), ומציאת שורשים של משוואות (שיטת ניוטון־לייבניץ).

8.1 הנגזרת בנקודה

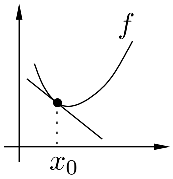
המוטיבציה העומדת מאחורי הנגזרת באה מפיזיקה ומגאומטריה. נתחיל בתיאור הפיזיקלי. נתבונן בחלקיק הנע בקו ישר ושמרחקו מהראשית בזמן t הוא $s(t)$. המרחק שהחלקיק עובר בין זמן t_1 לזמן t_2 הוא $s(t_2) - s(t_1)$, ופרק הזמן שחלף הוא $t_2 - t_1$. על כן אפשר לומר שהמהירות הממוצעת של החלקיק באותו פרק זמן היא היחס בין המרחק שעבר לבין הזמן שלקח לו, דהיינו $\frac{s(t_2)-s(t_1)}{t_2-t_1}$.

במקום מהירות ממוצעת על פני פרק זמן היינו רוצים לדעת מה המהירות של החלקיק ברגע מסוים t_0. איננו יודעים כיצד למדוד מהירות רגעית אך מתקבל על הדעת שהיא תהיה שווה בקירוב למהירות הממוצעת בפרקי זמן קצרים המתחילים ב־t_0 ומסתיימים כעבור פרק זמן קצר h, כלומר בזמן $t_0 + h$. לפי הדיון למעלה,

המהירות הממוצעת בפרק הזמן הזה היא

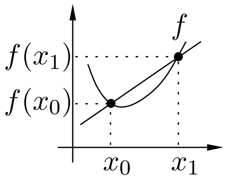
$$\frac{s(t_0+h)-s(t_0)}{h}$$

ולכן סביר להגדיר את המהירות הרגעית בזמן t_0 בתור הגבול של הביטוי הזה כאשר $h \to 0$.



איור 8.1.1 המשיק לגרף בנקודה

אותו גודל שהגענו אליו עתה מתעורר גם בבעיה הגאומטרית של מציאת ישר משיק לגרף. תהי f פונקציה המוגדרת בסביבה של x_0 ונתבונן בנקודה $P=(x_0,f(x_0))$ שעל הגרף של f. מבין כל הישרים העוברים דרך P היינו רוצים למצוא את הישר המשיק לגרף. מבחינה אינטואיטיבית, המשיק הוא הישר ש"נושק" לגרף ב־P במקום "לחצות" אותו. התבוננו למשל באיור 8.1.1.



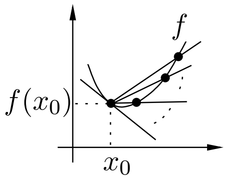
איור 8.1.2 מיתר בין נקודות בגרף

כדי למצוא את המשיק נתעניין בישרים העוברים דרך P ודרך נקודה נוספת על הגרף. ישר כזה נקרא **מיתר**. דוגמה למיתר מופיעה באיור 8.1.2. באופן כללי, אם $(x_0,y_0),(x_1,y_1)$ נקודות במישור עם $x_0 \neq x_1$ אז קיים ישר יחיד ℓ העובר דרכם. חישוב פשוט מראה שהשיפוע של הישר הוא $\frac{y_1-y_0}{x_1-x_0}$ והישר נתון על ידי הנוסחה

$$\ell(x)=\frac{y_1-y_0}{x_1-x_0}(x-x_0)+y_0$$

בפרט, עבור הנקודות $(x_0,f(x_0)),(x_1,f(x_1))$ על הגרף של f המיתר ℓ העובר דרכן הוא בעל שיפוע $\frac{f(x_1)-f(x_0)}{x_1-x_0}$ והנוסחה שלו היא

$$\ell(x)=\frac{f(x_1)-f(x_0)}{x_1-x_0}(x-x_0)+f(x_0)$$



איור 8.1.3 מיתרים המתקרבים למשיק

נשוב לבעיית המשיק. נתבונן במיתרים העוברים דרך הנקודה $(x_0,f(x_0))$ ונקודה קרובה לה, כלומר דרך נקודה מהצורה $(x_0+h,f(x_0+h))$ עבור מספר h שקרוב לאפס. סביר להניח שככל ש־h קרוב יותר ל־0, המיתרים המתאימים "קרובים" יותר לישר המשיק: ראו איור 8.1.3. בפרט אנו מצפים שכאשר h קרוב לאפס השיפוע של המיתר, דהיינו הגודל

$$\frac{f(x_0+h)-f(x_0)}{h}$$

יהיה קרוב לשיפוע של הישר המשיק. מכאן שהשיפוע של הישר המשיק הוא הגבול של המנה הזו כאשר h שואף ל־0.

בעקבות שתי הדוגמאות, שבהן הגענו בסוף לביטוי זהה, נגדיר:

הגדרה 8.1.1 תהי f פונקציה ממשית המוגדרת בסביבה מלאה של x_0. **הנגזרת** (derivative) של f בנקודה x_0 היא הגבול

$$\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}$$

בתנאי שהוא קיים, ומסומן $f'(x_0)$. אם הנגזרת קיימת ב־x_0 אומרים ש־f **גזירה** (differentiable) ב־x_0, או ש־x_0 היא **נקודת גזירות** של f.

פונקציית הנגזרת[1] היא הפונקציה f' המתאימה לכל נקודת גזירות x של f את הנגזרת $f'(x)$ באותה נקודה. התחום של פונקציית הנגזרת הוא אוסף נקודות הגזירות של f.

הערות

1. דרך שקולה לרשום את הנגזרת היא
$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$
השקילות נובעת מההצבה $h = x - x_0$ בהגדרת הנגזרת, או באופן מדויק יותר ממסקנה 7.8.4.

2. גזירות היא תכונה אינפיניטסימלית: אם f, g מסכימות בסביבה מלאה של x_0 אז f גזירה ב־x_0 אמ"מ g גזירה שם, ובמקרה זה הנגזרות שוות. זו מסקנה מידית מהגדרת הנגזרת ומכך שהגבול היא תכונה אינפיניטסימלית (אנו מפעילים שיקול זה על הגבול של הפונקציות $\frac{1}{h}(f(x_0+h) - f(x_0))$ ו־$\frac{1}{h}(g(x_0+h) - g(x_0))$ התלויות ב־h ומסכימות בסביבה מנוקבת של 0).

דוגמאות

1. תהי $\ell(x) = ax + b$ פונקציה המתארת ישר. מכיוון שכל המיתרים שלה הם ℓ עצמה, ולכן בעלי שיפוע a, נצפה ש־ $\ell' \equiv a$. ואמנם לכל $x_0 \in \mathbb{R}$ מתקיים
$$\frac{\ell(x_0+h) - \ell(x_0)}{h} = \frac{(a(x_0+h)+b) - (ax_0+b)}{h} = \frac{ah}{h} = a$$
לכן $\lim_{h \to 0} \frac{\ell(x_0+h) - \ell(x_0)}{h} = a$, והנגזרת ב־$x_0$ שווה ל־a.

2. מקרה פרטי של הדוגמה הקודמת הוא שהנגזרת של פונקציה קבועה היא 0 בכל נקודה, ואם $i(x) = x$ היא פונקציית הזהות על $\mathbb{R}$ אז $i'(x) = 1$ לכל $x \in \mathbb{R}$.

3. תהי $f(x) = x^2$. נחשב את הנגזרת ב־1:
$$\begin{aligned} \frac{f(1+h) - f(1)}{h} &= \frac{(1+h)^2 - 1^2}{h} \\ &= \frac{1 + 2h + h^2 - 1}{h} \\ &= 2 + h \end{aligned}$$

---

[1] ישנם בספרות סימונים נוספים עבור פונקציית הנגזרת, ביניהם $\frac{df}{dx}$, $\frac{d}{dx}f$, Df ו־$\dot{f}$. בסימונים אלה הנגזרת של f בנקודה x_0 מסומנת ב־ $\frac{df}{dx}|_{x=x_0}$, $\frac{d}{dx}f(x_0)$, $Df(x_0)$ או $\dot{f}(x_0)$, בהתאמה. בסימן $\frac{d}{dx}$ נדון בקצרה בהמשך.

לכן הגבול של הביטוי הזה שואף ל־2 כש־h שואף ל־0, ולכן $f'(1) = 2$. באופן דומה מראים שהנגזרת של f בכל נקודה כללית x היא $2x$.

4. תהי $f(x) = \frac{1}{x}$ ויהי $x_0 \neq 0$. נחשב את הנגזרת של f ב־x_0. מתקיים

$$\begin{aligned}\frac{f(x_0+h)-f(x_0)}{h} &= \frac{1}{h}\left(\frac{1}{x_0+h}-\frac{1}{x_0}\right)\\ &= \frac{1}{h}\cdot\frac{x_0-(x_0+h)}{x_0(x_0+h)}\\ &= -\frac{1}{x_0(x_0+h)}\end{aligned}$$

ולכן

$$f'(x_0) = \lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h} = -\frac{1}{x_0^2}$$

5. נחשב את הנגזרת של $f(x) = \sqrt{x}$ בנקודה $x_0 \neq 0$:

$$\begin{aligned}\lim_{h\to 0}\frac{\sqrt{x_0+h}-\sqrt{x_0}}{h} &= \lim_{h\to 0}\frac{\sqrt{x_0+h}-\sqrt{x_0}}{h}\cdot\frac{\sqrt{x_0+h}+\sqrt{x_0}}{\sqrt{x_0+h}+\sqrt{x_0}}\\ &= \lim_{h\to 0}\frac{(x_0+h)-x_0}{h(\sqrt{x_0+h}+\sqrt{x_0})}\\ &= \lim_{h\to 0}\frac{1}{\sqrt{x_0+h}+\sqrt{x_0}}\\ &= \frac{1}{2\sqrt{x_0}}\end{aligned}$$

כאשר במעבר האחרון השתמשנו ברציפות הפונקצית השורש ובאריתמטיקה של גבולות.

נשוב לענייננו ונשאל מתי פונקציה היא גזירה. תנאי הכרחי בסיסי לגזירות היא רציפות:

טענה 8.1.2 אם פונקציה גזירה ב־x_0 אז היא רציפה ב־x_0.

הוכחה תהי f הפונקציה הנתונה. די שנראה ש־ $\lim_{h\to 0}(f(x_0+h)-f(x_0)) = 0$. אבל

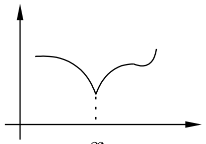
$$\begin{aligned}\lim_{h\to 0}(f(x_0+h)-f(x_0)) &= \lim_{h\to 0} h\cdot\frac{f(x_0+h)-f(x_0)}{h}\\ &= \lim_{h\to 0} h\cdot\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}\\ &= 0\cdot f'(x_0)\\ &= 0\end{aligned}$$

כנדרש. ■

בפרט, אם פונקציה אינה רציפה בנקודה אז היא אינה גזירה שם. כך אנו מקבלים שפע של דוגמאות לפונקציות שאינן גזירות בנקודה. למשל, הפונקציה sgn אינה רציפה ב־0 ולכן אינה גזירה שם, ופונקציית דירכלה (דוגמה (4) מעמוד 237) אינה רציפה באף נקודה ולכן אינה גזירה באף נקודה. מאידך, הדוגמאות הבאות מראות שרציפות אינה תנאי מספיק לגזירות.

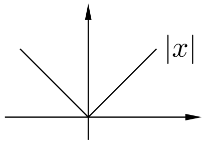
דוגמאות

1. סיבה אחת לאי־גזירות יכולה להיות שב־x_0 יש "שפיץ" בגרף. במקרה זה ברור באופן אינטואיטיבי שלא תהיה נגזרת כי אין משיק לגרף בנקודה. דוגמה כזאת מופיעה באיור 8.1.4.



איור 8.1.4 פונקציה שאינה גזירה ב־x_0

דוגמה קונקרטית היא הפונקציה $f(x)=|x|$ המופיעה באיור 8.1.5. זו פונקציה רציפה ב־0 ויש לה שם "פינה". נראה שהיא אינה גזירה ב־0: לכל $h\neq 0$ מתקיים

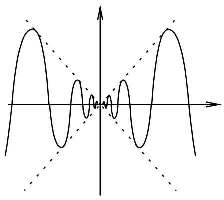


איור 8.1.5 הפונקציה $|x|$ אינה גזירה ב־0

$$\frac{f(0+h)-f(0)}{h}=\frac{|0+h|-|0|}{h}=\frac{|h|}{h}=\operatorname{sgn} h$$

ולכן למנה הזו אין גבול ב־0 והפונקציה $f(x)=|x|$ אינה גזירה ב־0.

2. סיבה נוספת לאי קיום הנגזרת בנקודה היא התנודדות חזקה של הפונקציה בקרבת הנקודה. תהי



איור 8.1.6 הפונקציה $x\sin(1/x)$

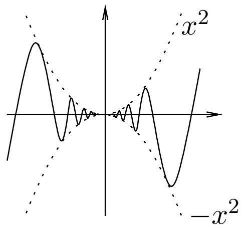
$$g(x)=\begin{cases} x\sin(1/x) & x\neq 0 \\ 0 & x=0 \end{cases}$$

פונקציה זו רציפה באפס (למה?) אך אינה גזירה שם. אפשר להבחין באי־גזירות אם נדמיין מיתר שקצהו האחד בנקודה $(0,0)$ שעל הגרף, וקצהו השני בנקודה הנעה על הגרף וקואורדינטת x שלה מתקרבת לאפס. כפי שרואים באיור 8.1.6, המיתרים האלה מקבלים לסירוגין שיפוע ± 1 (וגם את כל שיפועי הביניים) ולכן השיפוע שלהם אינו מתכנס. באופן פורמלי יותר, לכל $h\neq 0$ מתקיים

$$\frac{g(0+h)-g(0)}{h}=\sin\frac{1}{h}$$

ולפונקציה זו אין גבול ב־0.

3. האי־גזירות בדוגמה הקודמת נבעה מכך שהפונקציה הכילה תנודות שגרמו למיתרים לקבל כל מיני שיפועים כאשר נקודת הקצה מתקרבת לאפס. אולם



איור 8.1.7 הפונקציה $x^2\sin(1/x)$

"רעידות" בפונקציה אינן סיבה מספקת בפני עצמן כדי שלא תהיה נגזרת. נתבונן למשל בפונקציה

$$f(x) = \begin{cases} x^2\sin\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

הגרף שלה מופיע באיור 8.1.7. על אף שהגרף מתנודד ומשנה כיוון אינסוף פעמים בקרבת 0, הפונקציה גזירה באפס ונגזרתה אפס, כי לכל $h \neq 0$ מתקיים

$$\frac{f(0+h)-f(0)}{h} = h\sin\frac{1}{h}$$

והגבול של פונקציה זו באפס הוא אפס כי $0 \leq |h\sin\frac{1}{h}| \leq |h|$.

ראינו בדוגמאות שלפונקציה $f(x)=|x|$ אין נגזרת ב־0, אבל אפשר להבחין שיש לה "משיקים חד־צדדיים". ליתר דיוק,

הגדרה 8.1.3 תהי f פונקציה המוגדרת בסביבה ימנית של x_0 וב־x_0. **הנגזרת הימנית** (right derivative\) של f ב־x_0 היא הגבול

$$\lim_{h\to 0+}\frac{f(x_0+h)-f(x_0)}{h}$$

אם הוא קיים, ומסומנת ב־ $f'_+(x_0)$. באופן דומה מגדירים את הנגזרת השמאלית (left derivative), $f'_-(x_0)$.

עבור פונקציה המוגדרת בסביבת x_0, קיום גבול שקול לקיום הגבולות החד־צדדיים (טענה 7.4.8). על ידי הפעלת הטענה על המנה $\frac{1}{h}(f(x_0+h)-f(x_0))$ כאשר $h\to 0$, אנו מקבלים:

מסקנה 8.1.4 f גזירה ב־x_0 אמ"מ היא גזירה מימין ומשמאל ב־x_0 והנגזרות החד־צדדיות שוות.

דוגמאות

1. לפונקציה $f(x)=|x|$ יש נגזרות חד־צדדיות באפס ומתקיים

$$f'_+(0) = \lim_{h\to 0+}\frac{|h|}{h} = \lim_{h\to 0+}\frac{h}{h} = 1$$

$$f'_-(0) = \lim_{h\to 0-}\frac{|h|}{h} = \lim_{h\to 0-}\frac{-h}{h} = -1$$

שימו לב שהנגזרות החד־צדדיות שונות, ולכן אין ל־f נגזרת שם.

2. לפונקציה $f(x)=\sqrt{x}$ אין נגזרת ימנית ב־0, כי לפי החישוב בדוגמא 5 בעמוד 304, הגבול שמגדיר אותו שואף לאינסוף.
3. לפונקציה g מדוגמה (2) למעלה אין נגזרות חד־צדדיות ב־0, כי לפונקציה $\frac{f(h)-f(0)}{h}=\sin\frac{1}{h}$ אין גבולות חד־צדדיים ב־0 אפילו במובן הרחב.

כשגוזרים פונקציה מקבלים את פונקציית הנגזרת, ואם לזו יש נקודות גזירות אפשר לגזור גם אותה ולקבל את פונקציית הנגזרת שלה, וכן הלאה. כך מקבלים מפונקציה f את פונקציית הנגזרת f', את הנגזרת של הנגזרת $(f')'$, הנגזרת של הנגזרת של הנגזרת, וכן הלאה. באופן פורמלי:

הגדרה 8.1.5 **הנגזרת מסדר** k (או בקיצור הנגזרת ה־k) של f היא הפונקציה $f^{(k)}$ המוגדרת ברקורסיה על ידי $f^{(0)}=f$ ו־ $f^{(n+1)}=\left(f^{(n)}\right)'$. נאמר שפונקציה **גזירה** k **פעמים** בקבוצה D אם $f^{(k)}$ מוגדרת ב־D, ונאמר ש־f **גזירה** k **פעמים ברציפות** בקטע I אם $f^{(k)}$ מוגדרת ורציפה ב־I.

מסמנים בקיצור f'' במקום $f^{(2)}$ ו־ $f'''=f^{(3)}$.

הערות

1. אם f גזירה k פעמים בנקודה x_0 אז $f^{(k-1)}$ חייבת להיות מוגדרת בסביבה של x_0, דהיינו f גזירה $k-1$ פעמים בסביבה של x_0.
2. מכיוון שגזירות גוררת רציפות, אם f גזירה k פעמים בקטע אז היא גזירה m פעמים ברציפות בקטע לכל $m<k$, כי $f^{(m)}$ גזירה ולכן רציפה.
3. ישנן פונקציות שגזירות k פעמים בנקודה מסוימת אך לא $k+1$ פעמים. נראה דוגמאות לכך בהמשך (ראו תרגיל (16) בעמוד 341).

ישנו סימון נוסף לנגזרת שנעשה בו שימוש בעיקר כשנגזור פונקציה הנתונה בעזרת נוסחה. כאשר f פונקציה הנתונה על ידי נוסחה שמופיע בה המשתנה x הסימן $\frac{d}{dx}f$ או $\frac{df}{dx}$ פירושו פונקציית הנגזרת של f (ובאופן דומה, הנגזרת ה־k מסומנת לפעמים על ידי $\frac{d^k}{dx^k}f$). כך למשל $\frac{d}{dx}x^2$ מציין את הנגזרת של הפונקציה $f(x)=x^2$: כלומר $\frac{d}{dx}x^2=2x$.

התפקיד של "x" בסימן "$\frac{d}{dx}$" הוא להבהיר מיהו המשתנה בנוסחה המגדירה את הפונקציה: כך למשל $\frac{d}{dx}x^2=2x$, אבל עבור מספר קבוע y מתקיים $\frac{d}{dx}y^2=0$ כי y^2 היא הפונקציה הקבועה (היא אינה תלויה ב־x). לעומת זאת, $\frac{d}{dy}y^2=2y$.

סימון הנגזרת בעזרת $\frac{d}{dx}$ מעט מבלבל מהסיבה הבאה. אם $f(x)=x^2$ אז הנגזרת של f היא $f'(x)=2x$, אבל כשרושמים

$$\frac{d}{dx}x^2=2x$$

הרי ש־x מציין שני דברים שונים באגף ימין ובאגף ושמאל: בצד שמאל הוא משמש להגדרת הפונקציה שאותה גוזרים, ואילו בצד ימין הוא המשתנה של פונקציית הנגזרת. למשל, תהיה זו טעות להציב $x=1$ בשני האגפים: נקבל אז $\frac{d}{dx}1=2$ וזה כמובן לא נכון: 1 היא פונקציה קבועה ולכן הנגזרת שלה היא אפס.

כדי להסביר את הסימון $\frac{df}{dx}$, ניזכר ש־ $f'(x_0)=\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}$. אם נסמן Δx במקום h ונגדיר $\Delta f=f(x_0+\Delta x)-f(x_0)$ אז Δf מתאר את השינוי בפונקציה f שחל כתוצאה של שינוי בגודל Δx במשתנה, והוא פונקציה של Δx. כעת נוכל לרשום את הגדרת הנגזרת בתור

$$\frac{df}{dx}=\lim_{\Delta x\to 0}\frac{\Delta f}{\Delta x}$$

הרישום $\frac{df}{dx}$ נועד לרמוז על הגבול הזה. לפני המצאת הגבול התייחסו לסימן dx כעל גודל שמתקבל משינוי מזערי של x_0, ועל df כעל השינוי המזערי שחל ב־f כתוצאה מכך. הנגזרת אז הוגדרה בתור המנה של ה"מספרים המזעריים" האלה, אשר כונו **אינפיניטסימלים** (infinitesimals). אחד המניעים להגדרה המדויקת של הגבול היה חוסר הבהירות של המושג זה.

תרגילים

1. מצאו את נקודות הגזירות וחשבו את הנגזרות (דו־צדדיות או חד־צדדיות) של הפונקציות הבאות:
   (א) $f(x)=x^3$.
   (ב) $f(x)=1/x$.
   (ג) $f(x)=|x|$.
   (ד) $f(x)=[x]$ ($[x]$ הוא הערך השלם של x).
   (ה) $f(x)=\sqrt{|x|}$.

2. עבור אילו ערכים של a,b,c הפונקציה

$$f(x)=\begin{cases} ax & x<1 \\ bx^2+c & x\geq 1\end{cases}$$

   היא גזירה?

3. הראו שעבור $f(x)=x^k$ מתקיים $f'(x)=kx^{k-1}$ לכל k טבעי (היעזרו במשפט הבינום).

4. תהי f מוגדרת בסביבת x_0.
   (א) הראו שאם f גזירה ב־x_0 אז מתקיים

$$\lim_{h\to 0}\frac{f(x_0+h)-f(x_0-h)}{2h}=f'(x_0)$$

   ותנו פירוש גאומטרי לגבול הזה.

(ב) הראו שקיום הגבול מהסעיף הקודם אינו גורר בהכרח גזירות.

5. תהי $f:\mathbb{R}\to\mathbb{R}$ גזירה בכל נקודה, יהי $c\in\mathbb{R}$ ונגדיר $g(x)=f(x+c)$, $h(x)=f(cx)$. הוכיחו ש־ $g'(x)=f'(x+c)$ ו־ $h'(x)=cf'(cx)$. תנו פירוש גאומטרי לשוויונות אלה.

6. תהי f פונקציה גזירה ב־x_0.

(א) הוכיחו שאם $f(x_0)\neq 0$ אז $|f|$ גזירה ב־x_0, ושאם $f(x_0)=0$ אז $|f|$ גזירה ב־x_0 אמ"מ $f'(x_0)=0$.

(ב) נניח $|f|$ גזירה ב־x_0. האם נובע ש־f גזירה ב־x_0?

7. יהיו f,g פונקציות גזירות ב־x_0. הוכיחו שאם $f(x_0)\neq g(x_0)$ אז הפונקציה $h(x)=\max\{f(x),g(x)\}$ גזירה ב־x_0, ומצאו את נגזרתה. מה קורה אם $f(x_0)=g(x_0)$?

8. הוכיחו את הטענות הבאות:

(א) אם f פונקציה זוגית אז היא גזירה ב־x_0 אמ"מ היא גזירה ב־$-x_0$ ומתקיימת $f'(x_0)=-f'(-x_0)$.

(ב) אם f פונקציה אי־זוגית אז היא גזירה ב־x_0 אמ"מ היא גזירה ב־$-x_0$ ומתקיימת $f'(x_0)=f'(-x_0)$.

(ג) תהי $f:\mathbb{R}\to\mathbb{R}$ ונניח ש־ $\Delta>0$ הוא מחזור של f. הראו ש־f גזירה ב־x_0 אמ"מ f גזירה ב־ $x_0+\Delta$ ושהנגזרות שם שוות.

9. תהי

$$f(x)=\begin{cases} x^2 & x\notin\mathbb{Q} \\ -x^2 & x\in\mathbb{Q}\end{cases}$$

מצאו את נקודות הגזירות של f וחשבו את נגזרתה שם.

10. הוכיחו שפונקציית רימן, הנתונה על ידי

$$R(x)=\begin{cases} \frac{1}{q} & x=\frac{p}{q} \\ 0 & x\notin\mathbb{Q}\end{cases}$$

(p/q הוא שבר מצומצם עם $q>0$), אינה גזירה באף נקודה.

8.2 פונקציות אפסיות והנגזרת כקירוב לינארי

בסעיף זה נראה שעל מנת שפונקציה תהיה גזירה ב־x_0 הכרחי ומספיק שיהיה אפשר להציג את הפונקציה כסכום של ישר ℓ ופונקציה α באופן כזה שבקרבת x_0 התרומה של α לסכום היא מזערית.

הגדרה 8.2.1 נאמר שפונקציה α היא **אפסית** אם מתקיים $\alpha(x) = x \cdot \alpha^*(x)$ עבור פונקציה α^* שהיא רציפה ב־0 ו־ $\alpha^*(0) = 0$. במקרה זה גם נאמר ש־α היא $o(x)$ (קרי: α היא o־**קטן** של x).

אפסיות של פונקציה α פירושה שכאשר $x \to 0$, קצב ההתכנסות של $\alpha(x)$ ל־0 מהיר יותר מהשאיפה של כל פונקציה ליניארית $f(x) = cx$ ($c \neq 0$).

שימו לב שאם α אפסית אז α רציפה ב־0 ו־ $\alpha(0) = 0$, כי α היא מכפלה של שתי פונקציות רציפות ב־0 ששתיהן מתאפסות ב־0.

דוגמאות

1. הפונקציה $\alpha(x) = x^2$ היא אפסית כי $x^2 = x \cdot \alpha^*(x)$ כאשר $\alpha^*(x) = x$ וזו פונקציה רציפה המתאפסת ב־0. באופן כללי יותר, ל־ $\varepsilon > 0$ הפונקציה $\alpha(x) = |x|^{1+\varepsilon}$ היא אפסית (הוכיחו!).
2. הפונקציה $\alpha(x) = x$ אינה אפסית, כי אם $\alpha(x) = x \cdot \alpha^*(x)$ אז $\alpha^* \equiv 1$ בסביבה מנוקבת של 0 וממילא אז $\alpha^*(0) = 1 \neq 0$.

אפסיות היא תכונה אינפיניטסימלית בנקודה 0. אפשר לראות זאת מההגדרה, אך זה נובע מיד גם מהלמה הבאה:

למה 8.2.2 α אפסית אמ"מ $\alpha(0) = 0$ ומתקיים $\lim_{x\to 0} \frac{\alpha(x)}{x} = 0$.

הוכחה אם α אפסית אז $\alpha(0) = 0$ ומתקיים $\alpha(x) = x\alpha^*(x)$ כאשר α^* רציפה ב־0 ומקיימת $\alpha^*(0) = 0$. לכן

$$\lim_{x\to 0} \frac{\alpha(x)}{x} = \lim_{x\to 0} \frac{x\alpha^*(x)}{x} = \lim_{x\to 0} \alpha^*(x) = \alpha^*(0) = 0$$

להפך, אם $\alpha(0) = 0$ ו־ $\lim_{x\to 0} \frac{\alpha(x)}{x} = 0$ אז נגדיר

$$\alpha^*(x) = \begin{cases} \frac{\alpha(x)}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

ברור ש־ $\alpha(x) = x\alpha^*(x)$ וש־α^* רציפה באפס ומקיימת $\alpha^*(0) = 0$. ∎

אחד השימושים של מושג הפונקציה האפסית הוא לפשט ביטויים מסובכים. נאמץ את המוסכמה הבאה:

מוסכמה כאשר $o(x)$ מופיע בביטוי חשבוני הוא מייצג פונקציה אפסית (במשתנה x). בהקשרים שונים הסימון $o(x)$ יכול לייצג פונקציות שונות, אפילו בתוך אותו ביטוי. בפרט הסימון $\alpha(x) = o(x)$ פירושו ש־α היא אפסית.

בהמשך הסימון $o(x)$ יבוא לעתים קרובות במקום פונקציות אפסיות שכתיבתן מסובכת. ננהג כך במקרים שבהם אנו מנתחים ביטוי שבתוכו מופיע תת־ביטוי מסובך שהתכונה היחידה שחשובה לנו לגביו היא היותו אפסי.

למשל אפשר לרשום

$$x + x^3 \cos x + \frac{1}{1+x^2} = x + o(x) + \frac{1}{1+o(x)}$$

את השוויון יש לקרוא כך: הפונקציה באגף שמאל שווה לסכום של שלושה מחוברים, שאחד מהם הוא x, אחד מהם אפסי והשלישי שווה לאחד חלקי הסכום של אחד עם פונקציה אפסית. שימו לב שהסימון $o(x)$ מופיע פעמיים בצד ימין אך הוא אינו מציין את אותה הפונקציה בשתי הופעותיו: פעם הוא מציין את הפונקציה $x^3 \cos x$ ופעם את x^2 (וודאו שאלו פונקציות אפסיות!). שימו לב שהביטוי בצד ימין של השוויון נותן תיאור חלקי בלבד של הפונקציה המופיעה בצד שמאל, שכן הוא מכיל פחות מידע: החלפנו את הביטויים המדויקים $x^2, x^3 \cos x$ בסימון $o(x)$ שרק מרמז על כמה תכונות של הביטוי המקורי. ממילא בהינתן צד ימין, לא נוכל לשחזר את צד שמאל. ואמנם, מתקיים גם השוויון

$$x(1+x) + x^5 + \frac{1}{1+x\cos x} = x + o(x) + \frac{1}{1+o(x)}$$

הביטוי באגף ימין הוא אותו ביטוי כמו קודם אבל הוא מתאר פונקציה אחרת.

הסימון $o(x)$ הומצא כדי לפשט ביטויים מסובכים ויש כמה כללי אריתמטיקה פשוטים המסייעים בכך:

למה 8.2.3 (כללי תחשיב של פונקציות אפסיות) הסכום של פונקציות אפסיות הוא אפסי, והמכפלה של פונקציה אפסית עם פונקציות חסומה בסביבת 0 היא אפסית.

הוכחה נניח ש־α, β אפסיות ותהי $\gamma = \alpha + b$. אז $\gamma(0) = \alpha(0) + \beta(0) = 0$, וגם

$$\lim_{x\to 0} \frac{\gamma(x)}{x} = \lim_{x\to 0} \frac{\alpha(x)}{x} + \lim_{x\to 0} \frac{\beta(x)}{x} = 0$$

ולכן לפי הלמה הקודמת, γ אפסית. הוכחת הטענה השנייה דומה. ■

את מסקנת הלמה רושמים לפעמים כך: $o(x)+o(x) = o(x)$, ואם f חסומה בסביבת 0 אז $o(x) \cdot f(x) = o(x)$.

הלמה מאפשרת לפשט ביטויים המכילים פונקציות אפסיות. נדגים זאת בעזרת שרשרת השוויונות הבאה:

$$\begin{aligned}(x^2 e^x + x \cdot \sqrt{x} \cos x) \tan x &= (o(x) e^x + o(x) \cos x) \tan x \\ &= (o(x) + o(x)) \tan x \\ &= o(x) \tan x \\ &= o(x)\end{aligned}$$

(השוויון הראשון כי $x^2, x\sqrt{x}$ הן אפסיות, השני כי $e^x, \cos x$ חסומות בסביבת 0, השלישי כי הסכום של פונקציות אפסיות הוא אפסי, והאחרון כי $\tan x$ חסומה בסביבת 0). שרשרת השוויונות מראה שהביטוי המקורי הוא פונקציה אפסית של x.

האפיון החשוב הבא קובע שגזירות של פונקציה בנקודה פירושה שבסביבת הנקודה היא נבדלת מישר במידה אפסית כאשר $x \to x_0$. ליתר דיוק,

משפט 8.2.4 תהי f פונקציה ממשית. אז f גזירה ב־x_0 ונגזרתה שם שווה ל־a אמ"מ מתקיים

$$f(x_0+h) = f(x_0) + ah + o(h)$$

הערה השוויון במשפט הוא בין פונקציות של h. המספר x_0 הוא קבוע.

הוכחה נניח ש־ $f(x_0+h) = f(x_0) + ah + o(h)$. כלומר, נניח שקיימת פונקציה אפסית α כך ש־

$$f(x_0+h) = f(x_0) + ah + \alpha(h)$$

על ידי העברת אגפים מקבלים

$$\frac{\alpha(h)}{h} = \frac{f(x_0+h) - f(x_0)}{h} - a$$

הגבול של אגף שמאל הוא 0 כאשר h שואף ל־0 (כי α אפסית ולפי למה 8.2.2) ולכן גם הגבול של אגף ימין הוא אפס. פירוש הדבר שהנגזרת של f ב־x_0 קיימת ושווה ל־a.

להפך, נניח ש־f גזירה ב־x_0 ו־ $f'(x_0) = a$. עלינו להגדיר פונקציה אפסית α כך שיתקיים $f(x_0+h) = f(x_0) + ah + \alpha(h)$. למעשה שוויון זה מכתיב איך להגדיר את α: נגדיר

$$\alpha(h) = f(x_0+h) - f(x_0) - ah$$

נציב $h = 0$ ונקבל $\alpha(0) = 0$. כמו־כן מתקיים

$$\frac{\alpha(h)}{h} = \frac{f(x_0+h) - f(x_0)}{h} - a$$

ולכן מאחר ש־$f'(x_0) = a$ נקבל

$$\lim_{h\to 0} \frac{\alpha(h)}{h} = 0 = \alpha(0)$$

ולפי למה 8.2.2 פירוש הדבר ש־α אפסית. ■

הערות

1. כשרושמים $f(x_0+h) = f(x_0) + ah + o(h)$, המחובר $f(x_0)$ באגף ימין נקרא המחובר **מסדר אפס**, המחובר ah נקרא המחובר **מסדר ראשון** (first order) או **החלק הלינארי** (linear part) של f, והתוספת $o(h)$ נקראת **השארית** (מסדר ראשון).
2. אפשר לרשום את השוויון $f(x_0+h) = f(x_0) + ah + o(h)$ גם בצורה

$$f(x) = f(x_0) + a(x-x_0) + \alpha(x)$$

כאשר α היא פונקציה המקיימת $\alpha(x_0) = 0$ ו־ $\lim_{x\to x_0} \frac{\alpha(x)}{x-x_0} = 0$. כך או כך, פירוש הדבר הוא שבקרבת x_0 הפונקציה f שווה בקירוב לישר $\ell(x) = f(x_0) + a(x-x_0)$. עם זאת, אפסיות היא תכונה אינפיניטסימלית, ולכן הקרבה בין f ל־ℓ מובטחת רק בקרבת x_0.

דוגמאות

1. יהי $f(x) = ax+b$ ישר. לכל $x_0 \in \mathbb{R}$ ולכל h מתקיים

$$f(x_0+h) = a(x_0+h) + b = (ax_0+b) + ah = f(x_0) + ah$$

ולכן באמת אפשר לרשום את $f(x_0+h)$ כמו במשפט 8.2.4 כאשר התוספת האפסית היא פשוט אפס. אנו מסיקים ש־ $f'(x_0) = a$ בכל נקודה x_0.

2. תהי $f(x) = x^2$. אז לכל x_0, h מתקיים

$$f(x_0+h) = (x_0+h)^2 = x_0^2 + 2x_0h + h^2 = f(x_0) + 2x_0h + h^2$$

ומכיוון ש־ $h^2 = o(h)$ הרי ש־f גזירה ב־x_0 ומתקיים $f'(x_0) = 2x_0$.

נסיים את הדיון עם הגדרה של הישר המשיק ואפיון של הנגזרת במונחים שלו:

הגדרה 8.2.5 תהי f פונקציה ממשית המוגדרת בסביבה של x_0. נאמר שישר ℓ הוא **משיק** (tangent line) של f ב־x_0 אם $f(x_0+h) = \ell(x_0+h) + o(h)$.

טענה 8.2.6 ישר ℓ הוא ישר משיק ל־f ב־x_0 אמ"מ f גזירה ב־x_0 ו־ℓ נתונה על ידי $\ell(x) = f'(x_0)(x-x_0) + f(x_0)$.

הוכחה יהי $\ell(x) = a(x-x_0) + b$ ישר. אז מהגדרת המשיק, ℓ הוא משיק ל־f ב־x_0 אמ"מ

$$f(x_0+h) = \ell(x_0+h) + o(h) = ah + b + o(h)$$

על ידי הצבת $h = 0$ נקבל ששוויון כזה יכול להתקיים רק אם $b = f(x_0)$, כלומר הוא שקול לשוויון

$$f(x_0+h) = f(x_0) + ah + o(h)$$

ושוויון זה שקול לכך ש־ f גזירה ב־x_0 ומקיימת $f'(x_0) = a$, כנדרש. ■

תרגילים

1. אילו מהפונקציות הן אפסיות?
   (א) $|x|^\alpha$ עבור $\alpha > 0$.
   (ב) $x(1+\frac{1}{\sqrt{x}})$.
   (ג) $x+x^2+xe^x$.
   (ד) $\frac{x\sin x}{1+x}$.
2. נניח ש־α אפסית ו־β אינה אפסית. האם $\alpha+\beta$ יכולה להיות אפסית?
3. תהי α פונקציה אפסית. הוכיחו ש־ $\alpha(cx)$ אפסית לכל קבוע c.
4. פשטו את הביטויים הבאים בעזרת כללי התחשיב לפונקציות אפסיות:
   (א) $o(x)+o(x)\cdot 8$.
   (ב) $o(x)\cdot o(x)+\frac{x+e^x}{2}o(x)$.
   (ג) $\frac{o(x)+x^2}{e^x+o(x)}$.
   (ד) $\frac{1+o(x)\tan x+o(x)^2}{x+1}$.
5. תהי f גזירה ב־x_0 ונניח $g(x+h)=f(x+h)+o(h)$. הוכיחו ש־g גזירה ב־x_0 ומתקיים $g'(x_0)=f'(x_0)$.

8.3 כללי תחשיב של נגזרות

בעזרת אפיון הנגזרת מהסעיף הקודם (משפט 8.2.4) נפתח להלן כללי תחשיב לנגזרות. באופן כללי השיטה היא כדלקמן: נניח שנתונה פונקציה h המורכבת באופן כלשהו מפונקציות f,g אשר גזירות ב־x_0. נרשום כל אחת מהפונקציות f,g כסכום של ישר ותוספת אפסית, נציב זאת בהגדרה של h, וננסה על ידי מניפולציות אלגבריות להביא את h לצורה של ישר ועוד תוספת אפסית. אם נצליח בכך הרי ש־h גזירה והנגזרת שלה היא החלק הלינארי של הביטוי שקיבלנו.

נתחיל עם הכלל הפשוט ביותר, כלל הסכום. כאן אין הפתעות:

משפט 8.3.1 (כלל הסכום) אם f,g גזירות ב־x_0 אז גם $f+g$ גזירה ב־x_0 ומתקיים

$$(f+g)'(x_0)=f'(x_0)+g'(x_0)$$

כמו כן אם $c\in\mathbb{R}$ אז cf גזירה ב־x_0 ומתקיים

$$(cf)'(x_0)=c\cdot f'(x_0)$$

הוכחה קל להוכיח את הנוסחה $(f+g)' = f' + g'$ גם ישירות מההגדרה, אך לצורך הדגמה נוכיח אותה בשיטה שתיארנו למעלה. נסמן $a = f'(x_0)\,,\, b = g'(x_0)$. לפי משפט 8.2.4 אפשר לרשום את $f(x_0+h)$ כישר ועוד תיקון אפסי:

$$f(x_0+h) = f(x_0) + ah + o(h)$$

ובאופן דומה אפשר לרשום גם את g:

$$g(x_0+h) = g(x_0) + bh + o(h)$$

נחבר את השוויונות האלה, ונרשום את התוצאה גם־כן כסכום של ישר ותיקון אפסי:

$$\begin{aligned}(f+g)(x_0+h) &= f(x_0) + ah + o(h) + g(x_0) + bh + o(h)\\ &= (f+g)(x_0) + (a+b)h + (o(h)+o(h))\\ &= (f+g)(x_0) + (a+b)h + o(h)\end{aligned}$$

(במעבר מהשורה השנייה לשלישית השתמשנו בלמה 8.2.3). ממשפט 8.2.4 נסיק ש־$f+g$ גזירה ב־x_0 ונגזרתה שווה למקדם של h בביטוי שקיבלנו, כלומר ל־$a+b$, אשר שווה ל־$f'(x_0) + g'(x_0)$.

הטענה לגבי cf קלה עוד יותר ומושארת כתרגיל. ■

דוגמאות

1. ראינו ש־ $\frac{d}{dx}x^2 = 2x$ וש־ $\frac{d}{dx}\sqrt{x} = \frac{1}{2\sqrt{x}}$. לכן מכלל הסכום,

$$\frac{d}{dx}(x^2+\sqrt{x}) = \frac{d}{dx}(x^2) + \frac{d}{dx}(\sqrt{x}) = 2x + \frac{1}{2\sqrt{x}}$$

2. קל להראות באינדוקציה שכלל הסכום נכון לכל סכום סופי של פונקציות. כך למשל,

$$\begin{aligned}\frac{d}{dx}(1+x+x^2+x^3) &= \frac{d}{dx}1 + \frac{d}{dx}x + \frac{d}{dx}x^2 + \frac{d}{dx}x^3\\ &= 0 + 1 + 2x + 3x^2\end{aligned}$$

נעבור לכלל המכפלה. היה אפשר לצפות שיתקיים $(f\cdot g)'(x_0) = f'(x_0)\cdot g'(x_0)$, אך לא כך הדבר. נוסחה זו אינה מתקיימת אפילו כאשר f,g הם ישרים. ואמנם, אם f,g הן ישרים עם שיפועים a,b בהתאמה, אז אפשר לרשום

$$f(x_0+h) = ah + f(x_0) \qquad , \qquad g(x_0+h) = bh + g(x_0)$$

ואז

$$\begin{aligned}(f\cdot g)(x_0+h) &= (f(x_0)+ah)(g(x_0)+bh)\\ &= f(x_0)g(x_0)+(ag(x_0)+f(x_0)b)h+abh^2\end{aligned}$$

הפונקציה abh^2 היא אפסית ושני המחוברים האחרים מתארים ישר. אנו מסיקים לכן ש־fg גזירה ב־x_0 ומתקיים

$$(fg)'(x_0)=ag(x_0)+f(x_0)b=f'(x_0)g(x_0)+f(x_0)g'(x_0)$$

תוצאה זו נכונה באופן כללי:

משפט 8.3.2 (כלל המכפלה)[2] אם f,g גזירות ב־x_0 אז $f\cdot g$ גזירה ב־x_0 ומתקיים

$$(f\cdot g)'(x_0)=f'(x_0)g(x_0)+f(x_0)g'(x_0)$$

הוכחה החישוב דומה לחישוב שעשינו במקרה ש־f,g ישרים, אלא שיש לקחת בחשבון את התרומה של התוספות האפסיות. נסמן $a=f'(x_0)\,,\,b=g'(x_0)$. אז

$$\begin{aligned}(f\cdot g)(x_0+h) &= f(x_0+h)\cdot g(x_0+h)\\ &= (f(x_0)+ah+o(h))(g(x_0)+bh+o(h))\end{aligned}$$

נפתח את הסוגריים וננסה להציג את מה שמתקבל כסכום של ישר ותוספת אפסית:

$$\begin{aligned}= \;& f(x_0)g(x_0)\\ &+(ag(x_0)+f(x_0)b)h\\ &+\;(f(x_0)+ah)\,o(h)+(g(x_0)+bh)\,o(h)+o(h)\cdot o(h)\end{aligned}$$

יש לוודא שסכום שלושת המחוברים בשורה האחרונה הוא אפסי. לפי למה 8.2.3 די שנראה שכל אחד מהמחוברים הוא אפסי. ואכן, כל מחובר הוא מכפלה של פונקציה אפסית עם פונקציה שיש לה גבול כאשר $h\to 0$, ובפר חסומה בסביבת 0, ולכן אפסיות נובעת שוב לפי אותה למה. לכן קיבלנו

$$(f\cdot g)(x_0+h)=f(x_0)g(x_0)+(ag(x_0)+f(x_0)b)h+o(h)$$

ואנו מסיקים ש־fg גזירה ב־x_0 וש־ $(fg)'(x_0)=ag(x_0)+f(x_0)b$. נציב את ההגדרות של a,b ונקבל את המסקנה המבוקשת. ■

[2]כלל זה נקרא לפעמים כלל לייבניץ.

דוגמאות

1. נתבונן בפונקציה $x(1+x)$ ונציג אותה כ־ $f \cdot g$ כאשר $f(x) = x$ ו־ $g(x) = 1 + x$. אז $f'(x) = 1$ ו־ $g'(x) = 1$ ולכן לפי כלל המכפלה

$$\begin{aligned} \frac{d}{dx} x(1+x) &= f'(x)g(x) + f(x)g'(x) \\ &= 1 \cdot (1+x) + x \cdot 1 \\ &= 2x + 1 \end{aligned}$$

2. במכפלות עם יותר משני גורמים אפשר להשתמש בכלל המכפלה בשלבים. למשל, נתבונן בפונקציה $x(x+1)(1+x+x^2)$ ונציג אותה כ־ $f \cdot g \cdot h$ כאשר f, g כמו בדוגמה הקודמת ו־ $h(x) = 1 + x + x^2$. אז $h' = 1 + 2x$ (את f', g' חישבנו בדוגמה הקודמת) ושימוש בכלל המכפלה פעמיים נותן

$$\begin{aligned} ((fg)h)' &= (fg)'h + (fg) \cdot h' \\ &= (f'g + fg')h + (fg)h' \\ &= (1 \cdot (x+1) + x \cdot 1)(1 + x + x^2) + (x(x+1))(1 + 2x) \\ &= (2x+1)(1 + 2x + 2x^2) \end{aligned}$$

מתבקש בשלב זה לחשב את הנגזרת של מנה. נדון בכך במשפט 8.3.5 להלן, שמבוסס על הכלל לגזירה של הרכבה של פונקציות. כדי להבין את כלל הגזירה של פונקציה מורכבת נזדקק ללמה נוספת על פונקציות אפסיות.

למה 8.3.3 תהי α פונקציה מהצורה $\alpha(x) = cx + o(x)$, כאשר c מספר קבוע. אם γ פונקציה אפסית אז גם $\gamma \circ \alpha$ אפסית.

הוכחה לפי ההנחה יש פונקציה α^* שמתאפסת ב־0 ורציפה שם, כך ש־

$$\alpha(x) = cx + x\alpha^*(x) = x(c + \alpha^*(x))$$

כמו־כן לפי ההנחה $\gamma(x) = x\gamma^*(x)$ עבור פונקציה γ^* שרציפה ב־0 ומתאפסת שם. לכן

$$\begin{aligned} \gamma(\alpha(x)) &= \alpha(x) \cdot \gamma^*(\alpha(x)) \\ &= x\left((c + \alpha^*(x)) \cdot \gamma^*(\alpha(x))\right) \end{aligned}$$

כדי להסיק ש־ $\gamma \circ \alpha$ אפסית די שנראה ש־$(c+\alpha^*) \cdot (\gamma^* \circ \alpha)$ רציפה ב־0 ומתאפסת שם. אבל $c + \alpha^*$ היא פונקציה רציפה ב־0 כי היא סכום של פונקציות רציפות שם, ואילו $\gamma^* \circ \alpha$ היא הרכבה של פונקציות ששתיהן רציפות באפס ומתאפסות שם, ולכן גם ההרכבה מקיימת זאת. מכאן ש־ $(c+\alpha^*) \cdot \gamma^* \circ \alpha$ רציפה ומתאפסת באפס, כפי שרצינו. ∎

משפט 8.3.4 (כלל השרשרת) אם f גזירה ב־x_0 ו־g גזירה ב־ $y_0 = f(x_0)$ אז $g \circ f$ גזירה ב־x_0 ומתקיים $(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$.[3]

הוכחה כדאי לבדוק מה קורה כאשר f, g הם ישרים. אנו משאירים את הבדיקה הזו כתרגיל, ונוכיח את המקרה הכללי. נסמן $b = g'(y_0)$, $a = f'(x_0)$. כדי להעריך את $g(f(x_0+h))$ נשתמש כרגיל בעובדה ש־ $f(x_0+h) = y_0 + ah + o(h)$ ונקבל

$$\begin{aligned}(g \circ f)(x_0+h) &= g(f(x_0+h)) \\ &= g(y_0 + (ah + o(h)))\end{aligned}$$

לפי הנתון, קיימת פונקציה אפסית γ וסביבה של 0 כך שלכל t באותה סביבה מתקיים $g(y_0+t) = g(y_0) + bt + \gamma(t)$. נציב $t = ah + o(h)$ ונקבל בעזרת השוויון הקודם

$$\begin{aligned}(g \circ f)(x_0+h) &= g(y_0) + b \cdot (ah + o(h)) + \gamma(ah + o(h)) \\ &= g(y_0) + bah + (b \cdot o(h) + \gamma\,(ah + o(h)))\end{aligned}$$

לפי למה 8.3.3 מתקיים $\gamma(ah+o(h)) = o(h)$, לפי למה 8.2.3 מתקיים $b \cdot o(h) = o(h)$, ולפי אותה למה סכומם הוא $o(h)$. קיבלנו אפוא

$$(g \circ f)(x_0+h) = g(f(x_0)) + ba \cdot h + o(h)$$

ולכן $g \circ f$ גזירה ב־x_0 ונגזרתה $b \cdot a$, כלומר: $g'(f(x_0))f'(x_0)$, כנדרש. ■

דוגמאות

1. נתבונן בפונקציה $\sqrt{1+x}$. זו הרכבה $g \circ f$ כאשר $f(x) = 1+x$ ו־ $g(y) = \sqrt{y}$. מאחר ש־ $f'(x) = 1$ ו־ $g'(y) = \frac{1}{2\sqrt{y}}$ בכל $y > 0$, אנו מקבלים מכלל השרשרת שבכל נקודה x עבורה $f(x) > 0$ (דהיינו לכל $x > -1$) מתקיים

$$\begin{aligned}\frac{d}{dx}\sqrt{1+x} &= g'(f(x))f'(x) \\ &= \frac{1}{2\sqrt{1+x}} \cdot 1\end{aligned}$$

[3] אפשר לנסח את כלל השרשרת גם באופן הבא: יהי y גודל מספרי התלוי ב־x, ויהי z גודל מספרי התלוי ב־y, ולכן בעקיפין ב־x. הנגזרת $\frac{dz}{dx}$ של z כפונקציה של x אז מקיימת, לפי כלל השרשרת, $\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$. בסימון זה אנו לכאורה מבטאים את כלל השרשרת כ"צמצום" של dy לפי הכללים הרגילים של מנות. אין זה צמצום באמת כי לא נתנו כל משמעות לסימון dy לבדו ואין מדובר במנה של מספרים שלגביה חלים כללי האריתמטיקה של מספרים. נציין שאפשר לתת פירוש לסימנים dx, dy, dz באופן שפעולת הצמצום הזו תהיה מוצדקת, אך לא נעסוק בכך כאן.

2. נתבונן בפונקציה $(1+\sqrt{x})^{10}$. זו פונקציה מהצורה $g\circ f$ כאשר $f(x)=1+\sqrt{x}$ ו־ $g(y)=y^{10}$. מתקיים $f'(x)=\frac{1}{2\sqrt{x}}$ לכל $x>0$, ואילו g גזירה בכל הישר ומקיימת $g'(y)=10y^9$. לכן בכל נקודה $x>0$ מתקיים

$$\begin{aligned}\frac{d}{dx}(1+\sqrt{x})^{10} &= g'(f(x))f'(x)\\ &= 10(1+\sqrt{x})^9\cdot\frac{1}{2\sqrt{x}}\end{aligned}$$

3. שני מקרים פרטיים אך חשובים של כלל השרשרת הם

$$\begin{aligned}\frac{d}{dx}(f(x+c)) &= f'(x+c)\\ \frac{d}{dx}(f(cx)) &= cf'(cx)\end{aligned}$$

המתארות, עבור מספר c קבוע, את הנגזרת של הפונקציות המתקבלות מ־f על ידי הזזה ומתיחה (בדקו שהנוסחאות נובעות מכלל השרשרת! אפשר גם להוכיח אותן משיקולים אלמנטריים, כמו בתרגיל (5) בעמוד 309). נוסחאות אלה אינן מפתיעות: לפי הפירוש הגאומטרי של הנגזרת אנו מצפים שאם נזיז את גרף הפונקציה ימינה או שמאלה המשיק יזוז יחד עם הפונקציה ולא ישנה את שיפועו, ולכן השיפוע של משיק לאחר ההזזה לא יהיה שונה משיפועו לפני ההזזה. באופן דומה אם נמתח או נכווץ את הגרף בכיוון האופקי השיפוע של המשיק ישתנה באופן פרופורציוני למידת המתיחה.

מסקנה לא קשה מכלל השרשרת היא:

משפט 8.3.5 (כלל המנה) יהיו f,g פונקציות גזירות ב־x_0 ונניח ש־ $g(x_0)\neq 0$. אז $\frac{f}{g}$ גזירה ב־x_0 ומתקיים

$$(\frac{f}{g})'(x_0)=\frac{f'(x_0)g(x_0)-f(x_0)g'(x_0)}{g(x_0)^2}$$

בפרט,

$$(\frac{1}{g})'(x_0)=-\frac{g'(x)}{g(x)^2}$$

הוכחה תחילה נחשב את הנגזרת של $\frac{1}{g}$. יהי $r(y)=\frac{1}{y}$, כך ש־ $\frac{1}{g}=r\circ g$. לפי ההנחה $g(x_0)\neq 0$ ולכן r גזירה ב־$g(x_0)$ ולפי דוגמה (4) בעמוד 304 מתקיים $r'(y)=-\frac{1}{y^2}$ לכל $y\neq 0$. מכאן בעזרת כלל השרשרת אנו מקבלים

$$(\frac{1}{g})'(x_0)=r'(g(x_0))\cdot g'(x_0)=-\frac{g'(x_0)}{g(x_0)^2}$$

כעת נשים לב ש־ $\frac{f}{g} = f \cdot \frac{1}{g}$ ולכן לפי כלל המכפלה אנו מקבלים

$$\begin{aligned}(\frac{f}{g})'(x_0) &= f'(x_0)(\frac{1}{g})(x_0) + f(x_0)(\frac{1}{g})'(x_0)\\ &= \frac{f'(x_0)}{g(x_0)} - \frac{f(x_0)g'(x_0)}{g(x_0)^2}\\ &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}\end{aligned}$$

כנדרש. ■

דוגמאות

1. נתבונן בפונקציה $\frac{1}{1+x^2}$. אז לפי כלל המנה

$$\begin{aligned}\frac{d}{dx}(\frac{1}{1+x^2}) &= -\frac{\frac{d}{dx}(1+x^2)}{(1+x^2)^2}\\ &= -\frac{2x}{(1+x^2)^2}\end{aligned}$$

.

2. נתבונן בפונקציה $\frac{\sqrt{x}}{1+x}$. לפי כלל המנה,

$$\begin{aligned}\frac{d}{dx}(\frac{\sqrt{x}}{1+x}) &= \frac{(\frac{d}{dx}\sqrt{x})(1+x) - \sqrt{x}\cdot\frac{d}{dx}(1+x)}{(1+x)^2}\\ &= \frac{\frac{1}{2\sqrt{x}}(1+x) - \sqrt{x}}{(1+x)^2}\end{aligned}$$

נותר לנו למצוא כלל גזירה לפונקציה הפוכה. נניח ש־ $f : I \to \mathbb{R}$ חח"ע בעלת תמונה J ושיש ל־f פונקציה הפוכה $f^{-1} : J \to I$. נשים לב שאם $(x_0, f(x_0)), (x_1, f(x_1))$ הן נקודות על הגרף של f אז המיתר ביניהן הוא בעל שיפוע

$$a = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

לעומת זאת הנקודות $(f(x_0), x_0), (f(x_1), x_1)$ נמצאות על הגרף של f^{-1} והשיפוע של המיתר ביניהן הוא

$$\frac{x_1 - x_0}{f(x_1) - f(x_0)} = \frac{1}{a}$$

(אנו מניחים ש־$a \neq 0$). נניח כעת ש־f גזירה ב־x_0, ובפרט רציפה שם. אז כאשר $x_1 \to x_0$ מתקיים $f(x_1) \to f(x_0)$ ונצפה שהמנות למעלה שואפות לשיפוע של

המשיקים המתאימים, כלומר לשיפוע של המשיקים לגרף של f ב־x_0 ושל הגרף של f^{-1} ב־$f(x_0)$. לכן נצפה שהנגזרת של f^{-1} ב־$f(x_0)$ תהיה שווה ל־ $\frac{1}{f'(x_0)}$.

נעבור לדיון פורמלי יותר. נסמן $y_0 = f(x_0)$ ונניח ש־ $f'(x_0) \neq 0$. לפי הדיון למעלה אנו מצפים שהנגזרת של f^{-1} ב־y_0 תהיה קיימת ושיתקיים

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}$$

ואמנם, אילו ידענו ש־f^{-1} גזירה ב־y_0 היינו יכולים לחשב את הנגזרת בעזרת כלל השרשרת באופן הבא: נסמן $g = f^{-1}$. השוויון $g \circ f(x) = x$ מתקיים לכל $x \in I$ מהגדרת הפונקציה ההפוכה, וניתן לרשום שוויון זה גם כך: $g \circ f = \mathrm{id}_I$ כאשר $\mathrm{id}_I : I \to I$ היא פונקציית הזהות $\mathrm{id}_I(x) = x$. אם נגזור את $g \circ f$ ב־x_0 לפי כלל השרשרת נקבל

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0) = g'(y_0)f'(x_0)$$

ומצד שני לפי השוויון $g \circ f = \mathrm{id}_I$ מתקיים

$$(g \circ f)'(x_0) = \mathrm{id}'(x_0) = 1$$

על ידי השוואת הביטויים שקיבלנו והעברת אגפים מקבלים

$$(f^{-1})'(y_0) = g'(y_0) = \frac{1}{f'(x_0)}$$

בדיון האחרון חישבנו את $(f^{-1})'(y_0)$ תחת ההנחה ש־f^{-1} גזירה ב־y_0. כדי להוכיח את קיום הנגזרת אי אפשר להיפטר לגמרי מהנחות, אך אפשר להסתפק בהנחה לגבי הרציפות של f.

משפט 8.3.6 (כלל גזירה של פונקציה הפוכה) תהי f פונקציה ממשית הפיכה ו־f^{-1} הפונקציה ההפוכה לה. נניח ש־f גזירה ב־x_0 ו־ $f'(x_0) \neq 0$, ונניח בנוסף ש־f רציפה בסביבה מלאה של x_0. אז $g = f^{-1}$ גזירה בנקודה $y_0 = f(x_0)$ ומתקיים $g'(y_0) = \frac{1}{f'(x_0)}$.

הוכחה תהי V סביבה של x_0 שבה f רציפה. אז f חח"ע ב־V ורציפה שם, ולכן f^{-1} מוגדרת ורציפה בסביבה כלשהי U של y_0, ובפרט f^{-1} רציפה ב־y_0 (ראו משפט 7.11.6).

נסמן $g = f^{-1}$. לשם שינוי נחשב ישירות את g' מהגדרת הנגזרת:

$$\begin{aligned} g'(y_0) &= \lim_{t\to 0} \frac{g(y_0+t)-g(y_0)}{t} \\ &= \lim_{t\to 0}\left(\frac{t}{(g(y_0+t)-g(y_0))}\right)^{-1} \end{aligned}$$

מרציפות הפונקציה $\frac{1}{y}$ זה שווה ל־

$$= (\lim_{t\to 0} \frac{t}{g(y_0+t)-g(y_0)})^{-1}$$

בתנאי שהגבול בסוגריים קיים ושונה מאפס. מתכונות הפונקציה ההפוכה מתקיים $f(g(y))=y$ לכל y ולכן $t=f(g(y_0+t))-f(g(y_0))$. נציב זאת בביטוי האחרון ונקבל

$$\begin{aligned} &= (\lim_{t\to 0} \frac{f(g(y_0+t))-f(g(y_0))}{g(y_0+t)-g(y_0)})^{-1} \\ &= (\lim_{t\to 0} \frac{f(g(y_0+t))-f(x_0)}{g(y_0+t)-x_0})^{-1} \end{aligned}$$

(במעבר האחרון הצבנו את השוויון $g(y_0)=g(f(x_0))=x_0$, שגם־כן נובע מתכונות הפונקציה ההפוכה). נשים לב שמהרציפות של g נובע שכאשר $t\to 0$ הפונקציה $g(y_0+t)$ מתכנסת ל־$g(y_0)$, כלומר ל־x_0, לכן

$$= (\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0})^{-1}$$

(כאן השתמשנו במשפט 7.8.2, ויש לבדוק שתנאיו מתקיימים. ואמנם הפונקציה $g(y_0+t)$ מקבלת את הערך x_0 רק כאשר $t=0$, כי g חח"ע). אבל הביטוי האחרון הוא פשוט

$$= \frac{1}{f'(x_0)}$$

כנדרש (בשורה האחרונה השתמשנו בכך ש־ $f'(x_0)\neq 0$). ∎

מסקנה 8.3.7 אם f גזירה והפיכה בקטע פתוח I אז התמונה J היא קטע פתוח, f^{-1} גזירה ב־J ומתקיים $(f^{-1})'(y)=\frac{1}{f'(f^{-1}(y))}$.

דוגמאות

1. תהי $f(x)=x^2$ מוגדרת בקטע $[0,\infty)$. אז $f^{-1}(y)=\sqrt{y}$. מכיוון ש־ $f'(x)=2x$ נובע מהמשפט ש־

$$\begin{aligned} \frac{d}{dx}f^{-1}(y) &= (f^{-1})'(y) \\ &= \frac{1}{f'(f^{-1}(y))} \\ &= \frac{1}{2\sqrt{y}} \end{aligned}$$

2. תהי $f(x) = x^7 + 2x^5 + x^3 + 1$. קל לוודא ש־$f$ עולה ממש ולכן יש לה פונקציה הפוכה, שנסמנה g. גזירה מראה ש־ $f'(x) = 7x^6 + 10x^4 + 3x^2$ מתאפסת רק ב־0 (כי אם $x \neq 0$ כל המחוברים חיוביים). לכן g גזירה בכל נקודה y שבה $g(y) \neq 0$, דהיינו בכל נקודה $y \neq 0$, ואז מתקיים

$$g'(y) = \frac{1}{f'(g(y))} = \frac{1}{7g(y)^6 + 10g(y)^4 + 3g(y)^2}$$

בפרט מכיוון ש־ $f(1) = 5$, הרי $g(5) = 1$ ולכן $g'(5) = \frac{1}{7+10+3} = \frac{1}{20}$.

ללא הנחת הרציפות כלל המנה אינו נכון: יש דוגמאות לפונקציות f הפיכות וגזירות ב־x_0 כך ש־f^{-1} אינה גזירה ב־$f(x_0)$. בדוגמה כזו f כמובן רציפה ב־x_0 כי היא גזירה שם, אך אינה רציפה בסביבה של x_0.

נסיים בהערה על גזירה של פונקציות שאינן נתונות באופן מפורש אלא על ידי משוואה. פונקציות כאלה נקראות לפעמים **פונקציות סתומות** (implicit functions). נסביר את העניין בעזרת דוגמה: נניח שנתונה פונקציה u שגזירה בכל נקודה ונניח שאנו יודעים שהיא מקיימת את המשוואה $t^2 + u^2(t) = 1$ (כיוון שלא ביטאנו את u כפונקציה של t, אומרים שזו הגדרה סתומה של הפונקציה). ניתן לחשוב על שוויון זה כשוויון בין פונקציות: אם נגדיר $f(t) = t^2 + u^2(t)$ נקבל ש־f שווה לפונקציה קבועה. לכן $f'(t) = 0$ לכל t. מצד שני לפי כללי הגזירה, מתקיים

$$f'(t) = 2t + 2u(t) \cdot u'(t)$$

וקיבלנו את המשוואה $u(t) = -\frac{t}{u'(t)}$. אם למשל ידוע ש־ $u(\frac{1}{\sqrt{2}}) = \frac{1}{\sqrt{2}}$ נוכל להסיק ש־

$$u'(\frac{1}{\sqrt{2}}) = -\frac{1/\sqrt{2}}{1/\sqrt{2}} = -1$$

בדוגמה האחרונה היינו יכולים, על ידי העברת אגפים, למצוא נוסחה מפורשת עבור u. בדרך כלל בשאלות מעין אלה אי אפשר לקבל נוסחה פשוטה עבור u, ואז שיטה זו שימושית מאד.

תרגילים

1. מצאו את נקודות הגזירות של הפונקציות הבאות וגזרו אותן:
   (א) $f(x) = x + \sqrt{x}$.
   (ב) $f(x) = \frac{x+1}{x-1}$.
   (ג) $f(x) = \sqrt{x^2 + 1}$.
   (ד) $f(x) = \frac{1}{x+\sqrt{x^2-1}}$.

(ה) $\sqrt{x\sqrt{x\sqrt{x}}}$.

(ו) $(1+\sqrt{x})^{10}\cdot\sqrt{1+x}$.

2. הוכיחו באינדוקציה על n ובעזרת כלל המכפלה ש־ $\frac{d}{dx}x^n = nx^{n-1}$ לכל n טבעי.

3. תהי f גזירה ב־$\mathbb{R}$. התאימו לפונקציות בעמודה הימנית נגזרת מהעמודה השמאלית (לא לכל פונקציה יש בת־זוג!).

$f(x^2+1)$	$2f'(x)$
$f(x+1)$	$f'(x^2+1)$
$f(2x)$	$f'(x^2)+1$
$f(x)^2+1$	$2f'(2x)$
$(f(x)+1)^2$	$2xf'(x^2+1)$
$f(x/2)$	$2f(x)f'(x)$

4. הוכיחו את נוסחת הגזירה של לייבניץ: אם f,g גזירות n פעמים אז

$$(fg)^{(n)} = \sum_{k=0}^{n}\binom{n}{k}f^{(k)}g^{(n-k)}$$

(שימו לב לדמיון הרב לנוסחת הבינום).

5. מצאו נוסחה לנגזרת $(f_1\cdot f_2\cdot\ldots\cdot f_n)'$ של מכפלה של n פונקציות (כדאי תחילה לבדוק מה מקבלים עבור $n=3,4$).

6. יהיו f,g,h פונקציות גזירות. בטאו את הנגזרות של כל אחת מהפונקציות הבאות בעזרת f,g,h ונגזרותיהן (הניחו ש־ f הפיכה ו־f^{-1} גזירה):

(א) $h\circ g\circ f$.

(ב) $f^{-1}\circ g^{-1}$.

(ג) $f^{-1}\circ g\circ f$.

7. הוכיחו שאין פונקציות $f,g:\mathbb{R}\to\mathbb{R}$ שגזירות ב־0 כך ש־$f(x)=g(x)=0$ וגם $x=f(x)g(x)$.

8. (*) מצאו דוגמה לפונקציה חח"ע $f:[-1,1]\to\mathbb{R}$, הרציפה וגזירה ב־0 עם $f(0)=0$ ו־ $f'(0)\neq 0$, כך ש־ f^{-1} רציפה ב־0 אך אינה גזירה שם (שימו לב של־f חייבות להיות נקודות אי־רציפות בכל סביבה של 0, אחרת f^{-1} תהיה רציפה בסביבה של $f(0)$ וגזירה ב־0 לפי כלל המנה. כדאי לפתור קודם את תרגיל (8) בעמוד 286).

8.4 נגזרות הפונקציות האלמנטריות

הגיע הזמן לחשב כמה נגזרות. בעזרת חישוב ישיר ראינו ש־ $(x^n)' = nx^{n-1}$ (ראו סעיף 8.1, תרגיל (3) בעמוד 308 ותרגיל (2) בעמוד 324). מכאן

טענה 8.4.1 אם $p(x) = \sum_{k=0}^{d} a_k x^k$ פולינום אז

$$p'(x) = \sum_{k=0}^{d} k a_k x^{k-1} = \sum_{k=0}^{d-1} (k+1) a_{k+1} x^k$$

בפרט, הנגזרת של פולינום ממעלה $d \geq 1$ היא פולינום ממעלה $d-1$.

הוכחה נובע מהנוסחה לנגזרת של מונום ומכלל הסכום. ■

נעבור לפונקציות הלוגריתמיות, או ליתר דיוק לפונקציה $\ln$.

משפט 8.4.2 הפונקציה $\ln$ גזירה בכל תחום הגדרתה ומתקיים $\frac{d}{dx}\ln x = \frac{1}{x}$, ובאופן כללי יותר לכל $a > 0$ מתקיים $\frac{d}{dx}\log_a x = \frac{1}{x \ln a}$.

הוכחה נשתמש בעובדה ש־ $\ln(xy) = \ln x + \ln y$ ו־ $x \ln y = \ln(y^x)$ לכל $x, y > 0$ כדי לקבל שלכל $x_0 > 0$,

$$\frac{\ln(x_0 + h) - \ln x_0}{h} = \frac{1}{h}\ln(\frac{x_0 + h}{x_0}) = \ln\left((1 + \frac{1/x_0}{1/h})^{1/h}\right)$$

בפרק 5 הוכחנו שלכל סדרה $a_n \to \infty$ או $a_n \to -\infty$ מתקיים $(1 + \frac{y}{a_n})^{a_n} \to e^y$. לכן לכל סדרה $0 \neq h_n \to 0$ מתקיים

$$\lim_{n\to\infty}(1 + \frac{y}{1/h_n})^{1/h_n} = e^y$$

לכן לפי אפיון היינה לגבול,

$$\lim_{h\to 0}(1 + \frac{1/x_0}{1/h})^{1/h} = e^{1/x_0}$$

מכיוון ש־$\ln$ רציפה בנקודה e^{1/x_0} אנו מקבלים

$$\begin{aligned}\lim_{h\to 0}\frac{\ln(x_0 + h) - \ln x_0}{h} &= \lim_{h\to 0}\ln(1 + \frac{1/x_0}{1/h})^{1/h} \\ &= \ln e^{1/x_0} \\ &= \frac{1}{x_0}\end{aligned}$$

כפי שרצינו.

המסקנה לגבי $\log_a$ נובעת כעת מהעובדה ש־ $\log_a x = \frac{\ln x}{\ln a}$ ומכללי הגזירה. ■

מסקנה 8.4.3 הפונקציה e^x גזירה בכל תחום הגדרתה ומתקיים $\frac{d}{dx}e^x = e^x$. באופן כללי יותר לכל $a > 0$ מתקיים $\frac{d}{dx}a^x = a^x \cdot \ln a$.

הוכחה $\exp x = e^x$ גזירה לפי משפט 8.3.6 כי היא הפונקציה ההפוכה ל־$\ln$, ולפי אותו משפט,

$$\exp'(x) = \frac{1}{\ln'(e^x)} = \frac{1}{1/e^x} = e^x$$

כנדרש.

המסקנה עבור a^x נובעת מהזהות $a^x = e^{x \ln a}$ ומכללי הגזירה. ■

אנו מקבלים את ההכללה הבאה של המשפט על גזירת פולינומים:

מסקנה 8.4.4 לכל $\alpha \in \mathbb{R}$ מתקיים $\frac{d}{dx}x^\alpha = \alpha x^{\alpha-1}$ בתחום $(0,\infty)$.

הוכחה נשים לב שלכל $x > 0$ מתקיים $x^\alpha = \exp(\alpha \ln x)$ ונגזור לפי כלל השרשרת. נקבל

$$\frac{d}{dx}x^a = \frac{d}{dx}(e^{\alpha \ln x}) = e^{\alpha \ln x} \cdot \frac{d}{dx}(\alpha \ln x) = x^\alpha \cdot \alpha \cdot \frac{1}{x} = \alpha x^{\alpha-1}$$

כנדרש. ■

נותר לחשב את הנגזרות של הפונקציות הטריגונומטריות. נזכור שלפי התכונות של הפונקציות הטריגונומטריות בעמוד 227, לכל $x \in (0, \frac{\pi}{2})$ מתקיים

$$\sin x \le x \le \frac{\sin x}{\cos x}$$

נחלק ב־x ונעביר אגפים, ונקבל שבסביבה ימנית של 0 מתקיים

$$\cos x \le \frac{\sin x}{x} \le 1$$

ולכן מכלל הסנדוויץ' לגבולות,

$$\lim_{x\to 0+} \frac{\sin x}{x} = 1$$

הפונקציות x ו־$\sin x$ הן אי־זוגיות. לכן גם $\lim_{x\to 0-} \frac{\sin x}{x} = 1$, ומכאן נובע ש־$\lim_{x\to 0} \frac{\sin x}{x} = 1$. כעת

$$\lim_{h\to 0} \frac{\sin h - \sin 0}{h} = \lim_{h\to 0} \frac{\sin h}{h} = 1$$

ולכן $\sin$ גזירה ב־0 ומתקיים $\sin' 0 = 1$.

משפט 8.4.5 הפונקציות $\sin, \cos$ גזירות בכל תחום הגדרתן ומתקיים $\sin' x = \cos x$, $\cos' x = -\sin x$

הוכחה לכל $x \in \mathbb{R}$

$$\begin{aligned}\lim_{h\to 0}\frac{\sin(x+h)-\sin x}{h} &= \lim_{h\to 0}\frac{\sin x\cos h+\sin h\cos x-\sin x}{h}\\ &= \lim_{h\to 0}\frac{\sin h\cos x}{h}+\lim_{h\to 0}\frac{\sin x(\cos h-1)}{h}\end{aligned}$$

נחשב כל מחובר בנפרד. במחובר השמאלי $\cos x$ אינו תלוי ב־h, ולפי החישוב שקדם למשפט מתקיים

$$\lim_{h\to 0}\frac{\sin h\cos x}{h} = \cos x\lim_{h\to 0}\frac{\sin h}{h} = \cos x$$

באופן דומה במחובר הימני $\sin x$ קבוע ובעזרת הזהות $\cos h = 1-2\sin^2(h/2)$ נקבל

$$\begin{aligned}\lim_{h\to 0}\frac{\sin x(\cos h-1)}{h} &= \sin x\lim_{h\to 0}\frac{-2\sin^2(h/2)}{h}\\ &= -\frac{\sin x}{2}\lim_{h\to 0}\left(h\cdot(\frac{\sin(h/2)}{h/2})^2\right)\\ &= 0\end{aligned}$$

(העובדה ש־ $\lim_{h\to 0}\frac{\sin(h/2)}{h/2} = 1$ נובעת מהגבול $\lim_{h\to 0}\frac{\sin h}{h} = 1$). בסיכום קיבלנו

$$\lim_{h\to 0}\frac{\sin(x+h)-\sin x}{h} = \cos x$$

כפי שרצינו.

לגבי $\cos x$, מאחר ש־ $\cos x = \sin(x+\frac{\pi}{2})$ הרי שכלל השרשרת נותן

$$\cos'(x) = \sin'(x+\frac{\pi}{2}) = \cos(x+\frac{\pi}{2}) = -\sin x$$

כנדרש. ∎

מסקנה 8.4.6 $\tan, \mathrm{cotan}$ גזירות בכל תחום הגדרתן ומתקיים

$$\tan' x = \frac{1}{\cos^2 x} \qquad , \qquad \mathrm{cotan}' x = \frac{1}{\sin^2 x}$$

$\arcsin, \arccos$ גזירות ב־ $(-1,1)$ ומתקיים

$$\arcsin' x = \frac{1}{\sqrt{1-x^2}} \qquad , \qquad \arccos' x = \frac{\pi}{2}-\frac{1}{\sqrt{1-x^2}}$$

$\arctan, \text{arccotan}$ גזירות ב־$\mathbb{R}$ ומתקיים

$$\arctan' x = \frac{1}{1+x^2} \qquad , \qquad \text{arccotan}' x = \frac{\pi}{2} - \frac{1}{1+x^2}$$

הוכחה ההוכחות הן שימושים של כללי הגזירה וזהויות טריגונומטריות. מתקיים

$$\tan' x = (\frac{\sin}{\cos})'(x) = \frac{\sin' x \cos x - \sin x \cos' x}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}$$

כי $\sin^2 + \cos^2 \equiv 1$. כמו־כן לפי הכלל לגזירת פונקציה הפוכה,

$$\arcsin' x = \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-x^2}}$$

כאשר השוויון האחרון נובע מהנוסחה ל־ $\cos(\arcsin x)$ מעמוד 288. ההוכחה לגבי $\arccos$ נובעת מהשוויון $\arccos x = \frac{\pi}{2} - \arcsin x$. אנו מקבלים

$$\arctan' x = \frac{1}{\tan'(\arctan x)} = \cos^2(\arctan x) = \frac{1}{1+\tan^2(\arctan x)} = \frac{1}{1+x^2}$$

(השתמשנו בזהות $\cos^2 = \frac{1}{1+\tan^2}$, הוכיחו אותה!). לבסוף, $\text{arccotan} = \frac{\pi}{2} - \arctan$ והנוסחה ל־$\text{arccotan}'$ נובעת מכללי הגזירה. ■

משפט 8.4.7 הפונקציות האלמנטריות גזירות בכל נקודה פנימית בתחום הגדרתן.

הוכחה ראינו שהפולינומים, פונקציות החזקה, הפונקציות המעריכיות, הלוגריתמים, הפונקציות הטריגונומטריות והפונקציות הטריגונומטריות ההפוכות מקיימות זאת. ולכן לפי כללי האריתמטיקה כל סכום, הפרש, מכפלה, מנה או הרכבה שלהן מקיימים זאת. מכיוון שהפונקציות האלמנטריות הוגדרו להיות כל מה שמתקבל מהן על ידי פעולות אלה, סיימנו. ■

לפעמים מתעורר הצורך לגזור פונקציות מהצורה f^g כאשר f, g פונקציות. לכאורה f^g לא התקבלה מ־f, g על ידי פעולות שאנו יודעים לטפל בהן אך למעשה אפשר לרשום אותו כ־ $e^{g \ln f}$ (זהו אותו תכסיס שניצלנו כשחישבנו את הנגזרת של x^α) ולכן נוכל לגזור את f^g בעזרת הכללים הרגילים. נקבל:

$$\begin{aligned} (f^g)'(x) &= e^{g(x)\ln f(x)}(g \ln f)'(x) \\ &= f(x)^{g(x)} \cdot (g'(x) \ln f(x) + g(x)\frac{f'(x)}{f(x)}) \end{aligned}$$

למשל, לכל $x > 0$ מתקיים

$$\frac{d}{dx}(x^x) = x^x(1 \cdot \ln x + x \cdot \frac{1}{x}) = x^x(\ln x + 1)$$

תרגילים

1. מצאו באילו נקודות יש לפונקציות הבאות נגזרת, וגזרו אותן לפי כללי הגזירה:
   (א) $e^x \cdot \cos(x \cdot e^x)$.
   (ב) $\frac{x^2+\cos(x^2+\cdot e^x)}{2x+3}$.
   (ג) $a^{x+\cos x}$.
   (ד) $\ln(x+\sqrt{x^2+1})$.
   (ה) $\frac{1}{\sin x}$.
   (ו) $(\sqrt{1-x^{2/3}})^5$.
   (ז) $\arcsin e^{\sin\sqrt{x}}$.
   (ח) $x^{\sin^2 x}$.
   (ט) $(\ln x)^{\ln x}$.
   (י) $x^{(x^x)}$.
2. הוכיחו שאם p היא פולינום ממעלה d אז היא גזירה מכל סדר ו־ $p^{(d+1)} \equiv 0$.
3. אם p פולינום, הראו שיש פולינום q כך ש־ $q' = p$.
4. נניח ש־$a(x)$ היא פונקציה גזירה המקיימת $a(x)^x = x^{a(x)}$ בכל נקודה, ו־ $a(2) = 4$ (שימו לב שבאמת מתקיים $2^4 = 4^2$). חשבו את $a'(2)$.

8.5 פונקציות גזירות בקטע

כאשר פונקציה גזירה בכל נקודה בקטע, פונקציית הנגזרת מספקת מידע רב על התנהגות הפונקציה. אפשר בעזרתה לזהות תחומים בהם הפונקציה עולה או יורדת ולאתר נקודות מקסימום ומינימום של הפונקציה. יש לכך יישומים מעשיים שנדון בהם בסעיף הבא. כאן נכין את הבסיס התאורטי.

הגדרה 8.5.1 נאמר שפונקציה f **גזירה בקטע** I אם f רציפה ב־I וגזירה בכל נקודה פנימית של I.

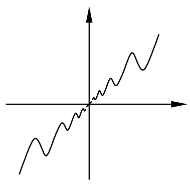
שימו לב שאף שרציפות ב־I גורר רציפות חד־צדדית בקצוות של I, במידה וקיימות כאלה, גזירות ב־I אינו מבטיח גזירות חד־צדדית בנקודות הקצה.

הניתוח של פונקציות גזירות בקטע מתחיל דווקא בגרסה מקומית של תכונת העלייה והירידה של הפונקציה:

הגדרה 8.5.2 תהי f פונקציה המוגדרת בסביבה מלאה של x_0. נאמר ש־f **עולה בנקודה** x_0 אם יש סביבה V של x_0 כך שלכל $x \in V$, אם $x > x_0$ אז $f(x) > f(x_0)$ ואם $x < x_0$ אז $f(x) < f(x_0)$. באופן דומה, נאמר ש־f **יורדת בנקודה** x_0 אם יש

סביבה V של x_0 כך שלכל $x \in V$, אם $x > x_0$ אז $f(x) < f(x_0)$ ואם $x < x_0$ אז $f(x) > f(x_0)$.

מגדירים עליה וירידה חד־צדדית בנקודה באותו אופן.



איור 8.5.1 פונקציה שעולה ב־0 אך אינה עולה באף סביבה של 0

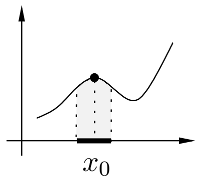
עליה וירידה בנקודה הן תכונות אינפיניטסימליות, וחשוב לשים לב שהן אינן גוררת עליה או ירידה במובן הרגיל באף סביבה של הנקודה. באיור 8.5.1 מופיעה פונקציה שעולה באפס אך אינה עולה באף סביבה של אפס. תמצאו דיון פורמלי יותר בתרגיל (1) בסוף הסעיף.

הקשר בין עליה וירידה בנקודה לבין הנגזרת בנקודה נתון על ידי הלמה הבאה:

למה 8.5.3 אם $f'(x_0) > 0$ אז f עולה בנקודה x_0 ואם $f'(x_0) < 0$ אז f יורדת ב־x_0.

הוכחה נניח $f'(x_0) > 0$, דהיינו $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} > 0$. אז יש סביבה V של x_0 שבה המנה בגבול חיובית ממש. אם $x \in V$ ו־ $x > x_0$ המכנה חיובי, ונובע של־x כזה מתקיים $f(x) - f(x_0) > 0$, דהיינו $f(x) > f(x_0)$. מאידך אם $x \in V$ ו־ $x < x_0$ אז המכנה שלילי ולכן גם המונה שלילי, ולכן $f(x) < f(x_0)$. בסיכום, קיבלנו ש־f עולה בנקודה. ההוכחה של המקרה השני דומה. ■

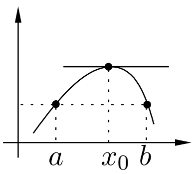
נעבור לפונקציות גזירות בקטע, ונתחיל עם השאלה כיצד לאתר מקסימום של פונקציה בקטע. משפט המקסימום של ויירשטראס מבטיח שלפונקציה רציפה בקטע סגור יש מקסימום אך אינו נותן כלים למצוא אותו. כאשר הפונקציה גזירה המצב טוב יותר.



איור 8.5.2 פונקציה עם מקסימום מקומי ב־x_0

הגדרה 8.5.4 תהי f פונקציה המוגדרת בסביבה מלאה של x_0. נאמר של־f יש **מקסימום מקומי** (local maximum) ב־x_0 אם יש סביבה מלאה של x_0 שבה x_0 היא נקודת מקסימום של הפונקציה. **מינימום מקומי** (local minimum) מוגדר באופן דומה. אם x_0 היא נקודת מקסימום מקומי או נקודת מינימום מקומי אומרים ש־x_0 היא **נקודת קיצון** (local extrimum) של f.

שימו לב שנקודת מקסימום של פונקציה f היא מקסימום מקומי אמ"מ f מוגדרת בסביבה מלאה של x_0, וכך לגבי מינימום.



איור 8.5.3 בנקודת מקסימום מקומי המשיק מאוזן

משפט 8.5.5 (משפט פרמה)[4] אם ל־f יש נקודת קיצון ב־x_0, ואם f גזירה ב־x_0, אז $f'(x_0) = 0$.

הוכחה נניח למשל ש־x_0 נקודת מקסימום מקומי. אם $f'(x_0) > 0$ אז f עולה ב־x_0, וזה סותר את העובדה ש־x_0 נקודת מקסימום מקומי (כי זה מבטיח קיום נקודות $x > x_0$ קרובות כרצוננו ל־x_0 בהן $f(x) > f(x_0)$). מאידך אם $f'(x_0) < 0$

[4] Pierre de Fermat, 1601-1665.

אז f יורדת ב־x_0 וגם זה סותר את העובדה ש־x_0 נקודת מקסימום מקומי. לכן $f'(x_0) = 0$. ■

משפט פרמה אינו אפיון, וייתכן שהנגזרת תתאפס ב־x_0 ובכל־זאת x_0 אינה נקודת קיצון. למשל, אם $f(x) = x^3$ אז $f'(x) = 3x^2$ ולכן $f'(0) = 0$ אבל f עולה ממש בכל הישר וממילא אין לה נקודות קיצון.

משפט 8.5.6 (משפט רול[5]) אם f רציפה בקטע $[a,b]$ וגזירה בקטע (a,b), ואם $f(a) = f(b)$, אז קיים $x_0 \in (a,b)$ כך ש־ $f'(x_0) = 0$.

הוכחה f רציפה בקטע הסגור $[a,b]$ ולכן יש לה מקסימום ומינימום בקטע (משפט המקסימום של ויירשטראס). אם נקודת המקסימום או המינימום מתקבלים בנקודה פנימית $x_0 \in (a,b)$ של הקטע $[a,b]$ אז יש ל־f נקודת קיצון ב־x_0. מכיוון שנתון שהיא גזירה ב־(a,b) היא גזירה בפרט ב־x_0, וממשפט פרמה אנו מקבלים $f'(x_0) = 0$.

אחרת המינימום והמקסימום של f מתקבלים בקצוות הקטע. מכיוון שהערכים של f בשני הקצוות שווים אנו מסיקים שהמקסימום והמינימום של f ב־$[a,b]$ שווים, ולכן הפונקציה קבועה. אבל הנגזרת של פונקציה קבועה בקטע היא 0 בכל נקודה, כלומר: $f'(x) = 0$ לכל $x \in (a,b)$, וסיימנו. ■

אי אפשר להחליש את דרישת הגזירות במשפט פרמה. למשל, $f(x) = |x|$ רציפה ב־$[-1,1]$ וגזירה שם למעט בנקודה אחת, ומקיימת $f(1) - f(-1)$. למרות זאת, אין נקודה $x \in [-1,1]$ עבורו $f'(x) = 0$.

דוגמאות

1. אחד השימושים העיקריים של משפט רול הוא לאיתור נקודות קיצון ונקודות מינימום ומקסימום של פונקציה. נדון בכך ביתר הרחבה בסעיף הבא ובינתיים נסתפק בדוגמה פשוטה: חישוב נקודות המינימום של פרבולה כללית. תהי $f(x) = ax^2 + bx + c$ עם $a > 0$. אנו יודעים שהגבול של f ב־$\pm\infty$ הוא ∞ ולכן יש לו מינימום ב־$\mathbb{R}$ (ראו תרגיל (15) בעמוד 296). מינימום זה הוא בהכרח נקודת קיצון ומכיוון ש־f מוגדרת וגזירה בכל נקודה ב־$\mathbb{R}$, לפי משפט פרמה בנקודת המינימום הנגזרת מתאפסת. לכל x מתקיים

$$f'(x) = 2ax + b$$

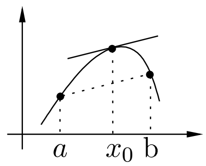
ולכן הנקודה היחידה המקיימת $f'(x) = 0$ היא $x = -\frac{b}{2a}$. מכאן שהמינימום מתקבל בהכרח בנקודה $-\frac{b}{2a}$.

[5] Michel Rolle, 1652-1719.

2. נוכיח שלפולינום ממעלה $d > 0$ יש לכל היותר d שורשים (הוכחה אחרת ניתנה בתרגיל (5) בעמוד 229). ההוכחה היא באינדוקציה על d. פולינום ממעלה $d = 1$ הוא מהצורה $p(x) = ax + b$ ו־$a \neq 0$, ובוודאי יש לו רק שורש אחד.

נניח שהוכחנו את הטענה עבור d. יהי $p(x) = \sum_{k=0}^{d+1} a_k x^k$ פולינום ממעלה $d+1$. יהיו $x_1 < x_2 < \ldots < x_N$ שורשים של p. עלינו להראות ש־$N \leq d+1$. ואמנם נתון ש־ $p(x_i) = 0$ לכל $i = 1, \ldots, N$, ולכן ממשפט רול בין כל שני x_i־ים הנגזרת מתאפסת, דהיינו יש נקודות $y_1 < \ldots < y_{N-1}$ כך ש־ $x_1 < y_i < x_{i+1}$ ו־ $p'(y_i) = 0$. אבל p' פולינום ממעלה d ולכן מהנחת האינדוקציה יש לו לכל היותר d שורשים. לכן $N - 1 \leq d$, כלומר $N \leq d+1$, כפי שרצינו.

נעיר שפולינומים ממעלה 0 הם פולינומים קבועים, ולכן גם־כן מקיימים את הטענה, למעט פולינום האפס.



איור 8.5.4 מיתר ונקודה בגרף שבה המשיק מקביל למיתר

אפשר לפרש את משפט רול באופן הבא: אם יש לגרף של f מיתר מאוזן אז אפשר להזיז את המיתר עד שישיק לפונקציה. בנקודת ההשקה הנגזרת כמובן מתאפסת כי לישר המשיק יש שיפוע אפס. תופעה דומה מתקיימת עבור מיתרים שאינם מאוזנים: אם יש מיתר עם שיפוע a אז אפשר להזיז אותו עד שישיק לגרף. בנקודת ההשקה הנגזרת תהיה שווה לשיפוע המשיק, כלומר ל־a. באופן מדויק,

משפט 8.5.7 (משפט לגרנג')[6] אם f גזירה ב־$[a,b]$ אז יש נקודה $x_0 \in (a,b)$ כך ש־

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}$$

הוכחה הרעיון הוא "להטות" את הפונקציה עד שהערכים שלה ב־a ו־b יהיו שווים ולהפעיל את משפט רול.

תהי ℓ הפונקציה המתארת את הישר העובר דרך הנקודות $(a, f(a))$ ו־ $(b, f(b))$ שעל גרף של f. הפונקציה ℓ מקיימת $\ell(a) = f(a)$ ו־ $\ell(b) = f(b)$ והשיפוע שלה הוא

$$\ell' \equiv \frac{f(b) - f(a)}{b - a}$$

נתבונן בפונקציה $g = f - \ell$ המוגדרת ב־$[a,b]$. אז g רציפה ב־$[a,b]$ וגזירה ב־(a,b) (למה?), ומתקיים $g(a) = g(b) = 0$. ממשפט רול מקבלים שיש נקודה $x_0 \in (a,b)$ שבה $g'(x_0) = 0$. אבל

$$g'(x) = f'(x) - \ell'(x)$$

[6] Joseph Lagrange, 1736־1813.

לכל $x \in (a,b)$ ולכן

$$f'(x_0) = g'(x_0) + \ell'(x_0) = 0 + \frac{f(b)-f(a)}{b-a} = \frac{f(b)-f(a)}{b-a}$$

כפי שרצינו. ■

דוגמאות

1. נראה שלכל $x', x'' \in \mathbb{R}$ מתקיים $|\sin x' - \sin x''| \leq |x' - x''|$. אם $x' = x''$ זה ברור. אחרת, נניח למשל ש־ $x' < x''$. אז ממשפט לגרנג' יש $c \in (x', x'')$ כך ש־

$$\frac{\sin x'' - \sin x'}{x'' - x'} = \sin' c = \cos c$$

הפעלת ערך מוחלט על שני הצדדים של השוויון נותנת

$$\frac{|\sin x'' - \sin x'|}{|x'' - x'|} = |\cos c| \leq 1$$

כנדרש.

2. ניתן הוכחה נוספת לכך שהטור $\sum_{k=0}^{\infty} \frac{1}{n^\alpha}$ מתכנס לכל $\alpha > 1$. ההוכחה מבוססת על ההערכה הבאה: לכל $n \in \mathbb{N}$ יש מספר $t \in (n, n+1)$ כך ש־

$$\frac{1}{(n+1)^{\alpha-1}} - \frac{1}{n^{\alpha-1}} = \frac{-\alpha+1}{t^\alpha}$$

משפט לגרנג' מבטיח את קיומו של t כזה, שכן עבור הפונקציה $f(x) = x^{1-\alpha}$ אגף שמאל הוא $\frac{f(n+1)-f(n)}{(n+1)-n}$ ואגף ימין הוא $f'(t)$. מכיוון ש־ $t \geq n$ אנו קבלים ש־

$$\frac{1}{(n+1)^{\alpha-1}} - \frac{1}{n^{\alpha-1}} \leq \frac{-\alpha+1}{n^\alpha}$$

לאחר העברת אגפים אנו מקבלים ש־

$$\frac{1}{n^\alpha} \leq \frac{1}{\alpha-1}\left(\frac{1}{n^{\alpha-1}} - \frac{1}{(n+1)^{\alpha-1}}\right)$$

(השתמשנו כאן בעובדה ש־ $\alpha - 1 > 0$). מהאי־שוויון הזה נובע שהטור החיובי $\sum_{n=1}^{\infty} \frac{1}{n^a}$ נשלט על ידי הטור

$$\frac{1}{\alpha-1} \sum_{n=1}^{\infty} \left(\frac{1}{n^{\alpha-1}} - \frac{1}{(n+1)^{\alpha-1}}\right)$$

אבל הטור המופיע למעלה הוא טור טלסקופי: לכל N מתקיים

$$\sum_{n=1}^{N}(\frac{1}{n^{\alpha-1}}-\frac{1}{(n+1)^{\alpha-1}})=(1-\frac{1}{N^{\alpha-1}})\xrightarrow{N\to\infty}\frac{1}{\alpha-1}$$

(בגבול האחרון השתמשנו שוב בכך ש־ $\alpha>1$). לכן גם הטור המקורי מתכנס, כפי שרצינו.

שימו לב לדמיון בין הוכחה זו להוכחה ש־ $\sum\frac{1}{n^2}$ מתכנס: אז השווינו את הטור לטור $\sum\frac{1}{n(n+1)}$ והשתמשנו בזהות $\frac{1}{n(n+1)}=\frac{1}{n}-\frac{1}{n+1}$ כדי להוכיח שהטור האחרון מתכנס.

משפט לגרנג' הוא בעל חשיבות רבה כי הוא מקשר בין תכונות של ערכי הפונקציה לתכונות של הנגזרת. למשל:

מסקנה 8.5.8 אם f גזירה בקטע $[a,b]$ ו־ $f'\equiv 0$ ב־(a,b) אז f קבועה ב־$[a,b]$.

הוכחה יהיו $x',x''\in[a,b]$ זוג נקודות שרירותי. אם $x'\neq x''$ אז יש נקודה c בין x' ל־x'' (וממילא $c\in(a,b)$) כך ש־ $f(x'')-f(x')=(x''-x')\cdot f'(c)$ (זהו משפט לגרנג' לאחר העברת אגפים). אבל $f'(c)=0$ ולכן $f(x'')-f(x')=0$, כנדרש. ■

דוגמאות

1. נראה שאם $f:\mathbb{R}\to\mathbb{R}$ גזירה ומקיימת $f'(x)=f(x)$ לכל $x\in\mathbb{R}$ אז יש קבוע c כך ש־ $f(x)=c\cdot e^x$. משמעות הדבר שהתכונה $f'=f$ ו־ $f(0)=1$ מאפיינת לחלוטין את הפונקציה e^x.

   תהי f פונקציה כך ש־ $f'=f$. נסמן $g(x)=e^x$, כך ש־ $g'=g$ ומתקיים $g(x)\neq 0$ לכל x. נגדיר $h=\frac{f}{g}$: אנו רוצים להראות ש־h קבועה. לכן די לגזור אותה ולבדוק שהנגזרת היא 0. ואמנם

$$h'(x)=\frac{f'(x)g(x)-f(x)g'(x)}{g^2(x)}=\frac{f(x)g(x)-f(x)g(x)}{g^2(x)}=0$$

   כפי שרצינו.

2. נראה שאם $f:\mathbb{R}\to\mathbb{R}$ ואם f' הוא פולינום אז f פולינום. לפי תרגיל (3) בעמוד 329, יש פולינום p כך ש־ $p'=f'$. לכן $(f-p)'=f'-f'\equiv 0$. מכאן שההפרש $f-p$ קבוע, דהיינו יש קבוע c כך ש־ $f=p+c$. אבל פירוש הדבר ש־f פולינום.

באופן אינטואיטיבי ברור שאם בקטע מסוים כל הישרים המשיקים לפונקציה עולים, אז גם הפונקציה עולה. במילים אחרות, אם הנגזרת חיובית הפונקציה עולה. זו אחת המסקנות החשובות ממשפט לגרנג':

מסקנה 8.5.9 תהי f פונקציה גזירה בקטע $[a,b]$. אם $f' \geq 0$ ב־(a,b) אז f עולה, ואם $f' \leq 0$ ב־(a,b) אז f יורדת. אם הנגזרת מקיימת $f' > 0$ או $f' < 0$ ב־(a,b) אז f עולה ממש או יורדת ממש ב־$[a,b]$, בהתאמה.

הוכחה נניח למשל ש־$f' \geq 0$ ב־(a,b). לכל $x' < x''$ ב־$[a,b]$ הפונקציה גזירה ב־$[x',x'']$ ולפי משפט לגרנג' מתקיים

$$f(x'') - f(x') = f'(x_0)(x'' - x') \geq 0$$

וקיבלנו ש־f עולה. אם $f' > 0$ ב־(a,b) אז האי־שוויון למעלה הוא חזק ו־f עולה ממש. המקרה של נגזרת שלילית מוכח באופן דומה. ■

למעשה, התנאי $f' \geq 0$ הוא תנאי הכרחי ומספיק לעליה חלשה של f בקטע. זה תנאי מספיק לפי המסקנה האחרונה, והוא הכרחי כי אם f עולה חלש בקטע לא תתכן נקודה שבה הנגזרת שלילית, כי זה היה גורר ירידה בנקודה, וזה לא ייתכן.

מצד שני חיוביות ממש של הנגזרת אינה הכרחית כדי שפונקציה תהיה עולה ממש. הדוגמה כרגיל היא הפונקציה $f(x) = x^3$, שעולה ממש בכל הישר אבל $f'(0) = 0$.

דוגמאות

1. נראה שאם p פולינום ו־$\deg p \geq 2$ אז יש מספר r כך ש־p מונוטונית בקרן (r,∞). ואמנם p' היא פולינום ממעלה לפחות 1 ולכן $\lim_{x\to\infty} p' = \pm\infty$. לכן יש קרן (r,∞) שבה p' אינה מתאפסת ובעלת סימן קבוע. לפי המסקנה האחרונה פירוש הדבר ש־p מונוטונית באותה קרן.
2. נוכיח את האי־שוויון $e^x \geq 1 + x$ (הוכחנו אותו בעבר בשיטות אחרות. ראו תרגיל (7) בעמוד 157). נתבונן בהפרש $f(x) = e^x - 1 - x$ ונראה שהוא אי־שלילי בכל נקודה. נשים לב ש־ $f(0) = 0$, ולכן די להראות ש־f עולה בקטע $[0,\infty)$ ויורדת בקטע $(-\infty,0]$. לשם כך די להראות ש־ $f'(x) > 0$ אם $x > 0$ ו־ $f'(x) < 0$ אם $x < 0$. ואמנם,

$$f'(x) = e^x - 1$$

והאי־שוויונות נובעים מתכונות החזקה.

למשפט לגרנג' יש גרסה חזקה יותר שנזדקק לה בהמשך. נקדים לדיון הפורמלי דיון גאומטרי. לפי משפט לגרנג', אם לפונקציה יש מיתר עם שיפוע a אז אפשר להזיז את המיתר עד שישיק לגרף הפונקציה, כלומר יש נקודה בגרף שבה הנגזרת היא a. אפשר להכליל זאת לעקומות במישור שאינם גרפים של פונקציות. לצורך העניין, עקומה היא קבוצת נקודות U במישור שמתקבלת משתי פונקציות $f,g : [a,b] \to \mathbb{R}$ על ידי

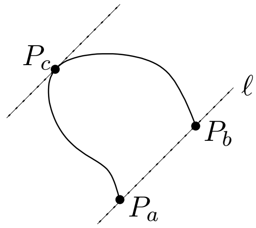
$$U = \{(g(x), f(x)) \ : \ x \in [a,b]\}$$

למשל מעגל היחידה הוא קבוצת הנקודות במישור שמרחקן מהראשית הוא 1. זו אינה אלא העקומה $\{(\cos x, \sin x) \ : \ x \in [0, 2\pi]\}$, אשר מתאימה לפונקציות $\sin, \cos : [0, 2\pi] \to \mathbb{R}$.

אם כן תהי $U = \{(g(x), f(x)) \ : \ x \in [a, b]\}$ עקומה במישור ויהיו

$$P_a = (g(a), f(a)) \, , \ P_b = (g(b), f(b))$$

הנקודות בעקומה המתאימות ל־a, b. נניח ש־$g(a) \neq g(b)$, כך ש־P_a, P_b נבדלות לפחות בקואורדינטת x שלהן. עבור $x \in [a, b]$ נסמן גם ב־$P_x = (g(x), f(x))$ את הנקודה בעקומה המתאימה ל־x.



איור 8.5.5 מיתר ומשיק המקביל לו

יהי ℓ הישר העובר דרך P_a, P_b, כך שהשיפוע של ℓ הוא $\frac{f(b)-f(a)}{g(b)-g(a)}$. מבחינה גאומטרית ברור שאם לעקומה U יש "משיק" בכל נקודה אפשר להזיז את הישר ℓ בצורה מקבילה לעצמה (כלומר בלי לסובב) עד שהוא ישיק לעקומה U בנקודה P_c כלשהי. בפרט, שיפוע המשיק לעקומה בנקודה P_c יהיה שווה לשיפוע של ℓ, כלומר ל־$\frac{f(b)-f(a)}{g(b)-g(a)}$. על מנת לחשב את השיפוע של העקומה בנקודה P_c נפעל כמו במקרה של גרף, ונחשב אותו בתור הגבול של השיפועים של המיתרים בין P_c לנקודות P_x כאשר x שואף ל־c. השיפוע של הישר בין P_c ל־P_{c+h} עבור $h \neq 0$ הוא

$$\frac{f(c+h) - f(c)}{g(c+h) - g(c)} = \frac{f(c+h) - f(c)}{h} \cdot \frac{h}{g(c+h) - f(c)}$$

ולכן בהנחה ש־f, g גזירות מקבלים שהשיפוע ב־P_c הוא

$$\lim_{h \to 0} \frac{f(c+h) - f(c)}{g(c+h) - g(c)} = \frac{f'(c)}{g'(c)}$$

ומכיוון שהשיפוע שווה לשיפוע של ℓ, קיבלנו שקיימת נקודה $c \in [a, b]$ כך ש־

$$\frac{g'(c)}{f'(c)} = \frac{g(b) - g(a)}{f(b) - f(a)}$$

עד כאן המוטיבציה שמאחורי ניסוח המשפט, ונעבור לדון בהוכחה שלו. לשם כך נזדקק למעט ידע מתורת הגאומטריה הווקטורית של המישור. מי שלא למד מעולם נושא זה יכול לדלג ישירות למשפט 8.5.10 ולהוכחה שלו, שהיא הוכחה מלאה ואינה תלויה בדיון הבא, שנועד רק לתת מוטיבציה להוכחה.

כאמור, אנו מחפשים נקודה c כך שהמשיק לעקומה ב־P_c מקביל למיתר בין P_a ל־P_b, וכמו במשפט רול ננסה לבחור את c כך שהמרחק של P_c מהישר ℓ העובר דרך P_a, P_b יהיה מקסימלי (או מינימלי). ניעזר בעובדה הגאומטרית הבאה: אם $\overrightarrow{u} = (u_1, u_2) \, , \ \overrightarrow{v} = (v_1, v_2)$ הם וקטורים אז המרחק בין $\overrightarrow{u}$ לישר הנקבע על ידי $\overrightarrow{v}$ שווה ל

$$\frac{\overrightarrow{u} \cdot \overrightarrow{v}^{\perp}}{\sqrt{\overrightarrow{v} \cdot \overrightarrow{v}}}$$

כאשר $\vec{v}^{\perp} = (-v_2, v_1)$ הוא וקטור הניצב ל־$\vec{v}$ ו־$\cdot$ מציין את המכפלה הפנימית בין וקטורים, דהיינו הפעולה $\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2$. לכן המרחק של נקודה $\vec{u} = (u_1, u_2)$ מהישר העובר דרך נקודות $\vec{v} = (v_1, v_2)\,,\ \vec{w} = (w_1, w_2)$ שווה למרחק בין הווקטורים $\vec{u} - \vec{v}$ לישר שנקבע על ידי $w - v$, דהיינו ל־

$$\begin{aligned}\frac{(\vec{u}-\vec{v})\cdot(\vec{w}-\vec{v})^{\perp}}{\sqrt{(\vec{w}-\vec{v})\cdot(\vec{w}-\vec{v})}} &= \frac{(u_1-v_1, u_2-v_2)\cdot(w_1-v_1, w_2-v_2)}{\sqrt{(\vec{w}-\vec{v})\cdot(\vec{w}-\vec{v})}} \\ &= \alpha\cdot((u_1-v_1)(w_1-v_1)+(u_2-v_2)(w_2-v_2))\end{aligned}$$

כאשר α הוא הקבוע $\alpha = (\sqrt{(\vec{w}-\vec{v})\cdot(\vec{w}-\vec{v})})^{-1}$. אם נציב כאן $\vec{u} = P_x$ ו־$\vec{v} = P_a\,,\ \vec{w} = P_b$ יוצא שהמרחק של הנקודה P_x מהישר ℓ העובר דרך P_a, P_b הוא

$$D(x) = \alpha\cdot((g(x)-g(a))(f(a)-f(b)) + (f(x)-f(a))(g(b)-g(a)))$$

כאשר α קבוע שתלוי ב־P_a, P_b אבל אינו תלוי ב־x. אנו מעוניינים למצוא x כך ש־$D(x)$ מקסימלי או מינימלי ולכן נחפש x שמאפס את הנגזרת שלו (המניע לכך הוא משפט פרמה). זה יהיה ה־c המבוקש.

נוכל כעת לעבור לדיון פורמלי לגמרי שלא מערב כלל גאומטריה:

משפט 8.5.10 (משפט ערך הביניים של קושי) יהיו $f, g : [a,b] \to \mathbb{R}$ רציפות וגזירות ב־(a,b). נניח ש־ g' לא מתאפסת ב־(a,b), ובפרט $g(a) \neq g(b)$. אז קיים $c \in (a,b)$ כך ש־

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}$$

הוכחה תהי

$$D(x) = (g(x)-g(a))(f(a)-f(b)) + (f(x)-f(a))(g(b)-g(a))$$

(בדיון הקודם הופיע גם קבוע בהגדרה של D. השמטנו אותו כי הוא אינו משפיע על החישוב). מכיוון ש־ f, g רציפות ב־$[a,b]$ וגזירות ב־(a,b) כך גם D. כמו־כן $D(a) = D(b) = 0$ (בדקו!). לכן לפי משפט רול קיימת נקודה $c \in (a,b)$ שבה $D'(c) = 0$. מהנוסחה עבור D רואים ש־

$$D'(x) = g'(x)(f(a)-f(b) + f'(x)(g(b)-g(a))$$

ולכן $D'(c) = 0$ גורר

$$0 = g'(c)(f(a)-f(b)) + f'(c)(g(b)-g(a))$$

לפי ההנחה ש־g' לא מתאפסת ב־(a,b) ברור ש־ $g'(c) \neq 0$, ויתר על כן, לפי משפט רול (משפט 8.5.6), גם $g(b) - g(a) \neq 0$. לכן השוויון המבוקש מתקבל מהקודם על־ידי העברת אגפים וחלוקה ב־$g'(c)$ וב־ $g(b) - g(a)$. ■

נסיים בכמה תוצאות על התכונות של פונקציית הנגזרת עצמה. אם f גזירה בקטע אז פונקציית הנגזרת f' אינה חייבת להיות פונקציה רציפה. לדוגמה, תהי

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

לכל $x_0 \neq 0$ הפונקציה f מוגדרת בסביבה של x_0 על־ידי הנוסחה $f(x) = x^2 \sin \frac{1}{x}$ ולכן גזירה ב־x_0 ונגזרתה היא $f'(x_0) = 2x_0 \sin \frac{1}{x_0} - \cos \frac{1}{x_0}$ (בדקו!). כמו־כן מכיוון ש־ $|f(x)| \leq x^2$ מתקיים

$$|\frac{f(0+h) - f(0)}{h}| = |\frac{f(h)}{h}| \leq |h|$$

ועל־ידי השאפת h ל־0 אנו מקבלים ש־f גזירה ב־0 ומתקיים $f'(0) = 0$. קיבלנו בסיכום ש־f גזירה בכל הישר ומתקיים

$$f'(x) = \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

מכיוון של־$2x \sin \frac{1}{x}$ יש גבול ב־0 ול־$\cos \frac{1}{x}$ אין, אנו רואים שאף ש־f גזירה בכל נקודה, הנגזרת אינה בעלת גבול ב־0 וממילא f' אינה רציפה ב־0.

אף שאינה חייבת להיות רציפה, הנגזרת מקיימת את תכונת ערך הביניים (הגדרה 7.9.4):

משפט 8.5.11 (משפט דרבו)[7] תהי $f : [a,b] \to \mathbb{R}$ גזירה ב־(a,b) וגזירה חד־צדדית ב־a,b. אז לכל מספר s הנמצא בין $f'_+(a)$ לבין $f'_-(b)$ (ממש) יש נקודה $c \in (a,b)$ כך ש־ $f'(c) = s$.

הוכחה אפשר להניח ש־ $f'_+(a) \neq f'_-(b)$, אחרת הטענה טריוויאלית. נניח למשל ש־ $f'_+(a) < f'_-(b)$, המקרה השני דומה.

אם כן, אנו מניחים ש־ $f'_+(x) < s < f'_-(b)$. נגדיר $g(x) = f(x) - sx$ ונשים לב ש־ $f'(x) = s$ אמ"מ $g'(x) = 0$. לכן די שנראה של־g יש נקודת קיצון, ולשם כך די להראות שהיא מקבלת מינימום ומקסימום בפנים הקטע (a,b), ולא בקצה.

אם $g(a) = g(b)$ המסקנה מתקבלת ממשפט רול. אחרת, נשים לב ש־

$$g'_+(a) = f'_+(a) - s < 0 \qquad , \qquad g'_-(b) = f'_-(b) - s > 0$$

[7] Jean Darboux, 1842־1917.

ובפרט יש נקודות $a' < b'$ בקטע עם $g(a') < g(a)$ ו־ $g(b') > g(b)$ (זו מסקנה מכך ש־ g יורדת מימין לנקודה a ועולה משמאל לנקודה b). כעת אם $g(a) < g(b)$ אז ממשפט ערך הביניים הערך $g(a)$ מתקבל בקטע (a', b) בנקודה c כלשהי, ואז אפשר להפעיל את משפט רול על הקטע $[a, c]$. אחרת $g(a) > g(b)$, ואז הערך $g(b)$ מתקבל בנקודה c כלשהי בקטע $[a, b')$, ונפעיל את משפט רול על הקטע $[c, b]$. ■

מסקנה 8.5.12 אם לפונקציה g המוגדרת בקטע יש נקודות אי־רציפות סליקה או ממין ראשון אז g אינה הנגזרת של אף פונקציה, דהיינו אין פונקציה f כך ש־ $f' = g$.

הוכחה אם לפונקציה יש אי־רציפות סליקה או ממין ראשון ב־x_0 אז בסביבת x_0 אפשר למצוא קטע תכונת ערך הביניים נכשלת, ולכן לפי משפט דרבו בקטע זה g אינה יכולה להיות נגזרת. ■

תרגילים

1. בשאלה זו נבחן יותר לעומק את מושג העלייה והירידה בנקודה.
   (א) הראו שאם $f'(x_0) = 0$ לא ניתן להסיק דבר על עלייה או ירידה של f ב־x_0 (כלומר: כל דבר יכול לקרות).
   (ב) התבוננו בפונקציה
   $$f(x) = \begin{cases} \frac{1}{2}x + x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$
   הראו שהיא עולה בנקודה 0 אך אינה עולה באף סביבה של 0.
   (ג) (*) הוכיחו שאם $f : [a, b] \to \mathbb{R}$ עולה בכל נקודה ב־$[a, b]$ אז f עולה בקטע $[a, b]$ (אל תניחו ש־f גזירה!). סעיף זה מספק הוכחה חלופית למסקנה 8.5.9.
2. נניח ש־f גזירה ב־x_0 ויש לה מקסימום מקומי שם. הוכיחו או הפריכו:
   (א) אם f גזירה בסביבה של x_0 אז קיים $\delta > 0$ כך שהמיתר בגרף של f בין $x_0 - \delta$ ל־ $x_0 + \delta$ מאוזן.
   (ב) אם f רציפה בסביבה של x_0 אז יש $\delta', \delta'' > 0$ כך שהמיתר בגרף של f בין $x_0 - \delta$ ל־ $x_0 + \delta$ מאוזן.
   (ג) אם f מוגדרת בסביבה של x_0 אז יש $\delta', \delta'' > 0$ כך שהמיתר בגרף של f בין $x_0 - \delta$ ל־ $x_0 + \delta$ מאוזן.
3. הוכיחו כי $\frac{x}{1+x} \leq \ln(1 + x) \leq x$ לכל $x > -1$.
4. הוכיחו שלמשוואה $x = a + b \sin x$ יש פתרון והפתרון יחיד לכל בחירה $b \in (0, 1)$ ו־ $a \in \mathbb{R}$.

5. תהי f גזירה ב־$[a,b]$.

(א) הראו שאם $f'(x)\geq M$ לכל $x\in(a,b)$ אז $f(x)\geq f(a)+M(x-a)$ לכל $x\in[a,b]$. נסחו טענה דומה למקרה $f'(x)\leq M$ ולמקרה $|f'(x)|\leq M$.

(ב) נניח ש־ $f(a)=f(b)=0$. ו־ $|f'(x)|\leq M$ לכל $x\in(a,b)$. הראו ש־ $|f(x)|\leq\frac{M}{2(b-a)}$ לכל $x\in[a,b]$.

6. תהי $f:(0,\infty)\to\mathbb{R}$ גזירה. הוכיחו או הפריכו את הטענות הבאות:

(א) אם $f(1)=0$ ו־ $f'(1)>0$ אז $f(2)>0$.

(ב) אם $\lim_{x\to 0+}f'(x)=\infty$ אז $\lim_{x\to 0+}f(x)=-\infty$.

(ג) אם $\lim_{x\to 0+}f(x)=\infty$ אז $\lim_{x\to 0+}f'(x)=-\infty$.

(ד) אם $\lim_{x\to 0+}f'(x)=c\in\mathbb{R}$ אז קיים וסופי הגבול $\lim_{x\to 0+}f(x)$.

(ה) אם $\lim_{x\to\infty}f'(x)=0$ אז $\lim_{x\to\infty}f(x)=0$.

(ו) אם $\lim_{x\to\infty}f(x)=0$ אז $\lim_{x\to\infty}f'(x)=0$.

(ז) אם $\lim_{x\to\infty}f(x)$ לא קיים במובן הרחב אז $\lim_{x\to\infty}f'(x)$ לא קיים במובן הרחב.

(ח) אם $\lim_{x\to\infty}f'(x)=c<0$ אז $\lim_{x\to\infty}f'(x)=-\infty$.

7. יהיו $p\neq q$ פולינומים ממעלה m,n בהתאמה. הוכיחו שהגרפים שלהם נחתכים ב־$\max\{m,n\}$ נקודות שונות לכל היותר. הראו שיש פולינומים עם בדיוק מספר זה של נקודות חיתוך.

8. הראו שאם $f:\mathbb{R}\to\mathbb{R}$ ואם יש k כך ש־ $f^{(k)}\equiv 0$ אז f פולינום ממעלה k.

9. יהיו $\alpha,\beta<0$.

(א) הוכיחו שיש קבועים $C,D>0$ כך ש־ $\ln x+\alpha x^{\beta}\geq Cx^{\beta}-D$ בסביבה ימנית מנוקבת של 0, ולכן $\lim_{x\to 0+}(\ln x+\alpha x^{\beta})=\infty$.

(ב) הסיקו ש־ $\lim_{x\to 0+}x^{\alpha}e^{-x^{\beta}}=0$.

10. יהי $n>1$. הוכיחו ש־ $(x^n+y^n)=(x+y)^n$ רק אם $x=0$ או $y=0$ (רמז: קבעו $y\neq 0$ והתבוננו בהפרש הביטויים כפונקציה של x).

11. תהי $f:\mathbb{R}\to\mathbb{R}$ גזירה כך ש־ $f(x+y)=f(x)f(y)$. הראו שאם $f\not\equiv 0$ אז $f(x)=a^x$ למספר a כלשהו.

12. יהיו $s,c:\mathbb{R}\to\mathbb{R}$ פונקציות ממשיות המקיימות $s'=c$, $c'=-s$, וכן $c(0)=1$, $s(0)=0$. הוכיחו ש־ $s=\sin$, $c=\cos$ (רמז: הוכיחו ש־ $s^2+c^2\equiv 1$ ו־ $s\cdot\sin+c\cdot\cos\equiv 1$. היעזרו בתרגיל 5(ד)) מעמוד 35).

13. הראו שאם f גזירה בכל נקודה בקרן $[0,\infty)$ ואם מתקיים $\lim_{x\to\infty}f'(x)=0$ אז $\lim_{x\to\infty}\frac{f(x)}{x}=0$

14. יהי $\alpha > 0$. פונקציה f בקטע I נקראת **הלדר**[8] **מסדר** α (Hölder of or-der α) אם יש קבוע חיובי C כך שלכל שתי נקודות $x', x'' \in I$ מתקיים $|f(x') - f(x'')| \le C|x' - x''|^\alpha$.
   (א) הראו שפונקציה שהיא הלדר מסדר α בקטע I היא רציפה ב־I.
   (ב) הראו שאם f גזירה בקטע I ואם הנגזרת חסומה אז f היא הלדר מסדר 1.
   (ג) הראו ש־$\sqrt{x}$ היא הלדר מסדר $\frac{1}{2}$ בקטע $[0, \infty)$.
   (ד) (*) הראו שפונקציה שהיא הלדר מסדר $\alpha > 1$ היא קבועה.
15. תהי f רציפה בסביבה של x_0 וגזירה בסביבה מנוקבת של x_0. נניח שהגבול $\lim_{x \to x_0} f'(x)$ קיים. הראו ש־f גזירה ב־x_0.
16. יהי $k \ge 2$ טבעי ותהי $f : \mathbb{R} \to \mathbb{R}$ מוגדרת על ידי

$$f(x) = \begin{cases} x^k \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

   הוכיחו ש־f גזירה $[k/2] - 1$ פעמים ב־0 אך אינה גזירה שם $[k/2]$ פעמים.
17. תהי $f : \mathbb{R} \to \mathbb{R}$ פונקציה גזירה ונניח ש־ $f(f(x)) = f(x)$ לכל $x \in \mathbb{R}$. הוכיחו ש־f קבועה או ש־ $f(x) = x$. (*) הוכיחו זאת בהנחה ש־f רציפה אבל לאו דווקא גזירה.
18. (*) הראו שאם $f : \mathbb{R} \to \mathbb{R}$ גזירה, $f(0) = 0$ ומתקיים $|f'(x)| \le |f(x)|$ לכל $x \in \mathbb{R}$ אז $f \equiv 0$ (רמז: הראו שאם $f(x_0) = 0$ אז $f(x) = 0$ לכל $x \in (x_0 - \frac{1}{2}, x_0 + \frac{1}{2})$).

8.6 חקירת פונקציות

בהינתן פונקציה גזירה f, השיטות מהסעיף הקודם מאפשרות לתת תיאור איכותי וכמותי של הגרף שלה. תהליך זה מכונה **חקירה** של f. ביתר פירוט, נרצה למצוא את נקודות הקיצון של f, לתאר באילו קטעים היא עולה ויורדת, לזהות את צורתה באותם קטעים, ולתאר את ההתנהגות האסימפטוטית שלה.

תהי f פונקציה רציפה בקטע I וגזירה בפנים שלה. השלב הראשון בחקירתה הוא מציאת כל הנקודות ב־I שבהן מתאפסת הנגזרת f'. נסמן נקודות אלה ב־ $x_1 < x_2 < \ldots < x_n$ (לשם פשטות אנו מניחים שיש מספר סופי של נקודות כאלה. לרוב אפשר לטפל במקרה הכללי בצורה דומה). בכל קטע (x_{i-1}, x_i) הנגזרת אינה מתאפסת, ומכאן אפשר להסיק ש־f' בעלת סימן קבוע ב־(x_{i-1}, x_i) (זו מסקנה מכך שלפי משפט 8.5.11, f' מקיימת את תכונת ערך הביניים). לכן לפי מסקנה

[8] Otto Hölder, 1859–1937

8.5.9 הפונקציה f מונוטונית ממש בקטע (x_{i-1},x_i), והמגמה שלה נקבעת על ידי סימן הנגזרת בקטע, שאותו אפשר לחשב על ידי בדיקת הערך שלה בנקודה אחת או על ידי חישוב ערך הפונקציה בשתי נקודות. כך אנו מקבלים תמונה מלאה של תחומי העלייה והירידה של f.

היינו גם רוצים למצוא את נקודות הקיצון של f. הנקודות היחידות המועמדות להיות נקודות קיצון של f הן הנקודות x_i, כי לפי משפט פרמה (משפט 8.5.5) בנקודת קיצון הנגזרת מתאפסת ו־$x_1,\ldots,x_n$ הן כל נקודות ההתאפסות של הנגזרת. אולם אין הן חייבות להיות נקודות קיצון. כדי להכריע אילו הן נקודות קיצון נשתמש במידע שיש לנו על העלייה והירידה של f בכל אחד מהקטעים (x_{i-1},x_i). נשים לב שאם f עולה ממש ב־(x_{i-1},x_i) ויורדת ממש ב־ (x_i,x_{i+1}) אז היא מקבלת ב־x_i מקסימום מקומי, ואם היא יורדת ב־(x_{i-1},x_i) ועולה ב־ (x_i,x_{i+1}) אז יש לה ב־x_i מינימום מקומי. מאידך, אם המגמה של f זהה בשני הקטעים האלה אז x_i איננו נקודת קיצון.

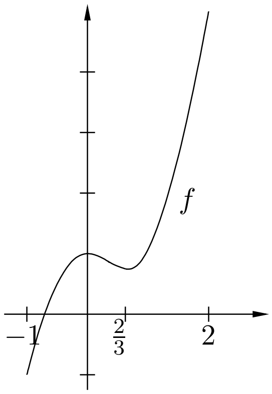
יתרה מזאת, אם I סגור, אז קיומם של מקסימום ומינימום של f ב־I מובטח כי f רציפה ו־I, וכעת יש לנו גם כלים לחשב אותם. אם $y\in I$ היא נקודת מקסימום של f אז או ש־y נקודה פנימית של I, ואז יש ל־f מקסימום מקומי שם, או ש־y היא נקודת קצה של I. אנו מסיקים שהמקסימום של f ב־I מתקבל באחת הנקודות $x_1,\ldots,x_n$ או בקצה של I (אם יש ל־I קצוות). באופן דומה המינימום מתקבל באחת הנקודות האלה.

הנה דוגמה: נחקור את הפונקציה $f(x)=x^3-x^2+1$ בקטע $[-1,2]$. חישוב מראה ש־

$$f'(x)=3x^2-2x=x(3x-2)$$

ולכן $f'(x)=0$ אמ"מ $x=0,\frac{2}{3}$. בדיקה מראה ש־ $f'(-\frac{1}{2}),f'(1)>0$ ואילו $f'(\frac{1}{2})<0$ ולכן אנו מסיקים ש־f עולה ממש בקטעים $(-1,0)$ וב־$(\frac{2}{3},2)$ ויורדת בקטע $(0,\frac{2}{3})$. משום כך יש ל־f מקסימום מקומי ב־0 ומינימום מקומי ב־$\frac{2}{3}$. נחשב ונקבל

$$f(-1)=-1 \quad , \quad f(0)=1 \quad , \quad f(\frac{2}{3})=\frac{23}{27} \quad , \quad f(2)=5$$



איור 8.6.1 הפונקציה $f(x)=x^3-x^2+1$

ובפרט המקסימום של f מתקבל ב־2 והמינימום ב־-1. אנו מקבלים את הסקיצה המופיעה באיור 8.6.1. שימו לב שהסקיצה מעידה על כך שהפונקציה מתאפסת בקטע $(-1,0)$ (זו מסקנה של משפט ערך הביניים).

ראינו שהתאפסות הנגזרת בנקודה x_0 אינה תנאי מספיק כדי ש־x_0 תהיה נקודת קיצון. דרך אחת לוודא זאת הייתה למצוא את מגמת הפונקציה משני צדי הנקודה. ישנה גם שיטה המשתמשת רק במידע על הנגזרות מסדר גבוה של הפונקציה בנקודה x_0.

הגדרה 8.6.1 נאמר ש־x_0 היא **מקסימום מקומי ממש** של f אם יש סביבה מנוקבת V של x_0 שבה $f(x) < f(x_0)$ לכל $x \in V$, או באופן שקול, אם יש סביבה מלאה שבה x_0 היא נקודת המקסימום היחידה. מגדירים **מינימום מקומי ממש** באופן דומה.

טענה 8.6.2 תהי f פונקציה גזירה בסביבת x_0 ונניח $f'(x_0) = 0$. אם f' גזירה גם היא ב־x_0 ואם $f''(x_0) > 0$ אז x_0 נקודת מינימום מקומי ממש של f, ואם $f''(x_0) < 0$ אז x_0 נקודת מקסימום מקומי ממש.

הוכחה נניח ש־ $f'(x_0) = 0$ ו־ $f''(x_0) > 0$. אז f' עולה בנקודה x_0, כלומר יש סביבה שמאלית של x_0 בה $f'(x) < f'(x_0) = 0$, ויש סביבה ימנית בה $f'(x) > f'(x_0) = 0$, ותנאים אלה מבטיחים שבאותה סביבה שמאלית f יורדת ממש ובאותה סביבה ימנית f עולה ממש, ומכאן ברור ש־x_0 נקודת מינימום מקומי של f.

המקרה $f''(x_0) < 0$ מוכח באופן דומה. ■

למשל, בדוגמה $f(x) = x^3 - x^2 + 1$ למעלה ראינו ש־ $f'(x) = x(3x-2)$ וש־ $f'(0) = f'(\frac{2}{3}) = 0$. גזירה נוספת מראה ש־ $f''(x) = 6x - 2$ ולכן $f''(0) = -2$ ו־ $f''(\frac{2}{3}) = 2$, ולכן הטענה נותנת הוכחה חדשה לכך ש־0 נקודת מקסימום מקומי ו־$\frac{2}{3}$ נקודת מינימום מקומי.

טענה 8.6.2 היא תנאי מספיק בלבד לקיום נקודות קיצון, ואינה מספקת כל מידע במקרה ש־ $f''(x_0) = 0$. למשל שתי הנגזרות הראשונות של x^3 מתאפסות ב־0 אך אין לה נקודת קיצון שם. המשפט הבא הוא הכללה של הטענה ומנצל גם מידע על נגזרות גבוהות יותר.

משפט 8.6.3 נניח ש־f גזירה $n+1$ פעמים ב־x_0 וש־n הנגזרות הראשונות מתאפסות אבל $f^{(n+1)}(x_0) \neq 0$. אזי

1. אם n זוגי אז x_0 אינה נקודת קיצון. במקרה זה אם $f^{(n+1)}(x_0) > 0$ אז f עולה ב־x_0 ואם $f^{(n+1)}(x_0) < 0$ אז f יורדת ב־x_0.
2. אם n אי־זוגי אז אם $f^{(n+1)}(x_0) > 0$ אז x_0 מינימום מקומי ממש של f, ואם $f^{(n+1)}(x_0) < 0$ אז x_0 מקסימום מקומי ממש של f.

הוכחה ההוכחה באינדוקציה על n.

עבור $n = 0$ זו למה 8.5.3 ועבור $n = 1$ זו טענה 8.6.2.

יהי $n \geq 1$ ונניח שהוכחנו את הטענה ל־$n-1$. נניח ש־ $f^{(k)}(x_0) = 0$ ל־ $k \leq n$ ו־ $f^{(n+1)}(x_0) \neq 0$. נוכיח את הטענה ל־n.

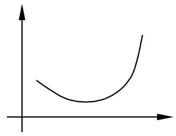
נסמן $g(x) = f'(x)$, כך ש־ $g^{(k)} = f^{(k+1)}$. הפונקציה g מקיימת את הנחת האינדוקציה, דהיינו $g^{(1)}(x_0) = \ldots = g^{(n-1)}(x_0) = 0$ אבל $g^{(n)}(x_0) \neq 0$.

כעת אנו צריכים להוכיח את (1) ו־(2), אך למעשה יש להוכיח רק את אותה טענה שרלוונטית לאור הזוגיות של n.

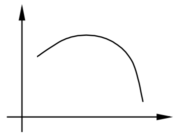
אם n זוגי נראה את (1) למקרה ש־$f^{(n+1)}(x_0) = g^{(n)}(x_0) < 0$ (מקרה האחר דומה). אז $n-1$ אי־זוגי. לכן מכיוון ש־ $n-1$ הנגזרות הראשונות של g מתאפסות, והנגזרת ה־n־ית של g שלילית, מהנחת האינדוקציה נובע ש־x_0 מקסימום מקומי ממש של g. כיוון ש־ $g(x_0) = f'(x_0) = 0$, פירוש הדבר שבסביבה מנוקבת של x_0 מתקיים $g(x) < 0$ (האי־שוויון החזק נובע מכך ש־x_0 היא מקסימום מקומי ממש של g). כיוון ש־ $g = f'$ נובע ש־f מונוטונית יורדת ממש בסביבה של x_0, ובפרט יורדת ב־x_0.

אם n אי־זוגי עלינו להוכיח את (2). מהנחת האינדוקציה הפונקציה $g = f'$ עולה או יורדת ב־x_0, ו־ $f'(x_0) = 0$. כעת ההוכחה ש־x_0 היא מקסימום מקומי ממש או מינימום מקומי ממש דומה למקרה $n = 1$ בטענה 8.6.2, ומושארת כתרגיל. ■

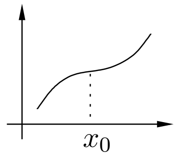
ייתכן ש־ $f^{(k)}(x_0) = 0$ לכל k ואז המשפט אינו מאפשר להכריע אם מדובר בנקודת קיצון או לא. פונקציות כאלה אכן קיימות ויכולות להתנהג ב־x_0 בכל מיני דרכים. ראו למשל תרגיל (8) בסוף בסעיף 8.7.



איור 8.6.2 פונקציה קמורה



איור 8.6.3 פונקציה קעורה



איור 8.6.4 פונקציה עם נקודת פיתול ב־x_0

על צורת הגרף. התבוננו למשל בשלושת הגרפים שבאיורים 8.6.2, 8.6.3, ו־8.6.4. ההבדל בין המקרים הוא שבראשון השיפוע הולך וגדל, בשני הוא הולך וקטן, ובשלישי הוא איננו בעל מגמה קבועה (תחילה הוא קטן ואחר־כך גדל). כאשר הנגזרת השנייה של הפונקציה קיימת בקטע אפשר לתאר תופעות אלה באמצעותה: פונקציית הנגזרת עולה כאשר הנגזרת השנייה חיובית, והנגזרת יורדת כאשר הנגזרת השנייה שלילית.

הגדרה 8.6.4 תהי $f : (a,b) \to \mathbb{R}$ גזירה פעמיים. אם $f''(x) > 0$ בכל נקודה ב־$x \in (a,b)$ אז f נקראת **קמורה ממש** (strictly convex) ב־(a,b). אם $f''(x) < 0$ ב־(a,b) אז f נקראת **קעורה ממש** (strictly concave) ב־(a,b). אם f קמורה בסביבה חד־צדדית של x_0 וקעורה בסביבה חד־צדדית בצד השני אז x_0 נקראת **נקודת פיתול** (inflection point) של f.

מכיוון שהנגזרת מקיימת את תכונת ערך הביניים (משפט 8.5.11) נקודת פיתול x_0 בהכרח מקיימת $f''(x_0) = 0$.

למשל, בדוגמה $f(x) = x^3 - x^2 + 1$ שבחנו קודם התקיים $f''(x) = 6x - 2$. לכן $f'' < 0$ כאשר $x < \frac{1}{3}$ ו־ $f''(x) > 0$ כאשר $x > \frac{1}{3}$, והנקודה $\frac{1}{3}$ היא נקודת פיתול. אפשר להתרשם מכך באיור 8.6.1 בעמוד 8.6.

שאלה מסוג אחר שאפשר לשאול על פונקציות היא מה ההתנהגות האסימפטוטית שלהן כאשר x שואף ל־∞ (או ל־$-\infty$), ומה התנהגות האינפיניטסימלית שלהן בנקודות בהן הגבול אינו קיים.

הגדרה 8.6.5 אם f מוגדרת בסביבה חד־צדדית או בסביבה מנוקבת של נקודה x_0 נאמר שיש ל־f **אסימפטוטה אנכית** אם $\lim_{x\to x_0} |f(x)| = \infty$, או במילים אחרות: הגבולות החד־צדדיים של f קיימים ואינסופיים ב־a (אבל אולי עם סימנים הפוכים).

לדוגמה, לפונקציה $f(x)=\frac{1}{x}$ יש אסימפטוטה אנכית ב־$x=0$.

אפשר גם להבחין בין צורות התנהגות שונות של פונקציה באינסוף. למשל, קיום גבול באינסוף אומר שעבור שכאשר x שואף לאינסוף, גרף הפונקציה קרוב לישר מאוזן. ואמנם, אם $\lim_{x\to\infty} f(x)=L$ אז לכל $\varepsilon>0$ מתקיים $L-\varepsilon \le f(x) \le L+\varepsilon$ לכל x מספיק גדול, כלומר יש קרן שבה הגרף של f קרוב עד כדי ε לישר האופקי בגובה L. באופן דומה, כאשר $\lim_{x\to\infty} f(x)=\infty$ אפשר לשאול האם הגרף קרוב לישר שאינו אופקי. ליתר דיוק,

הגדרה 8.6.6 נאמר שהישר $\ell(x)=Ax+B$ הוא **אסימפטוטה משופעת** של f ב־∞ אם מתקיים

$$\lim_{x\to\infty}(f(x)-\ell(x))=0$$

(כלומר: אם המרחק האנכי בין הגרף של f לגרף של הישר $Ax+B$ שואף ל־0 כש־x שואף ל־∞). מגדירים באופן דומה אסימפטוטה משופעת ב־$-\infty$.

כדי לחשב אסימפטוטה משופעת, יש למצוא A,B כנ"ל. נשים לב שאם היה קיים הגבול $\lim_{x\to\infty}(f(x)-(Ax+B))$ אז לפי אריתמטיקה של גבולות,

$$\lim_{x\to\infty}\frac{1}{x}(f(x)-(Ax+B))=0$$

ומצד שני

$$\begin{aligned}\lim_{x\to\infty}\frac{1}{x}(f(x)-(Ax+B)) &= \lim_{x\to\infty}(\frac{f(x)}{x}-A-\frac{B}{x})\\ &= \lim_{x\to\infty}(\frac{f(x)}{x}-A)\end{aligned}$$

ולכן אם יש A,B כאלה, מתקיים $\lim_{x\to\infty}(\frac{f(x)}{x}-A)=0$ כלומר

$$A=\lim_{x\to\infty}\frac{f(x)}{x}$$

כך נמצא את A, אם קיים. כמו־כן,

$$B=\lim_{x\to\infty}(f(x)-Ax)$$

(אם הגבול קיים וסופי).

מצד שני ברור שאם הגבולות $A=\lim_{x\to\infty}\frac{f(x)}{x}$ ו־ $B=\lim_{x\to\infty}(f(x)-Ax)$ קיימים אז השוויון האחרון אומר ש־ $Ax+B$ הוא אסימפטוטה של f ב־∞.

דוגמאות

1. לא לכל פונקציה יש אסימפטוטה משופעת. למשל, אם $f(x) = x^2$ אז $\lim_{x\to\infty} \frac{f(x)}{x} = \infty$, ולכן אין אסימפטוטה כזו. כאן הבעיה היא שקצב השאיפה לאינסוף מהיר מדי.
2. דוגמה טיפוסית נוספת היא $f(x) = \sqrt{x}$, שבה קצב השאיפה לאינסוף דווקא איטי מדי. במקרה זה מתקיים

$$\lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} \frac{1}{\sqrt{x}} = 0$$

ולכן אם יש אסימפטוטה משופעת אז $A = 0$. אבל כשננסה לחשב את B, נגלה ש־

$$\lim_{x\to\infty} (f(x) - Ax) = \lim_{x\to\infty} \sqrt{x} = \infty$$

ולכן אין ל־f אסימפטוטה, אף ש־ $\lim_{x\to\infty} \sqrt{x} = \infty$.

3. אם $\lim_{x\to\infty} f(x) = L$ אז $0x + L$ היא אסימפטוטה **אופקית** או **מאוזנת** של f.
4. הנה דוגמה לפונקציה שיש לה אסימפטוטה משופעת שאינה אופקית. נתבונן ב־ $f(x) = x + \frac{\cos x}{x}$. אז

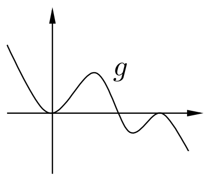
$$\lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} (1 + \frac{\cos x}{x^2}) = 1$$

וגם

$$\lim_{x\to\infty} (f(x) - Ax) = \lim_{x\to\infty} \frac{\cos x}{x} = 0$$

ולכן הישר $\ell(x) = x$ מהווה אסימפטוטה באינסוף.

תרגילים



איור 8.6.5

1. התבוננו בפונקציה g המתוארת באיור 8.6.5.
   (א) ציירו גרף של g'.
   (ב) תהי f פונקציה כך ש־ $f' = g$. ציירו סקיצה לגרף של f בהנחה ש־ $f(0) = 1$.
2. השלימו את הפרטים של הוכחת סעיף (2) בשלב האינדוקציה של הוכחת משפט 8.6.3.
3. חקרו את הפונקציות הבאות, ותארו את הגרפים שלהן:

(א) $f(x)=x+\frac{1}{x}$.

(ב) $f(x)=\sqrt{x^2+1}$.

(ג) $f(x)=1/(1+x^2)$.

(ד) $f(x)=x^x$.

(ה) $f(x)=\frac{x^2+1}{x^2-1}$.

4. הוכיחו:

(א) $x\le\tan(x)$ עבור $0\le x<\frac{\pi}{2}$.

(ב) $x^p-1\ge p\cdot(x-1)$ עבור $x>0$, $p>1$.

(ג) יהיו $p,q>0$ מספרים כך ש־ $\frac{1}{p}+\frac{1}{q}=1$. הוכיחו שלכל $x,y>0$ מתקיים $xy\le\frac{x^q}{q}+\frac{y^p}{p}$.

5. יהיו $a_1<\ldots<a_n$ מספרים.

(א) תהי $f(x)=\sum_{k=1}^{n}(x-a_k)^2$. מצאו את הערך המינימלי של f.

(ב) מה הערך המינימלי של $g(x)=\sum_{k=1}^{n}|x-a_k|$? (כאן לפונקציה יש נקודת אי־גזירות ולכן אי אפשר להשתמש בשיטות שלמדנו בסעיף הנוכחי, אבל אפשר להפעיל שיקול אלמנטרי יותר).

6. יהיו $a_1,...,a_n$ מספרים חיוביים. מצאו את הערך המינימלי של הפונקציה

$$f(x)=\frac{a_1+...+a_{n-1}+x}{n\cdot(a_1\cdot...\cdot a_{n-1}\cdot x)^{\frac{1}{n}}}$$

והיעזרו בו כדי למצוא הוכחה נוספת לאי־שוויון הממוצעים.

7. עבור n טבעי נגדיר את הפונקציה $e_n(x)=(1+\frac{x}{n})^n$ ויהי $r\ge 0$ מספר ממשי. הוכיחו שלכל $n\ge r$, בקרן $[-r,\infty)$ ההפרש $e_{n+1}-e_n$ יורד ב־$[-r,0]$ ועולה ב־$[0,\infty)$, ולכן הפונקציה מקבלת את הערך המינימלי שלה בנקודה 0. הסיקו מכך שלכל $x\in\mathbb{R}$ הסדרה $((1+\frac{x}{n})^n)_{n=1}^{\infty}$ עולה החל ממקום מסוים. (רמז: שימו לב שאפשר להציג את ערך הנגזרת e'_n בנקודה x בתור $e'_{n-1}(ax)$ עבור קבוע a מתאים, והוכיחו את הטענה באינדוקציה).

8. בשאלה זו נראה שכאשר כל הנגזרות של פונקציה בנקודה מתאפסות אי אפשר להסיק שהנקודה היא נקודת קיצון, וגם אי אפשר להסיק שלא.

(א) תהי $f:\mathbb{R}\to\mathbb{R}$ הפונקציה

$$f(x)=\begin{cases} e^{-1/x^2} & x\neq 0\\ 0 & x=0\end{cases}$$

הראו שלכל α הפונקציה $x^\alpha f(x)$ גזירה ב־0 ונגזרתה שם 0 (היעזרו בתרגיל (9) בעמוד 340). הסיקו ש־f גזירה אינסוף פעמים באפס, ש־ $f^{(n)}=0$ לכל n, ושל־f יש מינימום מקומי ב־0.

(ב) תהי $g(x)=xf(x)$. הראו ש־g גזירה אינסוף פעמים באפס, $g^{(n)}=0$ לכל n ו־g עולה ממש בסביבת 0.

(ג) תהי $h(x)=f(x)\cdot\sin\frac{1}{x}$ ו־ $h(0)=0$. הראו ש־h גזירה אינסוף פעמים באפס, $h^{(n)}=0$ לכל n, אבל ל־h אין נקודת קיצון ב־0 והיא גם אינה מונוטונית באף סביבה של 0.

8.7 כלל לופיטל

כלל המנה אינו מאפשר לחשב גבול של מנה $\frac{f}{g}$ אם המונה והמכנה מתכנסים ל־0. כאשר הפונקציות גזירות, אפשר להשתמש במקומו בכלל לופיטל. הרעיון מאחורי כלל זה הוא שבמקרים מסויימים גבול המנה $\frac{f}{g}$ נקבע לפי קצבי ההתכנסות של f,g, או במילים אחרות, על ידי מנת הנגזרות $\frac{f'}{g'}$.

לדוגמה, נתבונן בגבול $\lim_{x\to 0}\frac{(x+1)^2-1}{x+x^2}$. מכיוון שהמכנה שואף ל־0 לא ברור שהגבול קיים, אך חישוב ישיר מראה שהוא כן קיים:

$$\lim_{x\to 0}\frac{(x+1)^2-1}{x+x^2}=\lim_{x\to 0}\frac{2x+x^2}{x+x^2}=\lim_{x\to 0}\frac{2+x}{1+x}=2$$

אפשר להסביר את הגבול הזה באופן הבא. הפונקציה $(x-1)^2-1$ שווה עד כדי שגיאה אפסית לישר המשיק שלה ב־0, שהוא $2x$. הפונקציה $x+x^2$ שווה עד כדי שגיאה אפסית לישר המשיק שלה ב־0, שהוא הישר x. לכן המנה היא מהצורה $\frac{2x+o(x)}{x+o(x)}$, וכאשר $x\to 0$ הגבול של מנה זו שווה ל־$\frac{2}{1}$, שהוא מנת הנגזרות של הפונקציות ב־0.

שיטה זו עובדת גם באופן כללי:

טענה 8.7.1 אם f,g פונקציות כך ש־ $f(x_0)=g(x_0)=0$ ואם שתיהן גזירות ב־x_0 ו־ $g'(0)\neq 0$ אז

$$\lim_{x\to x_0}\frac{f(x)}{g(x)}=\frac{f'(x_0)}{g'(x_0)}$$

הוכחה לכל $x\neq x_0$ מתקיים

$$\frac{f(x)}{g(x)}=\frac{f(x)-f(x_0)}{g(x)-g(x_0)}=\frac{f(x)-f(x_0)}{x-x_0}\cdot\frac{x-x_0}{g(x)-g(x_0)}$$

ולכן מהגדרת הנגזרת ואריתמטיקה של גבולות,

$$\lim_{x\to x_0}\frac{f(x)}{g(x)}=\frac{f'(x_0)}{g'(x_0)}$$

כנדרש. ■

אפשר גם לנסח את הטענה עבור גבולות חד־צדדיים ונגזרות חד־צדדיות.

לדוגמה, נראה ש־ $\lim_{x\to 0}\frac{e^x-1}{\sin x}=1$. המונה והמכנה גזירים ב־$0$ ומתאפסים ב־0, והנגזרת של המכנה ב־0 היא $\cos 0=1\neq 0$, ולכן מהטענה

$$\lim_{x\to 0}\frac{e^x-1}{\sin x}=\frac{\frac{d}{dx}(e^x-1)|_{x=0}}{\frac{d}{dx}(\sin x)|_{x=0}}=\frac{e^0}{\cos 0}=1$$

בטענה 8.7.1 דרשנו גזירות של הפונקציות ב־x_0, אך היה רצוי לקבל גרסה שאינה דורשת זאת. למשל, אי אפשר לחשב בשיטה למעלה את הגבול $\lim_{x\to 0+}\frac{e^x-1}{\sqrt{x}}$ כי המכנה אינו גזיר מימין באפס. במצבים כאלה אפשר להיעזר במשפט הבא, שבו גזירות בנקודה מתחלפת בגזירות בסביבה מנוקבת של הנקודה. ננסח גרסה חד־צדדית, ונסיק ממנה את הגרסה המלאה:

משפט 8.7.2 (כלל לופיטל,[9] גרסה $\frac{0}{0}$ חד־צדדית) יהיו f,g פונקציות גזירות בסביבה ימנית מנוקבת של x_0 ונניח

$$\lim_{x\to x_0+}f(x)=\lim_{x\to x_0+}g(x)=0$$

נניח עוד ש־ $g(x)\neq 0$ בסביבה ימנית של x_0. אם הגבול $\lim_{x\to x_0+}\frac{f'(x)}{g'(x)}$ קיים אז גם $\lim_{x\to x_0+}\frac{f(x)}{g(x)}$ קיים, והם שווים. טענה דומה נכונה לגבולות שמאליים עם השינויים המתבקשים.

הוכחה היות ושינוי של ערך פונקציה בנקודה לא משפיע על ערך הגבול, נוכל להניח ש־ $f(x_0)=g(x_0)=0$, כלומר ש־f,g רציפות מימין ב־0.

נניח ש־ $\lim_{x\to x_0+}\frac{f'(x)}{g'(x)}$ קיים. יהי $r>0$ מספר כך ש־f,g גזירות בקטע (x_0,x_0+r). בפרט לכל $h\in(0,r)$ הפונקציות f,g מקיימות בקטע (x_0,x_0+h) את תנאי משפט ערך הביניים של קושי (משפט 8.5.10) ולכן אפשר לבחור מספר $c\in(x_0,x_0+h)$ כך ש־

$$\frac{f(x_0+h)}{g(x_0+h)}=\frac{f(x_0+h)-f(x_0)}{g(x_0+h)-g(x_0)}=\frac{f'(c)}{g'(c)}$$

ייתכן שיש כמה c־ים כאלה (התלויים ב־x), אבל נוכל לבחור אחד מהם ולסמן אותו $c(h)$. כך הגדרנו פונקציה $c:(0,r)\to\mathbb{R}$. מהאי־שוויון $x_0<c(h)<x_0+h$ נובע $\lim_{h\to 0+}c(h)=x_0$. ולכן

$$\lim_{x\to x_0+}\frac{f(x)}{g(x)}=\lim_{h\to 0+}\frac{f(x_0+h)}{g(x_0+h)}=\lim_{h\to 0+}\frac{f'(c(h))}{g'(c(h))}=\lim_{x\to x_0+}\frac{f'(x)}{g'(x)}$$

[9] Guillaume l'Hopital, 1661־1704.

כפי שרצינו (הצדיקו כל אחד מהשוויונות!) ∎

בהוכחת המשפט נעזרנו במשפט ערך הביניים של קושי, שבהוכחתו הסתמכנו על אינטואיציה גאומטרית לגבי העקום במישור המורכב מהנקודות $(g(x), f(x))$. נסו למצוא אינטואיציה דומה למשפט האחרון!

מסקנה 8.7.3 (כלל לופיטל, גרסה $\frac{0}{0}$ דו־צדדית) יהיו f, g פונקציות גזירות בסביבה מנוקבת של x_0 ונניח ש־ $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0$. נניח גם ש־ $g(x) \neq 0$ בסביבה מנוקבת של x_0. אם הגבול $\lim_{x\to x_0} \frac{f'(x)}{g'(x)}$ קיים אז גם $\lim_{x\to x_0} \frac{f(x)}{g(x)}$ קיים והם שווים.

הוכחה אם $\lim_{x\to x_0} \frac{f'(x)}{g'(x)} = L$ אז בפרט הגבולות החד־צדדיים של $\frac{f'}{g'}$ קיימים ב־x_0 ושווים ל־L, ולכן לפי משפט 8.7.2 הגבולות החד־צדדיים של $\frac{f}{g}$ ב־x_0 קיימים ושווים ל־L, ומכאן שהגבול הדו־צדדי קיים ושווה ל־L. ∎

דוגמאות

1. נחשב את $\lim_{x\to 0+} \frac{e^x-1}{\sqrt{x}}$. בסביבה ימנית של 0 הנגזרת של המונה היא e^x ושל המכנה היא $\frac{1}{2\sqrt{x}}$, ומתקיים

$$\lim_{x\to 0+} \frac{e^x}{1/(2\sqrt{x})} = \lim_{x\to 0+} 2\sqrt{x}\cdot e^x = 0$$

לכן לפי המשפט

$$\lim_{x\to 0+} \frac{e^x-1}{\sqrt{x}} = 0$$

2. נראה ש־ $\lim_{x\to 0} \frac{\sin^2 x}{1-\cos x} = 2$. המונה והמכנה גזירים בסביבת 0 ומקיימים

$$\begin{aligned} \frac{d}{dx}(\sin^2 x) &= 2\sin x\cos x \\ \frac{d}{dx}(1-\cos x) &= \sin x \end{aligned}$$

לכן

$$\lim_{x\to 0} \frac{\frac{d}{dx}\sin^2 x}{\frac{d}{dx}(1-\cos x)} = \lim_{x\to 0} \frac{2\sin x\cos x}{\sin x} = \lim_{x\to 0} \frac{2\cos x}{1} = 2$$

ולכן

$$\lim_{x\to 0} \frac{\sin^2 x}{1-\cos x} = \lim_{x\to 0} \frac{\frac{d}{dx}\sin^2 x}{\frac{d}{dx}(1-\cos x)} = 2$$

הבהירו לעצמכם מדוע אי אפשר להשתמש כאן בטענה 8.7.1!

3. לפעמים יש להפעיל את כלל לופיטל יותר מפעם אחת כדי לחשב גבול. לדוגמה, ננסה לחשב בעזרתו את $\lim_{x\to 0}\frac{e^x-x-1}{\sin^2 x}$. נסמן $f(x)=e^x-x-1$ ו־ $g(x)=\sin^2 x$. אז f,g גזירות בסביבת 0 ומתאפסות ב־0. נחשב נגזרות:

$$\begin{aligned} f'(x) &= e^x-1 \\ g'(x) &= 2\sin x\cos x \end{aligned}$$

מכלל לופיטל נובע שאם $\lim_{x\to 0}\frac{f'(x)}{g'(x)}$ קיים אז גם הגבול המקורי קיים, והם שווים. אבל שוב אין אנו יכולים לחשב את הגבול $\lim_{x\to 0}\frac{f'(x)}{g'(x)}$ ישירות בעזרת אריתמטיקה של גבולות כי המונה והמכנה שואפים ל־0 ב־0. נחשב נגזרות שוב:

$$\begin{aligned} f''(x) &= e^x \\ g''(x) &= 2\cos^2 x-2\sin^2 x \end{aligned}$$

לכן

$$\lim_{x\to 0}\frac{f''(x)}{g''(x)}=\frac{e^0}{2\cos^2 0-2\sin^2 0}=\frac{1}{2}$$

ולכן ממשפט לופיטל

$$\lim_{x\to 0}\frac{f'(x)}{g'(x)}=\frac{1}{2}$$

ולכן שוב לפי אותו משפט

$$\lim_{x\to 0}\frac{e^x-x-1}{\sin^2 x}=\lim_{x\to 0}\frac{f(x)}{g(x)}=\frac{1}{2}$$

לרוב מקצרים את הנימוק הזה לשוויון אחד ארוך, ורושמים

$$\lim_{x\to 0}\frac{e^x-x-1}{\sin^2 x}=\lim_{x\to 0}\frac{e^x-1}{2\sin x\cos x}=\lim_{x\to 0}\frac{e^x}{2\cos^2 x-2\sin^2 x}=\frac{1}{2}$$

כאשר שני השוויונות הראשונים מוצדקים בעזרת כלל לופיטל. בשרשרת שוויונות כזאת ההצדקה היא מהסוף להתחלה: קיום הגבול מימין מצדיק את קיום הגבול משמאלו וזה מצדיק את קיום הגבול השמאלי.

מסקנה 8.7.4 (כלל לופיטל, גרסה $\frac{0}{0}$ באינסוף) יהיו f,g פונקציות ממשיות המקיימות $\lim_{x\to\infty}f(x)=0$ ו־ $\lim_{x\to\infty}g(x)$ ונניח שיש קרן (r,∞) שבה f,g גזירות ו־g' אינה מתאפסת. אם הגבול $\lim_{x\to\infty}\frac{f'(x)}{g'(x)}$ קיים אז הגבול $\lim_{x\to\infty}\frac{f(x)}{g(x)}$ קיים, והם שווים.

הוכחה נגדיר

$$F(x)=f(\frac{1}{x})\ ,\ G(x)=g(\frac{1}{x})$$

לפי כלל השרשרת אם f,g גזירות ב־$\frac{1}{x}$ אז

$$F'(x)=-\frac{1}{x^2}f'(\frac{1}{x})\ ,\ G'(x)=-\frac{1}{x^2}g'(\frac{1}{x})$$

ולכן

$$\lim_{x\to\infty}\frac{f(x)}{g(x)}=\lim_{x\to 0+}\frac{F(x)}{G(x)}=\lim_{x\to 0+}\frac{F'(x)}{G'(x)}=\lim_{x\to 0+}\frac{f'(\frac{1}{x})}{g'(\frac{1}{x})}=\lim_{x\to\infty}\frac{f'(x)}{g'(x)}$$

כאשר השוויון השני משמאל נובע ממשפט 8.7.2. קיום הגבול מימין גורר את קיום הגבול משמאל, וסיימנו. ■

לדוגמה נחשב את $\lim_{x\to\infty}x\sin\frac{1}{x^2}$. נרשום $x\sin\frac{1}{x^2}=\frac{\sin(1/x^2)}{1/x}$, ואז הגבולות של המונה והמכנה באינסוף הם 0. כמו־כן המונה והמכנה גזירים בקרן $(0,\infty)$ ולכן

$$\lim_{x\to\infty}\frac{\sin(1/x^2)}{1/x}=\lim_{x\to\infty}\frac{\cos(1/x^2)\cdot(-2/x^3)}{-1/x^2}=-2\lim_{x\to\infty}\frac{\cos(1/x^2)}{x}=0$$

כי $\cos\frac{1}{x^2}$ חסומה ו־ $\lim_{x\to\infty}\frac{1}{x}=0$.

היחס בין קצבי הגידול של פונקציות קובע את המנה שלהן לא רק כאשר הגבול שלהן הוא 0, כפי שראינו עד כה, אלא גם כאשר הוא הגבול אינסופי:

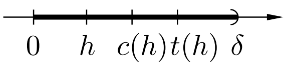
משפט 8.7.5 (כלל לופיטל, גרסה $\frac{*}{\infty}$ חד־צדדית) יהיו f,g פונקציות גזירות בסביבה ימנית כלשהי של x_0. נניח ש־ $\lim_{x\to x_0+}g(x)=\infty$. אם $\lim_{x\to x_0+}\frac{f'(x)}{g'(x)}$ קיים אז גם הגבול $\lim_{x\to x_0+}\frac{f(x)}{g(x)}$ קיים, והגבולות שווים. טענה דומה נכונה לגבול משמאל ולגבול הדו־צדדי לאחר השינויים המתבקשים בניסוח.

הערה בניגוד למקרה $\frac{0}{0}$, שם דרשנו התאפסות של המונה והמכנה (או של הגבולות שלהם), כאן דרשנו רק שהמכנה ישאף לאינסוף, ולא דרשנו דבר לגבי התנהגות המונה.

הוכחה כדי להקל על הקריאה נניח ש־ $x_0=0$. המקרה הכללי מוכח באופן דומה.

נבחר $r>0$ כך ש־f,g גזירות ב־$(0,r)$ ו־g אינה מתאפסת ב־$(0,r)$ (למה יש r כזה?). נשים לב שאם $t\in(0,r)$ ואם h קרוב מאד ל־0 אז המנה $\frac{f(h)}{g(h)}$ שווה בקירוב ל־$\frac{f(h)-f(t)}{g(h)-g(t)}$. ליתר דיוק, מתקיים

$$\frac{f(h)}{g(h)}=\frac{f(h)-f(t)}{g(h)-g(t)}\cdot(1-\frac{g(t)}{g(h)})+\frac{f(t)}{g(h)}$$



איור 8.7.1

ומכיון ש־$\lim_{h\to 0+} g(h)=\infty$ שתי המנות $\frac{g(t)}{g(h)}, \frac{f(t)}{g(h)}$ שואפות ל־0 כאשר $h\to 0+$. נוכל לכן לבחור פונקציה $t:(0,r)\to(0,r)$ באופן כזה שיתקיים

$$\lim_{h\to 0+} t(h)=\lim_{h\to 0+}\frac{f(t(h))}{g(h)}=\lim_{h\to 0+}\frac{g(t(h))}{g(h)}=0$$

הגדרה אפשרית לפונקציה t היא

$$t(h)=\min\{\tau\in(0,\frac{r}{2}]\ :\ \frac{g(\tau)}{g(h)}\le\tau\ ,\ \frac{f(\tau)}{g(h)}\le\tau\}$$

או $t(h)=\frac{r}{2}$ אם הקבוצה מימין ריקה. אנו משאירים כתרגיל את הבדיקה של־t יש את התכונות שרצינו.

נותר רק לשים לב שלפי משפט ערך הביניים של קושי, יש נקודה $c(h)\in(h,t(h))$ כך ש־

$$\frac{f(t(h))-f(h)}{g(t(h))-g(h)}=\frac{f'(c(h))}{g'(c(h))}$$

(כרגיל, ייתכן שיש יותר מנקודה אחת כזו. אז נבחר $c(h)$ אחת מביניהן). מכיוון ש־ $h\le c(h)\le t(h)$ וכיוון ש־ $\lim_{h\to 0+} t(h)=0$ גם $\lim_{h\to 0+} c(h)=0$. לכן

$$\begin{aligned}\lim_{h\to 0+}\frac{f(h)}{g(h)} &= \lim_{h\to 0+}\frac{f(h)-f(t(h))}{g(h)-g(t(h))}\cdot(1-\frac{g(t(h))}{g(h)})+\frac{f(t(h))}{g(h)}\\ &= \lim_{h\to 0+}\frac{f'(c(h))}{g'(c(h))}\cdot\lim_{h\to 0+}(1-\frac{g(t(h))}{g(h)})+\lim_{h\to 0+}\frac{f(t(h))}{g(h)}\\ &= \lim_{h\to 0+}\frac{f'(h)}{g'(h)}\end{aligned}$$

השתמשנו כאן בתכונות של t ובעובדה ש־ $\lim_{h\to 0+}\frac{f'(c(h))}{g'(c(h))}=\lim_{h\to 0+}\frac{f'(h)}{g'(h)}$ (למה זה נכון?). ■

דוגמאות

1. נחשב את הגבול $\lim_{x\to 0+} x^x$. נשים לב ש־ $x^x=e^{x\ln x}$ ולכן די לחשב את $\lim_{x\to 0+} x\ln x$. נרשום $x\ln x=\frac{\ln x}{1/x}$ ואז ממשפט לופיטל

$$\lim_{x\to 0+}\frac{\ln x}{1/x}=\lim_{x\to 0+}\frac{1/x}{-1/x^2}=-\lim_{x\to 0+} x=0$$

ולכן

$$\lim_{x\to 0+} x^x=\lim_{x\to 0+} e^{x\ln x}=\exp(\lim_{x\to 0+}\frac{\ln x}{1/x})=\exp 0=1$$

2. יהי $\alpha \in (-1,0)$. אז $\lim_{x\to 0+} x^\alpha = \infty$ וגם $\lim_{x\to 0+} e^{1/x^2} = \infty$ ולכן

$$\lim_{x\to 0}\frac{x^\alpha}{e^{1/x^2}} = \lim_{x\to 0}\frac{\alpha x^{\alpha-1}}{-2e^{1/x^2}/x^3} = -\frac{1}{2}\lim_{x\to 0}\frac{\alpha x^{\alpha+2}}{e^{1/x^2}} = 0$$

כי $\alpha + 2 > 0$. בפרט אנו מסיקים ש־

$$\lim_{x\to 0+} x^\alpha e^{-1/x^2} = 0$$

אפשר לנסח כללים מהסוג במשפט האחרון גם לגבול דו־צדדי, או כאשר $x \to \infty$. אנו משאירים את הניסוח וההוכחה שלהם כתרגיל.

לסיום, נציין שיש מקרים בהם משפט לופיטל אינו יעיל. נתבונן למשל בגבול של $\frac{x}{\sqrt{1+x^2}}$ ב־∞, שם המונה והמכנה שואפים שניהם לאינסוף. הנגזרת של המונה היא 1 והנגזרת של המכנה היא $\frac{2x}{2\sqrt{1+x^2}}$ ולכן אם ננסה להפעיל את כלל לופיטל נצטרך לחשב את הגבול $\lim_{x\to 0}\frac{\sqrt{1+x^2}}{x}$, ואם ננסה להפעיל את לופיטל גם על הביטוי מימין נחזור לגבול ממנו יצאנו (בדקו!). אפשר להפעיל את משפט לופיטל שוב ושוב ולעולם לא נגיע לביטוי חדש. מצד שני לא קשה לחשב את הגבול ישירות (חשבו!).

תרגילים

1. חשבו את הגבולות הבאים:
   (א) $\lim_{x\to a}\frac{x^n-a^n}{x-a}$.
   (ב) $\lim_{x\to 0}\frac{e^x-e^{-x}}{\sin x}$.
   (ג) $\lim_{x\to a}\frac{\arcsin x}{x}$.
   (ד) $x^{\sin x}$.
   (ה) $\lim_{x\to a}(1+x)^{1/x}$.
   (ו) $\lim_{x\to\pi/2}(\tan x)^{\sin 2x}$.
   (ז) $\lim_{x\to 0}(\cos x)^{1/x^2}$.

2. בשאלה זו נשווה את קצב ההתכנסות של פונקציות מעריכיות, פונקציות חזקה והלוגריתם.
   (א) הראו ש־ $\lim_{x\to\infty} x^\alpha e^{-x} = 0$ לכל α (והסיקו $\lim_{x\to\infty} x^\alpha e^x = 0$ לכל α).
   (ב) הראו ש־ $\lim_{x\to\infty}\frac{\ln x}{x^\alpha} = 0$ לכל $\alpha > 0$ וש־ $\lim_{x\to 0+} x^\beta \ln x = 0$ לכל $\beta > 0$.

3. יהי $a > 0$. הוכיחו $\lim_{x\to\infty} x(\sqrt[x]{a}-1) = \ln a$.

4. יהיו $f(x) = x + \sin(x)\cos(x)$, $g(x) = e^{\sin x} f(x)$. נסו לחשב את הגבול $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ בעזרת כלל לופיטל. מה משתבש?

5. נניח ש־ $f : \mathbb{R} \to \mathbb{R}$ מקיימת $\lim_{x\to\infty} f(x) + f'(x) = L$ כאשר $L \neq 0$ ממשי או $\pm\infty$. הראו ש־ $\lim_{x\to\infty} f(x) = L$ (רמז: התבוננו בפונקציה $e^x f(x)$, והיעזרו בכלל לופיטל).

6. (*) תהי f גזירה פעמיים בסביבת x_0 ו־f'' רציפה בסביבת x_0 ושונה מ־0 שם. לפי משפט לגרנג', לכל h קטן מספיק אפשר לבחור $c(h)$ כך ש־

$$\frac{f(x_0+h) - f(x_0)}{h} = f'(c(h))$$

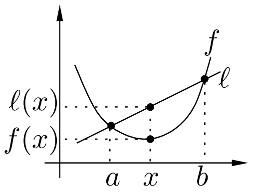
הוכיחו שבתנאים הנ"ל $c(h)$ נקבע ביחידות והפונקציה $c(h)$ מקיימת

$$\lim_{h\to 0} \frac{c(h) - x_0}{h} = \frac{1}{2}$$

מה המשמעות הגאומטרית של טענה זו?

8.8 פונקציות קמורות

בסעיף 8.6 הגדרנו קמירות בעזרת הנגזרת השנייה (ראו הגדרה 8.6.4 ואיור 8.6.2). בסעיף זה ניתן הגדרה כללית יותר. ההגדרה החדשה תאפשר לדון במושג הקמירות גם לפונקציות שאינן גזירות, אך יותר מכך, היא תקשר בין מושג הקמירות לאי־שוויון מסוים. קשר זה הופך את הקמירות לכלי מרכזי העומד מאחורי אי־שוויונות רבים, כמו למשל האי־שוויון האריתמטי־גאומטרי.



איור 8.8.1 מיתר בפונקציה קמורה נמצאת מעל הגרף

הגדרה 8.8.1 תהי f פונקציה ממשית בקטע I. אז f נקראת **קמורה** (convex) ב־I אם לכל $a < b$ ב־I המיתר ℓ העובר דרך $(a, f(a)), (b, f(b))$ נמצא מעל הגרף של f בקטע (a, b), דהיינו $\ell(x) \geq f(x)$ לכל $a < x < b$. אם לכל x כזה מתקיים $\ell(x) \leq f(x)$ אז f נקראת **קעורה** (concave) ב־I. אם האי־שוויונות הם חזקים נאמר שהפונקציה **קמורה ממש** או **קעורה ממש**, בהתאמה.

נתרכז תחילה בפונקציות קמורות, ולא בפונקציות קמורות ממש. באחרונות נדון בהמשך, ואז נראה שההגדרה של קמורות וקעירות ממש שניתנה בהגדרה 8.6.4 היא מקרה פרטי של ההגדרה לעיל.

נפתח בכמה הערות על ההגדרה. המיתר העובר דרך $(a, f(a)), (b, f(b))$ הוא

$$\ell(x) = \frac{f(b) - f(a)}{b - a}(x - a) + f(a)$$

אנו רואים ש־ f קמורה אמ"מ לכל $a<x<b$ ו־I מתקיים

$$\frac{f(b)-f(a)}{b-a}(x-a) \geq f(x)-f(a)$$

כל נקודה $x \in (a,b)$ היא מהצורה $x=\lambda a+(1-\lambda)b$ עבור מספר $\lambda \in (0,1)$ מתאים (חשבו את λ במפורש!) ולכל $\lambda \in (0,1)$ המספר $x=\lambda a+(1-\lambda)b$ נמצא בין a ל־b. לכן קמירות של f שקולה לכך שלכל $a,b \in I$ ולכל $\lambda \in (0,1)$ מתקיים

$$\ell(\lambda a+(1-\lambda)b) \geq f(\lambda a+(1-\lambda)b)$$

נשים לב ש־

$$\begin{aligned}\ell(\lambda a+(1-\lambda)b) &= \frac{f(b)-f(a)}{b-a}(\lambda a+(1-\lambda)b-a)+f(a)\\ &= \lambda f(a)+(1-\lambda)f(b)\end{aligned}$$

אנו מקבלים את האפיון הבא לקמירות:

טענה 8.8.2 f קמורה ב־I אמ"מ לכל $a,b \in I$ ולכל $\lambda \in (0,1)$ מתקיים

$$\lambda f(a)+(1-\lambda)f(b) \geq f(\lambda a+(1-\lambda)b)$$

והיא קעורה אמ"מ לכל a,b,λ כאלה מתקיים אי־השוויון ההפוך.

עבור $\lambda \in (0,1)$ המספר $\lambda a+(1-\lambda)b$ נקרא **צירוף קמור** (convex combination) של a,b, ואפשר לפרש אותו כממוצע משוקלל של a,b עם משקולת λ. למשל כאשר $\lambda=\frac{1}{2}$ זהו הממוצע הפשוט $\frac{1}{2}(a+b)$. שימו לב ש־a,b משחקים כאן תפקיד סימטרי, וכבר אין צורך לדרוש $a<b$ (אם כי אין בדרישה כזו שום נזק).

אם כן, טענה 8.8.2 מתארת את הקשר בין שתי פעולות: האחת היא פעולת המיצוע של של a,b ואחר־כך הפעלת f, והפעולה השניה היא הפעלת f על כל אחד מהנקודות a,b ואז מיצוע התוצאות. לפי הטענה, פונקציה היא קמורה אם לכל $0<\lambda<1$ תוצאת הפעולה הראשונה גדולה או שווה לתוצאה של הפעולה השניה.

דוגמאות

1. תהי $f(x)=Ax+B$. אז $f(\lambda a+(1-\lambda)b)=\lambda f(a)+(1-\lambda)f(b)$ לכל $a,b,\lambda \in \mathbb{R}$ (בדקו!), ובפרט זה נכון עבור $0<\lambda<1$. לכן f גם קמורה וגם קעורה.

2. תהי $f(x)=|x|$. אז f קמורה, כי אם $0 \le \lambda \le 1$ אז

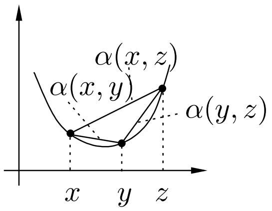
$$\begin{aligned} f(\lambda a+(1-\lambda)b) &= |\lambda a+(1-\lambda)b| \\ &\le |\lambda a|+|(1-\lambda)b| \\ &= \lambda|a|+(1-\lambda)|b| \\ &= \lambda f(a)+(1-\lambda)f(b) \end{aligned}$$

(במעבר מהשורה השנייה לשלישית השתמשנו בכך ש־ $\lambda, 1-\lambda \ge 0$).

בהמשך נוכיח משפטים על פונקציות קמורות, ונשאיר כתרגיל ניסוח והוכחה של גרסאות עבור פונקציות קעורות. המעבר בין הגרסאות קל אם משתמשים בטענה הבאה, שאף הוכחתה מושארת כתרגיל:

טענה 8.8.3 f קמורה ב־I אמ"מ $-f$ קעורה ב־I.

כדי לפשט את הסימונים נסכים שמעתה ועד סוף הסעיף אנו דנים בפונקציה ממשית f המוגדרת בקטע I, ולכל $x,y\in I$ כך ש־$x\neq y$ נסמן ב־ $\alpha(x,y)$ את השיפוע של המיתר העובר דרך הנקודות $(x,f(x))$ ו־$(y,f(y))$ (שתיהן על גרף הפונקציה).



איור 8.8.2

למה 8.8.4 יהיו $x<y<z$ ב־I. אז

$$\alpha(x,z)\le\alpha(y,z) \iff \alpha(x,y)\le\alpha(x,z) \iff \alpha(x,y)\le\alpha(y,z)$$

הוכחה באופן אינטואיטיבי הטענה ברורה: די להסתכל באיור 8.8.2 כדי להשתכנע שאם היחס בין שניים מהמיתרים הוא כמו בטענה אז היחס בין כל זוג אחר של מיתרים נקבע והוא כמו בטענה.

ההוכחה מבוססת על כך ש־$\alpha(x,z)$ ניתן לכתיבה כממוצע משוקלל (דהיינו צירוף קמור) של $\alpha(x,y)$ ו־$\alpha(y,z)$. ואמנם,

$$\begin{aligned} \alpha(x,z) &= \frac{f(z)-f(x)}{z-x} \\ &= \frac{f(z)-f(y)}{z-x}+\frac{f(y)-f(x)}{z-x} \\ &= \frac{y-x}{z-x}\frac{f(y)-f(x)}{y-x}+\frac{z-y}{z-x}\frac{f(z)-f(y)}{z-y} \\ &= \lambda\alpha(z,y)+(1-\lambda)\alpha(y,x) \end{aligned}$$

כאשר $\lambda=\frac{y-x}{z-x}$ ו־ $1-\lambda=\frac{z-y}{z-x}$ (שימו לב שמתקיים $0<\lambda<1$).

נוכיח למשל שאם $\alpha(x,y)\le\alpha(x,z)$ אז $\alpha(x,z)\le\alpha(y,z)$. ואמנם,

$$\alpha(x,z)=\lambda\alpha(y,z)+(1-\lambda)\alpha(x,y)\le\lambda\alpha(y,z)+(1-\lambda)\alpha(y,z)=\alpha(y,z)$$

הגרירות האחרות מוכחות באופן דומה. ■

טענה 8.8.5 f קמורה ב־I אמ"מ לכל $x<y<z$ ב־I מתקיים

$$\alpha(x,y)\le\alpha(x,z)\le\alpha(y,z)$$

הוכחה לפי הלמה הקודמת, התנאי $\alpha(x,y)\le\alpha(x,z)$ שקול לתנאי $\alpha(x,z)\le\alpha(y,z)$ ולכן די להראות שקמירות שקולה לכך שמתקיים $\alpha(x,y)\le\alpha(x,z)$ לכל $x,z\in I$, או באופן שקול, שלכל $x,z\in I$ ולכל $0<\lambda<1$ מתקיים $\alpha(x,\lambda x+(1-\lambda)z)\le\alpha(x,z)$.
עבור $x,z\in I$ ו־ $0<\lambda<1$ מתקיים

$$\begin{aligned}\alpha(x,\lambda x+(1-\lambda)z) &= \frac{f(\lambda x+(1-\lambda)z)-f(x)}{(\lambda x+(1-\lambda)z)-x}\\ &= \frac{f(\lambda x+(1-\lambda)z)-f(x)}{(1-\lambda)z+(1-\lambda)x}\end{aligned}$$

ומצד שני

$$\begin{aligned}\alpha(x,z) &= \frac{f(z)-f(x)}{z-x}\\ &= \frac{(1-\lambda)f(z)-(1-\lambda)f(x)}{(1-\lambda)z-(1-\lambda)x}\\ &= \frac{(\lambda f(x)+(1-\lambda)f(z))-f(x)}{(1-\lambda)z-(1-\lambda)x}\end{aligned}$$

ולכן האי־שוויון

$$\alpha(x,\lambda x+(1-\lambda)z)\le\alpha(x,z)$$

מתקיים לכל $x,z\in I$ ולכל $0<\lambda<1$, אמ"מ האי־שוויון

$$f(\lambda x+(1-\lambda)z)\le\lambda f(x)+(1-\lambda)f(z)$$

מתקיים לכל $x,z\in I$ ולכל $0<\lambda<1$. התנאי האחרון שקול לקמירות והראשון הוא התנאי בטענה, כנדרש. ■

מסקנה 8.8.6 אם f קמורה אז היא גזירה מימין ומשמאל בכל נקודה פנימית x של I, ומקיימת $f'_-(x)\le f'_+(x)$. כמו כן אם $x<y$ אז $f'_+(x)\le f'_-(y)$ (ייתכן שאחד מהאגפים באי־שוויון $f'_+(x)\le f'_-(y)$ אינו קיים במובן הצר, אך האי־שוויון נכון גם במקרה זה בתנאי שמפרשים בצורה המתבקשת אי־שוויונות בין מספרים ו־$\pm\infty$).

הוכחה תהי $x\in I$ נקודה פנימית, ו־ $\delta>0$ כך ש־f מוגדרת בקטע $[x-\delta,x+\delta]$. לכל $0\neq h\in(-\delta,\delta)$ נגדיר

$$s(h)=\frac{f(x+h)-f(x)}{h}=\alpha(x,x+h)$$

עלינו להוכיח ש־ $\lim_{h\to 0-} s(h) \le \lim_{h\to 0+} s(h)$, ובפרט שהגבולות קיימים. ואמנם, לפי טענה 8.8.5 הפונקציה s עולה, ולכן הטענה נובעת מהתכונות של גבולות חד־צדדיים של פונקציות מונוטוניות (משפט 7.10.2).

לגבי האי־שוויון השני, יהיו $x < y$ נקודות פנימיות ב־I. נגדיר $\delta = \frac{y-x}{2}$. אז מהשיקול הקודם,

$$f'_+(x) = \inf_{0<h\le\delta} \alpha(x, x+h) \le \alpha(x, x+\delta)$$

מאותה סיבה $f'_-(y) \ge \alpha(y-\delta, y)$ ולכן, מכיוון ש־ $x+\delta = y-\delta$, יוצא

$$f'_+(x) \le a(x, x+\delta) \le \alpha(y-\delta, y) \le f'_-(y)$$

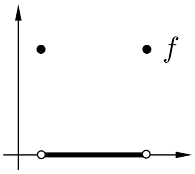
(האי־שוויון האמצעי היא מטענה 8.8.5), כפי שרצינו. ■

מסקנה 8.8.7 אם f קמורה אז היא רציפה בכל נקודה פנימית של I.

הוכחה לפי האמור למעלה f גזירה מימין ומשמאל בכל נקודה פנימית של I, ולכן רציפה שם מימין ומשמאל, ולכן רציפה שם. ■

בניגוד למקרה של נקודה פנימית של קטע, פונקציה קמורה המוגדרת בקטע אינה חייבת להיות רציפה או גזירה במובן החד־צדדי בקצות קטע (אם יש כאלה). למשל, תהי $f : [0,1] \to \mathbb{R}$ הפונקציה

$$f(x) = \begin{cases} 1 & x = 0, 1 \\ 0 & 0 < x < 1 \end{cases}$$



איור 8.8.3 פונקציה קמורה עם אי־רציפויות חד־צדדיות

הגרף של f מופיע באיור 8.8.3. קל לבדוק ש־f קמורה אך היא אינה גזירה או אפילו רציפה חד־צדדית ב־$0, 1$. כמו־כן הפונקציה $g : [0,1] \to \mathbb{R}$ הנתונה על ידי $g(x) = \sqrt{x}$ קמורה בכל הקטע (זה ינבע מהמשפט הבא) אך אינה גזירה חד־צדדית ב־0.

משפט 8.8.8 נניח ש־I קטע פתוח. אם f גזירה, אז f קמורה אמ"מ $f'(x)$ פונקציה עולה. בפרט אם f גזירה פעמיים היא קמורה אמ"מ $f'' \ge 0$.

הערה אם f רציפה בקטע I אז היא וקמורה בפנים של I אמ"מ היא קמורה בכל I. לכן התנאי שבמשפט שימושי גם לבדיקת קמירות בקטעים שאינם פתוחים.

הוכחה אם f גזירה וקמורה ב־I אז f גזירה דו־צדדית בכל נקודה $x \in I$ ולכן הנגזרות הדו־צדדיות והחד־צדדיות שוות בכל נקודה. אם $x < y$ אז לפי החלק השני של המסקנה הקודמת מתקיים $f'(x) = f'_+(x) \le f'_-(y) = f'(y)$, ולכן f' עולה.

להפך, אם f גזירה ו־f' עולה אז לפי טענה 8.8.5 די שנראה ש־ $\alpha(x,y) \le \alpha(y,z)$ לכל $x < y < z$. לפי משפט לגרנג', יש נקודות $c \in (x,y)$ ו־ $d \in (y,z)$ כך ש־

$$\alpha(x,y) = f'(c) \qquad , \qquad \alpha(y,z) = f'(d)$$

מכיוון ש־f עולה ו־ $c < y < d$ הרי $f'(c) \le f'(d)$, ומכאן ש־ $\alpha(x,y) \le \alpha(y,z)$, כנדרש.

לבסוף, אם f גזירה פעמיים אז עלייה של f' שקולה לכך שהנגזרת שלה (דהיינו הנגזרת השנייה של f) אי־שלילית. ■

למשל, הפונקציה $g : [0,1] \to \mathbb{R}$ הנתונה על ידי $g(x) = \sqrt{x}$ היא קעורה, כי היא גזירה ב־$(0,1)$ ומקיימת $g'(x) = \frac{1}{2\sqrt{x}} > 0$. לכן לפי המשפט היא קעורה בפנים של $[0,1]$ וכיוון שהיא רציפה ב־$[0,1]$ היא קעורה שם.

אם f גזירה אפשר לאפיין קמירות גם בעזרת תכונות של הישרים המשיקים לה:

משפט 8.8.9 נניח ש־I קטע פתוח ו־f גזירה. אז f קמורה ב־I אמ"מ לכל $x_0 \in I$ הישר המשיק ℓ ל־f ב־x_0 נמצא כולו מתחת לגרף של f, דהיינו מתקיים $f(x) \ge \ell(x)$ לכל $x \in I$.

הוכחה יהי ℓ הישר המשיק ל־f בנקודה $x_0 \in I$. נתבונן בפונקציית ההפרש $g(x) = f(x) - \ell(x)$. מכיוון ש־ℓ משיק ל־f ב־x_0 מתקיים

$$g(x_0) = f(x_0) - \ell(x_0) = 0 \qquad , \qquad g'(x) = f'(x) - \ell'(x) = f'(x) - f'(x_0)$$

ובפרט $g'(x_0) = 0$. אם f קמורה אז f' פונקציה עולה ולכן יש ל־g מינימום מקומי ב־x_0. לכן $g(x) \ge g(x_0)$ לכל $x \in (a,b)$, כלומר $f(x) \ge \ell(x)$.

מאידך נניח ש־f נמצאת מעל לכל משיק שלה. יהיו $x < z$ נקודות ב־I ויהי $y = \lambda x + (1-\lambda)z$ עבור מספר $0 < \lambda < 1$, כך ש־ $x < y < z$. יהי ℓ המשיק ל־f ב־y ויהי $\widetilde{\ell} : (x,z) \to \mathbb{R}$ המיתר בין $(x, f(x))$ ל־ $(z, f(z))$. לפי ההנחה $f \ge \ell$ מתקיים $\widetilde{\ell}(x) = f(x) \ge \ell(x)$ ו־ $\widetilde{\ell}(z) = f(z) \ge \ell(z)$ ולכן $\widetilde{\ell} \ge \ell$ בכל הקטע (x,z), ובפרט

$$\lambda f(x) + (1-\lambda) f(z) = \widetilde{\ell}(y) \ge \ell(y) = f(y) = f(\lambda x + (1-\lambda)z)$$

מכיוון ש־ x, z, λ היו שרירותיים, זה מוכיח ש־f קמורה ב־I. ■

נעיר, שעל מנת להסיק ש־f קמורה, לא די שתהיה נקודה P אחת על הגרף של f כך שכל הגרף נמצא מעל המשיק העובר דרך P (מצאו דוגמה!).

עד כאן חקרנו תכונות בסיסיות של פונקציות קמורות ומצאנו אפיונים שונים שלהן. נעבור לדון בשימושים.

קמירות של פונקציה בקטע היא תכונה המקשרת בין הערך של f בממוצע המשוקלל של x, y, לבין הממוצע המשוקלל של הערכים של f בנקודות x ו־y. אפשר להכליל את האי־שוויון למקרה של ממוצעים של מספר רב יותר של נקודות.

הגדרה 8.8.10 סדרה סופית של מספרים $(\lambda_1,\ldots,\lambda_n)$ נקראת **מערכת משקולות**[10] אם $\lambda_i \geq 0$ ו־ $\sum_{i=1}^n \lambda_i = 1$. בהינתן מערכת משקולות $\lambda_1,\ldots,\lambda_n$ ונקודות $x_1,\ldots x_n$, **הממוצע המשוקלל** (weighted average) של $x_1,\ldots,x_n$, או **הצירוף הקמור** (convex combination) שלהם, ביחס למשקולות הנתונות, הוא $\sum_{i=1}^n \lambda_i x_i$.

למשל, הממוצע הפשוט $\frac{1}{n}(x_1+x_2+\ldots+x_n)$ של n מספרים $x_1,\ldots,x_n$ הוא ממוצע משוקלל ביחס למערכת המשקולות $(\frac{1}{n},\frac{1}{n},\ldots,\frac{1}{n})$.

שימו לב שאם $0 \leq \lambda \leq 1$ אז הזוג $(\lambda, 1-\lambda)$ הוא מערכת משקולות, והצירוף הקמור $\lambda x + (1-\lambda)y$ הוא הממוצע המשוקלל של x,y ביחס למערכת משקולות זו.

משפט 8.8.11 (אי־שוויון ינסן)[11] אם f קמורה בקטע I אז לכל מערכת משקולות $(\lambda_1,\ldots,\lambda_n)$ ולכל סדרת נקודות $x_1,\ldots,x_n \in I$ מתקיים

$$f(\sum \lambda_i x_i) \leq \sum \lambda_i f(x_i)$$

הוכחה נוכיח את המשפט באינדוקציה על n. ליתר דיוק, אנו נראה באינדוקציה על $n \geq 2$ את הטענה שלכל מערכת משקולות $(\lambda_1,\ldots,\lambda_n)$ ולכל $x_1,\ldots,x_n \in I$ מתקיים האי־שוויון המבוקש. עבור $n=2$ האי שוויון אינו אלה הגדרת הקמורות ולכן מתקיימת לפי ההנחה ש־f קמורה.

נניח שהדבר נכון עבור $n-1$. תהי $(\lambda_i)_{i=1}^n$ מערכת משקולות ו־ $x_1,\ldots,x_n \in I$. נקודת המפתח בטיעון היא שאת המספר $y=\sum_{i=1}^n \lambda_i x_i$ ניתן להציג כצירוף קמור של x_n עם המספר $z=\sum_{i=1}^{n-1}\gamma_i x_i$, כאשר z הוא ממוצע משוקלל של $x_1,\ldots,x_{n-1}$ ביחס למערכת משקולות אחרת. ואמנם אם נגדיר

$$\gamma_i = \frac{\lambda_i}{\sum_{j=1}^{n-1}\lambda_j}$$

אז $(\gamma_1,\ldots,\gamma_{n-1})$ בבירור מערכת משקולות. כמו־כן, כיוון ש־ $\sum_{i=1}^{n-1}\lambda_i = 1-\lambda_n$ הרי אם נסמן $z=\sum_{i=1}^{n-1}\gamma_i x_i$ אז $(1-\lambda_n)z=\sum_{i=1}^{n-1}\lambda_i x_i$, ולכן

$$y=(1-\lambda_n)z+\lambda_n x_n$$

ומכיוון ש־f קמורה מקבלים

$$f(\sum_{i=1}^n \lambda_i x_i) = f((1-\lambda_n)z+\lambda_n x_n) \leq (1-\lambda_n)f(z)+\lambda_n f(x_n)$$

[10] לעתים קרובות מערכת משקולות נקראת גם **וקטור הסתברות**.
[11] Johan Jensen, 1859־1925.

אבל מהנחת האינדוקציה

$$f(z)=f(\sum_{i=1}^{n-1}\gamma_i x_i)\leq\sum_{i=1}^{n-1}\gamma_i f(x_i)$$

ויחד עם האי־שוויון הקודם קיבלנו

$$f(\sum_{i=1}^{n}\lambda_i x_i)\leq(1-\lambda_n)\sum_{i=1}^{n-1}\gamma_i f(x_i)+\lambda_n f(x_i)=\sum_{i=1}^{n-1}\lambda_i f(x_i)+\lambda_n f(x_n)$$

(בשוויון האחרון הצבנו את הגדרת γ_i). זהו האי־שוויון המבוקש. ■

ישנה גרסה של משפט ינסן לפונקציות קעורות: אם f קעורה אז לכל מערכת משקולות $(\lambda_1,\ldots,\lambda_n)$ ולכל סדרת נקודות $(x_1,\ldots x_n)$ מתקיים

$$f(\sum_{i=1}^{n}\lambda_i x_i)\geq\sum_{i=1}^{n}\lambda_i f(x_i)$$

כדוגמה נשתמש באי־שוויון ינסן כדי לתת הוכחה קצרה לאי־שוויון הממוצעים. כזכור, אי־שוויון הממוצעים קובע שאם $a_1,\ldots,a_n$ ממשיים חיוביים, אז

$$(a_1\cdot\ldots\cdot a_n)^{1/n}\leq\frac{a_1+\ldots+a_n}{n}$$

מכיוון ש־ln פונקציה עולה הרי שאי־שוויון זה נכון אמ"מ הוא נכון לאחר שמפעילים על שני אגפיו את הפונקציה ln (שימו לב ששני האגפים חיוביים!) לכן די להראות ש־

$$\ln\left((a_1\cdot\ldots\cdot a_n)^{1/n}\right)\leq\ln\left(\frac{a_1+\ldots+a_n}{n}\right)$$

ומתכונות הלוגריתם זה שקול ל־

$$\sum_{i=1}^{n}\frac{1}{n}\ln a_i\leq\ln(\sum_{i=1}^{n}\frac{1}{n}a_i)$$

נשים לב ש־ $(\lambda_1,\ldots,\lambda_n)=(\frac{1}{n},\ldots,\frac{1}{n})$ היא מערכת משקולות. לכן אם נראה ש־ln פונקציה קעורה האי־שוויון ינבע ממשפט ינסן. אבל $\ln''(x)=-\frac{1}{x^2}$ ולכן $\ln'' x<0$ לכל $x>0$. מכאן ש־ln קעורה (זו הגרסה לפונקציות קעורות של משפט 8.8.8).

באותו אופן אפשר להוכיח גרסה כללית של אי־שוויון הממוצעים: אם (λ_i) מערכת משקולות ואם $a_i\geq 0$ אז

$$a_1^{\lambda_1}\cdot\ldots\cdot a_n^{\lambda_n}\leq\lambda_1 a_1+\ldots+\lambda_n a_n$$

(השלימו את הפרטים!).

גרסאות של משפטים רבים מסעיף זה תקפים לפונקציות קמורות ממש במובן של הגדרה 8.8.1. לעתים ניסוחם זהה, ולעתים יש להפוך אי־שוויון חלש לחזק כדי להתאים אותם למקרה החדש. למשל, האי־שוויונות במשפט 8.8.5 הופכים לאי־שוויונות חזקים כאשר הפונקציה הנתונה קמורה ממש, ופונקציה גזירה היא קמורה ממש בקטע אמ"מ הנגזרת שלה עולה ממש (מכאן שקמירות ממש במובן של הגדרה 8.6.4 שקולה, עבור פונקציות גזירות פעמיים, להגדרה 8.8.1). גם אי־שוויון ינסן נכון בגרסה עם אי־שוויון חזק, בתנאי שדורשים שיהיו לפחות שתי נקודות שונות בעלות משקולות שאינן אפס. לעומת זאת לפעמים נשאר אי־שוויון חלש: למשל, באופן כללי האי־שוויון $f'_-(x) \leq f'_+(x)$ נשאר אי־שוויון חלש גם כש־f קמורה ממש. אנו משאירים לכם כתרגיל את המשימה לעיין בהוכחות ולהחליט מה הניסוח הנכון של המשפטים עבור פונקציות קמורות ממש.

נסיים בתכונה חשובה נוספת של פונקציות קמורות וקעורות:

טענה 8.8.12 תהי f פונקציה קעורה בקטע I. אז כל מקסימום מקומי של f הוא מקסימום גלובלי, ואם f קעורה ממש ויש לה מקסימום אז הוא מתקבל בנקודה יחידה. טענה דומה נכונה עבור נקודות מינימום של פונקציות קמורות.

הוכחה תהי f קעורה. יהיו $x_1 < x_2$ נקודות מקסימום מקומי ונניח בשלילה ש־$f(x_1) < f(x_2)$ (אם $f(x_1) > f(x_2)$ ההוכחה דומה). יהי ℓ המיתר העובר דרך הנקודות $(x_1, f(x_1))$ ו־$(x_2, f(x_2))$. אז ℓ עולה בקטע (x_1, x_2). מקעירות נובע שהמיתר נמצא מתחת לפונקציה, ויוצא שלכל $x \in (x_1, x_2)$ מתקיים

$$f(x) \geq \ell(x) \geq \ell(x_1) = f(x_1)$$

(האי־שוויון הראשון מקעירות והשני כי $x > x_1$ ו־ℓ עולה). אבל זה עומד בסתירה לכך ש־x_1 מקסימום מקומי.

ההוכחה של החלק השני של הטענה דומה: אילו היו שתי נקודות מקסימום אז הן היו שוות והמיתר ביניהן היה מאוזן. כיוון שהפונקציה קעורה ממש הפונקציה נמצאת ממש מעל המיתר, ולכן מקבלת ערך גדול יותר מהמקסימום, סתירה. ■

תרגילים

1. האם יש פונקציה שהיא גם קמורה וגם קעורה?
2. לאילו ערכים של α הפונקציה x^α קמורה ב־$(0, \infty)$, ולאילו קעורה?
3. הראו שאם f קמורה ואם g קמורה ועולה אז $g \circ f$ קמורה. מה קורה אם g יורדת? מה אם היא אינה מונוטונית?

4. הוכיחו שאם פונקציה קמורה בקטע (a,b) עולה בתת־קטע (a,c) כלשהו אז היא עולה בכל הקטע. האם ייתכן שהיא יורדת ב־(a,c) אך אינה יורדת בכל הקטע? מה קורה אם הפונקציה קעורה במקום קמורה?

5. הוכיחו שאם $f:\mathbb{R}\to\mathbb{R}$ קמורה וחסומה מלעיל אז f קבועה.

6. נניח ש־f רציפה בקטע I וש־ $f(\frac{1}{2}x+\frac{1}{2}y)\le\frac{1}{2}f(x)+\frac{1}{2}f(y)$ לכל $x<y$ ב־I. הוכיחו ש־f קמורה ב־I.

7. הראו שלפונקציה קעורה ממש יכולות להיות לכל היותר שתי נקודות מינימום בקטע (השוו עם טענה 8.8.12).

8. תהי $(\lambda_1,\ldots,\lambda_n)$ מערכת משקולות, יהיו $x_1,\ldots,x_n$ מספרים אי־שליליים ויהי $x=\sum\lambda_i x_i$ הממוצע המשוקלל שלהם. יהי $\alpha>0$, ותהי $I\subseteq\{1,\ldots,n\}$ קבוצת האינדקסים כך ש־ $x_i>\alpha$. הוכיחו את **אי־שוויון מרקוב** (Markov inequality): $\sum_{i\in I}\lambda_i<\frac{x}{a}$. הסיקו שאם $\sum_{i=1}^n\lambda_i x_i<\varepsilon$ אז

$$\sum_{i\,:\,x_i>\sqrt{\varepsilon}}\lambda_i<\sqrt{\varepsilon}$$

9. בשאלה זו נוכיח את **אי־שוויון הלדר** (Hölder inequality): לכל שתי n־יות של מספרים אי־שליליים $x_1,\ldots,x_n$ ו־$y_1,\ldots,y_n$ ולכל זוג של מספרים $\alpha,\beta>1$ כך ש־ $\frac{1}{\alpha}+\frac{1}{\beta}=1$ מתקיים

$$\sum_{i=1}^n x_i y_i\le(\sum_{i=1}^n x_i^{\alpha})^{1/\alpha}\cdot(\sum_{i=1}^n y_i^{\beta})^{1/\beta}$$

זו הכללה של אי־שוויון קושי־שוורץ (תרגיל (5) בעמוד 78). גם ההוכחה דומה: תחילה הוכיחו שלכל $x,y\ge 0$ מתקיים $xy\le\frac{1}{\alpha}x^{\alpha}+\frac{1}{\beta}y^{\beta}$ (זו הכללה של האי־שוויון $xy\le\frac{1}{2}x^2+\frac{1}{2}y^2$). הראו שדי להוכיח את אי־שוויון הלדר תחת ההנחה ש־ $\sum x_i^{\alpha}=\sum y_i^{\beta}=1$, הופכת את אי־שוויון הלדר ל־ $\sum x_i y_i\le 1$, והוכיחו זאת.

הסיקו גם את הגרסה הכללית הבאה של אי־שוויון הלדר: אם $(\lambda_1,\ldots,\lambda_n)$ מערכת משקולות אז

$$\sum_{i=1}^n\lambda_i\cdot x_i y_i\le(\sum_{i=1}^n\lambda_i\cdot x_i^{\alpha})^{1/\alpha}\cdot(\sum_{i=1}^n\lambda_i\cdot y_i^{\beta})^{1/\beta}$$

(אין צורך לעבוד קשה, זה נובע ממה שהוכחתם קודם).

10. יהי n טבעי. הוכיחו שלכל מערכת משקולות $p_1,\ldots,p_n$ מתקיים

$$-\sum_{i=1}^n p_i\ln p_i\le n\ln n$$

(נגדיר $0\ln 0 = 0$ כדי שהסכום יהיה מוגדר גם כשחלק מה־p_i־ים הם אפס). הראו שיש שוויון אמ"מ $p_i = \frac{1}{n}$ (רמז: התבוננו בפונקציה $x\ln x$ והיעזרו בטענה 8.8.12).

11. תהי $f : [0,1] \to \mathbb{R}$ פונקציה קמורה.
   (א) נגדיר $a_n = \frac{1}{2^n}\sum_{k=1}^{2^n} f(\frac{k}{2^n})$. הוכיחו ש־$(a_n)$ מתכנסת.
   (ב) (*) נגדיר $b_n = \frac{1}{n}\sum_{k=1}^{n} f(\frac{k}{n})$. הוכיחו שגם (b_n) מתכנסת ושמתקיים $\lim a_n = \lim b_n$.

12. תהי $f : I \to \mathbb{R}$ פונקציה. עבור $x \in I$ נגדיר

$$\widetilde{f}(x) = \sup\{\ell(x) \; : \; \ell \text{ ישר שמקיים } \ell \leq f \text{ ב־} I\}$$

$\widetilde{f}$ נקראת **המעטפת הקמורה** (convex hull) של f. הוכיחו ש־$\widetilde{f}$ קמורה ב־I, ושאם g פונקציה קמורה ב־I המקיימת $g \leq f$ אז $g \leq \widetilde{f}$.

13. הסיקו מטענה 8.8.6 שלפונקציה קמורה בקטע יש לכל היותר מספר בן מנייה של נקודות אי־גזירות (היעזרו בתרגיל (8) בעמוד 281).

14. (*) הוכיחו על סמך התכונות החשבוניות של האקספוננט (משפט 4.6.1) ש־e^x קמורה ב־$\mathbb{R}$. הסיקו בעזרת השאלה הקודמת שיש לפונקציה e^x נקודות גזירות, והראו שמכאן נובע שהיא גזירה בכל נקודה.

8.9 שיטת ניוטון־רפסון למציאת שורשים של פונקציה

אחת הבעיות הבסיסיות והקשות במתמטיקה היא כיצד למצוא שורש של פונקציה. כלומר, בהינתן פונקציה f מחפשים מספר r כך ש־ $f(r) = 0$. בעיה זו כללית מאד: למשל כל בעיה של "פתרון משוואה" היא מסוג זה, שכן על ידי העברת אגפים תמיד אפשר להביא את המשוואה לצורה $f(x) = 0$ לפונקציה f כלשהי. כך למשל מציאת x שמקיים $x^5 = e^{\sqrt{x}}$ שקולה לבעיה של מציאת שורש לפונקציה $f(x) = e^{\sqrt{x}} - x^5$.

בהינתן נוסחה מפורשת עבור f היינו רוצים לקבל נוסחה מפורשת לשורש של f. אם הדבר אפשרי אומרים שלבעיה יש **פתרון אנליטי**. אולם באופן כללי אין פתרון כזה, גם כאשר קיום השורש ידוע.[12] לאור המציאות הזו אנו נאלצים להסתפק בשיטות למציאת קירובים של שורשים, והבעיה הופכת לבעיה במתמטיקה יישומית: כיצד למצוא קירובים טובים באופן יעיל.

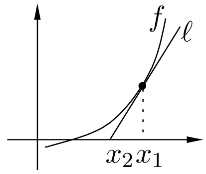
כבר פגשנו שיטה אחת לקירוב שורשים כחלק ממשפט ערך הביניים לפונקציות רציפות: זו שיטת החצייה (ראו תרגיל (15) בעמוד 275). בקצרה, מתחילים מקטע

[12] אחת התוצאות המפורסמות באלגברה היא שאף על פי שקיימת נוסחה כללית עבור השורשים של פולינומים ממעלה $1,2,3,4$, לא קיימת נוסחה כזאת לפולינומים ממעלה 5 ומעלה. לעומת זאת לכל פולינום ממעלה אי־זוגית יש שורש, גם אם לא נוכל לרשום לו נוסחה. מכאן שיש מקרים בהם אין ברירה אלא להסתפק בקירוב של השורש.

I_1 כך שהערכים של f בקצוות שלו בעלי סימנים מנוגדים, דבר המבטיח שיש ל־f שורש בקטע. נחצה כעת את הקטע לשני תת־קטעים. בקצוות של אחד מהתת־קטעים f ערכי הפונקציה בעלי סימנים מנוגדים, ולכן באותו תת־קטע יש שורש. נקרא לקטע זה I_2. נחצה את I_2 לשניים ונבחר את התת־קטע שבו מובטח שיהיה שורש. מקבלים כך סדרה יורדת של קטעים סגורים שאורכיהם קטנים פי שתיים בכל שלב, והחיתוך שלהם מכיל נקודה בודדת r, שהיא השורש המבוקש.

לשיטה זו היתרון שבכל שלב יש בידינו קטע המכיל את השורש, ולכן אנו יודעים להעריך אותו מלמעלה ומלמטה. אם נסמן $I_n = [a_n, b_n]$, נוכל לומר בוודאות שהמרחק של a_n מהשורש מקיים $|a_n - r| \leq 2^{-n+1}|I_1|$, וזאת כי r שייך לקטע $[a_n, b_n]$ שאורכו $|I_1| \cdot 2^{-n+1}$. פירוש הדבר שאם נייצג את a_n כשבר בינרי (בבסיס שתיים), אז במעבר מ־a_n ל־a_{n+1} אנו מוסיפים עוד ספרה אחת שלגביה יש לנו וודאות שהיא נכונה.

יש שיטות קירוב המבטיחות קצב התכנסות טוב בהרבה משיטת החצייה. אנו נתאר כאן אחת הוותיקות שבהן הנקראת שיטת ניוטון־רפסון[13] (Newton-Raphson).



איור 8.9.1 שיטת ניוטון־רפסון

הרעיון מאחורי השיטה פשוט מאד. נניח ש־f פונקציה המוגדרת בקטע I ויש לה שורש בנקודה פנימית r של הקטע, דהיינו $f(r) = 0$. יהי $x_1 \in I$. אנו רוצים למצוא נקודה x_2 קרובה יותר ל־r מאשר x_1. ובכן, נחשב את הישר המשיק ℓ לגרף בנקודה $(x_1, f(x_1))$. אנו יודעים ש־ℓ מקרב במידה מסוימת את f ובהיעדר מידע אחר תהיה זו הנחה סבירה שהשורש (היחיד) של ℓ קרוב לשורש r של f. ראו איור 8.9.1.

יהי ℓ המשיק לגרף של f בנקודה x_1. נסמן ב־x_2 את השורש של ℓ. קל למצוא נוסחה עבור x_2: הנוסחה של ℓ היא

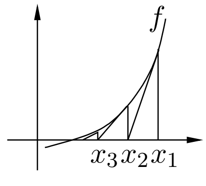
$$\ell(t) = f'(x_1)(t - x_1) + f(x_1)$$

ולכן אם $f'(x_1) \neq 0$ אז יש ל־ℓ שורש יחיד x_2 והוא נתון על ידי

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

לכן אם רוצים למצוא קירובים של r אפשר להציע את התהליך הבא: נקבע x_1 ונגדיר ברקורסיה

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$



איור 8.9.2 סדרת נקודות המיוצרת בשיטת ניוטון־רפסון

ראו איור 8.9.2.

באופן כללי לא מובטח שהסדרה הזו תתכנס. למשל אם בחרנו את a_1 בצורה גרועה יכול להיות ש־ $f'(x_1) = 0$ ואז הנוסחה לא מוגדרת (כי המשיק מאוזן וממילא אין

[13] Joseph Raphson, 1648-1716 בקירוב.

לו שורש). לא קשה למצוא דוגמאות בהן התהליך דווקא מרחיק אותנו מ־x_0 או שהסדרה (x_n) מתנהגת בצורה מחזורית ואינה מתקרבת ל־r. אך תחת תנאים חלשים למדי אפשר להראות שאם הסדרה מתכנסת אז הגבול הוא r:

טענה 8.9.1 אם f בעלת נגזרת רציפה ב־I ואם הסדרה (x_n) שהוגדרה למעלה מתכנסת לנקודה $x_0 \in I$ אז x_0 הוא שורש של f.

הוכחה נסמן $x_0 = \lim x_n$. אז

$$x_0 = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty}(x_n - \frac{f(x_n)}{f'(x_n)}) = x_0 - \lim_{n\to\infty}\frac{f(x_n)}{f'(x_n)}$$

כאשר קיום הגבול באגף שמאל מבטיח שהגבול מימין קיים. על ידי העברת אגפים רואים ש־ $\lim_{n\to\infty}\frac{f(x_n)}{f'(x_n)} = 0$. מהרציפות של f ו־f' אנו יודעים ש־ $f(x_n) \to f(x_0)$ ו־ $f'(x_n) \to f'(x_0)$, ולכן הגבול $\lim_{n\to\infty}\frac{f(x_n)}{f'(x_n)} = 0$ ייתכן רק אם $f(x_0) = \lim_{n\to\infty} f(x_n) = 0$, כנדרש. ■

נותר למצוא תנאים שיבטיחו שהסדרה (x_n) מתכנסת.

משפט 8.9.2 תהי f גזירה ברציפות, עולה ממש וקמורה בקטע I. יהי $r \in I$ כך ש־ $f(r) = 0$. יהי $x_1 > r$ ב־I ונגדיר את (x_n) לפי הנוסחה $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$. אז $x_n \to r$.

הוכחה אם נראה ש־(x_n) מתכנסת אז הגבול שלה הוא שורש של f לפי הטענה הקודמת. כיוון ש־f עולה ממש r הוא השורש היחיד של f, ונוכל אז להסיק ש־ $x_n \to r$.

על מנת להוכיח ש־(x_n) מתכנסת, די שנראה שהיא יורדת וחסומה מלרע ע"י איבר ב־I. לשם כך נראה באינדוקציה שאם $x_n > r$ ואם $x_{n+1} \neq r$ אז $r < x_{n+1} < x_n$. מצד שני מהגדרת הסדרה ברור שאם $x_n = r$ לאיזהשהו n אז $x_k = r$ לכל $k > n$. לכן יוצא שהסדרה (x_n) יורדת וחסומה מלרע על ידי r, ובפרט היא שייכת כולה לרטה $[r, x_1]$, ולכן ל־I, כפי שרצינו.

נניח אם כן ש־ $x_n > r$ וש־ $x_{n+1} \neq r$. מאחר ש־f עולה ממש הרי ש־ $f(x_n) > f(r) = 0$, וממילא $f'(x_n) > 0$ (שוב כי f עולה ממש). נסמן ב־ℓ את הישר המשיק לגרף של f ב־x_n. השיפוע של ℓ הוא $f'(x_n)$, שהוא מספר חיובי, ולכן ℓ עולה ממש. המספר x_{n+1} נבחר כך ש־ $\ell(x_{n+1}) = 0$. אבל $\ell(x_n) = f(x_n) > 0$, ולכן $x_{n+1} < x_n$ (כי ℓ עולה ממש). מצד שני, כיוון ש־f קמורה מתקיים $f \geq \ell$ ב־I (משפט 8.8.9), ובקטע (x_{n+1}, x_n) הפונקציה ℓ חיובית. לכן גם f חיובית באותו קטע. בפרט $r \notin (x_{n+1}, x_n)$, ויחד עם הנחת האינדוקציה $r < x_n$ נובע $r \leq x_{n+1}$. כיוון שהנחנו שלא מתקיים שוויון, חייב להתקיים האי־שוויון $x_{n+1} > r$, כנדרש. ■

במקרים רבים ההנחה $x_1 > r$ אינה חיונית. אם $x_1 < r$ אז תחת תנאי המשפט $x_2 > r$ (הוכיחו זאת!). לכן אם $x_2 \in I$ אז הסדרה $x_2, x_3, \ldots$ מתכנסת ל־r לפי

המשפט. נוכל אם כן לחזק את המשפט באופן הבא: אם $x_1 \in I$ ואם $x_2 \in I$ אז הסדרה (x_n) מתכנסת ל־r.

המשפט נכון גם עבור פונקציות קעורות עולות ממש עם השינויים המתבקשים בניסות. יש גרסאות מקבילות לפונקציות יורדות.

דוגמאות

1. תהי $f(x) = x^2 - 2$ ו־ $I = [0, \infty)$. השורש של f בקרן I הוא $\sqrt{2}$. כיוון ש־ $f'(x) = 2x$, $f''(x) = 2$ הפונקציה f מקיימת את תנאי המשפט, ולכל $x_1 > 0$ הסדרה (x_n) המוגדרת במשפט תתכנס ל־$\sqrt{2}$. הנוסחה עבור x_n במקרה זה היא

$$x_{n+1} = x_n - \frac{x_n^2 - 2}{2x_n} = \frac{1}{2}(x_n + \frac{2}{x_n})$$

קיבלנו הוכחה חדשה להתכנסות של הסדרה מדוגמה (3) בעמוד 124.

2. לכל $0 < x_1 < 1$, הסדרה $x_{n+1} = 2x_n - x_n \ln x_n$ מתכנסת לשורש של $\ln x - 1$, כלומר ל־e (בדקו את הפרטים!).

כפי שכבר ציינו, אחד המאפיינים החשובים של שיטת ניוטון־רפסון הוא שאפשר בתנאים מסוימים להעריך את קצב ההתכנסות של הסדרה (x_n) לשורש r, ולהראות שהקצב מהיר ביותר. אנו רוצים להעריך את ההפרש

$$x_{n+1} - r = x_n - r - \frac{f(x_n)}{f'(x_n)}$$

ואנו מקווים להראות שהוא קטן ביחס להפרש $x_n - r$. השוויון הקודם שקול ל־

$$x_{n+1} - r = (x_n - r) \cdot \frac{(x_n - r)f'(x_n) - f(x_n)}{(x_n - r)f'(x_n)}$$

ולכן די שהעריך את המנה

$$\frac{g(x)}{h(x)} = \frac{(x-r)f'(x) - f(x)}{(x-r)f'(x)}$$

נשים לב שהפונקציות g, h במונה ובמכנה מתאפסות ב־r, ולכן אנו בעצם רוצים להעריך את המנה $\frac{g(x)-g(r)}{h(x)-h(r)}$. בהנחה ש־$h'$ אינה מתאפסת בין r ל־x נוכל להיעזר במשפט ערך הביניים של קושי 8.5.10 ולהסיק שיש נקודה c בין x ל־r כך ש־

$$\frac{(x-r)f'(x) - f(x)}{(x-r)f'(x)} = \frac{g'(c)}{h'(c)}$$

חישוב הנגזרות של g, h נותן

$$\frac{g'(c)}{h'(c)} = \frac{1 \cdot f'(c) + (c-r)f''(c) - f'(c)}{1 \cdot f'(c) + (c-r)f''(c)} = (c-r)\frac{f''(c)}{f'(c) + (c-r)f''(c)}$$

וקיבלנו שיש c בין r ל־x_n כך ש־

$$\begin{aligned} |x_{n+1} - r| &= |x_n - r| \cdot |c - r| \frac{f''(c)}{f'(c) + (c-r)f''(c)} \\ &\leq |x_n - r|^2 \frac{f''(c)}{f'(c) + (c-r)f''(c)} \end{aligned}$$

כי $|c - r| \leq |x_n - r|$. מכאן אנו מקבלים

משפט 8.9.3 תהי f פונקציה גזירה, קמורה ועולה ממש בקטע I וגזירה פעמיים ברציפות ב־I. יהי $r \in I$ שורש של f. אז יש $\delta > 0$ וקבוע C כך שלכל $x_1 \in (r, r+\delta)$ הסדרה המוגדרת על ידי $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ מקיימת $|r - x_{n+1}| \leq C|r - x_n|^2$.

הוכחה נשים לב שהפונקציה $\frac{f''(x)}{f'(x)+(x-r)f''(x)}$ אינה מתאפסת ב־r ורציפה ב־r, ולכן אפשר לבחור $\delta > 0$ כך שהמספר

$$C = \max_{x \in (r, r+\delta)} \left| \frac{f''(x)}{f'(x) + (x-r)f''(x)} \right|$$

קיים (וסופי). עבור $x_1 \in (r, r+\delta)$ ראינו שהסדרה (x_n) יורדת ומוכלת ב־$(r, r+\delta)$ ואז האי־שוויון $|r - x_{n+1}| \leq C|r - x_n|^2$ נובע מהגדרת C ומההערכה שקיבלנו לפני המשפט. ■

נסמן

$$\varepsilon_n = x_0 - x_n$$

לפי משפט 8.9.2 מתקיים $\varepsilon_n \to 0$ ולכן עבור n גדול מספיק יתקיים $\varepsilon_n < (\frac{1}{C})^2$. אז

$$\begin{array}{ccccc} \varepsilon_{n+1} & \leq & C(\varepsilon_n)^2 & < & (\frac{1}{C})^3 \\ \varepsilon_{n+2} & \leq & C(\varepsilon_{n+1})^2 & < & (\frac{1}{C})^5 \\ \vdots & & \vdots & & \vdots \end{array}$$

ובאינדוקציה לכל k מתקיים

$$\varepsilon_{n+k} \leq (\frac{1}{C})^{1+2^k}$$

זהו קצב התכנסות מהיר ביותר. כדי לקבל תחושה למתרחש נניח לשם פשטות ש־$C = 10$ ושהאינדקס n למעלה הוא 1. אז האי־שוויון הוא $\varepsilon_n < (\frac{1}{10})^{1+2^{n-1}}$, ופירושו שהפיתוחים העשרונים של x_n ו־r מסכימים לפחות ב־2^{n-1} הספרות הראשונות! השווה זאת עם שיטת החצייה למציאת שורשים, שבה רק מובטח שאחרי n צעדים הקירוב מסכים עם r בערך ב־ $n \log_2 10$ ספרות עשרוניות.

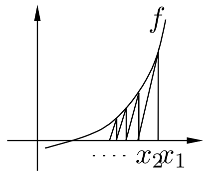
תרגילים

1. תהי f גזירה, עולה ממש וקמורה בקטע I. יהי $r \in I$ שורש, ונניח שמתחילים את תהליך ניוטון־רפסון ממספר $x_1 < r$. הראו ש־ $x_2 > r$.

2. חשבו את $\sqrt{2}$ בדיוק של שש ספרות אחרי הנקודה.

3. לכל k טבעי ולכל $x > 0$ מצאו הגדרה רקורסיבית לסדרה המתכנסת ל־$\sqrt[k]{x}$ (על ההגדרה להכיל רק את הפעולות האריתמטיות הבסיסיות: חיבור, חיסור, כפל וחילוק).

4. הנה גרסה של שיטת ניוטון־רפסון שהיא יעילה פחות אך קלה יותר לניתוח. תהי f פונקציה גזירה ברציפות בקטע I ונניח ש־ $M > 0$ הוא חסם של f' ב־I. נניח גם שיש שורש r של f ב־I. תהי $x_1 > r$ נקודה ב־I ונגדיר



איור 8.9.3

$$x_{n+1} = x_n - \frac{f(x_n)}{M}$$

(באיור 8.9.3 מתוארות שלושה נקודות הראשונות בסדרה. השוו זאת עם איור בעמוד 366).

(א) הוכיחו ש־(x_n) סדרה מונוטונית וחסומה מלרע על ידי r.

(ב) נסמן $x_0 = \lim x_n$. הראו ש־ $f(x_0) = 0$ (האם בהכרח $x_0 = r$?).

(ג) (*) האם קצב ההתכנסות של שיטה זו טוב כמו זה של שיטת ניוטון־רפסון?

8.10 מספרים אלגבריים ומספרים טרנסצנדנטיים

ראינו בפרקים הקודמים שיש מספרים ממשיים שאינם רציונליים: למשל $\sqrt{2}$ (סעיף 4.1) או המספר $\sum_{k=1}^{\infty} \frac{1}{10^{k^2}}$ (סעיף 6.7). אנו יודעים גם שיש הרבה מספרים אי־רציונליים: הם צפופים ב־$\mathbb{R}$, ויותר מכך, ראינו בסעיף 5.12 שקבוצת המספרים האי־רציונליים אינה בת מניה, ולכן במובן מסוים היא גדולה יותר מ־$\mathbb{Q}$.

למרות ריבוי המספרים האי־רציונליים עדיין אפשר היה לשער שקיים תיאור אלגברי למספרים הממשיים. בסעיף זה נראה שאין הגדרה כזו, ושיש מספרים ממשיים שאינם ניתנים לאפיון אלגברי. אפילו נצביע על מספר כזה.

הגדרה 8.10.1 פולינום $p(x) = \sum_{k=0}^{d} r_k x^k$ בעל מקדמים שלמים ($r_0, \ldots, r_d \in \mathbb{Z}$) נקרא **פולינום שלם** (integer polynomial). מספר $\alpha \in \mathbb{R}$ נקרא **אלגברי** (algebraic) אם יש פולינום שלם p שאיננו פולינום האפס כך ש־ $p(\alpha) = 0$.

דוגמאות

1. כל מספר שלם הוא אלגברי כי אם $n \in \mathbb{Z}$ אז n הוא שורש של הפולינום השלם $p(x) = x - n$.
2. כל מספר רציונלי $\frac{m}{n}$ עם $m, n \in \mathbb{Z}$ הוא אלגברי כי הוא שורש של הפולינום השלם $p(x) = nx - m$
3. $\sqrt{2}$ הוא אלגברי כי הוא שורש של הפולינום השלם $p(x) = x^2 - 2$.
4. $\sqrt{\frac{2}{3} + \sqrt{5}}$ הוא אלגברי כי הוא שורש של הפולינום $p(x) = (3x^2 - 2)^2 - 45$ (לאחר פתיחת סוגריים מתקבל פולינום שלם).

כבר ביוון העתיקה ידעו המתמטיקאים שקיימים מספרים אי רציונליים, אך ככל הנראה לא דמיינו שיש מספרים שאינם אלגבריים. הראשון שהוכיח שיש מספרים כאלה היה ליוביל,[14] שמצא מספר כזה בשנת 1844. המספר הזה נקרא על שמו:

הגדרה 8.10.2 מספר ליוביל הוא המספר $\lambda = \sum_{n=1}^{\infty} 10^{-n!}$. במילים אחרות, λ הוא המספר שבפיתוח העשרוני שלו הספרות במקום ה־ $1!, 2!, 3!, 4!, \ldots$ הם 1, וכל שאר הספרות הן 0.

שימו לב שהסדרה $n!$ גדלה במהירות עצומה, מהר יותר מכל סדרה מעריכית (איבריה הראשונים הם $1, 2, 6, 24, 120, 720, \ldots$). פירוש הדבר שהופעות הספרה 1 בפיתוח העשרוני של λ דלילות ביותר.

מספר ליוביל אינו מספר אלגברי. כדי להוכיח זאת נזדקק לשתי למות.

למה 8.10.3 אם α אלגברי אז יש פולינום שלם p כך ש־ $p(\alpha) = 0$ אבל $p'(\alpha) \neq 0$.

הוכחה יהי $d \geq 0$ המספר הטבעי המינימלי כך ש־α הוא שורש של פולינום שלם לא טריביאלי ממעלה d (כלומר, של פולינום ממעלה d שאינו שווה זהותית אפס). לא ייתכן $d = 0$ כי פולינום לא טריביאלי ממעלה 0 הוא פולינום קבוע שאין לו שורשים, ולכן $d \geq 1$. יהי p פולינום שלם ממעלה d ש־α מאפס אותו. אז p' פולינום שלם לא טריביאלי (כי $r_d \neq 0$) ממעלה קטנה מ־d, ולכן $p'(\alpha) \neq 0$. ∎

הלמה הבאה מעניינת בפני עצמה. היא קובעת שאם מנסים לקרב מספר אלגברי α בעזרת מספר רציונלי x, הקירוב לא יכול להיות טוב מדי, ולמעשה הוא חסום מלמעלה על ידי ביטוי שתלוי במכנה של x.

למה 8.10.4 אם α מספר אלגברי אז יש קבועים $M, d > 0$ וסביבה V של α כך שלכל מספר רציונלי $\alpha \neq \frac{m}{n} \in V$ מתקיים

$$|\frac{m}{n} - \alpha| > \frac{1}{M \cdot |n|^d}$$

[14] Joseph Liouville, 1809־1882.

הוכחה לפי הלמה הקודמת קיים פולינום שלם $p(x)=\sum_{k=0}^{d} r_k x^k$ כך ש־ $p(\alpha)=0$ אבל $p'(\alpha)\neq 0$. מכיוון של־p יש לכל היותר d שורשים, קיימת סביבה V' של α שבה α הוא השורש היחיד. מכיוון ש־ $p'(\alpha)\neq 0$ ו־p' רציפה, יש סביבה V'' וקבוע חיובי M כך ש־ $|p'(x)|<M$ לכל $x\in V''$. לכן בסביבה $V=V'\cap V''$ של α שתי התכונות האלה מתקיימות יחד.

יהי $\alpha\neq x\in V$. לפי משפט לגרנג' יש נקודה c בין α ל־x (ובפרט $c\in V$) כך ש־

$$\frac{p(x)-p(\alpha)}{x-\alpha}=p'(c)$$

אם נציב את השוויון הנתון $p(\alpha)=0$ ונפעיל ערך מוחלט על שני האגפים נקבל

$$\frac{|p(x)|}{|x-\alpha|}=|p'(c)|<M$$

או במילים אחרות

$$|x-\alpha|>\frac{|p(x)|}{M}$$

כעת יהי $\alpha\neq x\in V$ מספר רציונלי ויהיו $m,n\in\mathbb{Z}$ כך ש־ $x=\frac{m}{n}$. נציב זאת באי־שוויון ונקבל

$$|\frac{m}{n}-\alpha|>\frac{|p(m/n)|}{M}=\frac{1}{M}|\sum_{k=0}^{d} r_k(\frac{m}{n})^k|=\frac{1}{M|n|^d}|\sum_{k=0}^{d} r_k m^k n^{d-k}|$$

מכיוון של־p אין שורשים ב־V פרט ל־α הרי ש־ $p(\frac{m}{n})\neq 0$, ולכן אגף ימין שונה מאפס. אבל הסכום $\sum_{k=0}^{d} r_k m^k n^{d-k}$ הוא מספר שלם, ולכן אם הערך המוחלט שלו אינו 0 הוא לפחות 1. קיבלנו אם כן שלכל $\alpha\neq\frac{m}{n}$ רציונלי ב־ V מתקיים

$$|\frac{m}{n}-\alpha|>\frac{1}{M\cdot|n|^d}$$

כפי שרצינו. ■

מקרה פרטי של הלמה הוא שאם $x=0.x_1x_2\ldots x_n$ הוא שבר עשרוני קרוב מספיק ל־α אז $|x-\alpha|>\frac{1}{M\cdot 10^{nd}}$, כי x הוא בעצם שבר עם מכנה 10^n. אבחנה זו היא הלב של ההוכחה של המשפט הבא:

משפט 8.10.5 (ליוביל, 1844)המספר $\lambda=\sum_{n=1}^{\infty}10^{-n!}$ אינו אלגברי.

הוכחה נניח בשלילה ש־λ אלגברי. יהי $\lambda_n=\sum_{k=1}^{n}10^{-k!}$ (כלומר λ_n הוא הסכום החלקי ה־n־של הטור המגדיר את λ). אז $\lambda_n\to\lambda$. מצד שני אפשר לרשום את λ_n בצורה

$$\lambda_n=\frac{\sum_{k=0}^{n}10^{n!-k!}}{10^{n!}}$$

והמונה הוא שלם. לכן לפי הלמה קיימים קבועים M, d (שאינם תלויים ב־n) כך שלכל n גדול מספיק

$$|\lambda_n - \lambda| > \frac{1}{M \cdot 10^{d \cdot n!}}$$

מצד שני,

$$\begin{aligned} |\lambda_n - \lambda| &= |\sum_{k=0}^{n} 10^{-k!} - \sum_{k=0}^{\infty} 10^{-k!}| \\ &= \sum_{k=n+1}^{\infty} 10^{-k!} \\ &\leq \sum_{k=(n+1)!}^{\infty} 10^{-k} \\ &= 10^{-(n+1)!} \cdot \frac{1}{10-1} \end{aligned}$$

ומצירוף האי־שוויונות יוצא שלכל n גדול מספיק מתקיים

$$\frac{1}{9} 10^{-(n+1)!} > \frac{1}{M \cdot 10^{d \cdot n!}}$$

או, לאחר העברת אגפים,

$$10^{-(n+1-d)n!} > \frac{9}{M} > 0$$

וזו סתירה, כי אגף שמאל שואף ל־0 כאשר n שואף לאינסוף. ■

הגדרה 8.10.6 מספר ממשי שאינו אלגברי נקרא **טרנסצנדנטי** (transcendental).

מספר שנים לאחר שליוביל נתן דוגמה ראשונה למספר טרנסצנדנטי, תורת העוצמות האינסופיות של קנטור הניבה תוצאה מפתיעה עוד יותר: לא רק שיש מספרים לא אלגבריים אלא שרוב המספרים הם כאלה.

משפט 8.10.7 (קנטור, 1873) קבוצת המספרים האלגבריים היא בת מנייה.

הוכחה ההוכחה אינה משתמשת כלל בכלים אנליטיים או אלגבריים אלא רק בשיקולי עוצמות, כמו ההוכחה בסעיף 5.12 לקיום מספרים אי־רציונליים. ראשית, אנו טוענים שקבוצת הפולינומים השלמים היא בת־מניה. ואמנם, עבור $n = 1, 2, 3 \ldots$ תהי P_n קבוצת הפולינומים השלמים $p(x) = \sum_{k=0}^{d} r_k x^k$ כך ש־ $d \leq n$ ו־ $|r_k| \leq n$ לכל $k = 0, 1, \ldots, d$. קל להראות ש־P_n סופית, ומאידך כל פולינום שלם שייך לאחת הקבוצות P_n. לכן הקבוצה $P = \cup_{n=1}^{\infty} P_n$ של פולינומים שלמים היא בת־מניה (למה 5.12.2), ואפשר למצוא סדרה (p_n) של פולינומים כך ש־ $P = \{p_n \,:\, n \in \mathbb{N}\}$.

כעת לכל n תהי R_n קבוצת השורשים של ההפולינומים ב־P_n. לפי דוגמה (2) בעמוד 332 מספר השורשים של פולינום הוא סופי, לכן R_n סופית (למעשה כל $p \in P_n$ הוא פולינום ממעלה n, לכן יש לו לכל היותר n שורשים, ולכן $|R_n| \leq n|P_n|$). לכן הקבוצה $R = \bigcup_{n=1}^{\infty} R_n$ היא בת־מניה (שוב לפי למה 5.12.2). אבל R היא קבוצת המספרים האלגבריים, כי אם α מספר אלגברי אז יש פולינום שלם p כך ש־α שורש של p. כיוון ש־ $p \in P_n$ עבור n כלשהו, הרי $\alpha \in R_n$, ולכן $\alpha \in R$. ■

מסקנה 8.10.8 קבוצת המספרים הטרנסצנדנטיים אינה בת מניה, ובפרט אינה ריקה.

הוכחה קבוצת המספרים הטרנסצנדנטיים היא ההפרש בין קבוצת כל המספרים $\mathbb{R}$, שאינה בת מניה, וקבוצת המספרים האלגבריים, שהיא כן בת מניה. העובדה שההפרש שוב אינו בן מניה נובעת ממסקנה 5.12.6. ■

המספרים הטרנסצנדנטיים הם תחום מחקר פעיל במתמטיקה, והשאלות בתחום זה רבות מהתשובות. ידוע, למשל, שהמספרים e ו־π הם טרנסצנדנטיים (ההוכחה לכך קשה יותר מההוכחה למעלה, אם כי קיימות לכך הוכחות שמשתמשות רק בכלים אלמנטריים. ראו למשל [5],[17]). כמו־כן ידוע שהמספר $\ln r$ טרנסצנדנטי לכל $0 < r \neq 1$ רציונלי. אך גם כיום נותרו שאלות פתוחות רבות בתחום זה. למשל, אף שידוע ש־ e^π טרנסצנדנטי, לא יודעים אם π^e טרנסצנדנטי, או אפילו אם $\pi + e$ טרנסצנדנטי.

פרק 9

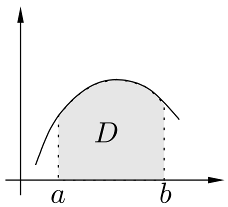
האינטגרל

האינטגרל המסוים של פונקציה ממשית בקטע $[a,b]$ הוא מספר המתאר את השטח הכלוא בין גרף הפונקציה לציר ה־x בקטע $[a,b]$. החלק הראשון של הפרק מוקדש להגדרת האינטגרל, הכוללת למעשה סוג חדש של גבול, ולחקירת תכונותיו.

הגדרת האינטגרל מקורה בבעיה גאומטרית אך מסתבר שבמובן מסוים חישוב האינטגרל הוא הפעולה ההפוכה לגזירה. זה תוכנו של אחד המשפטים החשובים בפרק זה, המשפט היסודי של החשבון האינפיניטסימלי. משפט זה מאפשר לחשב אינטגרל מסוים של פונקציה f בתנאי שיודעים למצוא פונקציה F שמקיימת $F' = f$. פונקציה F כזו נקראת גם אינטגרל לא מסוים של f, ונקדיש את סעיף 9.7 לתיאור שיטות לחישוב האינטגרל הלא מסוים.

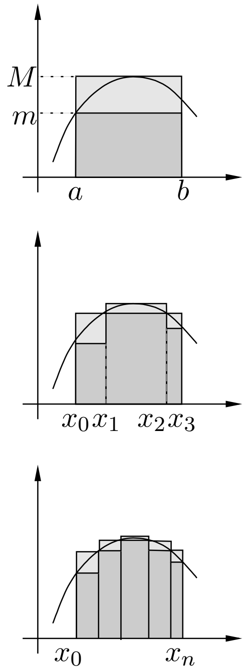
לאחר מכן נעסוק בהכללות ושימושים של האינטגרל: נראה כיצד הוא מאפשר לחשב נפחים ואורכים של גרפים. נראה הכללות של האינטגרל וקשרים שלו עם תורת הטורים. נראה גם שיטות נומריות לחישוב האינטגרל.

9.1 האינטגרל המסוים לפי דרבו



איור 9.1.1 השטח שמתחת לגרף בקטע $[a,b]$

תהי f פונקציה אי־שלילית ממשית וחסומה המוגדרת בקטע $[a,b]$, ויהי D התחום הכלוא בין ציר ה־x לגרף של הפונקציה בקטע $[a,b]$, כמתואר באיור 9.1.1. היינו רוצים לחשב את השטח של D, או ליתר דיוק להגדיר מספר ממשי $s(D)$ שייקרא השטח של D. המובן האינטואיטיבי של מושג השטח ברור, אך גם בלי להתעמק על הגדרות מדויקות יש להודות שאיננו יודעים לחשב שטחים, למעט עבור הצורות הפשוטות ביותר, כמו מלבנים ומשולשים. התחום D שמתחת לגרף של f בדרך כלל אינו בעל צורה פשוטה כזו.



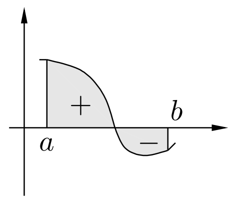
איור 9.1.2

אנו ננהג לכן במיטב המסורת וננסה למצוא חסמים מלמעלה ומלמטה עבור $s(D)$, במטרה להיעזר בהם כשנרצה להגדיר את $s(D)$. כהערכה ראשונה נשים לב לעובדה הבאה: אם m הוא הערך המינימלי של f ב־$[a,b]$ אז המלבן שבסיסו $[a,b]$ וגובהו m נמצא כולו מתחת לגרף של f, ולכן המלבן מוכל ב־D ונצפה ששטח המלבן אינו גדול יותר מ־$s(D)$. כמובן שלגבי מלבן אנו יודעים בדיוק מה אמור להיות שטחו: מכפלת רוחב המלבן בגובהו (ניתן להתייחס לנוסחה זו כאל ההגדרה של שטח של מלבן). מכאן אנו מקבלים את האי־שוויון $s(D) \geq m(b-a)$.

באופן דומה, אם M הוא הערך המקסימלי של f ב־$[a,b]$ אז המלבן שבסיסו $[a,b]$ וגובהו M מכיל את D ונצפה ש־$s(D)$ אינו גדול משטח המלבן, כלומר: $s(D) \leq M(b-a)$. קיבלנו בסיכום

$$m(b-a) \leq s(D) \leq M(b-a)$$

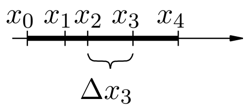
ההערכה האחרונה היא הערכה גסה מאד עבור השטח של D, אך נוכל לשפר אותה אם נחלק את הקטע $[a,b]$ לכמה תת־קטעים ונחזור על אותו שיקול בכל אחד מהם. אז נקבל שתי סדרות של מלבנים, סדרה אחת של מלבנים המוכלים ב־D וסדרה שנייה שאיחוד המלבנים בה מכיל את D. אנו מצפים ש־$s(D)$ יהיה חסום מלמטה על־ידי סכום שטחי המלבנים הקטנים, ומלמעלה על־ידי סכום שטחי המלבנים הגדולים, כפי שמודגם באיור 9.1.2. מהאיור נדמה שככל שנבחר חלוקה של $[a,b]$ לקטעים קצרים יותר, כך ההערכות הללו יתקרבו אחת לשנייה ובגבול יתנו מספר משותף שאפשר לפרש אותו בתור השטח שמתחת לגרף.



איור 9.1.3 השטח שמתחת לציר ה־x הוא שלילי

עד כה התייחסנו לפונקציות חיוביות, אך נרצה להגדיר שטח גם עבור פונקציות שאינן בהכרח חיוביות. לשם כך נתייחס לכל אותו שטח שמתחת לציר ה־x כשטח בעל סימן שלילי, כפי שמודגם באיור 9.1.3. עם מוסכמה זו עדיין מתקיים $m(b-a) \leq s(D) \leq M(b-a)$, כאשר שוב M ו־m הם המקסימום והמינימום של f, בהתאמה, וכן כאשר מחלקים את $[a,b]$ לכמה תת־קטעים.

נזדקק לכמה מושגים חדשים על מנת לנתח את תהליך הקירוב שתיארנו למעלה. נתחיל עם מושג החלוקה. אנו רוצים לתאר את חלוקתו של הקטע $[a,b]$ לתת־קטעים. הדרך הפשוטה ביותר לעשות זאת היא לתאר את הקצוות של התת־קטעים. למשל, אם מחלקים את הקטע $[0,1]$ לשלושה תת־קטעים שווי־אורך מקבלים תת־קטעים $[0,\frac{1}{3}],[\frac{1}{3},\frac{2}{3}],[\frac{2}{3},1]$. אפשר לתאר חלוקה זו בקצרה על־ידי מניית קצוות הקטעים: $\{0,\frac{1}{3},\frac{2}{3},1\}$. באופן כללי יותר:



איור 9.1.4 חלוקה של קטע לארבע תת־קטעים

הגדרה 9.1.1 חלוקה (partition) של קטע $[a,b]$ היא תת־קבוצה סופית $P \subseteq [a,b]$ המכילה את a, b.

עבור חלוקה P של $[a,b]$ נרשום $P = \{x_0, x_1, \ldots, x_k\}$, ונתכוון תמיד שהמספרים x_i מסודרים בסדר עולה: $a = x_0 < x_1 < \ldots < x_k = b$. לעתים קרובות נחשוב על

P כאילו היא מתארת את אוסף הקטעים $[x_0,x_1],[x_1,x_2],\ldots,[x_{k-1},x_k]$. נסמן ב־$\Delta x_i = x_i - x_{i-1}$ את האורך של הקטע ה־i בחלוקה. שימו לב שהסכום $\sum_{i=1}^{k} \Delta x_i$ הוא סכום אורכי כל הקטעים בחלוקה, ולכן נצפה שהוא יהיה שווה לאורך הקטע $[a,b]$. ואמנם, מתקיים

$$\sum_{i=1}^{k} \Delta x_i = \sum_{i=1}^{k} (x_i - x_{i-1}) = x_k - x_0 = b - a$$

(השוויון האמצעי מכיוון שמדובר בסכום טלסקופי, והימני כי $x_0 = a \, , \, x_k = b$).

תהי f פונקציה חסומה ב־$[a,b]$ ו־P חלוקה של $[a,b]$. אפשר להתאים לכל קטע בחלוקה שני מלבנים שחוסמים בצורה ההדוקה ביותר את גרף הפונקציה באותו קטע. הגבהים של הקטעים יהיו שווים למינימום והמקסימום של הפונקציה בקטע, ובהיעדר מינימום ומקסימום נסתפק בחסם תחתון ועליון:

הגדרה 9.1.2 עבור פונקציה חסומה f ב־$[a,b]$ וחלוקה $P = \{x_0,\ldots,x_k\}$ של $[a,b]$, לכל $1 \le i \le k$ נסמן

$$\begin{aligned} m_i &= \inf\{f(x) \, : \, x \in [x_{i-1},x_i]\} \\ M_i &= \sup\{f(x) \, : \, x \in [x_{i-1},x_i]\} \end{aligned}$$

מאחר שהגדלים m_i, M_i תלויים ב־f וב־P היה מוטב ש־f,P יופיעו בסימון. כשנרצה להדגיש את תפקידם נכתוב $m_i(f,P), M_i(f,P)$ במקום m_i, M_i, אבל בסעיף הנוכחי לא יהיה צורך בכך.

לכל קטע $[x_{i-1},x_i]$ בחלוקה $P = \{x_0,\ldots,x_k\}$ אפשר לבנות מלבן r_i שבסיסו $[x_{i-1},x_i]$ וגובהו m_i. כך נקבל סדרת מלבנים $r_1,\ldots,r_k$ שלכל אחד מהם יש צלע אחת על ציר ה־x והשנייה נושקת לגרף הפונקציה מלמטה. השטח של r_i אינו עולה על השטח שבין הגרף לציר ה־x בקטע $[x_{i-1},x_i]$, ולכן נצפה ש־ $s(r_1)+\ldots+s(r_k)$ לא יעלה על $s(D)$ (שימו לב שהמלבנים נפגשים רק במספר סופי של קטעים אנכיים ששטחם הכולל הוא אפס). במילים אחרות, מאחר שהשטח של r_i הוא $m_i \cdot \Delta x_i$ אנו מצפים שיתקיים האי־שוויון

$$s(D) \ge \sum_{i=1}^{k} s(r_i) = \sum_{i=1}^{k} m_i \Delta x_i$$

באותו אופן, לכל $i = 1,\ldots,k$ נבנה מלבן R_i שבסיסו $[x_{i-1},x_i]$ וגובהו M_i, כך שהצלע המקבילה לציר ה־x נושקת לגרף של f מלמעלה. נצפה ששטחו של כל R_i אינו קטן מהשטח שבין הגרף לציר ה־x בקטע $[x_{i-1},x_i]$, ושהשטח $s(D)$ אינו גדול מסכום השטחים של המלבנים R_i, דהיינו,

$$s(D) \le \sum_{i=1}^{k} s(R_i) = \sum_{i=1}^{k} M_i \Delta x_i$$

הגדרה 9.1.3 תהי f פונקציה חסומה ב־$[a,b]$ ותהי $P=\{x_0,\dots,x_k\}$ חלוקה של $[a,b]$. **הסכום התחתון** של f המתאים ל־P הוא

$$\underline{s}(f,P)=\sum_{i=1}^{k} m_i \Delta x_i$$

והסכום העליון של f המתאים ל־P הוא

$$\overline{s}(f,P)=\sum_{i=1}^{k} M_i \Delta x_i$$

יש לצפות שכל סכום תחתון יהיה קטן או שווה לכל סכום עליון, כי אנו מאמינים שהשטח של D הוא מספר הנמצא ביניהם. תכונה זו של הסכומים התחתונים והעליונים אינה מובנת מאליה וידרשו כמה פסקאות להוכחתה. השלב הראשון בהוכחה הוא להשוות בין סכומי המלבנים המתאימים לאותה חלוקה P.

למה 9.1.4 תהי f פונקציה חסומה המוגדרת על קטע $[a,b]$ ותהי $P=\{x_0,\dots,x_k\}$ חלוקה של $[a,b]$. אז $m_i \le M_i$ לכל $1\le i\le k$ ומתקיים $\underline{s}(f,P)\le\overline{s}(f,P)$.

הוכחה האי־שוויון $m_i \le M_i$ מידי מההגדרה של m_i, M_i. כיוון ש־

$$\overline{s}(f,P)-\underline{s}(f,P)=\sum_{i=1}^{k}(M_i-m_i)\Delta x_i$$

הרי שהסכום הזה מכיל רק מחוברים אי־שליליים וקיבלנו ש־ $\overline{s}(f,T)-\underline{s}(f,T)\ge 0$, כנדרש. ■

הגדרה 9.1.5 יהי $[a,b]$ קטע ויהיו P,Q חלוקות של $[a,b]$. אז Q נקראת **עידון** (refinement) של P אם כל נקודת חלוקה ב־P היא גם נקודת חלוקה ב־Q, כלומר: $P\subseteq Q$.

לא קשה להשתכנע ש־Q היא עידון של P אמ"מ כל קטע חלוקה של Q מוכל באחד מקטעי החלוקה של P (הוכיחו!)

למה 9.1.6 תהי f פונקציה חסומה ב־$[a,b]$ ויהיו P,Q חלוקות של $[a,b]$ כך ש־Q עידון של P. אז מתקיים

$$\underline{s}(f,P)\le\underline{s}(f,Q)\le\overline{s}(f,Q)\le\overline{s}(f,P)$$

הוכחה האי־שוויון האמצעי נובע מלמה 9.1.4. נוכיח את האי־שוויון השמאלי. האי־שוויון הימני מוכח בצורה דומה.

נוכיח תחילה את הלמה במקרה ש־Q מתקבלת מ־P על־ידי הוספת נקודה יחידה: נסמן $P = \{x_0, \ldots, x_k\}$ ו־ $Q = P \cup \{y\}$, כאשר $y \notin P$. קיים אז $1 \leq i \leq k$ כך ש־ $x_{i-1} < y < x_i$. לכן הקטעים בחלוקה Q הם בדיוק הקטעים בחלוקה P למעט הקטע $[x_{i-1}, x_i]$ שהוחלף בשני הקטעים $[x_{i-1}, y], [y, x_i]$. מהתבוננות בהגדרה של $\underline{s}(f,P), \underline{s}(f,Q)$ רואים שהם נבדלים רק בכך שבסכום המגדיר את $\underline{s}(f,P)$ מופיע מחובר $m_i(x_i - x_{i-1})$ ואילו בסכום המגדיר את $\underline{s}(f,Q)$ מופיע במקומו הסכום$m_i'(y - x_{i-1}) + m_i''(x_i - y)$ כאשר

$$m_i' = \inf\{f(x) \, : \, x \in [x_{i-1}, y]\} \qquad , \qquad m_i'' = \inf\{f(x) \, : \, x \in [y, x_i]\}$$

לכן כדי להוכיח $\underline{s}(f,P) \leq \underline{s}(f,Q)$ די להראות

$$m_i(x_i - x_{i-1}) \leq m_i'(y - x_{i-1}) + m_i''(x_i - y)$$

אבל $[x_{i-1}, y], [y, x_i] \subseteq [x_{i-1}, x_i]$ ולכן

$$m_i' \geq m_i \qquad , \qquad m_i'' \geq m_i$$

(למה?) ומכאן שמתקיים

$$m_i'(y - x_{i-1}) + m_i''(x_i - y) \geq m_i(y - x_{i-1}) + m_i(x_i - y) = m_i(x_i - x_{i-1})$$

כפי שרצינו להראות.

כאשר Q עידון כללי של P אפשר למצוא סדרת חלוקות $P = P_1, P_2, \ldots, P_k = Q$ כך ש־ P_i מתקבלת מ־P_{i-1} על־ידי הוספת נקודה אחת (זכרו כי P, Q סופיות). בפירוט, אם $Q \setminus P = \{y_1, \ldots, y_n\}$ אז נקבל את הסדרה המבוקשת אם נגדיר $P_0 = P$ ולכל $1 \leq i \leq n$ נגדיר $P_i = P_{i-1} \cup \{y_i\}$. מהמקרה שהוכחנו למעלה מתקיים $\underline{s}(f,P_i) \leq \underline{s}(f,P_{i+1})$ לכל $0 \leq i < n$ ולכן

$$\underline{s}(f,P) = \underline{s}(f,P_1) \leq \underline{s}(f,P_2) \leq \ldots \leq \underline{s}(f,P_k) = \underline{s}(f,Q)$$

כפי שרצינו. ■

מסקנה 9.1.7 אם P, Q חלוקות של $[a,b]$ אז

$$\underline{s}(f,P) \leq \overline{s}(f,Q)$$

הוכחה תהי R החלוקה $R = P \cup Q$. אז $P, Q \subseteq R$ (קרי: R היא עידון משותף של P ו־Q) ולכן

$$\underline{s}(f,P) \le \underline{s}(f,R) \le \overline{s}(f,R) \le \overline{s}(f,Q)$$

(כל האי־שוויונות נובעים מלמה 9.1.6). ■

הגדרה 9.1.8 עבור פונקציה חסומה f בקטע $[a,b]$ נגדיר את $\underline{S}(f)$ ואת $\overline{S}(f)$ להיות קבוצות המספרים המתקבלים כסכומים תחתונים ועליונים של f בהתאמה, דהיינו ש־

$$\begin{aligned}\underline{S}(f) &= \{\underline{s}(f,P) : [a,b] \text{ חלוקה של } P\}\\ \overline{S}(f) &= \{\overline{s}(f,P) : [a,b] \text{ חלוקה של } P\}\end{aligned}$$

בסימונים $\underline{S}(f), \overline{S}(f)$ אין זכר לקטע $[a,b]$ שבו אנו עוסקים, והיה מוטב שיהיה אזכור כזה, אך הקטע שבו מדובר תמיד יהיה ברור מההקשר ולא ייווצר בלבול.

נשוב לבעיה המקורית, שבה D מסמנת את הצורה הכלואה מתחת לגרף של f. אנו רוצים למצוא מספר $s(D)$ הנמצא בין כל סכום תחתון לכל סכום עליון, כלומר עליו לקיים $x \le s(D) \le y$ לכל $x \in \underline{S}(f)\,,\, y \in \overline{S}(f)$. על פי הטענה הקודמת, לכל $x \in \underline{S}(f)\,,\, y \in \overline{S}(f)$ מתקיים $x \le y$, ולכן

$$\sup \underline{S}(f) \le \inf \overline{S}(f)$$

מכאן נובע שישנם מספרים s שמקיימים $x \le s \le y$ לכל $x \in \underline{S}(f)\,,\, y \in \overline{S}(f)$, כמו למשל המספרים $\sup \underline{S}(f)$ או $\inf \overline{S}(f)$ עצמם. אם מתקיים אי־שוויון חזק $\sup \underline{S}(f) < \inf \overline{S}(f)$ אז כל מספר ביניהם מקיים את התכונה המבוקשת, ונתקשה לבחור מביניהם אחד להיות $s(D)$. לעומת זאת, אם $\sup \underline{S}(f) = \inf \overline{S}(f)$ אז יש רק מועמד אחד עבור $s(D)$, והוא המספר המבוקש.

הגדרה 9.1.9 (האינטגרל לפי דרבו) אומרים שפונקציה f **אינטגרבילית** (integrable) ב־$[a,b]$ אם היא מוגדרת וחסומה שם ומתקיים $\sup \underline{S}(f) = \inf \overline{S}(f)$. במקרה זה הערך המשותף נקרא **האינטגרל** (integral) של f ב־$[a,b]$, ומסומן $\int_a^b f$, $\int_a^b f(x)dx$ או $\int_{[a,b]} f$. הפונקציה f נקראת **האינטגרנד** (integrand) של האינטגרל $\int_a^b f$.

שימו לב שהגדרנו אינטגרביליות ואת האינטגרל רק לפונקציות חסומות ורק ביחס לקטעים סגורים.

כשמסמנים את האינטגרל ב־$\int_a^b f(x)dx$, הסימן dx מציין את המשתנה של הפונקציה שלגביו מבצעים את האינטגרציה, ודומה בתפקידו לסימן dx בסימון $\frac{d}{dx}f(x)$ של הנגזרת. יש בכך צורך במיוחד כאשר מביעים את הפונקציה בעזרת נוסחה, כי

אז אין תמיד דרך לדעת מי המשתנה הרלוונטי. למשל, האם $\int_0^1 x^2y^3$ מתאר את האינטגרל של הפונקציה $f(x)=y^3x^2$ שבהגדרתה y מספר קבוע, או של הפונקציה $g(y)=x^2y^3$ שבהגדרתה x מספר קבוע? אם רושמים $\int_0^1 x^2y^3\,dx$ כבר אין דו משמעות: מדובר במקרה הראשון, ואנו חושבים על הנוסחה x^2y^3 כפונקציה של x. אומרים אז ש־x הוא **משתנה האינטגרציה**. עם זאת, כאשר האינטגרנד אינו נתון על ידי נוסחה נעדיף את הסימון הקצר יותר $\int_a^b f$.

לסימן dx, אשר "כופל" את הפונקציה באינטגרל $\int_a^b f(x)dx$, יש היסטוריה מעניינת. דרך אחת לחשוב על התחום D שמתחת לגרף של f בקטע $[a,b]$ הוא כאיחוד של אינסוף מלבנים בעלי רוחב "מזערי": מעל לכל נקודה $x\in[a,b]$ אפשר לחשוב על המלבן בגובה $f(x)$ ובעל בסיס ברוחב "מזערי" (אינפיניטסימלי) dx. מפתה לומר שהשטח של D הוא סכום שטחי המלבנים האלה, כלומר: "$\sum_{x=a}^b f(x)dx$". כמובן שאין הצדקה פורמלית לכך, אך הסימון $\int_a^b f(x)dx$ בא בדיוק לרמוז על סכום זה. בנוסף, אם מפרשים את dx בצורה זו הוא מקבל משמעות דומה מאד לפירוש שלו בסימן $\frac{d}{dx}$ של הנגזרת, כפי שהסברנו לאחר הגדרת הנגזרת (עמוד 308). נציין ש־$\int$ הוא גלגול של האות האנגלית S, שהיא האות הראשונה במילה Sum, שפירושה סכום.

האם יש פונקציות אינטגרביליות והאם כל הפונקציות הן כאלה? הדוגמאות הבאות מראות שיש פונקציות משני הסוגים.

דוגמאות

1. תהי $f\equiv c$ פונקציה קבועה. הגרף של f הוא קו מאוזן והתחום מתחת לגרף שלה בקטע $[a,b]$ הוא מלבן שבסיסו באורך $b-a$ וגובהו c. לכן נצפה שהאינטגרל של f ב־$[a,b]$ יהיה שווה ל־$c\cdot(b-a)$.

   הנה ההוכחה הפורמלית. c הוא הערך המינימלי והמקסימלי של f בכל קטע שהוא, ולכל חלוקה $P=\{x_1,\dots,x_k\}$ של $[a,b]$ מתקיים $m_i=M_i=c$ לכל $i=1,\dots,k$. לכן

$$\underline{s}(f,P)=\sum_{i=1}^k c\Delta x_i=c\sum_{i=1}^k \Delta x_i=c\cdot(b-a)$$

   באופן דומה מראים ש־

$$\overline{s}(f,P)=c(b-a)$$

   קיבלנו אפוא ש־ $c(b-a)\in\underline{S}(f)$ וגם $c(b-a)\in\overline{S}(f)$ ומכאן אנו מסיקים ש־ $\inf\overline{S}(f)\le c(b-a)\le\sup\underline{S}(f)$. מאידך ראינו שתמיד מתקיים $\overline{S}(f)\ge\underline{S}(f)$, ונובע ש־ $\sup\underline{S}(f)=\inf\overline{S}(f)=c(b-a)$. לכן f אינטגרבילית ומתקיים $\int_a^b f=c(b-a)$, כפי שציפינו.

2. תהי

$$D(x)=\begin{cases}1 & x\in\mathbb{Q}\\ 0 & x\notin\mathbb{Q}\end{cases}$$

פונקציית דירכלה. הגרף של D מזכיר "אבקה" אשר מפוזרת בצורה צפופה על הישרים המאוזנים בגובה 1 ו־0. התחום בין הגרף של D לציר ה־x מורכב מקוים מאונכים בגובה 1 המתחילים בנקודות הרציונליות על ציר ה־x. קשה לומר מה השטח של צורה כזו, או אם יש טעם בכלל לדבר על שטח לפונקציה כזו.

ואמנם, D אינה אינטגרבילית באף קטע לא מנוון $[a,b]$. כדי להוכיח זאת נקבע חלוקה $P=\{x_0,\ldots,x_k\}$ של $[a,b]$. בכל קטע בחלוקה P יש מספרים רציונליים ואי־רציונליים, ולכן מתקיים $m_i=0, M_i=1$ לכל $i=1,\ldots,k$. מכאן

$$\begin{array}{rcccl}\underline{s}(D,P) &=& \sum_{i=1}^k 0\cdot\Delta x_i &=& 0\\ \overline{s}(D,P) &=& \sum_{i-1}^k 1\cdot\Delta x_i &=& b-a\end{array}$$

מכיוון שזה נכון לכל חלוקה של $[a,b]$ מקבלים ש־ $\sup\underline{S}(D)=0$ ואילו $\inf\overline{S}(D)=b-a>0$, ולכן D אינה אינטגרבילית על $[a,b]$.

האפיון הבא הוא הראשון בסדרה של אפיונים של האינטגרל שיהיו שימושיים בהמשך. למעשה זו אינה אלא גרסה של משפט 4.4.5 עם כמה התאמות להקשר הנוכחי.

משפט 9.1.10 תהי f פונקציה חסומה בקטע $[a,b]$. התנאי הבאים שקולים:

1. f אינטגרבילית ב־$[a,b]$.
2. לכל $\varepsilon>0$ קיימת חלוקה P של $[a,b]$ כך ש־ $\overline{s}(f,P)-\underline{s}(f,P)<\varepsilon$.
3. קיימת סדרת חלוקות $(P_n)_{n=1}^\infty$ של $[a,b]$ כך ש־

$$\lim_{n\to\infty}\overline{s}(f,P_n)=\lim_{n\to\infty}\underline{s}(f,P_n))$$

בנוסף, אם (P_n) סדרת חלוקות כמו ב־(3) אז $\int_a^b f=\lim\underline{s}(f.P_n)=\lim\overline{s}(f.P_n)$.

הוכחה נוכיח ש־(2) גורר את (1). נניח ש־f אינטגרבילית ב־$[a,b]$, כלומר שמתקיים $\sup\underline{S}(f)=\inf\overline{S}$. מכאן שלכל $\varepsilon>0$ קיימים איברים $u\in\underline{S}(f)$ ו־ $v\in\overline{S}(f)$ כך ש־ $v-u<\varepsilon$. פירוש הדבר שקיימות חלוקות Q,R של $[a,b]$ כך ש־ $\overline{s}(f,R)-\underline{s}(f,Q)<\varepsilon$ (כי u מהצורה $\underline{s}(f,Q)$ ו־v מהצורה $\overline{s}(f,R)$). כעת לפי למה 9.1.6, החלוקה $P=Q\cup R$ מקיימת את (2).

להפך, נניח ש־(1) מתקיים, ונראה את (2). יהי $\varepsilon > 0$. עלינו למצוא חלוקה R של $[a,b]$ כך ש־ $\overline{s}(f,R) - \underline{s}(f,R) < \varepsilon$. לפי הנחת האינטגרביליות קיימות חלוקות P,Q של $[a,b]$ כך ש־ $\overline{s}(f,Q) - \underline{s}(f,P) < \varepsilon$. תהי $R = P \cup Q$, כך ש־R והיא עידון משותף של P,Q. מכאן ש־ $\underline{s}(f,R) \geq \underline{s}(f,P)$ וגם $\overline{s}(f,R) \leq \overline{s}(f,Q)$ ולכן

$$\overline{s}(f,R) - \underline{s}(f,R) \leq \overline{s}(f,Q) - \underline{s}(f,P) = y - x \leq \varepsilon$$

ו־R היא החלוקה המבוקשת.

השקילות של (2) ו־(3) קלה. ראשית אם יש סדרה (P_n) כמו ב־(3) ברור ש־(2) מתקיים. בכיוון ההפוך, בהנחה ש־(2) מתקיים תהי P_n החלוקה המובטחת על ידי (2) עבור $\varepsilon = \frac{1}{n}$. אז

$$0 \leq |\overline{s}(f,P_n) - \underline{s}(f,P_n)| < \frac{1}{n}$$

ולכן $\overline{s}(f,P_n) - \underline{s}(f,P_n) \to 0$. אבל מתקיים גם

$$\underline{s}(f,P_n) \leq \int_a^b f \leq \overline{s}(f,P_n)$$

ולכן $\lim_{n\to\infty} \underline{s}(f,P_n) = \lim_{n\to\infty} \overline{s}(f,P_n) = \int_a^b f$ ובפרט הגבולות קיימים, כפי שרצינו. ■

ההגדרות והאפיונים של האינטגרל שראינו בשלב זה אינם נוחים לשימוש, שכן אינם מכילים נוסחה מפורשת עבור הערך של האינטגרל. בהמשך נפתח כלים משוכללים יותר, אך כרגע נחשב אינטגרלים ישירות מהגדרה.

דוגמאות

1. נחשב את האינטגרל של $f(x) = x$ ב־$[0,1]$. מאחר שהשטח תחת הגרף הוא משולש ישר זווית שניצביו באורך 1, נצפה שהאינטגרל יהיה $\frac{1}{2}$.

   הנה ההוכחה. תהי P_n החלוקה של $[0,1]$ ל־n קטעים שווים:

$$P_n = \{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}$$

   נקבע את n. לכל $1 \leq k \leq n$ הפונקציה $f(x) = x$ עולה בקטע $[\frac{k-1}{n}, \frac{k}{n}]$ ולכן המינימום מתקבל בקצה $\frac{k-1}{n}$ והמקסימום בקצה $\frac{k}{n}$, ולכן

$$m_k = f(\frac{k-1}{n}) = \frac{k-1}{n} \qquad , \qquad M_k = f(\frac{k}{n}) = \frac{k}{n}$$

מכאן

$$\begin{aligned}\underline{s}(f,P_n) &= \sum_{k=1}^{n} m_k(P_n)\cdot(\frac{k}{n}-\frac{k-1}{n})\\ &= \sum_{k=1}^{n}\frac{k-1}{n}\cdot\frac{1}{n}\\ &= \frac{1}{n^2}\sum_{k=1}^{n}(k-1)\\ &= \frac{1}{n^2}\cdot\frac{(n-1)n}{2}\\ &= \frac{n-1}{2n}\end{aligned}$$

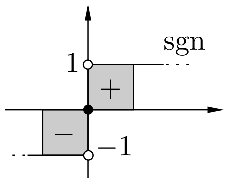
חישוב דומה מאד מראה ש־

$$\overline{s}(f,P_n)=\frac{n+1}{2n}$$

לכן

$$\lim_{n\to\infty}\underline{s}(f,P_n)=\lim_{n\to\infty}\overline{s}(f,P_n)=\frac{1}{2}$$

ולפי משפט 9.1.10 אנו מקבלים ש־f אינטגרבילית ומתקיים $\int_0^1 f=\frac{1}{2}$.



איור 9.1.5

2. הנה דוגמה לפונקציה אינטגרבילית שאינה רציפה. תהי

$$f(x)=\begin{cases}-1 & x\le 0\\ 1 & x>0\end{cases}$$

נראה שהיא אינטגרבילית ב־$[-1,1]$ ושהאינטגרל שלה שם הוא 0. אין זה מפתיע, כי אנו מצפים שהשטח החיובי מימין ל־0 יצטמצם בדיוק עם השטח השלילי משמאל ל־0. ראו איור 9.1.5.

ואמנם, תהי P_n חלוקה של $[-1,1]$ ל־$2n$ קטעים באורך n:

$$P_n=\{-1,\frac{-n+1}{n},\ldots,-\frac{1}{n},0,\frac{1}{n},\ldots,\frac{n-1}{n},1\}$$

נסמן $x_i=\frac{i}{n}-1$. אם $1\le i\le n$ אז $f(x)=-1$ לכל $x\in[x_{i-1},x_i]$ ולכן $m_i=M_i=-1$. אם $i>n+1$ אז $f(x)=1$ לכל $x\in[x_{i-1},x_i]$. לבסוף כיוון שבקטע $[x_{n+1},x_{n+2}]=[0,\frac{1}{n}]$ הפונקציה מקבלת את הערכים ±1 מתקיים $m_{n+1}=-1$ אבל $M_{n+1}=1$. לכן

$$\begin{aligned}\underline{s}(f,P_n) &= \textstyle\sum_{i=1}^{n}(-1)\cdot\frac{1}{n}+(-1)\frac{1}{n}+\sum_{i=n+2}^{2n}1\cdot\frac{1}{n} &= -\frac{2}{n}\\ \overline{s}(f,P_n) &= \textstyle\sum_{i=1}^{n}(-1)\cdot\frac{1}{n}+1\cdot\frac{1}{n}+\sum_{i=n+2}^{2n}1\cdot\frac{1}{n} &= 0\end{aligned}$$

לכן $\lim_{n\to\infty}\underline{s}(f,P_n)=\lim_{n\to\infty}\overline{s}(f,P_n)=0$, ולכן f אינטגרבילית ב־$[-1,1]$ והאינטגרל שלה שם הוא 0.

תרגילים

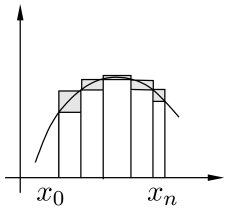
1. תהי $P=\{0,\frac{1}{4},\frac{1}{3},\frac{2}{3},\frac{8}{9},1\}$ חלוקה של $[0,1]$ ותהי $f(x)=x^2$. חשבו את $\underline{s}(f,P)$ ואת $\overline{s}(f,P)$.
2. הוכיחו או הפריכו:
   (א) חלוקה Q היא עידון של חלוקה P אמ"מ כל קטע בחלוקה Q מוכל בקטע בחלוקה P.
   (ב) אם חלוקה Q היא עידון של חלוקה P אז כל קטע בחלוקה Q הוא קטע בחלוקה P.
   (ג) אם R הוא עידון של חלוקה P וגם של חלוקה Q, אז הוא עידון של החלוקה $P\cup Q$.
3. תהי f פונקציה המוגדרת ב־$[a,b]$ ותהי P חלוקה של $[a,b]$. מה אפשר להסיק מהשוויון $\underline{s}(f,P)=\overline{s}(f,P)$?
4. תהי f פונקציה מונוטונית ממש בקטע $[a,b]$. הראו שאם P חלוקה של $[a.b]$ ואם Q עידון ממש של P (כלומר $P\subseteq Q$ אבל $P\neq Q$) האי־שוויונות בלמה 9.1.6 הם חזקים.
5. חשבו את האינטגרלים הבאים:
   (א) $\int_0^1 x^2dx$.
   (ב) $\int_0^1 e^xdx$.
   (ג) $\int_0^1 \sqrt{x}dx$.
   (ד) $\int_1^2 \frac{1}{x}dx$.

   (רמז: בשני הסעיפים האחרונים, לא פשוט להעריך את הסכומים המתאימים לחלוקות $\{0,\frac{1}{n},\frac{2}{n},\dots,\frac{n-1}{n}\}$. חפשו חלוקה אחרת).
6. הראו שאם בקטע $[a,b]$ יש מספרים m,M כך ש־ $m\le f\le M$ אז לכל חלוקה P של $[a,b]$ מתקיים
$$m(b-a)\le\underline{s}(f,P)\le\overline{s}(f,P)\le M(b-a)$$
   הסיקו שאם בנוסף f אינטגרבילית ב־$[a,b]$ אז
$$m(b-a)\le\int_a^b f\le M(b-a)$$
   ובפרט אם $f\ge 0$ אז $\int_a^b f\ge 0$.
7. הראו שאם $f\le g$ ב־$[a,b]$ אז $\underline{s}(f,P)\le\underline{s}(g,P)$ ו־ $\overline{s}(f,P)\le\overline{s}(g,P)$ לכל חלוקה P של $[a,b]$. הסיקו שאם $f\le g$ אינטגרביליות אז $\int_a^b f\le\int_a^b g$.

9.2 התנודה והפרמטר של חלוקה

אינטגרביליות של פונקציה f מבטיחה שישנן חלוקות טובות שעבורן הסכומים העליונים והתחתונים קרובים לערך האינטגרל, ושיש סדרת חלוקות כך שהסכום העליונים והתחתונים המתאימים שואפים לאינטגרל. בדוגמאות ראינו שסדרות כאלה מורכבות באופן טיפוסי מחלוקות אשר הולכות ונעשות עדינות יותר ככל שמתקדמים בסדרה. יש לצפות שזהו גם תנאי מספיק, כלומר, שכל חלוקה עדינה מספיק תיתן קירוב טוב לאינטגרל. עובדה זו אינה מובנת מאליה, ונקדיש את הסעיף הנוכחי להוכחתה.

לפני שנעסוק בכך נגדיר את מושג התנודה של פונקציה ונברר את הקשר שלו לאינטגרביליות. התנודה תהיה כלי מרכזי בהוכחות אינטגרביליות גם בסעיפים הבאים.



איור 9.2.1 השטח המסומן הוא ההפרש בין הסכום העליון לתחתון

אם f אינטגרבילית אז אפשר למצוא חלוקות שעבורן ההפרש $\overline{s}(f,P)-\underline{s}(f,P)$ קטן כרצוננו. מבחינה גאומטרית, ההפרש הזה מתאר את שטח המלבנים הנוצרים כ"הפרש" בין המלבנים התחתונים למלבנים העליונים. ראו איור 9.2.1.

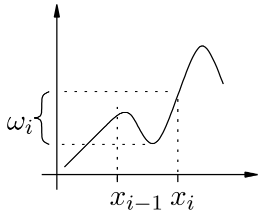
כדי להוכיח שפונקציה אינטגרבילית יש למצוא חלוקות כך ש־ $\overline{s}(f,P)-\underline{s}(f,P)$ קטן. לכן ננסה להבין הפרש זה טוב יותר. תהי $P=\{x_0,\ldots,x_k\}$ חלוקה של $[a,b]$. מתקיים

$$\overline{s}(f,P)-\underline{s}(f,P)=\sum_{i=1}^{k}M_i\Delta x_i-\sum_{i=1}^{k}m_i\Delta x_i=\sum_{i=1}^{k}(M_i-m_i)\Delta x_i$$

משוויון זה אנו רואים שהגורם שמשפיע על $\overline{s}(f,P)-\underline{s}(f,P)$ הם הגדלים M_i-m_i. לא קשה לראות שלכל i מתקיים

$$\begin{aligned} M_i-m_i &= \sup\{f(x)\,:\,x\in[x_{i-1},x_i]\}-\inf\{f(x)\,:\,x\in[x_{i-1},x_i]\} \\ &= \sup\{|f(x')-f(x'')|\,:\,x',x''\in[x_{i-1},x_i]\} \end{aligned}$$

(ודאו שאתם מבינים מדוע השורה השנייה שווה לראשונה!). הגודל האחרון מבטא את מידת הפיזור של ערכי הפונקציה בקטע $[x_{i-1},x_i]$. לגודל זה קוראים תנודה (פגשנו מושג דומה בתרגיל (10) בעמוד 281).



איור 9.2.2 התנודה של פונקציה בקטע החלוקה $[x_{i-1},x_i]$

הגדרה 9.2.1 תהי f פונקציה חסומה ב־$[a,b]$ ו־ $P=\{x_0,\ldots,x_k\}$ חלוקה של $[a,b]$. **התנודה** (oscilation) של f על הקטע $[x_{i-1},x_i]$ היא

$$\omega_i=\sup\{|f(x')-f(x'')|\,:\,x',x''\in[x_{i-1},x_i]\}$$

התנודה (הכוללת) של f ביחס לחלוקה P היא

$$\omega=\sum_{1=1}^{k}\omega_i\Delta x_i$$

כאשר נרצה להדגיש את הפונקציה והחלוקה נשתמש בסימונים המפורטים יותר $\omega(f,P),\omega_i(f,P)$.

לפי הדיון הקודם, $\omega(f,P)$ אינו אלא הגודל $\overline{s}(f,P)-\underline{s}(f,P)$, המתאר את שטח המלבנים באיור 9.2.1. אנו מסיקים ממשפט 9.1.10 את המסקנה הבאה:

מסקנה 9.2.2 תהי f חסומה על $[a,b]$. התנאים הבאים שקולים:

1. f אינטגרבילית על $[a,b]$.
2. לכל $\varepsilon>0$ קיימת חלוקה P של $[a,b]$ כך ש־ $\omega(f,P)<\varepsilon$.
3. יש סדרת חלוקות (P_n) של $[a,b]$ כך ש־ $\omega(f,P_n)\to 0$.

ואם (P_n) סדרה כמו בסעיף (3) אז $\int_a^b f=\lim\underline{s}(f,P_n)=\lim\overline{s}(f,P_n)$.

לפי האי־שוויון הבסיסי שהשתמשנו בו לחלוקות, הסכום התחתון גדל כשמעדנים את החלוקה, והסכום העליון קטן. הלמה הבאה היא התרגום של עובדה זו לשפה של תנודות:

למה 9.2.3 תהי f פונקציה חסומה ב־$[a,b]$ ו־P,Q חלוקות של $[a,b]$ כך ש־Q עידון של P. אז $\omega(f,Q)\le\omega(f,P)$.

הוכחה זו אינה אלא טענה 9.1.6: מכיוון ש־Q מעדן את P,

$$\omega(f,Q)=\overline{s}(f,Q)-\underline{s}(f,Q)\le\overline{s}(f,P)-\underline{s}(f,P)=\omega(f,P)$$

כנדרש. ■

נשוב לעניין שהצגנו בתחילת הסעיף, דהיינו הקשר בין מידת העדינות של חלוקה לבין המידה שבה הסכומים העליונים והתחתונים מקרבים את האינטגרל. באופן אינטואיטיבי חלוקה P של $[a,b]$ היא עדינה אם היא מכילה רק קטעים "קצרים". לכן נגדיר:

הגדרה 9.2.4 אם P חלוקה של $[a,b]$ אז **הפרמטר** (parameter או mesh) של החלוקה $P=\{x_1,\ldots,x_k\}$ הוא האורך המקסימלי של קטע בחלוקה. הפרמטר של P מסומן $\lambda(P)$, דהיינו,

$$\lambda(P)=\max\{\Delta x_1,\Delta x_2,\ldots,\Delta x_k\}$$

אם כן התנאי $\lambda(P)<\delta$ פירושו שכל הקטעים בחלוקה P הם בעלי אורך שאינו עולה על δ. ייתכן שבחלוקה יהיו קטעים קצרים עוד יותר, אך אנו מודדים את העדינות של חלוקה לפי הקטע הארוך ביותר שלה.

דוגמאות

1. אם P היא החלוקה של $[a,b]$ ל־k קטעים שווי אורך אז כל הקטעים בחלוקה הם באורך $\frac{b-a}{k}$ ולכן $\lambda(P)=\frac{b-a}{k}$.
2. אם $n>1$ ו־ $P=\{0,\frac{1}{n},\frac{1}{n-1},\ldots,\frac{1}{2},1\}$ היא החלוקה של $[0,1]$ אז קטע החלוקה הארוך ביותר הוא $[\frac{1}{2},1]$ שאורכו $\frac{1}{2}$ ולכן $\lambda(P)=\frac{1}{2}$ לכל n.

ראינו שהוספת נקודות לחלוקה מקטינה את התנודה הכוללת (למה 9.2.3). הלמה הבאה מבטיחה שאם הפרמטר של החלוקה המקורית קטן אז הוספת מספר קטן של נקודות אינה מקטינה את התנודה במידה ניכרת. עד כמה התנודה קטן תלוי במספר הנקודות שנוספו, בחסם של הפונקציה, ובפרמטר של החלוקה המקורית.

למה 9.2.5 תהי f פונקציה חסומה בקטע $[a,b]$. יהי L חסם של f, תהי P חלוקה של $[a,b]$ ותהי Q עידון של P המתקבלת מ־P על־ידי הוספת n נקודות. אז

$$\omega(f,P)\leq\omega(f,Q)+2nL\cdot\lambda(P)$$

הוכחה הרעיון פשוט: הוספת k נקודות ל־P משפיעה על התנודה רק בקטעים אליהם נוספו נקודות חדשות. אפשר לחסום את השינוי בתנודה בקטע כזה במונחים של החסם L של הפונקציה ושל אורך הקטע. לכן אפשר לחסום את השינוי בתנודה הכוללת במונחים של k,L ו־$\lambda(P)$.

הנה הפרטים. נסמן $P=\{x_1,\ldots,x_k\}$ ונניח תחילה ש־Q היא עידון של P שהתקבל על־ידי הוספת נקודה אחת: $Q=P\cup\{y\}$. במקרה זה כל קטע בחלוקה P הוא גם קטע בחלוקה Q פרט לקטע חלוקה אחד $[x_{i-1},x_i]$, שאינו שייך ל־Q ובמקומו מופיעים שני קטעים $[x_{i-1},y],[y,x_i]$.

נסמן $\omega_j=\omega_j(f,P)$. כעת אם נתבונן בסכומים המגדירים את $\omega(f,P)$ ואת $\omega(f,Q)$, נראה שהם נבדלים רק בכך שבמקום המחובר $\omega_i\Delta x_i$ המופיע ב־$\omega(f,P)$ מופיעים ב־$\omega(f,Q)$ שני מחוברים $\omega'\cdot(y-x_{i-1})$ ו־ $\omega''\cdot(x_i-y)$, כאשר ω',ω'' הם התנודות של f ב־$[x_{i-1},y],[y,x_i]$ בהתאמה. מכאן נובע

$$\begin{aligned}|\omega(f,P)-\omega(f,Q)| &= |\omega_i\Delta x_i-\omega'(y-x_{i-1})-\omega''(x_i-y)|\\ &= |(\omega_i-\omega')(y-x_{i-1})+(\omega_i-\omega'')(x_i-y)|\\ &\leq |\omega_i-\omega'|(y-x_{i-1})+|\omega_i-\omega''|(x_i-y)\end{aligned}$$

כיוון ש־f חסומה על־ידי L נקבל ששלוש התנודות $\omega_i,\omega',\omega''$ חסומות מלמעלה על־ידי $2L$ (וודאו שברור לכם מדוע!). לכן אפשר להמשיך את האי־שוויון האחרון ולקבל

$$|\omega(f,P)-\omega(f,Q)|\leq 2L(y-x_{i-1})+2L(x_i-y)=2L(x_i-x_{i-1})\leq 2L\cdot\lambda(P)$$

(כי $x_i - x_{i-1} = \Delta x_i \le \lambda(P)$). הוכחנו אם כן את הטענה במקרה $n = 1$.

נוכיח את המקרה הכללי באינדוקציה על מספר הנקודות ב־Q שאינן ב־P. מקרה הבסיס הוכח למעלה. נניח ש־Q התקבל מ־P על־ידי הוספת $n+1$ הנקודות $y_1, \ldots, y_{n+1}$. יהי $Q' = P \cup \{y_1, \ldots, y_n\}$. לפי הנחת האינדוקציה המופעלת על החלוקות P ו־Q' מתקיים

$$\omega(f, P) \le \omega(f, Q') + 2Ln \cdot \lambda(P)$$

ומכיוון ש־ $Q = Q' \cup \{y_{n+1}\}$ מתקיים

$$\omega(f, Q') \le \omega(f, Q) + 2L\lambda(Q') \le \omega(f, Q) + 2L \cdot \lambda(P)$$

(למה האי־שוויון האחרון נכון?) שילוב האי־שוויונות נותן

$$\omega(f, P) \le \omega(f, Q) + 2L(n+1) \cdot \lambda(P)$$

כנדרש. ■

רצינו להראות שכל חלוקה עדינה מספיק נותנת קירוב טוב לאינטגרל. המשפט הבא משיג את המטרה:

משפט 9.2.6 תהי f פונקציה חסומה ב־$[a, b]$. אז f אינטגרבילית ב־$[a, b]$ אמ"מ לכל $\varepsilon > 0$ קיים $\delta > 0$ כך שלכל חלוקה Q של $[a, b]$ כך ש־ $\lambda(Q) < \delta$ מתקיים $\omega(f, Q) < \varepsilon$.

הערה הניסוח של המסקנה מזכיר הגדרה של גבול, ולעתים כותבים אותה כך:

$$\int_a^b f = \lim_{\lambda(P) \to 0} \underline{s}(f, P) = \lim_{\lambda(P) \to 0} \overline{s}(f, P)$$

לא נשתמש בסימון זה אך כדאי לתת עליו את הדעת.

הוכחה כיוון אחד של המשפט אינו חדש: אם לכל $\varepsilon > 0$ קיים δ כאמור אז בפרט לכל $\varepsilon > 0$ קיימת חלוקה P עם $\omega(f, P) < \varepsilon$, וזה מספיק על מנת ש־$f$ תהיה אינטגרבילית (מסקנה 9.2.2).

הכיוון השני הוא החדש. נניח ש־f אינטגרבילית, ויהי $\varepsilon > 0$. עלינו למצוא $\delta > 0$ כך שלכל חלוקה P המקיימת $\lambda(P) < \delta$ מתקיים $\omega(f, P) < \varepsilon$. מאינטגרביליות של f קיימת חלוקה P כך ש־ $\omega(f, P) < \frac{\varepsilon}{2}$. נניח ש־$P$ מכילה k נקודות. תהי Q חלוקה שרירותית ונסמן ב־ $R = P \cup Q$ את העידון המשותף של P, Q. מכיוון ש־R עידון של P אנו יודעים ש־

$$\omega(f, R) \le \omega(f, P)$$

מצד שני R התקבלה מ־Q על ידי הוספת k נקודות (אלה הנקודות בחלוקה P) ולכן מלמה 9.2.5 נובע

$$\omega(f,Q) \le \omega(f,R) + 2Lk \cdot \lambda(Q)$$

כאשר L הוא חסם של f ב־$[a,b]$. משילוב האי־שוויונות מקבלים

$$\omega(f,Q) \le \omega(f,P) + 2Lk\lambda(Q) \le \frac{\varepsilon}{2} + 2Lk \cdot \lambda(Q)$$

ולכן אם $\lambda(Q) < \frac{\varepsilon}{4Lk}$ אז $\omega(f,Q) < \varepsilon$. לכן $\delta = \frac{\varepsilon}{4Lk}$ מתאים. ■

מהמשפט נובע שכל סדרת חלוקות שהולכת ומתעדנת "בלי סוף" היא טובה במובן זה שהסכומים התחתונים והעליונים המתאימים מתכנסים לאינטגרל. ליתר דיוק,

הגדרה 9.2.7 אם סדרה (P_n) של חלוקות של $[a,b]$ מקיימת $\lambda(P_n) \to 0$ אומרים שהיא **מפרידה נקודות** (separates points).

מקור השם נובע מהעובדה שאם (P_n) סדרת חלוקות כמו בהגדרה אז לכל $x<y$ בקטע ולכל n גדול מספיק יש נקודה בחלוקה P_n בין x ל־y. נציין שלכל קטע קיימות סדרות של חלוקות כאלה. למשל, אם P_n החלוקה של $[a,b]$ ל־n תת־קטעים שווי־אורך אז $\lambda(P_n) = \frac{b-a}{n} \to 0$ ולכן (P_n) מפרידה נקודות.

מסקנה 9.2.8 תהי f אינטגרבילית על $[a,b]$. אם (P_n) סדרת חלוקות כך ש־$\lambda(P_n) \to 0$ אז $\omega(f,P_n) \to 0$ ומתקיים

$$\int_a^b f = \lim_{n\to\infty} \underline{s}(f,P_n) = \lim_{n\to\infty} \overline{s}(f,P_n)$$

ההוכחה היא מסקנה קלה מהמשפט ומושארת כתרגיל.

תרגילים

1. הוכיחו בפירוט את מסקנה 9.2.2.
2. הראו שאם P חלוקה של $[a,b]$ ואם $f:[a,b]\to\mathbb{R}$ אינטגרבילית אז

$$|\underline{s}(f,P) - \int_a^b f| < \omega(f,P) \qquad , \qquad |\overline{s}(f,P) - \int_a^b f| < \omega(f,P)$$

3. יהיו f,g פונקציות ב־$[a,b]$, ותהי P חלוקה של $[a,b]$. הוכיחו או הפריכו:
   (א) אם $f > g$ בקטע $[a,b]$ אז $\omega(f,P) > \omega(g,P)$.

(ב) אם f רציפה ואם $g(x)=f(x)$ לכל $x\in\mathbb{Q}$ אז $\omega(f,P)\geq\omega(g,P)$.
(ג) אם אף אחד מהיחסים $g\geq f$ או $f\leq g$ אינו נכון אז $\omega(f,P)=\omega(g,P)$.
(ד) אם f,g גזירות ב־$[a,b]$ ואם $0\leq f'<g'$ אז $\omega(f,P)<\omega(g,P)$.

4. תהי $f:[a,b]\to\mathbb{R}$ פונקציה רציפה. נניח שלכל שתי חלוקות P,Q של $[a,b]$ מתקיים $\omega(f,P)=\omega(f,Q)$. מה ניתן להסיק מכך?

5. נניח ש־f גזירה ב־$[a,b]$ ומתקיים $|f'(x)|\leq M$ לכל $x\in[a,b]$. הוכיחו שלכל חלוקה P של $[a,b]$ מתקיים $\omega(f,P)\leq M\lambda(P)(b-a)$. הסיקו ממשפט 9.2.6 שפונקציות כאלה אינטגרביליות.

6. תהי $f:[a,b]\to\mathbb{R}$ פונקציה שהיא הלדר מסדר α, דהיינו יש קבוע C כך ש־ $|f(x)-f(y)|<C|x-y|^\alpha$. עבור איזה δ תוכלו להבטיח שאם $\lambda(P)<\delta$ אז $\omega(f,P)<\varepsilon$?

7. הוכיחו ש־f אינטגרבילית ב־$[a,b]$ אמ"מ קיים מספר I כך שלכל $\varepsilon>0$ ולכל חלוקה P של $[a,b]$ יש עידון Q של P כך ש־ $|I-\underline{s}(f,Q)|,|I-\overline{s}(f,Q)|<\varepsilon$.

8. הוכיחו את תנאי קושי לאינטגרלים: f אינטגרבילית ב־$[a,b]$ אמ"מ לכל $\varepsilon>0$ יש $\delta>0$ כך שלכל שתי חלוקות P,Q של $[a,b]$, אם $\lambda(P),\lambda(Q)<\delta$ אז

$$|\underline{s}(f,P)-\underline{s}(f,Q)|<\varepsilon \qquad , \qquad |\overline{s}(f,P)-\overline{s}(f,Q)|<\varepsilon$$

9.3 משפחות של פונקציות אינטגרביליות

סעיף זה נועד למטרה אחת בלבד: להראות שיש הרבה פונקציות אינטגרביליות.

כדי להוכיח אינטגרביליות של פונקציה, די להראות שיש חלוקות בעלות תנודה כוללת קטנה כרצוננו. לשם כך מספיק למצוא חלוקות כך שבכל קטע חלוקה התנודה קטנה. אם f פונקציה רציפה אפשר לצפות שהתנודה בכל קטע קצר דיו תהיה קטנה, כי רציפות פירושה שלנקודות קרובות יש תמונות קרובות תחת פעולת f. זה הרעיון מאחורי המשפט הבא:

משפט 9.3.1 אם f רציפה על $[a,b]$ אז היא אינטגרבילית שם.

הוכחה ממשפט החסימות של ויירשטראס אנו יודעים ש־f חסומה. לכן די לבדוק את התנאי השני במסקנה 9.2.2, כלומר, שלכל $\varepsilon>0$ קיימת חלוקה P של $[a,b]$ כך ש־ $\omega(f,P)\leq\varepsilon$.

נשים לב שאם נצליח למצוא חלוקה $P=\{x_0,\ldots,x_k\}$ כך ש־ $\omega_i\leq\frac{\varepsilon}{b-a}$ לכל $i=1,\ldots,k$ אז יתקיים

$$\omega=\sum_{i=1}^k\omega_i\Delta x_i<\frac{\varepsilon}{b-a}\sum_{i=1}^k\Delta x_i=\frac{\varepsilon}{b-a}(b-a)=\varepsilon$$

נחפש אם־כן חלוקה P שמקיימת $\omega_i \le \frac{\varepsilon}{2(b-a)}$ לכל i.

ניעזר במשפט קנטור על רציפות במידה שווה (משפט 7.14.2), הקובע שהרציפות של f על $[a,b]$ גוררת רציפות במידה שווה שם. ביתר פירוט, המשפט מבטיח שקיים $\delta > 0$ כך שלכל $x', x'' \in [a,b]$ המקיימים $|x' - x''| < \delta$ מתקיים $|f(x') - f(x'')| < \frac{\varepsilon}{b-a}$.

נבחר חלוקה P כך ש־ $\Delta x_i < \delta$ (בוודאי יש חלוקות כאלה). אז לכל $i = 1, \ldots, k$ ולכל $x', x'' \in [x_{i-1}, x_i]$ מתקיים

$$|x' - x''| \le |x_i - x_{i-1}| = \Delta x_i < \delta$$

ולכן $|f(x') - f(x'')| \le \frac{\varepsilon}{b-a}$. מכאן ש־

$$\omega_i = \sup\{|f(x') - f(x'')| \,:\, x', x'' \in [x_{i-1}, x_i]\} \le \frac{\varepsilon}{b-a}$$

כנדרש. ■

נפנה למחלקה רחבה נוספת של פונקציות: הפונקציות המונוטוניות. כאן ההוכחה לאינטגרביליות מבוססת על כך שהתנודה ω_i של פונקציה f בקטע $[x_{i-1}, x_i]$ היא בעלת צורה פשוטה: $\omega_i = f(x_i) - f(x_{i-1})$ אם f עולה, או $\omega_i = f(x_{i-1}) - f(x_i)$ אם f יורדת. הסיבה היא שממונוטוניות נובע שהערכים המינימליים והמקסימליים של f מתקבלים בקצוות הקטע.

משפט 9.3.2 פונקציה מונוטונית על קטע $[a,b]$ אינטגרבילית שם.

הוכחה נניח למשל ש־f עולה. אז f חסומה, מלעיל על־ידי $f(b)$ ומלרע על־ידי $f(a)$. די שנראה שלכל $\varepsilon > 0$ יש חלוקה $P = \{x_0, \ldots, x_k\}$ של $[a,b]$ כך ש־ $\sum_{i=1}^{k} \omega_i \Delta x_i < \varepsilon$. מכיוון ש־$f$ עולה, בהנתן קטע חלוקה $[x_{i-1}, x_i]$ מתקיים

$$\omega_i = f(x_i) - f(x_{i-1})$$

ולכן עלינו למצוא חלוקה P כך ש־

$$\sum_{i=1}^{k} (f(x_i) - f(x_{i-1}))\Delta x_i < \varepsilon$$

יהי $\varepsilon > 0$. אם נבחר חלוקה $P = \{x_0, \ldots, x_n\}$ של $[a,b]$ ל־n תת־קטעים שווים שכל אחד באורך $\frac{b-a}{n}$ יתקיים $\Delta x_i = \frac{b-a}{n}$, ואז

$$\begin{aligned}
\sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))\Delta x_i &= \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \cdot \frac{b-a}{n} \\
&= \frac{b-a}{n} \sum_{i=1}^{n} f(x_i) - f(x_{i-1}) \\
&= \frac{b-a}{n} \cdot (f(x_k) - f(x_0)) \\
&= \frac{b-a}{n} \cdot (f(b) - f(a))
\end{aligned}$$

ואם n מספיק גדול הביטוי האחרון קטן מ־ε. ■

דוגמאות

1. תהי f הפונקציה

$$f(x) = \begin{cases} -1 & x \le 0 \\ 1 & x > 0 \end{cases}$$

בדוגמה (2) בעמוד 384 ראינו שהיא אינטגרבילית ב־$[-1,1]$, אף על פי שאינה רציפה שם. אפשר להסיק זאת גם מהמשפט האחרון, שכן f היא פונקציה עולה.

2. תהי $(r_n)_{n=1}^{\infty}$ סדרה שמופיעים בה אינסוף מספרים מהקטע $[0,1]$. נגדיר $f:[0,1] \to \mathbb{R}$ על ידי

$$f(x) = \sum_{n:r_n<x} 2^{-n}$$

כפי שראינו בדוגמא שבעמוד 279, f עולה ויש לה קבוצה אינסופית של נקודות אי־רציפות (ואפשר שקבוצה זו גם צפופה בקטע). למרות זאת, f מונוטונית ולכן אינטגרבילית.

דרך נוספת לקבל פונקציות אינטגרביליות היא להתחיל מפונקציה אינטגרבילית ולשנות אותה בצורה עדינה מספיק כך שהאינטגרביליות שלה תישמר. אינטגרביליות היא תכונה עמידה הרבה יותר מרציפות או גזירות. למשל, שינוי של פונקציה במספר סופי של נקודות יכול להפוך אותה מרציפה ללא רציפה, אך לא תשפיע על האינטגרביליות שלה.

משפט 9.3.3 תהי f אינטגרבילית על $[a,b]$. אם g התקבלה מ־f על־ידי שינוי ערכיה במספר סופי של נקודות, כלומר: אם $f(x)=g(x)$ לכל $x \in [a,b]$ פרט למספר סופי של x־ים חריגים, אזי g אינטגרבילית ב־$[a,b]$ ו־ $\int_a^b g = \int_a^b f$.

הוכחה די להוכיח את הטענה בהנחה ש־g התקבלה מ־f על ידי שינוי הערך של f בנקודה אחת p (הטענה הכללית נובעת אז באינדוקציה). מכיוון ש־f אינטגרבילית היא חסומה ולכן גם g חסומה. יהי L חסם משותף של f ו־g.

תהי $P=\{x_1,\ldots,x_k\}$ חלוקה של $[a,b]$. נשים לב שאם $[x_{i-1},x_i]$ הוא קטע חלוקה שאינו מכיל את p אז f,g מסכימות ב־$[x_{i-1},x_i]$ ולכן $m_i(f,P)=m_i(g,P)$.

תהי $I \subseteq \{1,\ldots,k\}$ קבוצת האינדקסים כך ש־$[x_{i-1},x_i]$ לא מכיל את p, ותהי $J=\{1,\ldots,k\} \setminus I$, כך ש־$J$ מכילה לכל היותר שני אינדקסים (ייתכן ש־p שייכת לשני קטעי חלוקה אם היא נקודת קצה משותפת שלהם, אחרת היא שייכת לקטע

יחיד). בביטויים המגדירים את $\underline{s}(f,P),\underline{s}(g,P)$ מופיעים בסכום מחוברים זהים עבור $i \in I$, ובהפרש $\underline{s}(f,P)-\underline{s}(g,P)$ הם מתבטלים ונשארים רק המחוברים מ־J:

$$\underline{s}(f,P)-\underline{s}(g,P)=\sum_{j\in J}(m_j(f,P)-m_j(g,P))\Delta x_j$$

מכיוון ש־L חסם של f,g הרי שלכל $j \in J$ מתקיים $|m_j(f,P)|,|m_j(g,P)| \le L$, ולכן אנו מקבלים

$$|\underline{s}(f,P)-\underline{s}(g,P)| \le \sum_{j\in J}(|m_j(f,P)|+|m_j(g,P)|)\Delta x_j \le 2L\sum_{j\in J}\Delta x_j$$

כעת תהי (P_n) סדרת חלוקות המפרידה נקודות, דהיינו $\lambda(P_n)\to 0$. ממסקנה 9.2.8 אנו יודעים ש־ $\underline{s}(f,P_n)\to\int_a^b f$. לכל n תהי J_n קבוצת האינדקסים j כך שהקטע ה־j בחלוקה P_n מכיל את p. כאמור, $|J_n|\le 2$ לכל n. אם נסמן $P_n=\{x_1,\ldots,x_k\}$ אז מתקיים

$$|\underline{s}(f,P_n)-\underline{s}(g,P_n)| \le 2L\sum_{j\in J_n}\Delta x_j \le 2L\cdot 2\cdot\lambda(P_n)\to 0$$

(השתמשנו כאן בעובדה שבסכום יש לכל היותר שני מחוברים וכל קטע חלוקה של P_n הוא בעל אורך שלא עולה על $\lambda(P_n)$). הגבול שהוכחנו עתה, יחד עם העובדה שהגבול $\lim_{n\to\infty}\underline{s}(f,P_n)$ קיים גוררים שהגבול $\lim_{n\to\infty}\underline{s}(g,P_n)$ קיים גם־כן ויתר על כן

$$\lim_{n\to\infty}\underline{s}(g,P_n)=\lim_{n\to\infty}\underline{s}(f,P_n)=\int_a^b f$$

באופן דומה מראים שמתקיים

$$\lim_{n\to\infty}\overline{s}(g,P_n)=\lim_{n\to\infty}\overline{s}(f,P_n)=\int_a^b f$$

ומכאן נובע ש־g אינטגרבילית ב־$[a,b]$ והאינטגרל שלה הוא $\int_a^b f$. ■

דוגמאות

1. הפונקציה

$$f(x)=\begin{cases}1 & x=0\\ 0 & x\neq 0\end{cases}$$

אינטגרבילית בכל קטע והאינטגרל שלה בכל קטע הוא 0, כי היא נבדלת (בכל קטע, ובכלל) מפונקציית האפס רק בנקודה אחת.

2. אין למשפט 9.3.3 גרסה עבור המקרה שבו משנים פונקציה אינטגרבילית בסדרה אינסופית של נקודות. למשל, פונקציית דירכלה

$$D(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

התקבלה מפונקציית האפס על־ידי שינוי ערכה במספרים הרציונליים, שהם קבוצה קטנה (למשל היא בת־מנייה, ראו סעיף 5.12). בכל־זאת, D אינה אינטגרבילית באף קטע.

נציין שאפשר להראות באופן כללי שאם g התקבלה מ־f על־ידי שינוי בסדרת נקודות ואם f, g שתיהן אינטגרביליות אז האינטגרלים שלהם שווים. לא נוכיח זאת כאן.

אנו מכירים כעת פונקציות אינטגרביליות רבות, ביניהן כל הפונקציות הרציפות אבל גם פונקציות עם הרבה אי־רציפויות. האמת היא שיש קשר בין אינטגרביליות לרציפות. אפשר להוכיח שפונקציה היא אינטגרבילית אמ"מ היא רציפה ב"רוב" הנקודות בקטע, כאשר את המילה "רוב" יש לפרש במובן מדויק שלא הגדרנו. טענה מעט חלשה יותר תמצאו בתרגיל (3) למטה. אפשר למצוא פרטים נוספים בכרך השני של [9].

תרגילים

1. הראו שפונקציית רימן אינטגרבילית בכל קטע $[a, b]$ אף שיש קבוצת נקודות צפופה שבה איננה רציפה (פונקציית רימן הוגדרה בעמוד 245).

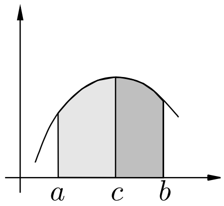
2. תהי f פונקציה אינטגרבילית ב־$[a, b]$. יהי $\varepsilon > 0$ ותהי $P = \{x_1, \ldots, x_k\}$ חלוקה של $[a, b]$ כך ש־ $\omega(f, P) < \varepsilon$. הוכיחו שקיימת קבוצת אינדקסים $I \subseteq \{1, 2, \ldots, k\}$ כך ש־ $\omega_i < \sqrt{\varepsilon}$ לכל $i \in I$ ו־

$$\sum_{i \in I} \Delta x_i \geq (1 - \sqrt{\varepsilon}) \cdot (b - a)$$

כלומר, אוסף הקטעים $\{\Delta x_i : i \in I\}$ מכסה את כל הקטע $[a, b]$ פרט לחלק יחסי $\sqrt{\varepsilon}$ של הקטע (רמז: הניחו תחילה שאורך הקטע הוא 1, והיעזרו בתרגיל (8) בעמוד 364).

3. הוכיחו שאם f אינטגרבילית ב־$[a, b]$ אז קבוצת נקודות הרציפות שלה צפופה ב־$[a, b]$, כלומר כל תת־קטע פתוח של $[a, b]$ מכיל נקודות רציפות של f (רמז: הראו שדי למצוא נקודת רציפות אחת. בעזרת השאלה הקודמת מצאו סדרה יורדת של קטעים סגורים I_n ללא נקודות קצה משותפות, כך שהתנודה של f ב־I_n שואפת לאפס עם n. הראו שכל נקודה בחיתוך היא נקודת רציפות. אפשר גם להיעזר במשפט 9.4.1 להלן).

9.4 משפטי תחשיב



איור 9.4.1 השטח הכולל שווה לסכום השטחים

מתקבל על הדעת שאם מחלקים צורה במישור לשניים אז השטח הכולל שווה לסכום השטחים של החלקים. לכן נצפה שכשמחלקים קטע לשני תת־קטעים משלימים, האינטגרל של פונקציה בקטע כולו שווה לסכום האינטגרלים בתת־הקטעים (ראו איור 9.4.1). לפני שנוכיח זאת עלינו לברר את הקשר בין אינטגרביליות בקטע לאינטגרביליות בתת־הקטעים.

משפט 9.4.1 אם f אינטגרבילית ב־$[a,b]$ אז היא אינטגרבילית בכל תת־קטע של $[a,b]$

הוכחה יהי $[a',b'] \subseteq [a,b]$. מספיק שנראה שלכל $\varepsilon > 0$ יש חלוקה P' של $[a',b']$ כך ש־ $\omega(f,P') < \varepsilon$.

יהי $\varepsilon > 0$. לפי ההנחה יש חלוקה Q של $[a,b]$ כך ש־ $\omega(f,Q) < \varepsilon$. נגדיר $P = Q \cup \{a',b'\}$. אז P היא חלוקה של $[a,b]$ והיא עידון של Q (ייתכן גם ש־ $Q = P$). לכן מלמה 9.2.3 מתקיים $\omega(f,P) \leq \omega(f,Q)$.

תהי $P' = P \cap [a',b']$, כלומר P' מכילה את כל הנקודות מ־P שנמצאות בקטע $[a',b']$. מכיוון ש־$a',b' \in P$ גם $a',b' \in P'$ ולכן P' חלוקה של $[a',b']$. כמו־כן כל קטע בחלוקה P' הוא קטע בחלוקה P, ולכן אם נשווה את הסכום המגדיר את $\omega(f,P')$ עם הסכום המגדיר את $\omega(f,P)$ נגלה שהראשון מכיל רק חלק מהאיברים המופיעים בשני, ומכיוון שבסכום כל המחוברים אי־שליליים, נובע ש־

$$\omega(f,P') \leq \omega(f,P) \leq \omega(f,Q) < \varepsilon$$

וסיימנו. ■

המשפט הבא עולה על השאלה שהצגנו בתחילת הסעיף, וגם מספק כיוון הפוך למשפט האחרון.

משפט 9.4.2 תהי f מוגדרת ב־$[a,b]$ ויהי $c \in (a,b)$. אז f אינטגרבילית ב־$[a,b]$ אמ"מ היא אינטגרבילית ב־$[a,c]$ וב־ $[c,b]$, ובמקרה זה מתקיים

$$\int_a^b f = \int_a^c f + \int_c^b f$$

הוכחה האינטגרביליות של f ב־$[a,b]$ גוררת אינטגרביליות בתת־קטעים $[a,c],[c,b]$ לפי המשפט הקודם. לכן מספיק להוכיח שאם f אינטגרבילית ב־$[a,c]$ וב־$[c,b]$ אז היא אינטגרבילית ב־$[a,b]$ ו־ $\int_a^b f = \int_a^c f + \int_c^b f$.

לכל חלוקה P' של $[a,c]$ ולכל חלוקה P'' של $[c,b]$ האיחוד $P = P' \cup P''$ הוא חלוקה של $[a,b]$. קטעי החלוקה P הם כלל קטעי החלוקות של P' ושל P'', ולכן קל לראות ש־

$$\underline{s}(f,P) = \underline{s}(f,P') + \underline{s}(f,P'') \qquad , \qquad \overline{s}(f,P) = \overline{s}(f,P') + \overline{s}(f,P'')$$

כעת אם נבחר סדרות של חלוקות $(P'_n),(P''_n)$ של $[a,c],[c,b]$ בהתאמה כך ש־ $\underline{s}(f,P'_n),\overline{s}(f,P'_n) \to \int_a^c f$ ו־ $\underline{s}(f,P''_n),\overline{s}(f,P''_n) \to \int_c^b f$, וניצור את $P_n = P'_n \cup P''_n$. אז

$$\lim_{n\to\infty} \underline{s}(f,P_n) = \lim_{n\to\infty} \underline{s}(f,P'_n) + \lim_{n\to\infty} \underline{s}(f,P''_n) = \int_a^c f + \int_c^b f$$

ובאותו אופן

$$\lim_{n\to\infty} \overline{s}(f,P_n) = \lim_{n\to\infty} \overline{s}(f,P'_n) + \lim_{n\to\infty} \underline{s}(f,P''_n) = \int_a^c f + \int_c^b f$$

ולכן f אינטגרבילית ב־$[a,b]$ ו־ $\int_a^b f = \int_a^c f + \int_c^b f$, כפי שרצינו. ■

גם המשפט הבא הוא מעין כיוון שני של משפט **9.4.1**. אנו משאירים את הוכחתו כתרגיל:

משפט 9.4.3 תהי f פונקציה חסומה בקטע $[a,b]$ ואינטגרבילית בקטע $[a+\delta, b-\delta]$ לכל $\delta > 0$. אז f אינטגרבילית ב־$[a,b]$, ומתקיים $\int_a^b f = \lim_{\delta\to 0+} \int_{a+\delta}^{b-\delta} f$.

נעבור למשפטי תחשיב ביחס לאינטגרנד. המצב כאן פחות טוב מאשר במקרה של גבולות או נגזרות, וברוב המקרים לא נוכל לקבל משפט תחשיב של ממש אלא רק להבטיח אינטגרביליות. נציין גם שלעתים אינטגרביליות נובעת משיקולי רציפות, והמשפטים המבטיחים אינטגרביליות מעניינים רק מכיוון שהם נכונים גם לפונקציות שאינן רציפות.

מקרה אחד שבו בכל־זאת ישנו משפט תחשיב מלא נתון במשפט הבא:

משפט 9.4.4 יהיו f,g מוגדרות על $[a,b]$ ו־$c \in \mathbb{R}$. אם f,g אינטגרביליות על $[a,b]$ אז כך גם הפונקציות $-f, f+g, cf$, ומתקיים

$$\begin{aligned}
\int_a^b (f+g) &= \int_a^b f + \int_a^b g \\
\int_a^b (-f) &= -\int_a^b f \\
\int_a^b cf &= c\int_a^b f
\end{aligned}$$

הוכחה מאחר ש־ $-f = (-1)\cdot f$, טענת המשפט עבור $-f$ נובעת מהטענה על cf לקבוע כללי c. אנו נוכיח רק את הטענה על $f+g$. את הטענה על cf אנו משאירים כתרגיל.

תהי $R=\{x_0,\ldots,x_k\}$ חלוקה של $[a,b]$. אם $[x_{i-1},x_i]$ קטע בחלוקה אז

$$\begin{aligned} m_i(f+g,R) &= \inf\{(f+g)(x)\ :\ x\in[x_{i-1},x_i]\} \\ &\geq \inf\{f(x)\ :\ x\in[x_{i-1},x_i]\}+\inf\{g(x)\ :\ x\in[x_{i-1},x_i]\} \\ &= m_i(f,R)+m_i(g,R) \end{aligned}$$

אנו מסיקים לכן ש־

$$\begin{aligned} \underline{s}(f+g,R) &= \sum_{i=1}^{k} m_i(f+g,R)\Delta x_i \\ &\geq \sum_{i=1}^{k} m_i(f,R)\Delta x_i+\sum_{i=1}^{k} m_i(g,R)\Delta x_i \\ &= \underline{s}(f,R)+\underline{s}(g,R) \end{aligned}$$

באופן דומה מתקיים

$$M_i(f+g,R)\leq M_i(f,R)+M_i(g,R)$$

ולכן

$$\overline{s}(f+g,R)\leq \overline{s}(f,R)+\overline{s}(g,R)$$

תהי (R_n) סדרת חלוקות המפרידה נקודות. קיבלנו למעלה שלכל n מתקיים

$$\underline{s}(f,R_n)+\underline{s}(g,R_n)\leq \underline{s}(f+g,R_n)\leq \overline{s}(f+g,R_n)\leq \overline{s}(f,R_n)+\overline{s}(g,R_n)$$

הביטויים בקצוות שואפים שניהם ל־ $\int_a^b f+\int_a^b g$ (כי (R_n) מפרידה נקודות ו־f,g אינטגרביליות) ולכן כך גם שני הביטויים המרכזיים. אבל מכך כבר נובע ש־ $f+g$ אינטגרבילית ושהאינטגרל שלה הוא $\int_a^b f+\int_a^b g$, כנדרש. ■

לגבי המכפלה והמנה של פונקציות, לא נוכל לחשב את האינטגרל שלהן אלא נצטרך להסתפק בידיעה שהן אינטגרביליות.

משפט 9.4.5 יהיו f,g אינטגרביליות על $[a,b]$. אז $f\cdot g$ אינטגרבילית ב־$[a,b]$, ואם קיים $c>0$ כך ש־ $|f(x)|>c$ לכל $x\in[a,b]$ אז $\frac{1}{f}$ אינטגרבילית ב־$[a,b]$.

הוכחה תהי $P=\{x_0,\ldots,x_k\}$ חלוקה של $[a,b]$. מאינטגרביליות אנו מקבלים ש־f,g חסומות ב־$[a,b]$. יהי L חסם משותף ל־f,g. כדי להעריך את $\omega(fg,P)$

במונחים של $\omega(f,P),\omega(g,P)$ נשתמש בשיטה הקבועה להערכת מכפלות: נשים לב שאם $x',x''\in[x_{i-1},x_i]$ אז

$$\begin{aligned}
&|f(x')g(x')-f(x'')g(x'')| \\
&\quad\le |f(x')g(x')-f(x')g(x'')|+|f(x')g(x'')-f(x'')g(x'')| \\
&\quad= |f(x')|\cdot|g(x')-g(x'')|+|g(x'')|\cdot|f(x')-f(x'')| \\
&\quad\le L|g(x')-g(x'')|+L|f(x')-f(x'')|
\end{aligned}$$

ומכאן

$$\omega_i(fg,P)\le L\cdot(\omega_i(f,P)+\omega_i(g,P))$$

לכן

$$\begin{aligned}
\omega(fg,P) &= \sum_{i=1}^{k}\omega_i(fg,P)\Delta x_i \\
&\le \sum_{i=1}^{k}L\omega_i(f,P)\Delta x_i+\sum_{i=1}^{k}L\omega_i(g,P)\Delta x_i \\
&= L(\omega(f,P)+\omega(g,P))
\end{aligned}$$

כעת אם (P_n) סדרת חלוקות של $[a,b]$ המפרידה נקודות אז $\omega(f,P_n)\to 0$ וגם $\omega(g,P_n)\to 0$, והאי־שוויון הקודם וכלל הסנדוויץ' נותנים $\omega(fg,P_n)\to 0$, ומזה נובע ש־fg אינטגרבילית.

ההוכחה עבור $\frac{1}{f}$ דומה, ומושארת כתרגיל (תצטרכו להעריך את התנודה של $\frac{1}{f}$ בקטע במונחים של התנודה של f. לשם כך אפשר להיעזר בהוכחה של כלל המנה לסדרות). ■

נעבור לאינטגרציה של פונקציות מורכבות. כאן המצב דומה במקצת למשפט על הגבול של פונקציה מורכבת: באופן כללי לא ניתן להבטיח דבר אודות $g\circ f$, אלא אם g רציפה.

משפט 9.4.6 תהי f פונקציה אינטגרבילית ב־$[a,b]$ עם ערכים בקטע $[c,d]$ ותהי g רציפה ב־$[c,d]$. אז $g\circ f$ אינטגרבילית ב־$[a,b]$.

הוכחה הרעיון, בקיצור, הוא שאם $I\subseteq[a,b]$ הוא קטע ו־ω התנודה של f ב־I, אז לכל $x',x''\in I$ מתקיים $|f(x')-f(x'')|<\omega$. מכיוון ש־g רציפה נצפה שכאשר ω קטן זה יגרור ש־ $|g(f(x'))-g(f(x''))|$ היא גם־כן קטן. כך ניתן לשלוט על התנודה של $g\circ f$ בקטע I או ביחס לחלוקה של I.

ליתר דיוק, עלינו להראות שלכל $\varepsilon>0$ יש חלוקה P של $[a,b]$ כך ש־ $\omega(g\circ f,P)<\varepsilon$. יהי $\varepsilon>0$. מכיוון ש־g רציפה ב־$[c,d]$ היא רציפה שם במידה שווה, וקיים $\delta>0$ כך שלכל $y',y''\in[c,d]$, אם $|y'-y''|<\delta$ אז $|g(y')-g(y'')|<\varepsilon$.

מכיוון ש־f אינטגרבילית ב־$[a,b]$ יש חלוקה $P=\{x_0,\dots,x_k\}$ של $[a,b]$ כך ש־$\omega(f,P)<\delta^2$. אנו טוענים שתנאי זה גורר שלמרבית ה־i־ים, התנודה של f בקטע ה־i קטנה. ליתר דיוק, אנו נראה שסכום אורכי הקטעים בהם התנודה של f גדולה מ־δ בעצמו אינו עולה על δ. לשם כך נגדיר

$$I=\{1\le i\le k\,:\,\omega_i(f,P)<\delta\}$$

ויהי $J=\{1,\dots,k\}\setminus I$. אז

$$\begin{aligned}\omega(f,P)&=\sum_{i=1}^{k}\omega_i(f,P)\Delta x_i\\&=\sum_{i\in I}\omega_i(f,P)\Delta x_i+\sum_{i\in J}\omega_i(f,P)\Delta x_i\\&\ge\sum_{i\in J}\omega_i(f,P)\Delta x_i\\&\ge\delta\sum_{i\in J}\Delta x_i\end{aligned}$$

כאשר האי־שוויון האחרון נובע מכך שלכל $i\in J$ מתקיים לפי הגדרת J ש־ $\omega_i\ge\delta$. אנו מסיקים ש־

$$\sum_{i\in J}\Delta x_i\le\frac{\omega(f,P)}{\delta}<\frac{\delta^2}{\delta}=\delta$$

כעת נחסום את התנודה של $g\circ f$ ביחס לחלוקה P. ישנם כמה קטעים בחלוקה שבהם התנודה של f גדולה, אך סך־כל האורך של קטעים אלו קטן, ולא משפיע הרבה על התנודה הכוללת. לעומת זאת, ביתר הקטעים התנודה של f קטנה ולכן התנודה של $g\circ f$ גם־כן קטנה. באופן מדויק, יהי L חסם של g ב־$[c,d]$, אז התנודה של $g\circ f$ בכל קטע חלוקה אינה עולה על $2L$, ולכן

$$\begin{aligned}\omega(g\circ f,P)&=\sum_{i=1}^{k}\omega_i(g\circ f,P)\Delta x_i\\&=\sum_{i\in I}\omega_i(g\circ f,P)\Delta x_i+\sum_{i\in J}\omega_i(g\circ f,P)\Delta x_i\\&\le\sum_{i\in I}\omega_i(g\circ f,P)\Delta x_i+2L\sum_{i\in J}\Delta x_i\\&=\sum_{i\in I}\omega_i(g\circ f,P)\Delta x_i+2L\delta\end{aligned}$$

מצד שני לכל $i\in I$ ולכל $x',x''\in[x_{i-1},x_i]$ מתקיים

$$|f(x')-f(x'')|\le\omega_i(f,P)<\delta$$

ולכן לפי בחירת δ,

$$|g(f(x')) - g(f(x''))| < \varepsilon$$

דהיינו $\omega_i(g \circ f, P) \leq \varepsilon$, וקיבלנו

$$\begin{aligned} \omega(g \circ f, P) &\leq \sum_{i \in I} \varepsilon \Delta x_i + 2L\delta \\ &\leq \varepsilon(b-a) + 2L\delta \end{aligned}$$

מהגדרת δ ברור שאם נקטין אותו עוד הוא ימשיך לקיים את התכונות שלשמו הוא נבחר, ולכן נוכל להניח שהוא מקיים $\delta < \frac{\varepsilon}{2L}$ (שימו לב ש־L תלוי רק ב־g, ו־δ נבחר אחרי ε, ולכן מותר ש־δ יהיה תלוי ב־ε, L). אז מה שהראינו הוא שלכל $\varepsilon > 0$ יש חלוקה P כך ש־ $\omega(g \circ f, P) \leq \varepsilon(b-a) + \varepsilon$, וזה מספיק כדי להסיק ש־$g \circ f$ אינטגרבילית. ■

נסיים את הסעיף בכמה אי־שוויונות. הראשונים קלים מאד, וכבר הוכחתם אותם בתרגילים בסוף סעיף 9.1, אך הם חשובים ולכן ננסח אותם במלואם:

טענה 9.4.7 יהיו f, g אינטגרביליות ב־$[a,b]$. אם $f(x) \leq g(x)$ לכל $x \in [a,b]$ אז $\int_a^b f \leq \int_a^b g$. בפרט אם $m \leq f \leq M$ אז $m(b-a) \leq \int_a^b f \leq M(b-a)$, ואם $f \geq 0$ אז $\int_a^b f \geq 0$.

מהטענה האחרונה נובע שהאינטגרל של פונקציה אי־שלילית הוא אי־שלילי. אילו תנאים יש להוסיף להנחות הטענה כדי להבטיח שהאינטגרל חיובי ממש? לא די לדרוש שהפונקציה לא תתאפס זהותית. למשל, אם $f(0) = 1$ ו־ $f(x) = 0$ לכל $x \neq 0$ אז f אינה מתאפסת זהותית ב־$[-1,1]$, אבל $\int_{-1}^1 f = 0$, כי f נבדלת מפונקציית האפס רק בנקודה אחת. בכל־זאת, בתנאים מסוימים חיוביות של f תגרור חיוביות של $\int_a^b f$.

טענה 9.4.8 תהי $f : [a,b] \to \mathbb{R}$.

1. אם f רציפה ואי־שלילית ואם f אינה זהותית אפס אז $\int_a^b f > 0$.
2. אם f אינטגרבילית וחיובית ממש בכל $[a,b]$ אז $\int_a^b f > 0$.

הוכחה נוכיח את החלק הראשון, החלק השני מושאר כתרגיל (ראו תרגיל (8) בסוף הסעיף). נניח ש־f רציפה, אי־שלילית ואינה זהותית 0 ב־$[a,b]$, כלומר יש $x_0 \in [a,b]$ עם $f(x_0) > 0$. נניח ש־x_0 נקודה פנימית של הקטע (ההוכחה של המקרה האחר דומה). אז מרציפות יש $M > 0$ וסביבה מלאה $[a,b] \supseteq (a', b')$ של x_0 שבה $f \geq M$. מצד שני,

$$\int_a^b f = \int_a^{a'} f + \int_{a'}^{b'} f + \int_{b'}^b f$$

המחובר הראשון והשלישי אי־שליליים כי $f \geq 0$, ואילו האינטגרנד במחובר האמצעי הוא לפחות M בקטע האינטגרציה (למה זה נכון בקצוות?) ולכן $\int_{a'}^{b'} f \geq M(b'-a')$. קיבלנו ש־

$$\int_a^b f \geq M(b'-a') > 0$$

כנדרש. ■

את הטענה הבאה אפשר להוכיח כבר עכשיו, אך נוכל לתת הוכחה פשוטה יותר בהמשך (ראו עמוד 408):

טענה 9.4.9 (אי־שוויון המשולש) תהי $f : [a,b] \to \mathbb{R}$ פונקציה אינטגרבילית. אז $|f|$ אינטגרבילית ב־$[a,b]$ ומתקיים $|\int_a^b f| \leq \int_a^b |f|$.

תרגילים

1. תהי $f : [-1,1] \to \mathbb{R}$ פונקציה אינטגרבילית. הראו שאם f אי־זוגית אז $\int_{-1}^1 f = 0$ ואם f זוגית אז $\int_{-1}^1 f = 2\int_0^1 f$.
2. הוכיחו ש־ $\int_a^b f(x+c)dx = \int_{a+c}^{b+c} f(x)dx$ ו־ $\int_a^b f(cx)dx = \frac{1}{c}\int_{a/c}^{b/c} f(y)dy$ (כלומר, הראו בכל אחד מהמקרים שקיום האינטגרל באגף ימין שקול לקיום האינטגרל באגף שמאל, ושאם הם קיימים אז הם שווים).
3. תהי $f : \mathbb{R} \to \mathbb{R}$ פונקציה מחזורית עם מחזור Δ ונניח שהיא אינטגרבילית ב־$[0,\Delta]$. הראו שהיא אינטגרבילית בכל קטע סגור ושמתקיים $\int_a^{a+\Delta} f = \int_0^\Delta f$ לכל $a \in \mathbb{R}$.
4. הוכיחו את משפט הסנדוויץ' הבא. יהיו פונקציה f,g,h המקיימות $f \leq g \leq h$ ב־$[a,b]$. נניח ש־f,h אינטגרביליות ב־$[a,b]$ וש־ $\int_a^b f = \int_a^b h$. אז גם g אינטגרבילית ב־$[a,b]$, וכל שלושת האינטגרלים שווים.
5. נכון או לא נכון: אם לפונקציה יש אינסוף נקודות אי־רציפות שאינן סליקות, אז היא אינה אינטגרבילית.
6. תהי $f : [a,b] \to [c,d]$ ותהי $g : [c,d] \to \mathbb{R}$ ונניח ששתיהן אינטגרביליות. נניח עוד של־g יש רק נקודת אי רציפות אחת. האם $g \circ f$ אינטגרבילית?
7. הוכיחו את כלל המנה (משפט 9.4.5).
8. הוכיחו את הסעיף השני במשפט 9.4.8: אם f חיובית ממש ואינטגרבילית ב־$[a,b]$ אז $\int_a^b f > 0$. (רמז: הניחו בשלילה ש־ $\int_a^b f = 0$. הראו שלכל $\varepsilon > 0$ ולכל קטע סגור $I \subseteq [a,b]$, קיים תת־קטע סגור $J \subseteq I$ כך ש־ $\sup_{x\in J} f(x) < \varepsilon$. הסיקו שיש נקודה $x_0 \in [a,b]$ עם $f(x_0) = 0$).
9. תהי $f : [a,b] \to \mathbb{R}$ פונקציה רציפה ואי־שלילית. הראו שאם $\int_a^b fg = 0$ לכל פונקציה אינטגרבילית g, אז $f \equiv 0$.

10. תהי f פונקציה רציפה ב־$[a,b]$. הוכיחו ש־

$$\lim_{n\to\infty}(\int_a^b |f|^n)^{1/n} = \max_{x\in[a,b]} |f(x)|$$

11. הוכיחו את משפט קושי־שוורץ: אם f,g אינטגרביליות ב־$[a,b]$ אז

$$\int_a^b fg \le \sqrt{\int_a^b f^2}\cdot\sqrt{\int_a^b g^2}$$

תנו שתי הוכחות:

(א) בעזרת משפט קושי־שוורץ לסכומים, תרגיל (5) בעמוד 78.

(ב) התבוננו בפונקציה $\varphi(t)=\int_a^b (f+tg)^2$. שימו לב ש־ $\varphi\ge 0$ (למה?). מצאו את נקודת המינימום הגלובלית t_0 של φ, והתבוננו באי־שוויון $0\le\varphi(t_0)$.

12. נניח ש־f פונקציה בקטע $[a,b]$ ושיש חלוקה $P=\{x_0,x_1,\ldots,x_k\}$ של $[a,b]$ כך ש־f קבועה בכל קטע (x_{i-1},x_i) (פונקציה כזאת נקראת פונקציית מדרגה (step function)). במקרה זה נאמר שהחלוקה P מתאימה ל־f. שימו לב שיש הרבה חלוקות המתאימות ל־f: אם P מתאימה אז כל עידון של P מתאים גם־כן. ברור שאם f פונקציית מדרגה אז היא מקבלת מספר סופי של ערכים ושאם P חלוקה מתאימה ל־f אז כל נקודת אי־רציפות של f היא אחת מנקודות החלוקה.

(א) הוכיחו שאם f פונקציית מדרגה ב־$[a,b]$ אז היא אינטגרבילית. הסיקו שאם $P=\{x_0,x_1,\ldots,x_k\}$ חלוקה מתאימה ל־f ו־f מקבלת את הערך c_i בקטע (x_{i-1},x_i) אז $\int_a^b f=\sum_{i=1}^k c_i\Delta x_i$.

(ב) הוכיחו שפונקציה כללית f אינטגרבילית ב־$[a,b]$ אמ"מ לכל $\varepsilon>0$ יש פונקציות מדרגה φ,ψ ב־$[a,b]$ כך ש־ $\varphi\le f\le\psi$ ו־ $\int|\psi-\varphi|<\varepsilon$, ובמקרה זה מתקיים

$$\begin{aligned}\int_a^b f &= \sup\{\int_a^b \varphi : \varphi\le f \text{ עם } \varphi \text{ פונקציית מדרגות}\}\\ &= \inf\{\int_a^b \psi : f\le\psi \text{ עם } \psi \text{ פונקציית מדרגות}\}\end{aligned}$$

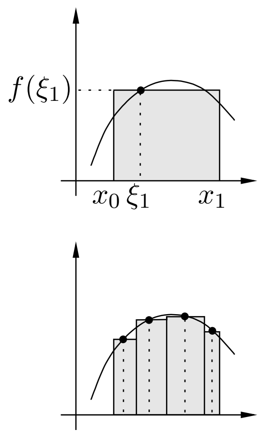
(ג) הוכיחו שאם f אינטגרבילית ב־$[a,b]$ אז

$$\begin{aligned}\int_a^b f &= \sup\{\int_a^b \varphi : \varphi\le f \text{ עם } \varphi \text{ פונקציה רציפה}\}\\ &= \inf\{\int_a^b \psi : f\le\psi \text{ עם } \psi \text{ פונקציה רציפה}\}\end{aligned}$$

9.5 האינטגרל המסוים לפי רימן

הגדרת האינטגרל שהצגנו בתחילת הפרק היא מאוחרת יחסית מבחינה היסטורית. אחת ההגדרות המוקדמות היא השיטה של סכומי רימן, שנדון בה להלן. השיטה שונה מהשיטה של דרבו, אף שבדיעבד יסתבר שההגדרות שקולות.

האינטגרל של דרבו מבוסס על קירוב האינטגרל על ידי סכומים עליונים ותחתונים, כלומר סכומים מהצורה $\sum_{i=1}^{k} m_i\Delta x_i$ ו־$\sum_{i=1}^{k} M_i\Delta x_i$. סכומים אלה חוסמים את האינטגרל מלמטה ומלמעלה. הרעיון מאחורי אינטגרל רימן הוא להחליף את המספרים m_i, M_i בערך $f(\xi_i)$ של הפונקציה בנקודה ξ_i כלשהי בקטע $[x_{i-1}, x_i]$, כלומר, עם "דגימה" של הפונקציה בנקודה כלשהי.



איור 9.5.1 סכומי רימן

הגדרה 9.5.1 תהי $f : [a,b] \to \mathbb{R}$ פונקציה ותהי $P = \{x_0, \dots, x_k\}$ חלוקה של $[a,b]$. **בחירת נקודות** עבור P היא סדרה $\xi = (\xi_i)_{i=1}^{k}$ כך ש־ $\xi_i \in [x_{i-1}, x_i]$. **סכום רימן** (Riemann sum) של f ביחס לחלוקה P ובחירת הנקודות ξ הוא המספר

$$\sigma(f, P, \xi) = \sum_{i=1}^{k} f(\xi_i)\Delta x_i$$

שימו לב שסכום רימן מוגדר גם לפונקציות לא חסומות.

ניתן לפרש סכום זה כניסיון להעריך את השטח מתחת לגרף של f בעזרת מלבנים שבסיסם $[x_{i-1}, x_i]$ וגובהם $f(\xi_i)$. סכום רימן הוא אז סכום שטחי המלבנים.

יש לצפות שסכומי רימן של פונקציה יקרבו את האינטגרל ככל שהחלוקה עדינה יותר, ורעיון זה מוביל להגדרת האינטגרל של רימן:

הגדרה 9.5.2 (האינטגרל לפי רימן) תהי f פונקציה המוגדרת על $[a,b]$. נאמר ש־f **אינטגרבילית לפי רימן** אם קיים מספר I כך שלכל $\varepsilon > 0$ קיים $\delta > 0$ כך שלכל חלוקה P של $[a,b]$ עם $\lambda(P) < \delta$ ולכל בחירת נקודות ξ עבור P מתקיים

$$|\sigma(f, P, \xi) - I| < \varepsilon$$

במקרה זה I נקרא **אינטגרל רימן** (Riemann integral) של f בקטע $[a,b]$.[1]

---

[1]אינטגרל רימן מוגדר כסוג של גבול, ואם I הוא אינטגרל רימן של f לעתים רושמים

$$I = \lim_{\lambda(P)\to 0} \sigma(f, P, \xi)$$

הערות

1. יש לנו כעת שני סוגי אינטגרלים, ועל פניו המושגים שונים זה מזה. בעמודים הקרובים נוכיח את שקילותם. עד אז חשוב להבחין ביניהם, ונקפיד להבחין בין אינטגרל דרבו ואינטגרביליות דרבו לבין אינטגרל רימן ואינטגרביליות רימן. הסימן $\int_a^b f$ ימשיך לציין את אינטגרל דרבו.
2. לא קשה לבדוק שאם f אינטגרבילית לפי רימן אז יש מספר יחיד I שהוא אינטגרל רימן שלה (זו חזרה נוספת על השיטה שבה הוכחנו כמה פעמים את יחידות הגבול. השלימו את הפרטים!).

נרצה לנתח את הקשר בין סכומי רימן לסכומים העליונים והתחתונים של פונקציה. תהי f חסומה על $[a,b]$, תהי $P = \{x_0, \ldots, x_k\}$ חלוקה של $[a,b]$ ותהי ξ בחירת נקודות עבור P. מההגדרה של החסמים m_i, M_i ברור ש־ $m_i \leq f(\xi_i) \leq M_i$ ואנו מקבלים מיד את האי־שוויון

$$\underline{s}(f,P) \leq \sigma(f,P,\xi) \leq \overline{s}(f,P)$$

אשר תקף לכל בחירת נקודות ξ עבור P. מכיוון ששני הקצוות של האי־שוויון שואפים ל־$\int_a^b f$ "כאשר $\lambda(P) \to 0$" כך צריך להיות גם לסדרה הכלואה ביניהם. ואמנם,

משפט 9.5.3 אם f אינטגרבילית דרבו בקטע $[a,b]$ אז היא אינטגרבילית רימן שם, והאינטגרלים שווים.

הוכחה יהי $\varepsilon > 0$. מכיוון ש־f אינטגרבילית דרבו יש $\delta > 0$ כך שלכל חלוקה P עם $\lambda(P) < \delta$ מתקיים $\omega(f,P) < \varepsilon$, כלומר: $\overline{s}(f,P) - \underline{s}(f,P) < \varepsilon$. אם P חלוקה כזו ו־ξ בחירת נקודות עבור P אז לפי האמור למעלה

$$\underline{s}(f,P) \leq \sigma(f,P,\xi) \leq \overline{s}(f,P)$$

אבל גם מתקיים

$$\underline{s}(f,P) \leq \int_a^b f \leq \overline{s}(f,P)$$

חיסור האי־שוויונות זה מזה נותן

$$-\omega(f,P) \leq \int_a^b f - \sigma(f,P,\xi) \leq \omega(f,P)$$

דהיינו $|\sigma(f,P,\xi) - \int_a^b f| < \varepsilon$. בסיכום, הראינו שלכל $\varepsilon > 0$ יש $\delta > 0$ כך שלכל חלוקה P עם $\lambda(P) < \delta$ ולכל בחירת נקודות ל־P סכום רימן המתאים קרוב כדי ε ל־$\int_a^b f$, ומכאן ש־f אינטגרבילית רימן ב־$[a,b]$ ואינטגרל רימן שלה הוא $\int_a^b f$. ■

כדי להוכיח את הכיוון השני, דהיינו שאינטגרביליות רימן גוררת אינטגרביליות דרבו, נצטרך ראשית להראות שאינטגרביליות רימן גוררת חסימות. ואמנם,

למה 9.5.4 אם f אינטגרבילית רימן ב־$[a,b]$ אז f חסומה ב־$[a,b]$.

הוכחה די שנראה שאם f אינה חסומה אז לכל חלוקה $P=\{x_0,\ldots,x_k\}$ של $[a,b]$ יש שתי בחירות של נקודות ξ',ξ'' עבור P כך ש־ $|\sigma(f,P,\xi')-\sigma(f,P,\xi'')|>1$. זה מספיק כי אם f הייתה אינטגרבילית רימן אז הייתה קיימת חלוקה P כך שלכל שתי בחירות ξ',ξ'' עבור P מתקיים $|\sigma(f,P,\xi')-I|<\frac{1}{2}$ וגם $|\sigma(f,P,\xi'')-I|<\frac{1}{2}$, ומכאן $|\sigma(f,P,\xi')-\sigma(f,P,\xi'')|<1$, סתירה.

ובכן תהי $P=\{x_0,\ldots,x_k\}$ חלוקה של $[a,b]$ ותהי ξ' בחירת נקודות שרירותית עבור P. מכיוון שאנו מניחים ש־f אינה חסומה ב־$[a,b]$ והאיחוד של קטעי החלוקה P הם כל הקטע $[a,b]$, יש קטע $[x_{i-1},x_i]$ ב־P שבו f אינה חסומה. לכן יש $t\in[x_{i-1},x_i]$ כך ש־ $f(t)-f(\xi'_i)>\frac{1}{\Delta x_i}$. נגדיר בחירת נקודות ξ'' ל־P על־ידי $\xi''_i=t$, ולכל $j\neq i$ נגדיר את ξ''_j באופן שרירותי (למשל $\xi''_j=\xi'_j$). אז

$$|\sigma(f,P,\xi')-\sigma(f,P,\xi'')|\geq|f(\xi'_i)-f(\xi''_i)|\Delta x_i=|f(\xi'_i)-f(t)|\Delta x_i>1$$

כפי שרצינו. ∎

נצטרך גם לקבל הערכה לסכומים תחתונים ועליונים של חלוקה בעזרת מידע על סכומי רימן של אותה חלוקה. הלמה הבאה מספקת את הקשר הדרוש:

למה 9.5.5 אם f חסומה על $[a,b]$ ואם $P=\{x_0,\ldots,x_k\}$ חלוקה של $[a,b]$ אז מתקיים

$$\begin{aligned}\underline{s}(f,P)&=\inf\{\sigma(f,P,\xi)\,:\,\xi\text{ בחירת נקודות עבור }P\}\\ \overline{s}(f,P)&=\sup\{\sigma(f,P,\xi)\,:\,\xi\text{ בחירת נקודות עבור }P\}\end{aligned}$$

ובפרט

$$\omega(f,P)=\sup\{\sigma(f,P,\xi')-\sigma(f,P,\xi'')\,:\,\xi',\xi''\text{ בחירת נקודות עבור }P\}$$

הוכחה נוכיח את השוויון הראשון. כפי שראינו קודם, לכל בחירת נקודות ξ עבור P מתקיים $\underline{s}(f,P)\leq\sigma(f,P,\xi)$ ומכאן

$$\underline{s}(f,P)\leq\inf\{\sigma(f,P,\xi)\,:\,\xi\text{ בחירת נקודות עבור }P\}$$

כדי להוכיח את האי־שוויון ההפוך, די שנראה שלכל $\varepsilon>0$ יש בחירת נקודות ξ ל־P כך ש־ $\sigma(f,P,\xi)\leq\underline{s}(f,P)+\varepsilon$.

יהי $\varepsilon>0$. בכל קטע חלוקה $[x_{i-1},x_i]$ אפשר לבחור נקודה $\xi_i\in[x_{i-1},x_i]$ כך ש־

$$f(\xi_i)\leq m_i+\varepsilon$$

זה מגדיר סדרת נקודות שמהווה בחירה עבור P, ומתקיים

$$\begin{aligned}\sigma(f,P,\xi) &= \sum_{i=1}^{k} f(\xi_i)(x_i-x_{i-1})\\ &\le \sum_{i=1}^{k}(m_i+\varepsilon)(x_i-x_{i-1})\\ &= \underline{s}(f,P)+\varepsilon\sum_{i=1}^{k}(x_i-x_{i-1})\\ &= \underline{s}(f,P)+\varepsilon(b-a)\end{aligned}$$

ואם היינו בוחרים מראש $\frac{\varepsilon}{b-a}$ במקום ε היינו מקבלים בחירה ξ כנדרש.

ההוכחה של השוויון עבור $\overline{s}(f,P)$ דומה. כעת הטענה על התנודה נובעת מהשוויון $\omega(f,P)=\overline{s}(f,P)-\underline{s}(f,P)$ (כתבו הוכחה מפורטת!). ■

משפט 9.5.6 אם f אינטגרבילית רימן אז היא אינטגרבילית דרבו, והאינטגרלים שווים.

הוכחה יהי I אינטגרל רימן של f ב־$[a,b]$ ויהי $\varepsilon>0$. קיים $\delta>0$ כך שלכל חלוקה P של $[a,b]$ עם $\lambda(P)<\delta$ ולכל בחירת נקודות ξ עבור P מתקיים $|\sigma(f,P,\xi)-I|<\varepsilon$. נובע מכאן שלכל שתי בחירות של נקודות ξ',ξ'' עבור P מתקיים

$$|\sigma(f,P,\xi')-\sigma(f,P,\xi'')|<2\varepsilon$$

ולכן לפי הלמה הקודמת,

$$\begin{aligned}\omega(f,P) &= \sup\{\sigma(f,P,\xi')-\sigma(f,P,\xi'')\ :\ \xi',\xi''\ \text{בחירת נקודות עבור}\ P\}\\ &\le 2\varepsilon\end{aligned}$$

בסיכום, ראינו שלכל $\varepsilon>0$ קיים $\delta>0$ כך שאם P חלוקה המקיימת $\lambda(P)<\delta$ אז $\omega(f,P)\le 2\varepsilon$, וזה גורר אינטגרביליות במובן של דרבו.

לבסוף, מאחר ש־f אינטגרבילית דרבו, השוויון בין האינטגרלים נובע ממשפט 9.5.3 (גם לא קשה להסיק זאת ישירות). ■

בסיכום,

מסקנה 9.5.7 אינטגרל דרבו ואינטגרל רימן שקולים.

ישנן טענות שהוכחתן קלה יותר בעזרת סכומי רימן מאשר בעזרת ההגדרה של דרבו. לפני שניתן דוגמה כזאת נוכיח את הלמה הבאה, המקבילה למסקנה 9.2.8.

למה 9.5.8 תהי f אינטגרבילית ב־$[a,b]$. תהי $P^{(n)}=\{x_1^{(n)},\ldots,x_{k(n)}^{(n)}\}$ סדרה של חלוקות המפרידה נקודות (כלומר $\lambda(P^{(n)})\to 0$), ולכל n נקבע בחירת נקודות $\xi^{(n)}=(\xi_1^{(n)},\ldots,\xi_{k(n)}^{(n)})$ עבור $P^{(n)}$. אזי

$$\int_a^b f=\lim_{n\to\infty}\sigma(f,P^{(n)},\xi^{(n)})$$

הוכחה נסמן $I=\int_a^b f$. לכל $\varepsilon>0$ יש δ כך שאם P חלוקה עם $\lambda(P)<\delta$ ואם ξ בחירת נקודות ל־P אז $|\sigma(f,P,\xi)-I|<\varepsilon$. מכיוון ש־ $\lambda(P^{(n)})\to 0$ הרי ש־ $\lambda(P^{(n)})<\delta$ לכל n גדול מספיק ולכן לכל n גדול מספיק $|\sigma(f,P^{(n)},\xi^{(n)})-I|<\varepsilon$, ופירוש הדבר הוא ש־ $\sigma(f,P^{(n)},\xi^{(n)})\to I$ כנדרש. ■

הנה דוגמה לשימוש בלמה. נוכיח למשל את אי־שוויון המשולש לאינטגרלים, הקובע שאם f אינטגרבילית ב־$[a,b]$ אז כך גם $|f|$ ומתקיים $|\int_a^b f|\leq\int_a^b|f|$ (זו טענה 9.4.9). העובדה ש־$|f|$ אינטגרבילית נובעת ממשפט האינטגרביליות של פונקציה מורכבת, כי פונקציית הערך המוחלט רציפה. לגבי האי־שוויון, נבחר סדרת חלוקות $P^{(n)}$ של $[a,b]$ עם $\lambda(P^{(n)})\to 0$, ולכל $P^{(n)}$ נבחר סדרת נקודות מתאימה $\xi^{(n)}=(\xi_i^{(n)})$. לכל n מתקיים

$$\begin{aligned}|\sigma(f,P^{(n)},\xi^{(n)})| &= |\sum_i f(\xi_i^{(n)})\Delta x_i^{(n)}|\\ &\leq \sum_i |f(\xi_i^{(n)})|\Delta x_i^{(n)}\\ &= \sigma(|f|,P,\xi^{(n)})\end{aligned}$$

האי־שוויון כאן נובע מאי־שוויון המשולש הרגיל לסכומים סופיים. כעת

$$\begin{aligned}|\int_a^b f| &= |\lim_{n\to\infty}\sigma(f,P_n,\xi^{(n)})|\\ &= \lim_{n\to\infty}|\sigma(f,P_n,\xi^{(n)})|\\ &\leq \lim_{n\to\infty}\sigma(|f|,P_n,\xi^{(n)})\\ &= \int_a^b|f|\end{aligned}$$

כנדרש (הצדיקו כל שלב באי־שוויון!).

תרגילים

1. אם P חלוקה של $[a,b]$, האם תמיד קיימת בחירת נקודות ξ ל־P כך ש־ $\sigma(f,P,\xi)=\underline{s}(f,P)$?

2. הראו שאם $f:[a,b]\to\mathbb{R}$ אינטגרבילית, P חלוקה של $[a,b]$ ו־ξ בחירת נקודות ל־P אז
$$|\sigma(f,P,\xi)-\int_a^b f|\le\omega(f,P)$$
3. הוכיחו בעזרת אינטגרל רימן את כלל הסכום (משפט 9.4.4).
4. חשבו את הגבולות של הסדרות הבאות:
   (א) $a_n=\sum_{k=1}^n \frac{k}{n^2}$.
   (ב) $b_n=\sum_{k=1}^n \frac{1}{n}e^{k/n}$.
   (ג) $c_n=\sum_{k=1}^n \frac{\sqrt{k}}{n^{3/2}}$.
5. תהי f פונקציה ממשית שאינה אינטגרבילית ב־$[a,b]$. הראו שיש חלוקות $P^{(n)}$ ובחירות של נקודות $\xi^{(n)}$ כך ש־ $(P^{(n)})$ היא סדרה מפרידה נקודות של חלוקות והסדרה $(\sigma(f,P^{(n)},\xi^{(n)}))_{n=1}^\infty$ מתבדרת.
6. הראו שאם $f:[a,b]\to\mathbb{R}$ אינה אינטגרבילית אז לכל מספר $\underline{S}(f)\le c\le\overline{S}(f)$ יש סדרת חלוקות המפרידה נקודות $(P^{(n)})$ וסדרה של בחירות של נקודות $(\xi^{(n)})$ כך ש־ $c=\lim_{n\to\infty}\sigma(f,P^{(n)},\xi^{(n)})$.

9.6 המשפט היסודי

הנגזרת והאינטגרל נראים כיצורים שונים לגמרי, אך יש קשר הדוק ביניהם. נטיב להבין זאת אם נשוב לדוגמה המוכרת של החלקיק שנע בקו ישר, בזמן 0 הוא נמצא בראשית, ומיקומו בזמן t הוא $s(t)$ (בפרט, $s(0)=0$). נניח גם שמהירות החלקיק בזמן t נתונה על־ידי $v(t)$. בפרק על נגזרות ראינו שאם פונקציית המרחק s נתונה אז אפשר לקרב את $v(t)$ על ידי המנה $\frac{1}{h}(s(t+h)-s(t))$ עבור h־ים קטנים, ו־$v(t)$ הוא הגבול של מנות אלה, כלומר $v=s'$ (עיינו בדיון בעמוד 301).

נתבונן כעת בבעיה ההפוכה: נניח שפונקציית המהירות v של החלקיק נתונה לנו, ואנו רוצים לשחזר ממנה את פונקציית המרחק s (באופן מוחשי יותר: אתם נוסעים במכונים ורוצים להעריך את המרחק שנסעתם על ידי התבוננות במד־המהירות בזמן הנסיעה). ננסה להעריך למשל את $s(1)$, שהוא המרחק שהחלקיק עבר מזמן 0 עד זמן 1.

אם המהירות v קבועה ושווה זהותית למספר v_0 אז המרחק שהחלקיק עובר בפרק זמן שאורכו Δt הוא $v_0\cdot\Delta t$. גם אם המהירות אינה קבועה, משיקולים פיזיקליים סביר להניח שבפרקי זמן קצרים המהירות תהיה בקירוב קבועה. לכן המהירות בין הזמנים t ו־ $t+\Delta t$ היא בקירוב קבועה ושווה ל־$v(t)$. מכאן שבפרק הזמן בין t ל־ $t+\Delta t$, שאורכו Δt, המרחק שהחלקיק עבר שווה בקירוב ל־ $v(t)\cdot\Delta t$.

נבחר סדרת זמנים $0 = t_0 < t_1 < \ldots < t_{k-1} < t_k = 1$ ונסמן ב־ $\Delta t_i = t_i - t_{i-1}$ את פרק הזמן שבין t_{i-1} ל־t_i. כפי שראינו, אם פרקי הזמן קצרים אז בפרק הזמן $[t_{i-1}, t_i]$ החלקיק עובר בקירוב מרחק של $v(t_i) \cdot \Delta t_i$ ולכן ההערכה שלנו למרחק הכולל היא

$$\sum_{i=1}^{k} v(t_i) \cdot \Delta t_i$$

קל לזהות שזהו סכום רימן של הפונקציה v בקטע $[0,1]$. מכאן שאנו מצפים שהמרחק שהחלקיק עבר בין זמן 0 לזמן 1 הוא $\int_0^1 v$, ומאותם שיקולים נסיק שהמרחק שהחלקיק עבר בין זמן 0 לזמן t הוא $s(t) = \int_0^t v$.

אם כן, השתכנענו ש־ $s(t) = \int_0^t v(t)$ וש־ $s'(t) = v(t)$. זהו רמז ראשון לכך שאינטגרציה היא פעולה הפוכה לגזירה. לפני שננסח עובדה זו במדויק, נשכלל מעט את הסימונים שלנו. עד כה הסימן $\int_a^b f$ היה מוגדר רק עבור $a < b$. נרחיב את ההגדרה:

הגדרה 9.6.1 אם $s < t$ ו־f אינטגרבילית ב־$[s,t]$ אז המספר $\int_t^s f$ מוגדר על־ידי $\int_t^s f = -\int_s^t f$. כמו כן אם $s = t$ נגדיר $\int_s^t f = 0$.

הסימון נבחר באופן כזה שנוכל להכליל את השוויון $\int_a^b f = \int_a^c f + \int_c^b f$ (שהוכח במשפט 9.4.2) למקרה שבו c אינו בין a ל־c. למשל אם $a < b < c$ אז

$$\begin{aligned} \int_a^c f + \int_c^b f &= (\int_a^b f + \int_b^c f) + \int_c^b f \\ &= \int_a^b f + \int_b^c f + (-\int_b^c f) \\ &= \int_a^b f \end{aligned}$$

כאשר השוויון הראשון נובע ממשפט 9.4.2 עצמו והשני מהגדרה 9.6.1. שוויון דומה מתקיים כאשר $c < a < b$. בסיכום,

משפט 9.6.2 השוויון $\int_a^b f = \int_a^c f + \int_c^b f$ מתקיים לכל $a,b,c \in \mathbb{R}$ ללא תלות בסדר היחסי בין a,b,c, בתנאי שכל שלושת האינטגרלים מוגדרים (קרי: בתנאי ש־f אינטגרבילית בקטע $[\min\{a,b,c\}, \max\{a,b,c\}]$).

ההוכחה כרוכה בבדיקת מקרים. באחד המקרים טיפלנו למעלה. את האחרים אנו משאירים כתרגיל.

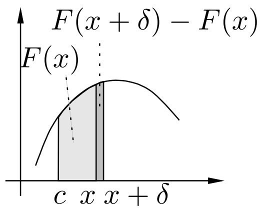
למה 9.6.3 יהיו $a \neq b$ מספרים ממשיים ותהי f פונקציה אינטגרבילית בקטע שקצותיו a,b. אם $m \leq f(x) \leq M$ לכל x בין a ל־b אז

$$m \leq \frac{1}{b-a} \int_a^b f \leq M$$

הוכחה אם $a < b$ המסקנה מתקבלת מטענה 9.4.7 על ידי העברת אגפים. אם $b < a$ נשים לב ש־ $\frac{1}{b-a}\int_a^b f = \frac{1}{a-b}\int_b^a f$. אנו יודעים מהמקרה הקודם שמתקיים $m \le \frac{1}{a-b}\int_a^b f \le M$, והטענה נובעת. ■

מסקנה 9.6.4 אם f אינטגרבילית בקטע שקצותיו a, b וחסומה בקטע על ידי L, אז $|\int_a^b f| \le L|b-a|$.

נשוב לענייננו. נניח שוב ש־f אינטגרבילית בקטע $[a,b]$, יהי $c \in [a,b]$, ונגדיר $F(t) = \int_c^t f(x)dx$. אז F מוגדרת בקטע $[a,b]$. הפונקציה F מתארת את השטח המסומן בין הגרף של f לציר ה־x בין c ל־t (אם $t < c$ מחליפים סימן לפי הגדרה 9.6.1).



איור 9.6.1

מכיוון ש־f אינטגרבילית היא חסומה, נניח על־ידי L. שינוי של t בכמות קטנה δ משנה את השטח במידה שאינה עולה על שטח של המלבן שבסיסו δ וגובהו L, ולכן שינוי קטן של t גורר שינוי קטן של F, ונצפה ש־f תהיה רציפה במשתנה t. ואמנם, זו כבר כמעט הוכחה. כל שנותר הוא לבצע את החישוב בדיוק:

משפט 9.6.5 אם f אינטגרבילית ב־$[a,b]$ ואם $c \in [a,b]$ אז הפונקציה $F(t) = \int_c^t f$ רציפה ב־$[a,b]$.

הוכחה יהי L חסם של f ב־$[a,b]$, תהי t_0 נקודה ב־ $[a,b]$ ותהי $(t_n) \subseteq [a,b]$ כך ש־ $t_n \to t_0$. אז

$$|F(t_n) - F(t_0)| = |\int_c^{t_n} f - \int_c^{t_0} f| = |\int_{t_0}^{t_n} f| \le L|t_n - t_0| \to 0$$

כאשר $n \to \infty$. לכן מאפיון היינה לרציפות, F רציפה ב־t_0. ■

דוגמאות

1. תהי $f(x) = \alpha$ קבועה. אז $F(t) = \int_0^t f = \alpha t$, ו־$F$ באמת רציפה.
2. תהי $F(t) = \int_0^t \mathrm{sgn}$, כאשר sgn היא פונקציית הסימן המקבלת ערך -1 על השליליים וערך 1 על החיוביים (הערך ב־0 אינו בעל חשיבות מכיוון שהערך של האינטגרנד בנקודה אחת אינו משפיע על ערך האינטגרל). אם $t > 0$ מתקיים $F(t) = \int_0^t 1 = t$ ואילו אם $t < 0$ מתקיים

$$F(t) = \int_0^t \mathrm{sgn} = -\int_{-|t|}^0 \mathrm{sgn} = -\int_{-|t|}^0 (-1) = -(0-(-|t|)) = |t|$$

(הצדיקו כל שוויון!). בסיכום, לכל t מתקיים $F(t) = |t|$. מכאן שאמנם F רציפה, אך אינה גזירה ב־0. לעומת זאת היא כן גזירה בכל נקודה אחרת.

כפי שרואים בדוגמה האחרונה, הפונקציה $F(t) = \int_c^t f$ אינה תמיד גזירה. אולם מסתבר שכאשר האינטגרנד f רציף, F גזירה ויתר על כן הנגזרת שלה היא f. זהו הקשר בין גזירה לאינטגרציה אותו הזכרנו בתחילת הסעיף.

משפט 9.6.6 (המשפט היסודי של החשבון האינפיניטסימלי) תהי f אינטגרבילית ב־$[a,b]$, יהי $c \in [a,b]$ ותהי $F(t) = \int_c^t f$. אם $t_0 \in [a,b]$ ואם t_0 נקודת רציפות של f אז F גזירה ב־t_0 ומתקיים $F'(t_0) = f(t_0)$ (אם t_0 נקודת קצה אנו מפרשים גזירות במובן החד־צדדי).

הוכחה תהי $t_0 \in (a,b)$ נקודת רציפות של f. כמו בהוכחה של המשפט הקודם אפשר לרשום

$$\frac{F(t_0+h)-F(t_0)}{h} = \frac{1}{h}(\int_c^{t_0+h} f - \int_c^{t_0} f) = \frac{1}{h}\int_{t_0}^{t_0+h} f$$

עלינו לחשב את הגבול הזה כש־h שואף לאפס. יהי $\varepsilon > 0$. מאחר ש־f רציפה ב־t_0 יש $\delta > 0$ כך שאם $x \in (t_0-\delta, t_0+\delta)$ אז $f(t_0)-\varepsilon \leq f(x) \leq f(t_0)+\varepsilon$. לכן אם $|h| < \delta$ אז האינטגרנד ב־$\int_{t_0}^{t_0+h} f$ חסום בין $f(t_0)-\varepsilon$ ל־ $f(t_0)+\varepsilon$ ואנו רואים מלמה 9.6.3 ש־

$$f(t_0)-\varepsilon \leq \frac{1}{h}\int_{t_0}^{t_0+h} f \leq f(t_0)+\varepsilon$$

משמעות הדבר היא שהגבול המבוקש קיים ושווה ל־$f(t_0)$, כפי שרצינו.

הוכחות המקרים $t_0 = a,b$ דומה. ■

מסקנה 9.6.7 אם f רציפה בקטע $[a,b]$ אז $F(t) = \int_a^t f$ רציפה וגזירה ב־$[a,b]$ ומתקיימת $F'(t) = f(t)$ לכל $t \in (a,b)$. נוסף, היא גזירה חד־צדדית בקצוות ומקיימת $F'_+(a) = f(a)$ ו־ $F'_-(b) = f(b)$.

הגדרה 9.6.8 תהי f פונקציה ממשית המוגדרת בקבוצה D. פונקציה ממשית F מוגדרת ב־D נקראת **פונקציה קדומה** (primitive function או anti-derivative) של f ב־D אם F גזירה ב־D ובכל נקודה $x \in D$ מתקיים $F'(x) = f(x)$ (נפרש נגזרת במובן החד־צדדי עבור x־ים שהם קצה של קטע).

אם כן, מסקנה 9.6.7 פירושה שלכל פונקציה רציפה בקטע יש פונקציה קדומה. החשיבות הרבה של פונקציות קדומות נובעת מהמשפט הבא, שמבטא את האינטגרל של פונקציה במונחים של פונקציה קדומה שלה:

משפט 9.6.9 (נוסחת ניוטון־לייבניץ) אם f רציפה ב־$[a,b]$ ו־F פונקציה קדומה ל־f ב־$[a,b]$, אז $\int_a^b f = F(b) - F(a)$.

הוכחה לפי הנתון, $f = F'$. לפי המשפט היסודי הפונקציה $G(t) = \int_a^t f$ גם־כן מקיימת $G' = f$. לכן $(G-F)' = 0$ בכל נקודה בקטע, ומכאן שההפרש $G - F$ קבוע. מכיוון ש־

$$(G-F)(a) = G(a) - F(a) = -F(a)$$

אנו מסיקים כי $G(t) - F(t) = -F(a)$ לכל $t \in [a,b]$, ואם מציבים $t = b$ יוצא זה בדיוק אומר ש־ $\int_a^b f = F(b) - F(a)$, כנדרש. ■

לנוסחת ניוטון־לייבניץ יש חשיבות עצומה. בעוד האינטגרל $\int_a^b f$ מוגדר בצורה סתומה יחסית וקשה מאד לחשב אותו ישירות, עבור פונקציות רבות f ניתן למצוא פונקציה קדומה בעלת נוסחה פשוטה, ונוסחת ניוטון־לייבניץ מאפשרת לבטא את האינטגרל של f בעזרת הפונקציה הקדומה. בשיטה זו נוכל לחשב אינטגרלים רבים שקודם לכן היו בלתי אפשריים לחישוב.

סימון 9.6.10 אם $F : [a,b] \to \mathbb{R}$ נסמן את ההפרש $F(b)-F(a)$ ב־ $F|_a^b$, או לפעמים ב־ $F(x)|_{x=a}^{x=b}$.

בעזרת הסימון החדש אפשר לנסח את נוסחת ניוטון־לייבניץ כך: אם f פונקציה רציפה ב־$[a,b]$ ואם F פונקציה קדומה של f ב־$[a,b]$ אז $\int_a^b f = F|_a^b$.

דוגמאות

1. אם $f(x) = e^x$ אז $F(x) = e^x$ היא פונקציה קדומה של f, ולכן

$$\int_a^b e^x dx = F|_a^b = e^b - e^a$$

2. תהי $f(x) = \sqrt{1-x^2}$. הגרף של f הוא חצי מעגל ברדיוס 1 שמרכזו בראשית. השטח שמתחת לגרף הזה הוא שטח של חצי עיגול ברדיוס 1. נשים לב ש־

$$F(x) = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)$$

היא פונקציה קדומה של f ב־$[-1,1]$ (בסעיף הבא נראה כיצד הגענו לנוסחה הזו, אך לעת עתה תוכלו לבדוק שזו פונקציה קדומה של f על־ידי חישוב הנגזרת F'). מכאן מקבלים

$$\int_{-1}^1 \sqrt{1-x^2}dx = F|_{-1}^1 = F(1) - F(-1) = \frac{\pi}{2}$$

ומכאן שהשטח של עיגול ברדיוס 1 הוא π, כפי שציפינו.

נסיים בגרסה אינטגרלית של משפט ערך ביניים:

משפט 9.6.11 (משפט ערך הביניים האינטגרלי): תהי $f:[a,b]\to\mathbb{R}$ רציפה. אז קיים $c\in[a,b]$ כך ש־ $\frac{1}{b-a}\int_a^b f=f(c)$.

הוכחה נסמן $F(x)=\int_a^x f$. עלינו להראות שיש $c\in[a,b]$ כך ש־ $\frac{F(b)-F(a)}{b-a}=f(c)$. אבל זה נובע ממשפט לגרנג' ומהעובדה שלפי המשפט היסודי, F רציפה ב־$[a,b]$ ו־ $F'(x)=f(x)$ לכל $x\in(a,b)$. ■

תרגילים

1. הראו שאם f אינטגרבילית ב־$[a,b]$ אז $F(x)=\int_a^x f$ היא ליפשיצית ב־$[a,b]$ (ראו תרגיל (13) בעמוד 275).
2. גזרו את הפונקציות הבאות:
   (א) $F(x)=\int_x^1 \cos t dt$.
   (ב) $F(x)=\int_0^{e^x} \cos t dt$.
   (ג) $F(x)=\int_x^{x^2} e^t dt$.
3. תהי $f:\mathbb{R}\to\mathbb{R}$ פונקציה רציפה ותהי $F(x)=\int_0^x f$.
   (א) הראו ש־F זוגית אמ"מ f אי־זוגית, וש־F אי־זוגית אמ"מ f זוגית.
   (ב) הראו שאם F מחזורית אז f מחזורית. מצאו תנאי הכרחי ומספיק כדי שגם ההפך יהיה נכון.
   (ג) היכן בסעיפים הקודמים השתמשתם בעובדה ש־f רציפה?
4. תהי f רציפה ב־$\mathbb{R}$ ולכל $\delta>0$ נגדיר $F_\delta(x)=\int_{x-\delta}^{x+\delta} f$. הראו ש־$F_\delta$ רציפה בכל נקודה, ולכל $x\in\mathbb{R}$ מתקיים $\lim_{\delta\to0+}F_\delta(x)=f(x)$.
5. הוכיחו את ההכללה הבאה של משפט ערך הביניים האינטגרלי. בהינתן פונקציות רציפות $f,g:[a,b]\to\mathbb{R}$ כך ש־ $g\geq0$, יש $c\in[a,b]$ כך ש־ $\int_a^b fg=f(c)\int_a^b g$. (רמז: הראו שאפשר להניח ש־ $\int_a^t g>0$ לכל $t>a$, והיעזרו במשפט ערך הביניים של קושי 8.5.10).
6. תהי $f:[a,b]\to\mathbb{R}$ פונקציה אינטגרבילית ותהי $F(x)=\int_a^x f$. הראו שאם $c\in(a,b)$ היא נקודת אי רציפות מסוג ראשון של f אז F אינה גזירה ב־c.
7. תהי f מוגדרת כמו בתרגיל (6) בעמוד 280. היעזרו בתרגיל הקודם כדי להסיק ש־f אינטגרבילית, $F(x)=\int_0^x f$ רציפה אך קבוצת נקודות האי־רציפות שלה צפופות בתחום הגדרתה.
8. הוכיחו שאם $f:[a,b]\to\mathbb{R}$ אינטגרבילית ואם F פונקציה קדומה של f ב־$[a,b]$, אז $\int_a^b f=F(b)-F(a)$. זו הכללה של נוסחת ניוטון־לייבניץ, אך לא ניתן להוכיחה בעזרת המשפט היסודי כי אין מניחים ש־f רציפה. חפשו הוכחה ישירה.

9.7 האינטגרל הלא מסוים

כאשר f רציפה ב־$[a,b]$, המשפט היסודי ממיר את בעיית החישוב של האינטגרל $\int_a^b f$ בבעיה אחרת: מציאת פונקציה קדומה של f. כך אפשר לחשב אינטגרלים רבים שחישובם הישיר מסובך ביותר. השיטה הזו שימושית רק בתנאי שנוכל למצוא נוסחה מפורשת עבור פונקציה קדומה של f. ליתר דיוק, בהינתן פונקציה אלמנטרית f, נרצה למצוא פונקציה אלמנטרית F כך ש־ $F' = f$. בעיה זו נקראת **בעיית אינטגרציה לא מסוימת**. זו הבעיה ההפוכה לגזירה, אך היא קשה מבעיית הגזירה במידה ניכרת, ולא תמיד ניתנת לביצוע. כללי התחשיב של גזירה מבטיחים שאם f פונקציה אלמנטרית אז גם f' היא אלמנטרית, אך ההפך אינו נכון: יש פונקציות אלמנטריות שהפונקציה הקדומה שלהן אינה אלמנטרית. דוגמה לפונקציה כזו היא e^{x^2}. יש לה פונקציה קדומה (כי היא רציפה בכל הישר), אבל הפונקציה הקדומה אינה אלמנטרית (ההוכחה לכך מורכבת למדי, ולא נוכל להציג אותה כאן).

הגדרה 9.7.1 תהי f פונקציה ממשית המוגדרת בתחום כלשהו. קבוצת הפונקציות הקדומות של f נקראת **האינטגרל הלא־מסוים** (indefinite integral) של f, ומסומנת $\int f$ או $\int f(x)dx$ (דהיינו כמו אינטגרל רגיל אך ללא ציון גבולות אינטגרציה).

על מנת להבדיל אותו מהאינטגרל הלא־מסויים, האינטגרל שעסקנו בו בסעיפים הקודמים נקרא לפעמים **אינטגרל מסוים** (definite integral).

הערות

1. בבעיית אינטגרציה מסוימת מספיק לנו למצוא פונקציה קדומה אחת של f, ואם F פונקציה כזו נרשום $\int f = F$. יש בכך אי־דיוק כי פורמלית $\int f$ היא קבוצה של פונקציות, והיינו צריכים לכתוב $F \in \int f$. אולם זו כתיבה מסורבלת ולא נוחה, ונרשה לעצמנו להתייחס לסימון $\int f$ כאילו מציין את אחת הפונקציות הקדומות של f. כך נוכל לרשום $\int f + \int g$ כדי לציין סכום של פונקציות קדומות של f, g בהתאמה. באותו אופן נרשום $f = (\int f)'$ ו־ $f = \int (f')$.

2. אם f מוגדרת בקטע I אז כל שתי פונקציות קדומות F, G של f ב־I נבדלות בקבוע חיבורי. הסיבה לכך היא ש־ $(F-G)' = F' - G' = f - f \equiv 0$, ולכן בקטע I הפונקציה $F - G$ קבועה, כלומר, $F = G + C$ לקבוע C כלשהו.[2]

   כאשר תחום ההגדרה של f אינו קטע, המצב שונה. למשל, התחום של הפונקציה $f(x) = -\frac{1}{x^2}$ הוא $\mathbb{R} \setminus \{0\}$, והפונקציה $F(x) = 1/x$ היא פונקציה

[2] מסיבה זו יש ספרים בהם רושמים $\int f = F + C$ כדי לציין ש־F היא פונקציה קדומה של f ו־$\int f$ היא קבוצת הפונקציות הנבדלות מ־F בקבוע חיבורי. לא נשתמש בסימון זה.

קדומה של f, אבל גם

$$G(x) = \begin{cases} \frac{1}{x} & x < 0 \\ \frac{1}{x} + 1 & x > 0 \end{cases}$$

היא פונקציה קדומה שלה, והיא אינה נבדלת מ־F בקבוע.

3. לעתים, כאשר F פונקציה קדומה של f נרשום

$$F(x) = \int f(x)\,dx$$

שימו לב שהמשמעות של הסימן x באגף שמאל שונה לגמרי מהמשמעות שלו באגף ימין. אגף שמאל מתאר פונקציה שהמשתנה שלה הוא x, ולעומת זאת באגף ימין x הוא משתנה האינטגרציה. למשל, יש משמעות להצבת $x = 5$ בצד שמאל אך הדבר חסר משמעות בצד ימין (השוו זאת עם ההערה על סימון הנגזרת בעמוד 307). עם זאת, זו צורת כתיבה זו מקובלת, וגם אנו נשתמש בה לפעמים.

יתר הסעיף מוקדש לשיטות שונות שבעזרתן אפשר לפתור בעיות אינטגרציה לא מסוימת. כדי ללמוד להשתמש בשיטות האינטגרציה יש לתרגל אותן (הדבר נכון על כל נושא במתמטיקה, אך במיוחד בנושא שלפנינו). בסוף הסעיף תמצאו מגוון בעיות לפתרון. מומלץ מאד לפתור את כולן.

אינטגרלים מידיים

כצעד ראשון נעיין רשימת הפונקציות האלמנטריות הבסיסיות, נחשב את הנגזרות שלהן, ונפעיל את הכלל שאם $g = f'$ אז $\int g = f$. למשל, מהעובדה ש־ $\frac{d}{dx}e^x = e^x$ אנו מסיקים מיד ש־ $\int e^x dx = e^x$, ומהעובדה ש־ $\frac{d}{dx}\sin x = \cos x$ אנו מסיקים כי

$\int \cos x = \sin x$. באופן זה אנו מקבלים את הרשימה הבאה של אינטגרלים מידיים:

$$\begin{aligned}
\int x^\alpha dx &= \frac{1}{\alpha+1}x^{\alpha+1} \quad (\alpha \neq -1) \\
\int \frac{1}{x} dx &= \ln|x| \\
\int e^x dx &= e^x \\
\int \frac{1}{\sqrt{1-x^2}} dx &= \arcsin x \\
\int \frac{1}{1+x^2} dx &= \arctan x \\
\int \cos x dx &= \sin x \\
\int \sin x dx &= -\cos x \\
\int \frac{1}{\cos^2 x} dx &= \tan x \\
\int \frac{1}{\sin^2 x} dx &= \cot x
\end{aligned}$$

לינאריות האינטגרל

אנו מעוניינים במשפטי תחשיב עבור האינטגרל הלא מסויים. האפשרות העומדת לרשותנו היא לנתח את כללי התחשיב של הנגזרת ולנסות לתרגם אותם לשפה של אינטגרלים. למשל, נניח ש־ F, G הן פונקציות קדומות של f, g בהתאמה. מכיוון ש־ $(F+G)' = F' + G' = f + g$, אנו מסיקים ש־$F+G$ היא פונקציה קדומה של $f+g$. באופן דומה אנו רואים שלכל קבוע c, הפונקציה cF היא פונקציה קדומה של cf. בסיכום,

משפט 9.7.2 (לינאריות האינטגרל הלא מסוים) לכל זוג פונקציות f, g בעלות תחום משותף מתקיים $\int (f+g) = \int f + \int g$ ולכל קבוע c מתקיים $\int (cf) = c\int f$.

כללים פשוטים אלה מאפשרים לחשב כמה אינטגרלים שאינם מופיעים ברשימת האינטגרלים המידיים:

דוגמאות

1. נחשב את

$$\int \frac{2x^4}{1+x^2} dx$$

בעזרת משפט 9.7.2 אנו מקבלים

$$\begin{aligned}
\int \frac{2x^4}{1+x^2}dx &= \int 2(\frac{x^4-1}{1+x^2}+\frac{1}{1+x^2})dx \\
&= 2\left(\int \frac{x^4-1}{1+x^2}dx + \int \frac{dx}{1+x^2}\right) \\
&= 2\left(\int (x^2-1)dx + \int \frac{dx}{1+x^2}\right) \\
&= 2\left(\int x^2dx - \int dx + \int \frac{dx}{1+x^2}\right)
\end{aligned}$$

כל אחד מהאינטגרלים בשורה האחרונה הוא אינטגרל מידי, וקיבלנו ש־

$$\int \frac{2x^4}{1+x^2}dx = 2\left(\frac{x^3}{3} - x + \arctan(x)\right)$$

2. נחשב את

$$\int \frac{dx}{sin^2(x)\cos^2(x)}$$

נציב את השוויון $\sin^2 x + \cos^2 x = 1$ ונקבל

$$\begin{aligned}
\int \frac{dx}{sin^2(x)\cos^2(x)} &= \int \frac{sin^2(x)+\cos^2(x)}{sin^2(x)\cos^2(x)}dx \\
&= \\
&= \int \frac{1}{\cos^2(x)}dx + \int \frac{1}{sin^2(x)}dx \\
&= \tan(x) - \cot(x)
\end{aligned}$$

3. מכיוון ש־ $\int x^n = \frac{1}{n+1}x^{n+1}$ לכל n טבעי, אנו מקבלים נוסחה כללית לפונקציה הקדומה של פולינום:

$$\int \sum_{n=0}^{d} a_n x^n = \sum_{n=0}^{d} \frac{1}{n+1} a_n x^{n+1} = \sum_{n=1}^{d+1} \frac{1}{n} a_{n-1} x^n$$

מכיוון שפונקציות קדומות שונות של פולינום נבדלות בקבוע (ראו הערה (2) בעמוד 415), נובע שכל פונקציה קדומה של פולינום הוא גם־כן פולינום.

אינטגרציה בחלקים

התרגום של יתר כללי הגזירה למונחים של אינטגרלים מסוימים הוא מסובך יותר. נתחיל בכלל המכפלה לנגזרות, $(f \cdot g)' = f' \cdot g + f \cdot g'$. בעזרת כלל הסכום אנו מסיקים ממנו ש־

$$f \cdot g = \int (fg)' = \int (f'g + fg') = \int f'g + \int fg'$$

נעביר אגפים ונקבל

משפט 9.7.3 (אינטגרציה בחלקים) אם f, g הן פונקציות גזירות ברציפות אז מתקיים

$$\int f' \cdot g = f \cdot g - \int f \cdot g'$$

ההנחה שהנגזרות f', g' רציפות נועדה להבטיח שקיימות פונקציות קדומות ל־fg' ו־ $f'g$.

כתוצאה מהמשפט האחרון ומנוסחת ניוטון־לייבניץ מקבלים:

מסקנה 9.7.4 אם f, g גזירות ברציפות בקטע $[a, b]$ אז

$$\int_a^b f'g = fg|_a^b - \int_a^b fg'$$

הוכחה נסמן $F = \int f'g$ ו־ $G = \int fg'$. אז $F = fg - G$. אז

$$\begin{aligned}
\int_a^b f'g &= F(b) - F(a) \\
&= (f(b)g(b) - G(b)) - (f(a)g(a) - G(a)) \\
&= fg|_a^b - (G|_a^b) \\
&= fg|_a^b - \int_a^b fg'
\end{aligned}$$

כאשר השוויון הראשון והאחרון נובעים מנוסחת ניוטון־לייבניץ. ■

אינטגרציה בחלקים ממירה את בעיית החישוב של $\int f'g$ בבעיית החישוב של $\int fg'$. על מנת להשתמש בשיטה זו יש תחילה לזהות צורה של מכפלה באינטגרנד. בהנחה שזיהינו באינטגרנד כמכפלה $f'g$, אינטגרציה בחלקים תהיה כדאית בתנאי שהבעיה $\int fg'$ קלה יותר. זה יהיה המצב בעיקר כאשר הגורם g נהיה פשוט יותר לאחר גזירה. למשל, אם הפונקציות $\log(x)$ או $\arctan(x)$ הן גורמים באינטגרנד כדאי לנסות לבצע אינטגרציה בחלקים, משום שהנגזרות שלהן הן הפונקציות הרציונליות $\frac{1}{x}$ ו־$\frac{1}{1+x^2}$. גם פולינומים נעשים פשוטים יותר אחרי גזירה, כי המעלה שלהם קטנה בכל פעם שגוזרים, ואחרי מספר סופי של גזירות הן הופכות לקבוע.

דוגמאות

1. נחשב את

$$\int x\ln(x)dx$$

ניחוש סביר הוא שגזירת הגורם $\ln(x)$ תהפוך את הביטוי לפשוט יותר. אנו רוצים שהאינטגרנד יהיה מהצורה $f'\cdot g$ כאשר $f'(x)=x$ ו־ $g(x)=\ln x$. מכאן נובע ש־ $f(x)=\frac{x^2}{2}$ (בחרנו באופן שרירותי פונקציה קדומה אחת של f, אבל כל פונקציה קדומה אחרת הייתה טובה באותה מידה). כעת

$$\begin{aligned}\int x\ln(x)dx &= \int f'g \\ &= fg-\int fg' \\ &= \frac{x^2}{2}\ln(x)-\int\frac{x^2}{2}\cdot\frac{1}{x}dx \\ &= \frac{x^2}{2}\ln(x)-\frac{1}{2}\int xdx \\ &= \frac{x^2}{2}\ln(x)-\frac{1}{4}x^2\end{aligned}$$

בפרט, האינטגרל המסוים $\int_1^2 x\ln x$ שווה ל־

$$\int_1^2 x\ln x=(\frac{x^2}{2}\ln x-\frac{1}{4}x^2)|_1^2=(\frac{4}{2}\ln 2-\frac{4}{4})-(\frac{1}{2}\ln 1-\frac{1}{4})=2\ln 2-\frac{3}{4}$$

2. לעתים, כשהאינטגרנד אינו נראה כמו מכפלה, אפשר בכל זאת לרשום אותו כמכפלה של עצמו עם 1. נחשב למשל את

$$\int\ln(x)dx$$

אם נחשוב על האינטגרנד בתור $1\cdot\ln x$ נוכל לרשום $\ln x=f'g$ כאשר $f(x)=x$ ו־ $g(x)=\ln x$. ביצוע אינטגרציה בחלקים נותן

$$\begin{aligned}\int\ln(x)dx &= x\ln x-\int x\cdot\frac{1}{x}dx \\ &= x\ln x-x\end{aligned}$$

3. כאשר מופיעה באינטגרנד מכפלה של פולינום בפונקציה אחרת אפשר לנסות לגזור את הפולינום וכך להקטין את מעלתו. אסטרטגיה זו יעילה בעיקר כאשר הגורם השני הוא כזה שפונקציה קדומה שלו אינה מסובכת ממנו בהרבה, כמו במקרה של $e^x, \sin x, \cos x$. נחשב למשל את

$$\int xe^x dx$$

אנו רוצים לגזור את הגורם x, ולכן נרשום $xe^x = gf'$ כאשר $g(x) = x$ ו־ $f(x) = e^x$. אז

$$\begin{aligned} \int xe^x dx &= xe^x - \int 1 \cdot e^x dx \\ &= xe^x - e^x \end{aligned}$$

4. כשמופיע פולינום ממעלה גבוהה נצטרך לפעמים לבצע אינטגרציה בחלקים כמה פעמים. למשל נחשב את

$$\int x^2 \cos(x) dx$$

כמו בסעיף הקודם, נרצה להקטין את מעלת הפולינום. נרשום $x^2 \cos x = gf'$ כאשר $g(x) = x^2$ ו־ $f(x) = \sin x$. יוצא ש־

$$\int x^2 \cos(x) dx = x^2 \sin(x) - \int 2x \cdot \sin(x) dx$$

את המחובר האחרון נחשב על ידי אינטגרציה בחלקים. נרשום $2x \sin x = uv'$ כאשר $u(x) = 2x$ ו־ $v(x) = -\cos x$ ונקבל

$$\begin{aligned} \int 2x \sin(x) dx &= -2x \cos x - \int 2(-\cos(x)) dx \\ &= -2x \cos x + 2 \int \cos(x) dx \\ &= -2x \cos x + 2 \sin x \end{aligned}$$

ולכן

$$\int x^2 \cos(x) dx = x^2 \sin(x) + 2x \cos(x) - 2 \sin(x)$$

5. לעתים, לאחר ביצוע אינטגרציה בחלקים כמה פעמים אנו מקבלים ביטוי המכיל את הביטוי ממנו התחלנו (הדבר קורה בעיקר כאשר הגורמים במכפלה

הם בעלי התנהגות מחזורית תחת פעולת גזירה, כמו האקספוננט והפונקציות הטריגונומטריות). כך מקבלים משוואה המכילה את הביטוי המבוקש וניתן לחלץ אותו. לדוגמה, ננסה לחשב את

$$\int e^x \sin(x)dx$$

נסמן $e^x \sin x = f'g$ כאשר $f(x) = e^x$ ו־ $g(x) = \sin x$ ונקבל

$$\int e^x \sin(x)dx = e^x \sin(x) - \int e^x \cos(x)dx$$

את האינטגרנד במחובר הימני נרשום $e^x \cos x = u'v$ כאשר $u(x) = e^x$ ו־ $v(x) = \cos x$, ונקבל

$$\int e^x \cos(x)dx \quad = \quad e^x \cos x - \int e^x(-\sin x)$$

ובסך הכול קיבלנו ש־

$$\begin{aligned} \int e^x \sin(x)dx \quad &= \quad e^x \sin(x) - \left(e^x \cos(x) - \int e^x(-\sin(x))dx\right) \\ &= \quad e^x \sin(x) - e^x \cos(x) - \int e^x \sin(x)dx \end{aligned}$$

האינטגרל שאנו מעוניינים בו הוא $\int e^x \sin x$, אשר מופיע בשני האגפים. נעביר אגפים ונקבל

$$2\int e^x \sin(x)dx \quad = \quad e^x \sin(x) - e^x \cos(x)$$

ובסיכום,

$$\int e^x \cos(x)dx \quad = \quad \frac{1}{2}\left(e^x \sin(x) - e^x \cos(x)\right)$$

הצבה (החלפת משתנה)

שתי שיטות האינטגרציה הבאות הן גרסאות של כלל השרשרת לנגזרות.

משפט 9.7.5 (הצבה ישירה) נניח ש־ $F' = f$ בקטע I (סגור או פתוח, חסום או לא חסום), ותהי $\phi : J \to I$ פונקציה גזירה מקטע כלשהו J לתוך I (כך שהפונקציות $f \circ \phi$ ו־$F \circ \phi$ מוגדרות בקטע J). אז

$$\int f(\phi(x))\phi'(x)dx = F(\phi(x))$$

הוכחה כל שיש לעשות הוא להוכיח כי הפונקציה באגף ימין היא פונקציה קדומה של האינטגרנד באגף שמאל, ואכן בעזרת כלל השרשרת מקבלים

$$\begin{aligned}(F(\phi(x)))' &= F'(\phi(x)) \cdot \phi'(x) \\ &= f(\phi(x)) \cdot \phi'(x)\end{aligned}$$

כנדרש. ■

מסקנה 9.7.6 נניח ש־ $F' = f$ רציפה בקטע I (סגור או פתוח, חסום או לא חסום), ותהי $\phi : [a,b] \to I$ פונקציה גזירה ברציפות. אז

$$\int_a^b f(\phi(x))\phi'(x)dx = \int_{\phi(a)}^{\phi(b)} f(t)dt$$

הוכחה תהי F פונקציה קדומה של f, כך ש־ $F \circ \phi$ היא פונקציה קדומה של $(f \circ \phi) \cdot \phi'$. לפי המשפט היסודי,

$$\int_a^b f(\phi(x))\phi'(x)dx = F \circ \phi|_a^b = F(\phi(b)) - F(\phi(a)) = \int_{\phi(a)}^{\phi(b)} f(t)dt$$

כנדרש. ■

בפועל, שיטת ההצבה מאפשרת להפוך את האינטגרנד לפשוט יותר על ידי המרת בעיית האינטגרציה $\int f(\phi(x))\phi'(x)dx$ בבעיית האינטגרציה $\int f(y)dy$. אם פותרים את הבעיה האחרונה (למשל, בעזרת שיטות האינטגרציה שכבר למדנו, או באמצעות הצבה נוספת) ומקבלים פונקציה קדומה $F(y)$, אפשר להציב בחזרה את השוויון $y = \phi(x)$ ונקבל פתרון $F(\phi(x))$ לבעיה המקורית.

לעתים שיטת ההצבה מכונה **החלפת משתנה**. בשפה זו מתארים את המעבר מהבעיה $\int f(\phi(x))\phi'(x)dx$ לבעיה $\int f(y)dy$ כאילו המשתנה $\phi(x)$ הוחלף במשתנה y. כשמבצעים החלפה כזאת יש גם להחליף את $\phi'(x)dx$ ב־ dy. כדי לזכור את הכלל אפשר להיעזר בשוויון

$$\phi' = \frac{d\phi}{dx} = \frac{dy}{dx}$$

(בשוויון השני השתמשנו בהצבה $y = \phi(x)$). קיבלנו $\phi'(x) = \frac{dy}{dx}$, והעברת אגפים נותנת $dy = \phi'(x)dx$.

פעולות כמו כפל וחילוק בסימנים dx, dy לרוב מובילות לתוצאות נכונות, ויש ספרים רבים המעודדים חשבונות מסוג זה. אולם נדגיש שבמסגרת שלנו, ולאור ההגדרות שבחרנו, אין לחשבונות כאלה כל הצדקה פורמלית. הסימנים dx, dy אינם אלא חלק מהסימון של נגזרת ואינטגרל. לא הגדרנו אותם כישויות עצמאיות, וממילא אין

טעם להעביר אותם אגפים ולבצע בהם פעולות חשבון כאילו היו מספרים. קיימת מסגרת שבה יש לסימנים dx, dx משמעות עצמאית. אז הם נקראים **דיפרנציאלים** (differentials), ועם הגדרות מתאימות אפשר להצדיק פעולות כמו אלה שתיארנו. לא נעסוק בכך כאן, ובמידה שנעשה בהם חשבונות כאלה תמיד תהיה למסקנה גם הצדקה פורמלית.

בפועל, כשמשתמשים בשיטת ההצבה מתעלמים מהדרישות הטכניות של המשפט, כמו התאמה בין התמונה של ϕ לתחום של f. למרות ההתרשלות מקבלים בדרך־כלל תוצאה נכונה בסופו של דבר. אך אי אפשר לסמוך על כך, ולכן אם "מרמים" יש לבדוק בסוף על ידי גזירה שהתוצאה שהתקבלה היא באמת הפונקציה הקדומה שחיפשנו. אנו ננהג כך בדוגמאות, אך את הבדיקה אנו נשאיר לכם כתרגיל.

דוגמאות

1. נחשב את

$$\int \sin(x)\cos(x)dx$$

האינטגרנד הוא מהצורה $\phi^2(x)\cdot\phi'(x)$ עבור $\phi(x)=\sin x$, ולכן הוא מהצורה $f(\phi(x))\cdot\phi'(x)$, כאשר $f(t)=t^2$. פונקציה קדומה ל־ $f(t)$ היא $F(t)=\frac{1}{3}t^3$, ולכן על פי משפט ההצבה הישירה מקבלים

$$\begin{aligned}\int \sin^2(x)\cos(x)dx &= \int f(\phi(x))\phi'(x)dx\\ &= F(\phi(x))\\ &= \frac{1}{3}\sin^3(x)\end{aligned}$$

בפרט האינטגרל המסוים $\int_0^1 \sin^2(x)\cos(x)dx$ הוא

$$\int_0^1 \sin^2(x)\cos(x)dx = F|_{\sin 0}^{\sin 1} = \frac{1}{3}\sin^3 1 - \frac{1}{3}\sin^3 0 = \frac{1}{3}\sin^3 1$$

2. נחשב את

$$\int 2xe^{x^2}dx$$

נזהה שוב את התבנית $f(\phi(x))\cdot\phi'(x)$ באינטגרנד, כאשר $\phi(x)=x^2$ ו־ $f(t)=e^t$. הפונקציה הקדומה של $f(t)$ היא אותה הפונקציה: $F(t)=e^t$, ולכן

$$\int 2xe^{x^2}dx = F(\phi(x)) = e^{x^2}$$

3. נחשב את

$$\int \cot(x)dx$$

כזכור $\cot(x) = \frac{1}{\sin(x)} \cdot \cos(x)$. נשים לב ש־$\cot$ היא מהצורה $f(\phi(x)) \cdot \phi'(x)$ עבור הפונקציות $\phi(x) = \sin x$ ו־ $f(t) = \frac{1}{t}$. הפונקציה הקדומה ל־$f(t)$ היא $F(t) = \log|t|$, ומכאן:

$$\begin{aligned} \int \cot(x)dx &= \int \frac{\cos(x)}{\sin(x)}dx \\ &= \log|\sin(x)| \end{aligned}$$

במשפט הקודם ראינו כיצד פותרים בעיית אינטגרציה מהצורה $\int f(\phi(x))\phi'(x)dx$ על ידי פתרון האינטגרל $\int f(t)dt$. כעת נראה קרוב משפחה של השיטה, המאפשר לחשב את $\int f(t)dt$ על ידי חישוב האינטגרל $\int f(\phi(x))\phi'(x)dx$.

משפט 9.7.7 (הצבה הפוכה) נניח ש־f מוגדרת בקטע כלשהו I (סגור או פתוח, חסום או לא חסום), ותהי ϕ פונקציה גזירה מקטע J לתוך I (כך שהפונקציה $f \circ \phi$ מוגדרת בקטע J). נניח בנוסף כי ϕ הפיכה. אם F מקיימת

$$F = \int f(\phi(t))\phi'(t)dt$$

אז

$$\int f(x)dx = F(\phi^{-1}(x))$$

הוכחה יש לבדוק שהפונקציה האמורה היא פונקציה קדומה של f. נסמן $\psi = \phi^{-1}$ ו־ $G(x) = F(\psi(x))$. עלינו להראות ש־ $G' = f$. לפי כלל השרשרת

$$G'(x) = F'(\psi(x))\psi'(x)$$

מכיוון ש־ $F'(x) = f(\phi(x))\phi'(x)$, קיבלנו

$$G'(x) = f(\phi(\psi(x)))\phi'(\psi(x)) \cdot \psi'(x)$$

אבל $\phi(\psi(x)) = x$ (כי $\psi = \phi^{-1}$), וגזירת השוויון הזה בעזרת כלל השרשרת נותנת $\phi'(\psi(x)) \cdot \psi'(x) = 1$ (זו גרסה של הנוסחה $\psi'(x) = \frac{1}{\phi'(\psi(x))}$ לפונקציה הפוכה, אלא שלא צריך לדאוג שמא ψ' תתאפס). אם מציבים שוויונות אלה בביטוי שקיבלנו ל־G' נקבל

$$G'(x) = f(x)$$

כנדרש. ■

מסקנה 9.7.8 (שיטת ההצבה ההפוכה באינטגרל המסוים) בסימונים של המשפט, אם $I = [a, b]$ ואם F פונקציה קדומה של $f(\phi(x))\phi'(x)$ אז

$$\begin{aligned} \int_a^b f(x)dx &= \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\phi(t))\phi'(t)dt \\ &= F|_{\phi^{-1}(a)}^{\phi^{-1}(b)} \end{aligned}$$

ההוכחה דומה להוכחת מסקנה 9.7.6 ואנו משאירים אותה כתרגיל.

גם להצבה ההפוכה יש תיאור סמלי כהחלפת משתנה: עברנו מבעיית האינטגרציה $\int f(x)dx$ לבעיה $\int f(\phi(t))\phi'(t)dt$ בעזרת ההצבה $x = \phi(t)$. הכלל הוא שאז יש גם להחליף את dx ב־$\phi'(t)dt$. כדי לזכור כלל זה אפשר להיעזר בשוויון

$$\phi' = \frac{d\phi}{dt} = \frac{dx}{dt}$$

(בשוויון האחרון הצבנו $x = \phi(t)$). לאחר "העברת אגפים" נקבל $dx = \phi'(t)dt$. כמו במקרה של ההצבה הישירה, אין הצדקה פורמלית לפעולות אלה, ויש להתייחס אליהן בהתאם.

כמו במקרה של ההצבה הישירה, גם בדוגמאות לשימוש בכלל ההצבה ההפוכה לרוב לא נטרח לוודא שהפונקציה ϕ מקיימת את כל דרישות המשפט, ונעדיף ל"רמות" ולבדוק בסוף שהתוצאה שהתקבלה נכונה (זהו, כמובן, התפקיד שלכם).

דוגמאות

1. נחשב את

$$\int \frac{\sqrt{x}}{x^{3/2}+1}dx$$

על מנת להיפטר מהשורשים נחליף את x ב־t^2. נסמן $\phi(t) = t^2$. קיבלנו בעיית אינטגרציה חדשה:

$$\int \frac{\sqrt{\phi(t)}}{\phi(t)^{3/2}+1}\phi'(t)dt = \int \frac{t}{t^3+1}2tdt = 2\int \frac{t^2}{t^3+1}dt$$

בעזרת השיטות שכבר למדנו ניתן לפתור אינטגרל זה ולקבל שהפונקציה $G(t) = \frac{2}{3}\ln|t^3+1|$ היא פתרון (הוכיחו!). לכן הפונקציה

$$F(x) = G(\phi^{-1}(x)) = G(\sqrt{x}) = \frac{2}{3}\ln|x^{3/2}+1|$$

היא פתרון לבעיית האינטגרציה המקורית (אפשר לפתור בעיה זו גם בהצבה ישירה. איך?).

2. נחשב את

$$\int \sqrt{1-x^2}dx$$

מכיוון ש־ $\sqrt{1-\sin^2 t} = \cos t$, הצבת $\phi(t) = sin(t)$ במקום x תפשט את הביטוי במידה ניכרת. נציב ונקבל

$$\begin{aligned}
\int \sqrt{1-\phi(t)^2} \cdot \phi'(t)dt &= \int \sqrt{1-\sin^2(t)}\cos(t)dt \\
&= \int \cos^2(t)dt \\
&= \int \left(\frac{1}{2}+\frac{1}{2}\cos(2t)\right)dt \\
&= \frac{t}{2}+\frac{1}{4}\sin(2t)
\end{aligned}$$

(השתמשנו בזהות הטריגונומטרית $\cos^2 t = \frac{1}{2}+\frac{1}{2}\cos(2t)$). ניתן עדיין לפשט את הביטוי בעזרת הזהות

$$\sin(2t) = 2\sin(t)\cos(t)$$

ומכאן

$$\sin\left(2\arcsin(t)\right) = 2t\sqrt{1-t^2}$$

לכן הצבת $\phi^{-1}(x) = \arcsin x$ במקום t נותנת

$$\int \sqrt{1-x^2}dx = \frac{1}{2}\arcsin x + \frac{x}{2}\sqrt{1-x^2}$$

פונקציות רציונליות

באופן כללי לפונקציה אלמנטרית אין פונקציה קדומה אלמנטרית. יוצאת מן הכלל היא משפחת הפונקציות הרציונליות. לפונקציה כזו תמיד יש פונקציה קדומה אלמנטרית, ויתר על כן יש שיטה לחשב אותה. אנו לא נוכיח את השיטה אלא נסתפק בתיאור קצר שלה ובדוגמאות לשימוש בה.

הגדרה 9.7.9 פולינום p נקרא **אי־פריק** (irreducible) אם לכל שני פולינומים r, s המקיימים $p = r \cdot s$, אחד הפולינומים r, s הוא ממעלה 1.

קל לוודא שכל פולינום ניתן לכתיבה כמכפלה של פולינומים אי־פריקים (מוכיחים זאת באינדוקציה על המעלה, תוך שימוד באבחנה שאם $p = r \cdot s$ ואם r, s אינם קבועים, אז $\deg r < \deg p$ וגם $\deg s < \deg p$).

המשפט הבא מאפיין את הפולינומים האי־פריקים. זהו משפט חשוב ולא טריביאלי מתורת האלגברה של שדות, ונצטט אותו ללא הוכחה. אפשר למצוא הוכחות שלו ב־[5] או [17].

משפט 9.7.10 כל פולינום אי־פריק הוא ממעלה אחת או שתיים.

נובע מכאן שכל פולינום הוא מכפלה של פולינומים ממעלה אחת ושתיים, כי הוא שווה למכפלה של פולינומים אי־פריקים.

אפשר להראות שאם לפולינום ממעלה גדולה מאחת יש שורש אז הוא פריק (ראו תרגיל (5) בעמוד 229). לכן פולינום ריבועי אי־פריק הוא פולינום ללא שורשים. מכאן ש־ $ax^2 + bx + c$ הוא אי־פריק אמ"מ $b^2 - 4ac < 0$ (ראו תרגיל (3) בעמוד 77).

עובדה נוספת שנשתמש בה היא משפט הפירוק הבא לפונקציות רציונליות:

משפט 9.7.11 תהי $\frac{p}{q}$ פונקציה רציונלית. אז אפשר לרשום את f כסכום

$$\frac{p(x)}{q(x)} = r(x) + \sum_{i=1}^{n} \frac{s_i(x)}{t_i(x)^{k_i}}$$

כאשר r, s_i, t_i פולינומים ו־k_i מספרים טבעיים המקיימים

- $\deg r < \deg p - \deg q$. (אם $\deg p < \deg q$ אז $r = 0$).
- t_i גורם אי־פריק של q ו־$(t_i)^{k_i}$ גורם של q.
- $\deg s_i < \deg t_i$ (כלומר $\deg s_i \leq 1$, ואם $\deg t_i = 1$ אז s_i קבועה).

דוגמאות

1. נתבונן בפונקציה $f(x) = \frac{1}{x^2-1}$. אפשר לרשום את המכנה בתור המכפלה $x^2 - 1 = (x+1)(x-1)$, ולכן מהמשפט אנו יודעים שאפשר לרשום את f בצורה

$$\frac{1}{x^2-1} = \frac{A}{x-1} + \frac{B}{x+1}$$

כאשר A, B קבועים (הפולינום r מהמשפט הוא אפס כי המעלה של המונה של $\frac{1}{x^2-1}$ היא 0). נותר רק למצוא את A, B. נעבור למכנה משותף ונקבל

$$\frac{1}{x^2-1} = \frac{A}{x-1} + \frac{B}{x+1} = \frac{A(x+1) + B(x-1)}{x^2-1}$$

ולכן מהשוואת המונים בשני האגפים נקבל

$$1 = (A+B)x + (A-B)$$

כלומר, $A+B=0$ ו־ $A-B=1$. קל לפתור את מערכת המשוואות הזו ולהסיק ש־ $A=\frac{1}{2}$ ו־ $B=-\frac{1}{2}$, ולכן

$$\frac{1}{x^2-1} = \frac{1}{2(x-1)} - \frac{1}{2(x+1)}$$

זו ההצגה המבוקשת.

2. נתבונן בפונקציה הרציונלית $f(x) = \frac{x^4-2x^3-x^2-5}{(x^2+x+1)(x-3)}$. הגורמים האי־פריקים של המכנה הם $x-3$ ו־ x^2+x+1 ולכן לפי המשפט אפשר לרשום את f בתור

$$f(x) = (Ax+B) + \frac{Cx+D}{x^2+x+1} + \frac{E}{x-3}$$

נעבור למכנה משותף ונקבל

$$\begin{aligned} f(x) &= \\ &= \frac{(Ax+B)(x^2+x+1)(x-3) + (Cx+D)(x-3) + E(x^2+x+1)}{(x^2+x+1)(x-3)} \end{aligned}$$

אחרי פתיחת סוגריים במונה והשוואת המקדמים של הפולינום המתקבל עם המקדמים במונה של f מקבלים את המשוואות

$$\begin{aligned} 1 &= A \\ -2 &= -2A+B \\ -1 &= -2A-3B+C+E \\ 0 &= -3A-2B-3C+D+E \\ -5 &= -3B-3D+E \end{aligned}$$

פתרון המערכת וחישוב A,B,C,D,E נותן

$$\frac{x^4-2x^3-x^2-5}{(x^2+x+1)(x-3)} = x + \frac{1}{x-3} + \frac{2}{x^2+x+1}$$

הסיבה שפירוק כמו במשפט 9.7.11 מסייע בחישוב אינטגרלים היא שאפשר לחשב את האינטגרל של כל אחד מסוגי המחוברים המופיעים בפירוק. ליתר דיוק, עד כדי קבוע כפלי, שאפשר להתעלם ממנו, סוגי הפונקציות g שמופיעים כמחוברים בפירוק הם:

1. פולינומים.
2. פונקציות מהצורה $g(x) = \frac{1}{(x+a)^k}$.
3. פונקציות מהצורה $g(x) = \frac{1}{(x^2+ax+b)^k}$ כאשר $x^2 + ax + b$ אי־פריק.
4. פונקציות מהצורה $g(x) = \frac{x+c}{(x^2+ax+b)^k}$ כאשר $x^2 + ax + b$ אי־פריק.

אין קושי לחשב אינטגרל של פולינום, וגם של פונקציות מהסוג השני:

$$\int \frac{1}{(x+a)^k}dx = \begin{cases} \ln|x+a| & k=1 \\ \frac{1}{1-k}(x+a)^{1-k} & k>1 \end{cases}$$

בעזרת מניפולציות אלגבריות אפשר להמיר בעיית אינטגרציה של פונקציה מהסוג הרביעי בבעיה מהסוג השלישי, שכן

$$\int \frac{x+c}{(x^2+ax+b)^k}dx = \frac{1}{2}\int \frac{2x+a}{(x^2+ax+b)^k}dx + (c-\frac{a}{2})\int \frac{1}{(x^2+ax+b)^k}dx$$

האינטגרל השמאלי באגף ימין הוא מידי:

$$\int \frac{2x+a}{(x^2+ax+b)^k}dx = \begin{cases} \ln|x^2+ax+b| & k=1 \\ \frac{1}{1-k}(x^2+ax+b)^{1-k} & k>1 \end{cases}$$

ואילו האינטגרנד הימני הוא מהסוג השלישי ברשימה.

בסיכום, ראינו איך לפתור את בעיית האינטגרציה הלא מסוימת של כל סוגי הפונקציות המופיעות כמחוברים בפירוק של פונקציה רציונלית, למעט פונקציות מהצורה $\frac{1}{(x^2+ax+b)^k}$. נראה כעת שגם עם אלה אפשר להתמודד. ראשית נרצה להחליף משתנה באינטגרל ולהביא אותו לצורה $\int \frac{c}{(y^2+1)^k}dy$. נשים לב ש־

$$\begin{aligned} x^2+ax+b &= (x+\frac{a}{2})^2 + (b-\frac{a^2}{4}) \\ &= (b-\frac{a^2}{4})\left((\frac{x}{\sqrt{b-a^2/4}} + \frac{a}{2\sqrt{b-a^2/4}})^2 + 1\right) \end{aligned}$$

(שימו לב ש־ $b-\frac{a^2}{4} > 0$ כי x^2+ax+b אי פריק ולכן השורש מוגדר). לכן אם נסמן $c = \sqrt{b-\frac{a^2}{4}}$ נקבל

$$\int \frac{1}{(x^2+ax+b)^k}dx = \frac{1}{c^{2k}}\int \frac{1}{((\frac{x}{c}+\frac{a}{c})^2+1)^k}dx = \frac{1}{c^{2k-1}}\int \frac{1}{(y^2+1)^k}dy$$

כאשר הצבנו כאן $y = \frac{x}{c} + \frac{a}{c}$.

נותר לפתור את בעיית האינטגרציה $\int \frac{1}{(y^2+1)^k}dy$. עבור $k=1$ זהו אינטגרל מידי:

$$\int \frac{1}{y^2+1}dy = \arctan y$$

כלומר

$$\int \frac{1}{x^2+ax+b}dx = \frac{1}{\sqrt{b-a^2/4}}\arctan(\frac{x}{\sqrt{b-a^2/4}}+\frac{a}{\sqrt{b-a^2/4}})$$

לגבי המקרה $k>0$, אפשר לקבל נוסחת נסיגה עבור $\int \frac{1}{(x^2+ax+b)^k}dx$. אנו משאירים זאת כתרגיל (ראו גם שאלה (5) בסוף הסעיף).

דוגמאות

1. ראינו בדוגמה (1) בעמוד 428 ש־ $\frac{1}{x^2-1}=\frac{1}{2(x-1)}-\frac{1}{2(x+1)}$, ולכן

$$\int \frac{1}{x^2-1}dx = \frac{1}{2}\int\frac{1}{x-1}dx - \frac{1}{2}\int\frac{1}{x+1}dx = \frac{1}{2}\ln|x-1|+\frac{1}{2}\ln|x+1|$$

2. ראינו בדוגמה (2) בעמוד 429 ש־ $\frac{x^4-2x^3-x^2-5}{(x^2+x+1)(x-3)}=x+\frac{1}{x-3}+\frac{2}{x^2+x+1}$, ולכן

$$\begin{aligned}\int \frac{x^4-2x^3-x^2-5}{(x^2+x+1)(x-3)}dx &= \int xdx+\int\frac{1}{x-3}dx+\int\frac{2}{x^2+x+1}dx\\ &= \frac{1}{2}x^2+\ln|x-3|+\frac{4}{\sqrt{3}}\arctan(\frac{2}{\sqrt{3}}x+\frac{2}{\sqrt{3}})\end{aligned}$$

הצבות טריגונומטריות

כאשר האינטגרנד מכיל סכומים, מכפלות ומנות של פונקציות טריגונומטריות, אפשר להמיר את בעיית האינטגרציה בבעיית אינטגרציה של פונקציה רציונלית. השיטה מבוססת על הזהויות

$$\begin{aligned}\sin x &= \frac{2\tan(x/2)}{1+\tan^2(x/2)}\\ \cos x &= \frac{1-\tan^2(x/2)}{1+\tan^2(x/2)}\end{aligned}$$

(ראו תרגיל (11) בעמוד 230). כמו־כן מתקיים

$$\frac{d}{dx}\tan\frac{x}{2}=\frac{1}{2\cos^2(x/2)}=\frac{1}{1+\cos x}=\frac{1}{2}(1+\tan^2\frac{x}{2})$$

זהויות אלה שימושיות כאשר רוצים לחשב אינטגרל $\int f$ כאשר האינטגרנד $f(x)$ הוא ביטוי המורכב מפונקציות טריגונומטריות. בשלב ראשון ניעזר בזהויות המבטאות

את $\sin x, \cos x$ בעזרת $\tan\frac{x}{2}$ כדי להפוך את האינטגרנד לפונקציה שבה מופיעים סכומים, מכפלות ומנות של $\tan\frac{x}{2}$. על ידי מניפולציות אלגבריות נוכל להביא את f לצורה

$$\begin{aligned} f(x) &= g(\tan\frac{x}{2})\cdot\frac{1+\tan^2\frac{x}{2}}{2} \\ &= g(\tan\frac{x}{2})\cdot\frac{d}{dx}\tan\frac{x}{2} \end{aligned}$$

ולכן

$$\int f(x)dx = \int g(\phi)\phi'(x)dx$$

כאשר $\phi(x) = \tan\frac{x}{2}$. לכן לפי משפט ההצבה 9.7.5 אם נחשב את $G(t) = \int g(t)dt$ אז נקבל ש־ $\int f = G(\phi(x)) = G(\tan\frac{x}{2})$, וזה פותר את הבעיה המקורית שלנו.

לדוגמה, נחשב את $\int\frac{1}{\sin x}dx$. ראשית נביא את האינטגרנד לצורה המבוקשת על ידי הצבת הזהות המבטאת את $\sin x$ בעזרת $\tan\frac{x}{2}$:

$$\begin{aligned} \int\frac{1}{\sin x}dx &= \int\frac{1+\tan^2(x/2)}{2\tan(x/2)}dx \\ &= \int\frac{1}{\tan(x/2)}\cdot\frac{1}{2}(1+\tan^2\frac{x}{2})dx \\ &= \int\frac{1}{\phi(x)}\phi'(x)dx \end{aligned}$$

כאשר $\phi(x) = \tan\frac{x}{2}$. לכן אם נמצא פתרון $F(t)$ לבעיית האינטגרציה $\int\frac{1}{t}dt$ הרי ש־ $F(\tan\frac{x}{2})$ הוא פתרון לבעיה המקורית. אבל זהו אינטגרל מידי: $F(t) = \ln|t|$. לכן

$$\int\frac{1}{\sin x}dx = \ln|\tan\frac{x}{2}|$$

הצבת אוילר

הצבה חשובה נוספת היא **הצבת אוילר** שמאפשרת במקרים מסוימים להיפטר מביטויים באינטגרנד המכילים שורשים. ראינו בדוגמה (2) בעמוד 427 שהצבה של פונקציה טריגונומטרית יכולה לעזור להיפטר מביטוי מהסוג $\sqrt{1-x^2}$ באינטגרנד, אך אם תבדקו תגלו שהצבה כזו אינה מועילה לחישוב האינטגרל $\int\sqrt{1+x^2}dx$. כדי להיפטר מהשורש נחפש הצבה $x = \phi(t)$ כך שהביטוי $\sqrt{1+\phi(t)^2}$ באינטגרנד יהיה פשוט ככל האפשר. ניסיון ראשון הוא ההצבה $\sqrt{y^2-1}$. הצבה זו אינה מובילה לבעיה פשוטה יותר (נסו!). במקומה ניעזר בפונקציות הבאות:

הגדרה 9.7.12 הפונקציות

$$\begin{aligned}\sinh(t) &= \frac{1}{2}\left(e^t - e^{-t}\right)\\ \cosh(t) &= \frac{1}{2}\left(e^t + e^{-t}\right)\end{aligned}$$

נקראות **הסינוס ההיפרבולי והקוסינוס ההיפרבולי** (hyperbolic sine and cosine) בהתאמה.

דמיון מסוים בין פונקציות אלה והפונקציות הטריגונומטריות הרגילות ניכר בתכונות הבאות שלהן:

$$\begin{aligned}\sinh 0 &= 0\\ \cosh 0 &= 1\\ \cosh^2(t) - \sinh^2(t) &= 1\\ \sinh(s+t) &= \sinh(t)\cosh(s) + \sinh s \cosh t\\ \cosh(s+t) &= \cosh s \cosh t + \sinh s \sinh t\\ \sinh'(t) &= \cosh(t)\\ \cosh'(t) &= \sinh(t)\end{aligned}$$

(הוכיחו את התכונות!). בתרגיל (11) בסוף הסעיף נצביע על דמיון נוסף בין פונקציות אלה לפונקציות הטריגונומטריות הרגילות.

נתעניין גם בפונקציות היפרבוליות הפוכות. קל לבדוק ש־sinh עולה ממש בכל הישר והתמונה שלה היא $\mathbb{R}$, ולכן יש לה פונקציה הפוכה המוגדרת בכל $\mathbb{R}$. הפונקציה ההפוכה ל־sinh היא הפונקציה שערכה ב־x הוא t אמ"מ $\sinh t = x$. כלומר, אנו רוצים לפתור את המשוואה

$$x = \sinh t = \frac{1}{2}\left(e^t - e^{-t}\right)$$

ולחלץ מהביטוי את t. נשים לב שעל ידי כפל ב־e^{2x} מקבלים משוואה ריבועית ב־e^t:

$$\left(e^t\right)^2 - (2x)e^t - 1 = 0$$

ומכאן:

$$e^t = x + \sqrt{x^2+1}$$

(מדוע ניתן להתעלם מהפתרון בו השורש בסימן שלילי?). נפעיל ln ונקבל

$$\sinh^{-1}(x) = t = \ln(x + \sqrt{x^2+1})$$

אנו משאירים כתרגיל את החישוב של $\cosh^{-1}(x)$ (במקרה זה יש לחפש פונקציה הפוכה בתחום מצומצם כי $\cosh$ אינה חד־חד־ערכית בכל תחומה).

נשוב לבעיית האינטגרציה שבה פתחנו. נשים לב שלו היינו מחליפים את x בביטוי $\sqrt{1+x^2}$ ב־$\sinh(t)$, על פי הזהות $1+\sinh^2 x = \cosh^2 x$ היינו מקבלים

$$\sqrt{1+\sinh^2(t)} = \sqrt{\cosh^2(t)} = \cosh(t)$$

(מדוע אין צורך לעטוף את אגף ימין בערך מוחלט?). ננסה, אם כך, את ההצבה הזו. נחליף את x ב־$\sinh(t)$ ואת dx ב־ $\cosh(t)dt$, ונקבל את הבעיה החדשה

$$\int \sqrt{1+\sinh^2(t)}\cosh(t)dt \quad = \quad \int \cosh^2(t)dt$$

מהתכונות של $\cosh(t)$ קל להוכיח את הזהות $\cosh 2t = 1 + 2\cos^2 t$, ומקבלים

$$\begin{aligned}\int \cosh^2(t)dt &= \frac{1}{2}\int \cosh(2t)dt - \frac{1}{2}\int 1dt \\ &= \frac{1}{4}\sinh(2t) - \frac{1}{2}t \\ &= \frac{1}{2}\sinh(t)\sqrt{1+\sinh^2 t} - \frac{1}{2}t\end{aligned}$$

כל שנותר לעשות הוא להמיר את הביטוי הרשום לביטוי במשתנה המקורי x, כלומר לבצע את ההצבה $t = \operatorname{arcsinh}(x) = \ln(x+\sqrt{1+x^2})$, ומכאן

$$\int \sqrt{1+x^2}dx = \frac{1}{2}x\sqrt{1+x^2} - \frac{1}{2}\ln(x+\sqrt{1+x^2})$$

תרגילים

כפי שתגלו בתרגילים, הקושי העיקרי בהפעלת הכלים שפיתחנו למעלה הוא לבחור את השיטה המתאימה. בניגוד לכללי התחשיב הקודמים שהכרנו, בהם הבחירה לרוב הייתה ברורה, כאן יש חופש פעולה רב, ולפעמים יש לנסות כמה גישות לפני שמגיעים לפתרון. מסיבה זו אינטגרציה לא מסוימת היא בחלקה אומנות, ויש לתרגל אותה שוב ושוב עד שמקבלים את האינטואיציה הדרושה.

1. פתרו את האינטגרלים המידיים הבאים:

(א) $\int(e^x+\cos x)dx$.

(ב) $\int(x^2+100x+\ln x)dx$.

(ג) $\int \ln(x^{10})dx$.

2. פתרו את האינטגרלים הבאים על ידי אינטגרציה בחלקים:

(א) $\int x^2 \log(x)dx$.

(ב) $\int x^2 \log^2(x)dx$.

(ג) $\int \sin(\log(x))dx$.

(ד) $\int x \arctan(x)dx$.

(ה) $\int \frac{xdx}{\sin^2(x)}$.

3. בטאו את $\int x^4 e^{x^2}dx$ בעזרת $\int e^{x^2}dx$ (האינטגרל $\int e^{x^2}$ אינו ניתן לכתיבה כפונקציה אלמנטרית).

4. הוכיחו את הנוסחה

$$\int e^x \cdot p(x)dx = e^x \cdot (p(x) - p'(x) + p''(x) - \ldots + (-1)^n p^{(n)}(x))$$

כאשר $\deg p = n$.

5. נסמן

$$I_n = \int \frac{dx}{(1+x^2)^n}$$

I_0, I_1 הם אינטגרלים מידיים. מצאו ביטוי רקורסיבי ל־I_n.

6. פתרו את האינטגרלים הבאים על ידי הצבה ישירה:

(א) $\int e^{\sin(x)} \cos x dx$.

(ב) $\int x \cos(x^2)dx$.

(ג) $\int \frac{\cos(\ln x)}{x}dx$.

(ד) $\int e^x \ln(1+e^x)$.

7. פתרו את האינטגרלים הבאים בעזרת הצבה הפוכה:

(א) $\int \sqrt{1-x^2}\arcsin(x)dx$.

(ב) $\int \frac{1}{1-(e^x)^2}dx$.

(ג) $\int e^{\sin(x)}dx$.

(ד) $\int x\sqrt{1+x^2}dx$.

8. בתרגיל זה נלמד על הצבות טריגונומטריות פשוטות.

(א) יהיו $m, n \in \mathbb{N}$ ונתבונן באינטגרל $\int \sin^n(x)\cos^m(x)dx$. נניח כי אחד המעריכים, למשל n, הוא אי־זוגי. הראו שניתן לפתור את האינטגרל בעזרת ההצבה $t = \sin(x)$ (מה אם m אי־זוגי?).

(ב) פתרו את האינטגרלים הבאים:

i. $\int \frac{dx}{\cos(x)}$
ii. $\int \tan(x)dx$
iii. $\int x \cdot \tan^2(x)dx$ (הצעה: חשבו תחילה את $\int \tan^2(x)dx$)

9. חשבו את האינטגרלים הבאים (נסו לבחור את השיטה המתאימה. לפעמים יש כמה שלבים בפתרון שבהם יש להפעיל שיטות שונות).

(א) $\int a^{3x}dx$.
(ב) $\int \frac{x}{\sqrt{x^2+1}}dx$.
(ג) $\int \frac{e^{2x}}{\sqrt[4]{e^x+1}}dx$.
(ד) $\int \frac{xdx}{\sqrt{x^2-2x+5}}$.
(ה) $\int \arcsin x$.
(ו) $\int \frac{\sqrt{x}}{\sqrt{x}-\sqrt[3]{x}}dx$.
(ז) $\int \cot^3(x)dx$
(ח) $\int \frac{dx}{\sqrt{x^2-2x+5}}$
(ט) $\int \sin^5(x)dx$
(י) $\int x \cdot \arcsin(x)$.
(יא) $\int \frac{dx}{\sqrt{a^2+x^2}}$.

10. בשאלה זו ניתן הגדרה חלופית ללוגריתם ולאקספוננטה. בסעיפים הבאים עליכם להסתמך רק על תורת האינטגרציה והגזירה ללא חישוב מפורש של פונקציות הקדומות.

(א) עבור $x \in (0,\infty)$ נגדיר $L(x) = \int_1^x \frac{1}{t}dt$. הוכיחו ש־$L$ היא חח"ע ועל ומקיימת $L(xy) = L(x) + L(y)$ ו־$xL(y) = L(x^y)$ לכל $x, y > 0$ (רמז: $\int_1^{xy} = \int_1^x + \int_x^{xy}$).
(ב) תהי $E : \mathbb{R} \to (0,\infty)$ הפונקציה $E = L^{-1}$. הוכיחו ש־$E' = E$ ושלכל $x, y \in \mathbb{R}$ מתקיים $E(x+y) = E(x)E(y)$, $E(xy) = E(x)^y$.
(ג) הסיקו ש־ $E(1) = e$, $E(x) = e^x$, $L(x) = \ln x$.

הערה למעט בסעיף האחרון, לא השתמשנו כאן בהגדרת החזקה. מכיוון שאפשר לפתח את תורת הגזירה והאינטגרציה באופן בלתי תלוי מההגדרה של חזקות או שורשים, הרי שניתן לבסס את ההגדרה של חזקות ממשיות על ההגדרה לעיל.

11. בשאלה זו נצביע על הקבלה נוספת בין הפונקציות הטריגונומטריות הרגילות וההיפרבוליות. נזכור כי מעגל היחידה במישור הוא קבוצת הנקודות (x,y) במידות המקיימות את הנוסחה $x^2 + y^2 = 1$, ונשים לב קיימת פרמטריזציה טבעית שלו באמצעות הפונקציות הטריגונומטריות, כלומר, מעגל שיחידה הוא קבוצת הנקודות שצורתם $(x(t), y(t)) = (\cos t, \sin t)$ עבור $t \in [0, 2\pi)$.

נתבונן כעת בהיפרבולה, שהיא קבוצת הנקודות (x,y) במישור המקיימות את המשוואה $x^2 - y^2 = 1$. הראו שלהיפרבולה יש פרמטריזציה בעזרת הפונקציות הטריגונומטריות ההיפרבוליות, כלומר זו קבוצת הנקודות מהצורה $(x(t), y(t))$ כאשר $x(t) = \cosh(t)$, $y(t) = \sinh(t)$, ו־$t \in [0,\infty)$.

12. תהי f פונקציה גזירה בעלת פונקציה קדומה F. נניח ש־f הפיכה בכל תחום הגדרתה. הציגו את $\int f^{-1}(x)dx$ בעזרת F. (רמז: השתמשו תחילה בהצבה, ואז באינטגרציה בחלקים). השתמשו בשיטה זו כדי לחשב את $\int \arcsin(x)dx$.

13. לכל $n \geq 1$ טבעי תהי $r_n = \int_0^{\pi/4} \tan^{2n} x\, dx$.

(א) הראו ש־ $r_n \to 0$.

(ב) מצאו נוסחת נסיגה עבור r_n והסיקו כי $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots$.

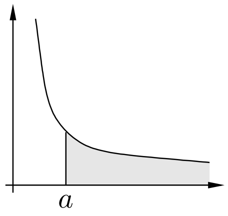
9.8 האינטגרל הלא אמתי

בסעיף זה נעסוק בהכללה של מושג האינטגרל למקרה שתחום האינטגרציה או האינטגרנד אינם חסומים (כפי שהוגדר עד כה, האינטגרל המסוים מוגדר רק עבור פונקציות חסומות, ורק בקטעים סגורים וחסומים). כפי שיתברר, היחס בין האינטגרל המוכלל, הנקרא אינטגרל לא אמתי, לבין האינטגרל המסוים דומה ליחס שבין סכימה סופית של מספרים לבין טורי מספרים (שהם סכומים אינסופיים).

נתבונן בגרף הפונקציה $f(x) = \frac{1}{x^2}$, ונשאל מהו השטח הכלוא בין הגרף לציר ה־x בקרן $[1,\infty)$. מכיוון שלא מדובר בקטע סגור לא נוכל להפעיל את הכלים מהסעיפים הקודמים, אבל לכל $t > 1$ השטח הכלוא מתחת לגרף בתחום $[1,t]$ נתון על ידי האינטגרל $S(t) = \int_1^t \frac{1}{x^2}dx$. גודל זה מקביל לסכום החלקי של טור, וכמו במקרה של טור אם $\lim_{t\to\infty} S(t)$ קיים אזי הוא מועמד טבעי לשטח שאנו מחפשים. ואמנם, לפי המשפט היסודי

$$S(t) = \int_1^t \frac{dx}{x^2} = (-\frac{1}{x})|_{x=1}^{x=t} = 1 - \frac{1}{t}$$

לכן כאשר $t \to \infty$ מתקיים $S(t) \to 1$, ונוכל לטעון שזה השטח שמתחת לגרף $\frac{1}{x^2}$ בקרן $[1,\infty)$. באופן כללי יותר,



איור 9.8.1 אינטגרל לא אמתי מ־a על ∞

הגדרה 9.8.1 תהי $f : [a,\infty) \to \mathbb{R}$ פונקציה, ונניח כי f אינטגרבילית בכל תת־קטע סגור $[a,b] \subseteq [a,\infty)$. אם קיים הגבול

$$\lim_{t\to\infty} \int_a^t f$$

אז הוא נקרא **האינטגרל הלא אמתי** (improper integral) של f בקטע $[a,\infty)$, ומסומן $\int_{[a.\infty)} f$. במקרה זה נאמר שהאינטגרל $\int_{[a,\infty)} f$ מתכנס וש־f **אינטגרבילית**

ב־$[a,\infty)$. אם הגבול אינו קיים, נאמר שהאינטגרל הלא אמתי אינו קיים. אם הגבול קיים במובן הרחב נאמר שהאינטגרל הלא אמתי קיים, או מתכנס, במובן הרחב.

האינטגרל הלא אמתי בקרן מהסוג $(-\infty,a]$ מוגדר באופן דומה.

לרוב נקרא לאינטגרל הלא אמתי פשוט אינטגרל. הסימון $\int_{[a,\infty)} f$ לאינטגרל הלא אמתי דומה לסימון $\int_{[a,b]} f$ של האינטגרל האמתי, שהופיע בהגדרה 9.1.9. בסעיף הנוכחי נעדיף סימון זה על פני הסימון $\int_a^b f$.

דוגמאות

1. כפי שראינו למעלה, הפונקציה $f(x)=\frac{1}{x^2}$ אינטגרבילית ב־ $[1,\infty)$ ומתקיים $\int_{[1,\infty)} f=1$.

   באופן כללי יותר, עבור $\alpha\neq -1$ קבוע מתקיים

$$\lim_{t\to\infty}\int_1^t x^\alpha dx=\lim_{t\to\infty}\frac{1}{\alpha+1}(t^{\alpha+1}-1)=\begin{cases}-\frac{1}{\alpha+1} & \alpha<-1\\ \infty & \alpha>-1\end{cases}$$

   ואילו עבור $\alpha=-1$,

$$\lim_{t\to\infty}\int_1^t \frac{1}{x}dx=\lim_{x\to\infty}\ln t=\infty$$

   נובע ש־x^α אינטגרבילית ב־$[1,\infty)$ אמ"מ $\alpha<-1$, ובמקרה זה מתקיים $\int_{[1,\infty)} x^\alpha dx=-\frac{1}{\alpha+1}$ (שימו לב שזה מספר חיובי).

2. תהי $g:[1,\infty)\to\mathbb{R}$ נתונה על־ידי $g(x)=e^{-x}$. אז

$$\lim_{t\to\infty}\int_1^t g(x)dx=\lim_{t\to\infty}\int_1^t e^{-x}dx=\lim_{t\to\infty}(-e^{-t}+e^{-1})=\frac{1}{e}$$

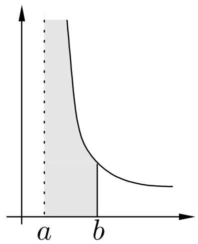
   ולכן g אינטגרבילית במובן הרחב ב־ $[1,\infty)$ ומתקיים $\int_{[1,\infty)} e^{-x}dx=\frac{1}{e}$.

רעיון דומה מאפשר להכליל את מושג האינטגרל למקרה שהאינטגרנד אינו מוגדר באחד מקצוות תחום האינטגרציה, כמו למשל הפונקציה $\frac{1}{\sqrt{x}}$ בקטע $(0,1]$:

הגדרה 9.8.2 תהי $f:(a,b]\to\mathbb{R}$ ונניח כי f אינטגרבילית בכל תת־קטע סגור של $(a,b]$. אם קיים הגבול

$$\lim_{t\to a+}\int_t^b f(x)dx$$

אז הוא נקרא **האינטגרל הלא אמתי** (improper intergral) של f ב־$(a,b]$, ומסומן $\int_{(a,b]} f$. במקרה זה אומרים שהאינטגרל $\int_{(a,b]} f$ מתכנס או ש־f **אינטגרבילית** ב־$(a,b]$. אם הגבול לא קיים נאמר ש־f אינה אינטגרבילית ב־$(a,b]$. האינטגרל הלא אמתי בקטע מהסוג $[a,b)$ מוגדר בצורה דומה.



איור 9.8.2 אינטגרל לא אמתי ב־$(a,b]$

דוגמאות

1. נברר האם $\int_{(0,1]} x^\alpha dx$ קיים. במקרה $\alpha=-1$ חישוב מראה ש־

$$\int_t^1 \frac{1}{x}dx = \ln x|_t^1 = -\ln t$$

מכיוון ש־ $\lim_{t\to 0+}\ln t = -\infty$ הרי שהאינטגרל לא קיים. עבור $\alpha\neq -1$ מתקיים

$$\lim_{t\to 0+}\int_t^1 x^\alpha dx = \lim_{t\to 0+}\left(\frac{1}{\alpha+1}-\frac{1}{\alpha+1}t^{\alpha+1}\right) = \begin{cases} \infty & \alpha<-1 \\ \frac{1}{\alpha+1} & \alpha>-1 \end{cases}$$

ולכן $\int_{(0,1]} x^\alpha dx$ קיים אמ"מ $\alpha>-1$, ובמקרה זה האינטגרל שווה ל־ $\frac{1}{\alpha+1}$.

שימו לב שהתלות בפרמטר α הפוכה מהתלות בטענה המקבילה לגבי הקרן $[1,\infty)$. השווה עם דוגמה (1) בעמוד 438.

2. אם f אינטגרבילית בקטע $[a,b]$ אז גם האינטגרל הלא אמתי שלה קיים ב־ $[a,b)$, ומתקיים $\int_{[a,b)} f = \int_{[a,b]} f$. זו מסקנה ממשפט 9.6.5, שכן לפי וההנחה ש־$\int_{[a,b]} f$ קיים נובע מהמשפט ש־

$$\lim_{t\to b-}\int_a^t f = \int_a^b f$$

וזו בדיוק הטענה.

3. הפונקציה $e^{-1/x}$ אינה מוגדרת ב־0 אבל האינטגרל הלא אמתי $\int_{(0,1]} e^{-\frac{1}{x}}dx$ קיים. שכן נגדיר פונקציה $f:[0,1]\to\mathbb{R}$ על־ידי $f(x)=e^{-1/x}$ אם $x\neq 0$, ו־ $f(0)=0$. אז f רציפה בקטע $[0,1]$ (בדקו!) ולכן אינטגרבילית שם. מכאן ש־ f אינטגרבילית גם ב־$(0,1]$. אבל בקטע $(0,1]$ הפונקציה f אינה אלא הפונקציה $e^{-1/x}$, ולכן האינטגרל הלא אמתי $\int_{(0,1]} e^{-1/x}dx$ קיים.

התכונות הבסיסיות של אינטגרל לא אמתי דומות לאלו של האינטגרל הרגיל. נסכם את עיקרן בטענה הבאה:

משפט 9.8.3 (תכונות של אינטגרלים לא אמתיים) יהיו f,g מוגדרות בקטע $[a,b)$, כאשר ייתכן ש־ $b=\infty$. אזי

1. יהי $c\in[a,b)$. אז האינטגרל של f קיים ב־$[a,b)$ אמ"מ הוא קיים ב־$[a,c]$ וב־$[c,b)$, ובמקרה זה מתקיים $\int_{[a,b)} f = \int_{[a,c]} f + \int_{[c,b)} f$ (שימו לב שהאינטגרל האמצעי הוא אינטגרל אמתי).

2. לכל קבוע $c\in\mathbb{R}$, אם f אינטגרבילית ב־$[a,b)$ אז cf אינטגרבילית ב־$[a,b)$ ומתקיים $\int_{[a,b)} cf = c\int_{[a,b)} f$

3. אם f,g אינטגרביליות ב־$[a,b)$ אז $f+g$ אינטגרבילית ב־$[a,b)$, ומתקיים $\int_{[a,b)}(f+g)=\int_{[a,b)}f+\int_{[a,b)}g$
4. אם f,g אינטגרביליות ב־$[a,b)$ ואם $f\le g$ אז $\int_{[a,b)}f\le\int_{[a,b)}g$. בפרט אם $f\ge 0$ אז $\int_{[a,b)}f\ge 0$.
5. אם $f,|f|$ אינטגרביליות ב־ $[a,b)$ אז $|\int_{[a,b)}f|\le\int_{[a,b)}|f|$.

הוכחה כל ההוכחות נובעות מהתכונות המקבילות לאינטגרלים אמתיים. נוכיח כאן כמה סעיפים, ואת היתר תוכלו להשלים בעצמכם.

נוכיח את (1). נניח שהאינטגרל הלא־אמתי של f קיים ב־$[a,b)$ ויהי $c\in[a,b)$. אינטגרביליות במובן הרחב של f ב־$[a,b)$ גוררת ש־f אינטגרבילית בכל תת־קטע סגור של $[a,b]$ ובפרט ב־$[a,c]$. כמו־כן לכל $t\in[a,b)$ מתקיים $\int_a^t f=\int_a^c f+\int_c^t f$ ולכן

$$\lim_{t\to b-}\int_c^t f=\lim_{t\to b-}(\int_a^t f-\int_a^c f)=(\lim_{t\to b-}\int_a^t f)-\int_a^c f=\int_{[a,b)}f-\int_{[a,c]}f$$

ובפרט הגבול קיים. נעביר אגפים ונקבל את השוויון הדרוש. ההוכחה של הכיוון השני דומה (השלימו את הפרטים!).

נוכיח את (3). לכל t מתקיים $\int_a^t(f+g)=\int_a^t f+\int_a^t g$ לפי התכונות של האינטגרל האמתי. לכן

$$\lim_{t\to b-}\int_a^t(f+g)=\lim_{t\to b-}\int_a^t f+\lim_{t\to b-}\int_a^t g=\int_{[a,b)}f+\int_{[a,b)}g$$

ולכן $f+g$ אינטגרבילית ב־$[a,b)$ ומתקיים $\int_{[a,b)}(f+g)=\int_{[a,b)}f+\int_{[a,b)}g$, כפי שרצינו. ∎

הערות

1. אפשר לנסח משפטי אריתמטיקה לאינטגרלים לא אמתיים המתכנסים במובן הרחב. הפרטים מושארים כתרגיל.
2. במשפט נעדר כלל המכפלה. ואמנם, קיום האינטגרל במובן הרחב של f,g ב־$[a,\infty)$ אינו גורר באופן כללי שהאינטגרל במובן הרחב של $f\cdot g$ קיים שם, אף שעל פי התכונות של האינטגרל האמתי, $f\cdot g$ אינטגרבילית בכל תת־קטע חסום של $[a,\infty)$. למשל, $\int_0^1\frac{1}{\sqrt{x}}dx$ קיים אבל $\int_0^1(\frac{1}{\sqrt{x}})^2dx=\int_0^1\frac{1}{x}dx$ לא קיים. נראה דוגמאות נוספות בתרגילים בסוף הסעיף.
3. בסעיף האחרון של המשפט דרשנו במפורש שהאינטגרל הלא אמתי של $|f|$ יהיה קיים. אי אפשר לוותר על הדרישה הזו: היא אינה נובעת מההנחה בדבר קיום האינטגרל הלא אמתי של f.

נותר להגדיר אינטגרל לא אמתי בקטעים הפתוחים בשני הקצוות.

הגדרה 9.8.4 תהי f פונקציה ממשית המוגדרת בקטע (a,b), כאשר ייתכן $a=-\infty$ או $b=\infty$. יהי $c\in(a,b)$. אם האינטגרלים הלא אמתיים של f קיימים בקטעים $(a,c]$ ו־$[c,b)$ אז נאמר שהאינטגרל הלא אמתי של f קיים ב־(a,b), והוא מוגדר על ידי $\int_{(a,b)} f = \int_{(a,c]} f + \int_{[c,b)} f$.

יש לבדוק שההגדרה אינה תלויה בבחירה של הנקודה c. ואמנם,

טענה 9.8.5 תהי f פונקציה ממשית על (a,b) (ייתכן $a=-\infty$ או $b=\infty$). יהיו $c<d$ נקודות בקטע. אז האינטגרלים של f ב־$(a,c],[c,b)$ קיימים אמ"מ האינטגרלים של f ב־$(a,d],[d,b)$ קיימים, ובמקרה זה מתקיים

$$\int_{(a,c]} f + \int_{[c,b)} f = \int_{(a,d]} f + \int_{[d,b)} f$$

בפרט, האינטגרל הלא אמתי בהגדרה 9.8.4 אינו תלוי בבחירה של נקודת הביניים c ולכן ההגדרה חד משמעית.

הוכחה נניח אינטגרביליות ב־ $(a,c],[c,b)$. לפי משפט 9.8.3 אינטגרביליות ב־ $[c,b)$ גוררת אינטגרביליות ב־ $[c,d]$ וב־ $[d,b)$, אינטגרביליות ב־ $(a,c],[c,d]$ גורר אינטגרביליות ב־ $(a.d]$, ומתקיים

$$\int_{(a,d]} f = \int_{(a,c]} f + \int_{[c,d]} f \qquad , \qquad \int_{[c,b)} f = \int_{[c,d]} f + \int_{[d,b)} f$$

ולכן

$$\begin{aligned}
\int_{(a,c]} f + \int_{[c,b)} f &= \int_{(a,c]} f + (\int_{[c,d]} f + \int_{[d,b)} f) \\
&= (\int_{(a,c]} f + \int_{[c,d]} f) + \int_{[d,b)} f \\
&= \int_{(a,d]} f + \int_{[d,b)} f
\end{aligned}$$

כנדרש. ■

דוגמאות

1. נחשב את $\int_{-\infty}^{\infty} e^{-|x|}dx$. חישוב מראה ש־

$$\int_0^\infty e^{-|x|}dx = 1 \qquad , \qquad \int_{-\infty}^0 e^{-|x|} = 1$$

(הוכיחו!) ולכן

$$\int_{-\infty}^{\infty} e^{-|x|}dx = 1 + 1 = 2$$

2. נראה ש־ $\int_0^\infty x^\alpha$ אינו קיים לאף מספר α. ואמנם, אם $\alpha \geq -1$ אז $\int_1^\infty x^\alpha dx$ אינו קיים, ואילו אם $\alpha \leq -1$ אז $\int_0^1 x^\alpha dx$ אינו קיים.

היה אפשר לחשוב שטבעי יותר להגדיר את האינטגרל הלא אמתי בקטעים מהסוג $(-\infty,\infty)$ על־ידי גבול מהסוג $\lim_{t\to\infty}\int_{-t}^{t} f$. ואולם, זו אינה הגדרה מוצלחת. בביטוי $\lim_{t\to\infty}\int_{-t}^{t} f$ יש בחירה נסתרת של נקודה מיוחדת: בחרנו ב־0 כמרכז תחום האינטגרציה $[-t,t]$, והבחירה אינה תמימה. היינו מצפים, למשל, שנקבל את אותו ערך אם במקום לבחור את 0 כמרכז תחום האינטגרציה נבחר את 1, כלומר דהיינו מצפים שיתקיים $\lim_{t\to\infty}\int_{-t}^{t} f = \lim_{t\to\infty}\int_{1-t}^{1+t} f$. אולם לא חייב להיות כאן שוויון. למשל, עבור $f = \operatorname{sgn}$ מתקיים $\int_{-t}^{t}\operatorname{sgn} = 0$ לכל t, ולכן $\lim_{t\to\infty}\int_{-t}^{t}\operatorname{sgn} = 0$, ומאידך $\int_{1-t}^{1+t}\operatorname{sgn} = 2$ ולכן $\lim_{t\to\infty}\int_{1-t}^{1+t}\operatorname{sgn} = 2$. נציין שלפי הגדרה 9.8.4, האינטגרל של sgn כלל אינו קיים ב־$(-\infty,\infty)$ (הוכיחו!).

תכונות האינטגרל הלא אמתי בקטע (a,b) דומות לתכונות בקטע חצי פתוח כפי שנוסחו במשפט 9.8.3. כמו־כן I הוא אחד הקטעים $[a,b)$, $(a,b]$ או $[a,b]$ ואם f אינטגרבילית ב־I אז היא אינטגרבילית ב־(a,b) ו־ $\int_{(a,b)} f = \int_I f$.

סיימנו להגדיר את האינטגרל על כל סוגי הקטעים השונים, וממה שראינו נובע שאם האינטגרל של פונקציה קיים בקטע $[a,b]$ אז קיימים גם האינטגרלים הלא אמתיים ב־$[a,b)$, $(a,b]$ ו־(a,b), וכולם שווים. לכן לא ייגרם בלבול אם נשוב ונסמן ב־ $\int_a^b f$ את האינטגרל של f בקטעים אלה מבלי להבחין ביניהם, ונתכוון לאינטגרל האמתי אם הוא קיים, ולאינטגרל הלא אמתי אחרת.

אפשר להכליל את מושג האינטגרל מעט יותר:

הגדרה 9.8.6 תהי f פונקציה ממשית המוגדרת בתחום D ונניח שאפשר להציג את D כאיחוד של קטעים $I_1,\ldots,I_n$, וכל שניים מהקטעים זרים. אם האינטגרל (האמתי או הלא אמתי) של f קיים בכל אחד מהקטעים נאמר שהאינטגרל של f קיים ב־I, ונסמן $\int_D f = \sum_{i=1}^n \int_{I_i} f$.

אנו משאירים כתרגיל את הבדיקה שהאינטגרל שהגדרנו אינו תלוי באופן שבו אנו מחלקים את D לקטעים ושהוא ומתלכד עם ההגדרות הקודמות במקרה ש־D הוא קטע (ראו תרגיל (5) להלן).

לעתים קרובות קל יותר להוכיח שקיים אינטגרל לא אמתי מאשר לחשב את ערכו: מצב עניינים זה דומה במידה רבה לעובדה שקל יותר להוכיח שטור מספרים מתכנס מאשר לחשב את ערכו. בהמשך הסעיף נפגוש מבחני התכנסות כאלה. רובם דומים למשפטי התכנסות שפגשנו בפרק על טורים. אנו נתעניין כאן באינטגרלים מהסוג

$\int_a^\infty f$, אבל אפשר לנסח גרסאות של רוב התוצאות לאינטגרלים מהסוג $\int_{-\infty}^a f$ או לאינטגרלים לא אמתיים מהסוג $\int_{[a,b)} f$ כאשר b ממשי.

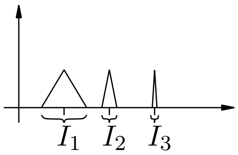
נתחיל בשאלה, האם קיום האינטגרל של f ב־ $[a,\infty)$ גורר ש־ $\lim_{x\to\infty} f(x) = 0$? כזכור, אחת התכונות הבסיסיות של טור מתכנס $\sum a_n$ היא ש־ $a_n \to 0$, וזה מוביל אותנו לשער שכך גם לפונקציות. אולם אין זה המצב.

דוגמאות

1. תהי

$$f(x) = \begin{cases} 0 & x \notin \mathbb{N} \\ 1 & x \in \mathbb{N} \end{cases}$$

ברור ש־ $\lim_{x\to\infty} f(x) \neq 0$. מצד שני לכל t מתקיים $\int_0^t f = 0$, ומכאן $\int_0^\infty f = 0$ ובפרט קיים.



איור 9.8.3

2. בדוגמה הקודמת האינטגרנד לא היה רציף, אבל לא קשה לבנות גם דוגמאות כאלה עם אינטגרנד רציף. הנה דרך אחת. נבחר קטעים $I_1, I_2, \ldots$ כאשר I_n באורך $\frac{1}{n^2}$ ומרכזו n (כלומר $I_n = [n - \frac{1}{2n^2}, n + \frac{1}{2n^2}]$) נגדיר את f להיות 0 אם x אינו שייך לאף אחד מהקטעים I_n, ובקטע I_n נגדיר את f להיות בעל גרף משולשי שגובהו 0 בקצוות של I_n ו־1 במרכז (אם תרצו תוכלו להגדיר את f במפורש).

מצד אחד, ל־f אין גבול ב־∞ כי הסדרה $(f(n))_{n=1}^\infty$ היא הסדרה הקבועה שערכה 1, ואילו $(f(n+\frac{1}{2}))_{n=1}^\infty$ היא הסדרה הקבועה 0.

מצד שני, השטח שמתחת לגרף של f בקטע I_n הוא $\frac{1}{2n^2}$ (כי הגרף הוא משולש שגובהו 1 ובסיסו $\frac{1}{n^2}$), ומכיוון ש־f מתאפסת חוץ מאשר בקטעים I_n, נקבל

$$\int_0^{n+\frac{1}{2}} f = \sum_{k=1}^n \frac{1}{2k^2}$$

נסמן $F(t) = \int_0^t f$. מכיוון ש־f אי־שלילית הרי ש־F עולה, ולכן הגבול $\lim_{t\to\infty} F(t)$ קיים במובן הרחב ולכל סדרת מספרים $t_n \to \infty$ מתקיים $\lim_{n\to\infty} F(t_n) = \lim_{t\to\infty} F(t)$. בפרט אם נתבונן בסדרה $t_n = n + \frac{1}{2}$ נקבל לפי האמור לעיל ש־

$$\lim_{n\to\infty} F(t_n) = \lim_{n\to\infty} \sum_{k=1}^n \frac{1}{2k^2} = \sum_{k=1}^\infty \frac{1}{2k^2}$$

והטור באגף ימין מתכנס. מכאן שהאינטגרל קיים.

אם כן, שאיפת האינטגרנד לאפס אינה תנאי הכרחי לקיום אינטגרל לא אמתי בקרן. אולם אם $\lim_{x\to\infty} f(x)$ בכל־זאת קיים, אז על מנת שהאינטגרל של f בקרן $[a,\infty)$ יהיה קיים, הגבול חייב להיות אפס:

טענה 9.8.7 תהי $f:[a,\infty)\to\mathbb{R}$. אם $\lim_{x\to\infty} f(x) = L \neq 0$ אז f אינה אינטגרבילית במובן הרחב ב־$[a,\infty)$.

הוכחה נניח למשל $L>0$. משמע שקיים t_0 כך שבקרן $[t_0,\infty)$ הפונקציה נמצאת כולה מעל $\frac{L}{2}$. לכן לכל $t>t_0$

$$\int_{t_0}^{t} f \geq \frac{L}{2}(t-t_0) \xrightarrow{t\to\infty} \infty$$

לכן $\int_{t_0}^{\infty} f$ אינו קיים, ולכן גם $\int_a^{\infty} f$ אינו קיים. ■

נפנה כעת למבחני התכנסות.

משפט 9.8.8 (תנאי קושי) נניח ש־ $f:[a.\infty)\to\mathbb{R}$ אינטגרבילית בכל תת־קטע סגור של $[a,\infty)$. אז האינטגרל $\int_a^{\infty} f$ קיים אמ"מ לכל $\varepsilon>0$ קיים t_0 כך שלכל $x,y>t_0$ מתקיים $|\int_x^y f| \leq \varepsilon$.

הוכחה האינטגרל $\int_a^{\infty} f$ קיים אמ"מ הגבול $\lim_{t\to\infty}\int_a^t f$ קיים וסופי. לפי תנאי קושי עבור גבולות של פונקציות באינסוף, הגבול הזה קיים אמ"מ לכל $\varepsilon>0$ קיים t_0 כך שלכל $x,y>t_0$ מתקיים $|\int_a^y f - \int_a^x f| < \varepsilon$. כעת נותר רק לשים לב לשוויון $\int_a^y f - \int_a^x f = \int_x^y f$. ■

ישנם כמה מבחני השוואה לאינטגרלים לא אמתיים. כאן המצב דומה מאד למצב בתורת ההתכנסות של טורים. השוו את המשפט הבא למשפט 6.3.3.

משפט 9.8.9 (מבחן ההשוואה) נניח $f,g:[a,\infty)\to\mathbb{R}$ אינטגרביליות בכל תת־קטע סגור של $[a,\infty)$ ומקיימות $0\leq f\leq g$. אז

1. אם $\int_a^{\infty} g$ קיים אז כך גם $\int_a^{\infty} f$ קיים, ומתקיים $0\leq\int_a^{\infty} f\leq\int_a^{\infty} g$
2. אם $\int_a^{\infty} f$ לא קיים אז גם $\int_a^{\infty} g$ לא קיים.

הוכחה נסמן $F(t)=\int_a^t f$ ו־ $G(t)=\int_a^t g$. אז F,G עולות (כי $f,g\geq 0$) ולכן מתכנסות במובן הרחב כאשר $t\to\infty$. כעת מהאי־שוויון $f\leq g$ נובע $F\leq G$ ולכן אם $\lim_{t\to\infty} G(t)$ סופית אז $\lim_{t\to\infty} F(t)$ סופית, ואם $\lim_{t\to\infty} F(t)=\infty$ אז $\lim_{t\to\infty} G(t)=\infty$, כפי שרצינו להראות. ■

נראה למשל שהאינטגרל $\int_0^{\infty} e^{-x^2}dx$ קיים. כבר ציינו שלאינטגרנד אין פונקציה קדומה, ולכן לא נוכל להוכיח את קיום האינטגרל על־ידי חישוב ישיר, כפי שעשינו

בדוגמאות מהסעיף הקודם. במקום חישוב ישיר נשים לב שדי להראות שהאינטגרל $\int_1^\infty e^{-x^2}dx$ קיים. אבל עבור $x \geq 1$ מתקיים $x^2 \geq x$ ולכן

$$0 \leq e^{-x^2} \leq e^{-x}$$

ראינו בסעיף הקודם ש־e^{-x} אינטגרבילית במובן הרחב בקטע $[1,\infty)$. לכן ממבחן ההשוואה מקבלים ש־e^{-x^2} אינטגרבילית בקטע זה, וממילא היא אינטגרבילית גם ב־$[0,\infty)$.

כדי להשתמש במבחן ההשוואה לטורים שאינם חיוביים, נזדקק למושג של התכנסות בהחלט (השוו עם מושג ההתכנסות בהחלט לטורים בהגדרה 6.4.1 ובמשפט 6.4.2):

הגדרה 9.8.10 פונקציה $f:[a,\infty)\to\mathbb{R}$ נקראת **אינטגרבילית בהחלט** (absolutely integrable) בקטע $[a,\infty)$ אם $|f|$ אינטגרבילית שם. במקרה זה נאמר שהאינטגרל $\int_a^\infty f$ מתכנס בהחלט.

משפט 9.8.11 אם f אינטגרבילית בהחלט ב־$[a,\infty)$ אז היא אינטגרבילית שם.

ההוכחה דומה להוכחת הטענה המקבילה לטורים, ומושארת כתרגיל.

כדוגמה לשימוש במשפט, יהי $f(x)=\frac{\cos x}{x^2}$. עבור $x\geq 1$ מתקיים $|f(x)|\leq\frac{1}{x^2}$ ולכן f אינטגרבילית בהחלט ב־ $[1,\infty)$ ובפרט היא אינטגרבילית שם.

קיימים גם מבחני התכנסות לפונקציות שהאינטגרל שלהן מתכנס בתנאי. ראו תרגיל (11) בסוף הסעיף.

תרגילים

1. אילו מהאינטגרלים הבאים מתכנסים?
   (א) $\int_1^\infty \frac{1}{1+x^2}dx$.
   (ב) $\int_1^\infty \frac{1}{1+x}dx$.
   (ג) $\int_1^\infty \sin(x)dx$.
   (ד) $\int_0^1 \ln x dx$.
   (ה) $\int_0^1 \frac{e^x}{e^x-1}dx$.
   (ו) $\int_0^1 \frac{e^x}{(e^x-1)^2}dx$.
   (ז) $\int_0^1 \frac{e^x}{(e^x-1)^{1/2}}dx$.
2. הוכיחו או הפריכו את הטענות הבאות:
   (א) אם f מחזורית ורציפה אז $\int_0^\infty f$ קיים.
   (ב) אם $\int_{-\infty}^\infty f(x)dx=0$ אז יש נקודה x_0 כך ש־ $\int_{x_0-x}^{x_0+x} f(t)dt=0$ לכל x גדול מספיק.

(ג) אם $\int_{-\infty}^{\infty} f(x)dx = 0$ אז קיים x_0 כך ש־ $\int_{-\infty}^{x_0} f(t)dt = -\int_{x_0}^{\infty} f(t)dt$.

3. הראו שאם האינטגרל הלא אמתי של f ב־$(a, b]$ קיים ואם f חסומה ב־$(a, b]$ אז f אינטגרבילית ב־$[a, b]$ במובן הרגיל (שימו לב: f אינה מוגדרת בהכרח ב־a. הטענה היא שכל הגדרה שלה ב־a תהפוך אותה לאינטגרבילית. ממילא שינוי בנקודה אחת אינו משנה את ערך האינטגרל).

4. הוכיחו את הגרסה הבאה של המשפט היסודי: אם $\int_a^{\infty} f$ קיים אז הפונקציה $F : [a, \infty) \to \mathbb{R}$ המוגדרת על ידי $F(t) = \int_t^{\infty} f$ רציפה ב־$[a, b]$, ואם f רציפה אז F גזירה שם.

5. הוכיחו שהגדרה 9.8.6 היא חד משמעית.

6. אילו מהאינטגרלים הבאים מתכנסים ואילו מתכנסים בהחלט?

(א) $\int_0^{\infty} \frac{1}{1+x^{\alpha}} dx$ (α קבוע חיובי).

(ב) $\int_1^{\infty} e^{-x} \ln x dx$.

(ג) $\int_1^{\infty} \frac{\ln x}{x^2} dx$.

(ד) $\int_0^{\infty} \frac{\sin(x)+1}{x} dx$.

(ה) $\int_0^{\infty} \frac{(1-\{x\}/2)}{x} dx$.

7. תהי $f : \mathbb{R} \to \mathbb{R}$. הוכיחו או הפריכו:

(א) אם $\int_0^{\infty} f$ מתכנס אז קיימת סדרה (x_n) עם $x_n \to \infty$ ו־ $f(x_n) \to 0$.

(ב) אם $\int_0^{\infty} f$ אינו מתכנס אז לכל סדרת נקודות (x_n) עם $x_n \to \infty$ מתקיים $f(x_n) \not\to 0$.

(ג) ייתכן ש־f רציפה והאינטגרל $\int_0^{\infty} f$ מתכנס, אבל קיימת סדרה (x_n) עם $x_n \to \infty$ ו־ $\lim_{n\to\infty} f(x_n) = \infty$.

(ד) אם $\int_0^{\infty} f(x)dx$ מתכנס ו־f מונוטונית אז $\lim_{x\to\infty} f(x) = 0$.

(ה) אם $\int_0^{\infty} f$ מתכנס ואם f גזירה אז $\lim_{x\to\infty} f'(x) = 0$.

(ו) אם f גזירה ואם $\lim_{x\to\infty} f'(x) = M \neq \infty$ אז $\int_0^{\infty} f$ לא מתכנס.

8. תהי $(a_n)_{n=1}^{\infty}$ סדרה. נגדיר פונקציה $\widehat{a} : [0, \infty) \to \mathbb{R}$ על־ידי

$$\widehat{a}(x) = a_n \iff x \in [n-1, n)$$

הוכיחו ש־$\sum_{n=1}^{\infty} a_n$ מתכנס אמ"מ $\int_0^{\infty} f$ קיים.

9. הוכיחו את משפט 9.8.10: אם f אינטגרבילית בהחלט בקטע כלשהו אז היא אינטגרבילית שם.

10. תהי f פונקציה רציפה ומחזורית ב־$\mathbb{R}$. נניח ש־f אי־שלילית ואינה זהותית 0. הוכיחו ש־ $\int_1^{\infty} \frac{f(x)}{x} dx$ אינו קיים.

11. בשאלה זו נוכיח מבחן התכנסות לאינטגרלים המקביל למשפט דירכלה 6.4.10 על טורים, ונראה כמה שימושים שלו.

(א) יהיו $f,g:[a,\infty)\to\mathbb{R}$ פונקציות רציפות ו־f גזירה. נניח ש־f יורדת ומקיימת $\lim_{x\to\infty}f(x)=0$, ונניח שהפונקציה $G(t)=\int_a^t g$ חסומה ב־$[a,\infty)$. הראו שהאינטגרל $\int_a^\infty f\cdot g$ קיים (רמז: תהי $H(t)=\int_a^t fg$. הראו ש־H מקיימת את תנאי קושי באינסוף ולכן $\lim_{t\to\infty}H(t)$ קיים. לשם כך העריכו את $\int_s^t fg$ על ידי אינטגרציה בחלקים, ושימו לב ש־$|f'|=-f'$ כי f יורדת. היעזרו במשפט היסודי).

(ב) הראו שהאינטגרלים $\int_1^\infty \frac{\cos x}{\sqrt{x}}dx$ ו־$\int_1^\infty \frac{\cos x}{x}dx$ קיימים אבל הם אינם מתכנסים בהחלט. הסיקו שמכפלה של פונקציות אינטגרביליות במובן הרחב לא חייבת להיות אינטגרבילית בעצמה.

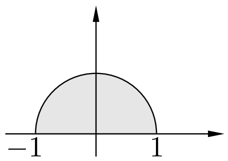
(ג) הראו שייתכן ש־$\int_1^\infty f$ קיים אבל $\int_1^\infty |f|$ אינו קיים (כלומר שהאינטגרל קיים אך לא מתכנס בהחלט).

9.9 שימושים של האינטגרל

לאינטגרל שימושים רבים. בסעיף זה נצביע על כמה מהם.

חישוב שטחים

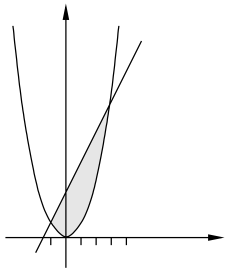
המוטיבציה בהגדרת האינטגרל הייתה לחשב שטחים, והגיע הזמן לתת דוגמה לחישוב כזה. נחשב למשל שטח של עיגול. ליתר דיוק, נתבונן בחצי העליון של מעגל היחידה, שנתון על ידי הגרף של הפונקציה $f(x)=\sqrt{1-x^2}$ בקטע $[-1,1]$. השטח בינו לבין ציר ה־x הוא השטח S של חצי עיגול מרדיוס 1. לכן



איור 9.9.1

$$S=\int_{-1}^{1}\sqrt{1-x^2}dx=(\frac{1}{2}\arcsin x+\frac{x}{2}\sqrt{1-x^2})|_{-1}^{1}=\frac{\pi}{2}$$

ולכן השטח של עיגול היחידה הוא π. שימו לב ־ זו עובדה לא טריביאלית, שכן π הוגדר בתור האורך של חצי מעגל, וללא קשר ישיר עם השטח שלו!


איור 9.9.2

הנה דוגמה נוספת: נחשב את השטח S של הגזרה הכלואה בין הישר $y=2x+3$ והפונקציה $y=x^2$, כמתואר באיור 9.9.2. ראשית נמצא את נקודות החיתוך של הגרפים: אלה הנקודות x המקיימות $2x+3=x^2$, דהיינו הנקודות $x=-1,3$. השטח שמתחת לישר בקטע זה הוא $\int_{-1}^{3}(2x+3)dx$ וממנו יש להחסיר את השטח $\int_{-1}^{3}x^2dx$ שמתחת לפרבולה בגזרה (כך נקבל את השטח שביניהם). לכן התשובה היא

$$S=\int_{-1}^{3}(2x+3-x^2)=(x^2+3x-\frac{1}{3}x^3)|_{-1}^{3}=5\frac{1}{3}$$

נפח של גוף סיבוב

גוף סיבוב הוא גוף תלת־ממדי שיש לו סימטריה סיבובית, דהיינו הוא סימטרי לסיבובים דרך ישר כלשהו במרחב. דרך אחת לקבל גוף כזה הוא לסובב גוף דו־ממדי סביב ציר במרחב. מטרתנו היא לתאר את הנפח של גוף כזה בעזרת אינטגרל.

לצורך הדיון נזהה כרגיל את המרחב התלת־ממדי עם הקבוצה

$$\mathbb{R}^3 = \{(x,y,z) \, : \, x,y,z \in \mathbb{R}\}$$

נתבונן למשל בכדור $B \subseteq \mathbb{R}^3$ בעל רדיוס 1, שמרכזו בראשית $(0,0,0)$:

$$B = \{(x,y,z) \in \mathbb{R}^3 \, : \, x^2+y^2+z^2 \leq 1\}$$

נסמן את הנפח של B ב־$v(B)$. לא הגדרנו מהו נפח אך עוד מעט נקבל ביטוי אשר יכול להיחשב הגדרה של $v(B)$.

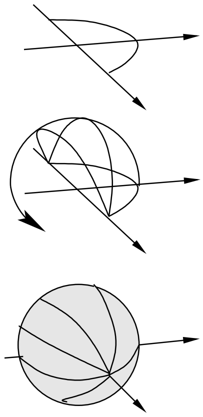
B ניתן לתיאור באופן הבא: נקבע במישור הדו־ממדי חצי עיגול עליון

$$C_0 = \{(x,y) \in \mathbb{R}^2 \, : \, y \geq 0 \, , \, x^2+y^2 \leq 1\}$$

ונזהה אותו עם קבוצת הנקודות במרחב

$$C = \{(x,y,0) \, : \, y \geq 0 \, , \, x^2+y^2 \leq 1\}$$

שמתקבל מזיהוי המישור הדו־ממדי עם המישור $\{(x,y,0) \, , \, x,y \in \mathbb{R}\}$ במרחב. אז הכדור B היא הקבוצה הנפרשת על ידי סיבוב של C סביב ציר ה־x (כלומר סביב הישר $\{(x,0,0) \, : \, x \in \mathbb{R}\}$), כמו באיור 9.9.3.



איור 9.9.3 ספירה מתקבלת כגוף סיבוב של חצי עיגול

אם נתבונן בקטע קצר $I = [a, a+h] \subseteq [-1,1]$ וב"פרוסה" של C שרכיב ה־x שלה שייך ל־I, כלומר ב

$$C_I = \{(x,y,z) \in \mathbb{R}^3 \, : \, x \in I \, , \, x^2+y^2+z^2 \leq 1\}$$

אז C_I הוא בקירוב גליל, שגבהו h ובסיסו מעגל בעל רדיוס $\sqrt{1-a^2}$. לכן הנפח של C_I הוא בקירוב $h \cdot \pi(1-a^2)$ (זו נוסחה יסודית מתורת הגאומטריה: הנפח של גליל שווה לשטח הבסיס כפול גובה הגליל). אם נחלק את $[-1,1]$ ל־$2n$ קטעים שווי אורך בעזרת החלוקה $P = \{-1, \frac{1-n}{n}, \frac{2-n}{n} \ldots, \frac{n-1}{n}, 1\}$, אנו מקבלים את הקירוב הבא לנפח של C: v_n

$$v_n = \sum_{k=0}^{2n-1} (C_{[\frac{k-n}{n}, \frac{k-n+1}{n}]} \text{ נפח מקורב של }) = \pi \sum_{k=0}^{2n-1} \frac{1}{n}(1-(\frac{k-n}{n})^2)$$

זהו סכום רימן של הפונקציה $1-x^2$, ולכן

$$\lim_{n\to\infty} v_n = \pi\int_{-1}^{1}(1-x^2) = \pi(x-\frac{1}{3}x^3)|_{-1}^{1} = \frac{4\pi}{3}$$

כאשר השוויון האמצעי נובע מהמשפט היסודי והעובדה ש־ $x+\frac{x^3}{3}$ היא פונקציה קדומה של $1-x^2$. לכן ניתן להסיק או להגדיר

$$v(B) = \lim_{n\to\infty} v_n = \frac{4\pi}{3}$$

אותו נימוק נותן את הטענה הכללית הבאה:

טענה 9.9.1 תהי $f:[a,b]\to\mathbb{R}$ פונקציה אי־שלילית, תהי D הגזרה במישור בין הקטע $[a,b]$ לגרף הפונקציה, ויהי R הגוף התלת־ממדי הכולל את כל הנקודות שפוגשות את D במהלך סיבובו סביב ציר ה־x, כלומר

$$R = \{(x,y,z)\in\mathbb{R}^3 \,:\, x\in[a,b]\,,\ \sqrt{y^2+z^2}\le f(x)\}$$

אז הנפח של R נתון על ידי

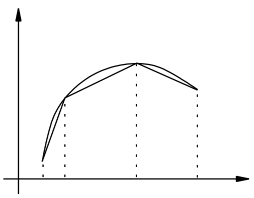
$$v(R) = \pi\int_a^b f^2(x)dx$$

ניתן לחלופין לראות זאת כהגדרה של הנפח R.

אורך של גרף

תהי f פונקציה ב־$[a,b]$. נניח שאנו רוצים לחשב את אורך הגרף שלה. אורך הוא מושג שלא הוגדר, אך נסכים בוודאי שהאורך של קטע במישור ניתן לחישוב בעזרת משפט פיתגורס: האורך של הקטע המחבר בין הנקודות (x_1,y_1) ו־(x_2,y_2) שווה ל־

$$\sqrt{(x_2-x_1)^2+(y_2-y_1)^2} = (x_2-x_1)\sqrt{1+(\frac{y_2-y_1}{x_2-x_1})^2}$$


איור 9.9.4 אורך הגרף שווה בקירוב לאורך המיתרים

אם $P=\{x_0,x_1,\dots,x_k\}$ היא חלוקה של $[a,b]$ אז היא קובעת סדרה של נקודות $(x_0,f(x_0)),(x_1,f(x_1)),\dots,(x_k,f(x_k))$ על גרף הפונקציה. אם נחבר במיתרים בין נקודות סמוכות, אז סכום אורכי המיתרים הללו הוא קירוב טבעי לאורך של הגרף. נגדיר אם כן את האורך המשוער $\ell(f,P)$ של f ביחס לחלוקה P על ידי

$$\ell(f,P) \;=\; \sum_{i=1}^{k}\sqrt{(x_i-x_{i-1})^2+(f(x_i)-f(x_{i-1}))^2}$$

ונגדיר את האורך של הגרף של f להיות ה"גבול" של $\ell(f,P)$ כאשר $\lambda(P)$ שואף ל־0, בתנאי שהוא קיים:

הגדרה 9.9.2 תהי f פונקציה ממשית ב־$[a,b]$. מספר ℓ_0 הוא **האורך של הגרף** של f בין a ל־b אם לכל $\varepsilon > 0$ קיים $\delta > 0$ כך שלכל חלוקה P של $[a,b]$ שמקיימת $\lambda(P) < \delta$ מתקיים $|\ell(f,P) - \ell_0| < \varepsilon$.

קל לראות שאם ℓ_0, ℓ_1 מקיימים את התנאי בהגדרה אז $\ell_0 = \ell_1$. אנו נסמן ב־$\ell(f)$ את האורך של הגרף, כאשר הוא קיים.

הגדרה זו אינה נותנת כלים לחשב את האורך של גרף ועבור פונקציה כללית לא ברור שהאורך קיים. אולם אם f היא פונקציה גזירה בעלת נגזרת f' רציפה אז אפשר לתת ביטוי פשוט לאורך של הגרף. נשים לב שאפשר לרשום את $\ell(f,P)$ גם כך:

$$\begin{aligned}
\ell(f,P) &= \sum_{i=1}^{k}(x_i - x_{i-1})\sqrt{1 + (\frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}})^2} \\
&= \sum_{i=1}^{k}\Delta x_i \cdot \sqrt{1 + (\frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}})^2}
\end{aligned}$$

לפי משפט לגרנג' יש נקודה $\xi_i \in [x_{i-1}, x_i]$ כך ש־

$$f'(\xi_i) = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$$

כך קיבלנו בחירת נקודות ξ עבור P שעבורה מתקיים

$$\ell(f,P) = \sum_{i=1}^{k}\Delta x_i \cdot \sqrt{1 + f'(\xi_i)^2} = \sigma(g,P,\xi)$$

כאשר g היא הפונקציה $g(t) = \sqrt{1 + f'(t)^2}$. הפונקציה g רציפה ב־$[a,b]$ כי f' רציפה שם, ולכן g אינטגרבילית ב־$[a,b]$, כלומר: לכל $\varepsilon > 0$ קיים $\delta > 0$ כך שאם P חלוקה של $[a,b]$ המקיימת $\lambda(P) < \delta$, ואם בוחרים את ξ כמו למעלה אז $|\sigma(g,P,\xi) - \int_a^b g| < \varepsilon$. בפרט אם P חלוקה כזאת אז מתקיים $|\ell(f,P) - \int_a^b g| < \varepsilon$. הוכחנו את התוצאה הבאה:

טענה 9.9.3 אם f גזירה ובעלת נגזרת רציפה ב־$[a,b]$ אז מתקיים

$$\ell(f) = \int_a^b \sqrt{1 + f'(t)^2}dt$$

למשל, תהי $f(x)=\sqrt{1-x^2}$ בקטע $[-1,1]$. הגרף של f מתאר חצי מעגל ברדיוס 1. הנגזרת של f היא הפונקציה

$$f'(x)=-\frac{x}{\sqrt{1-x^2}}$$

היא אינה מוגדרת ב־± 1, אך היא רציפה בכל תת־קטע סגור של $[-1,1]$. לכן אפשר לחשב את האינטגרל כאינטגרל לא אמתי:

$$\begin{aligned}
\ell(f) &= \int_{-1}^{1}\sqrt{1+f'(x)^2}dx \\
& \quad 2\int_0^1\sqrt{1+\frac{x^2}{1-x^2}}dx \\
&= 2\lim_{\delta\to 0+}\int_0^{1-\delta}\frac{dx}{\sqrt{1-x^2}} \\
&= 2\lim_{\delta\to 0+}(\arcsin(1-\delta)-\arcsin 0) \\
&= 2\cdot\frac{\pi}{2}
\end{aligned}$$

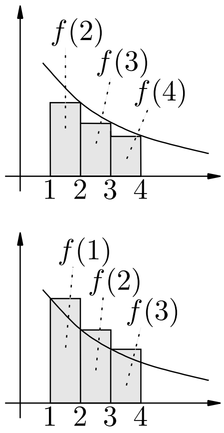
לכן האורך של חצי מעגל ברדיוס 1 הוא π.

שימו לב שבהגדרה הגיאומטרית של הפונקציות הטריגונומטריות, π הוגדר להיות אורך של חצי עיגול ברדיוס 1, ולכן אין הפתעה גדולה בחישוב הקודם. אולם חשוב להדגיש שזהו חישוב מדוייק שהתבסס רק על התכונות הפורמליות של הפונקציות הטריגונומטריות שהצגנו בעמוד 227 והכלים האחרים שפיתחנו. לכן התוצאה האחרונה מעניינת ולא טריביאלית, ומעידה שהתכונות הפורמליות של הפונקציות הטריגונומטריות אכן משקפות את הגיאומטריה באופן נאמן.

מבחן האינטגרל להתכנסות טורים

ראינו בסעיף 9.8 שיש הקבלה בין אינטגרלים מהסוג $\int_a^\infty f$ לטורי מספרים, וטבעי לשאול האם יש קשר הדוק יותר. למשל, האם יש קשר בין התנהגות הטור $\sum\frac{1}{n^2}$ לבין התנהגות האינטגרל $\int_1^\infty\frac{1}{x^2}dx$? האם מקרי ששניהם מתכנסים? בסעיף זה נראה שאין מדובר במקרה, ונראה כיצד להשתמש באינטגרלים כדי להכריע התכנסות של טורים מסוימים.

תהי f פונקציה המוגדרת בקטע $[n,n+1]$ למספר כלשהו $n\in\mathbb{N}$. יהי M חסם מלעיל ו־m חסם מלרע של f בקטע, כך ש־ $m\le\int_n^{n+1}f\le M$. אם f יורדת ב־$[n,n+1]$ אז הערך המקסימלי מתקבל בקצה השמאלי של הקטע, כלומר $M=f(n)$, ובאופן דומה הערך המינימלי מתקבל בקצה הימני של הקטע, דהיינו $m=f(n+1)$. לכן במקרה זה מתקיים $f(n+1)\le\int_n^{n+1}f\le f(n)$.



איור 9.9.5 השוואת האינטגרל של f עם סכום הערכים שלו בנקודות $1,2,3,\ldots$

אם f יורדת בקטע $[n,n+k]$ אז נוכל לפרק את $[n,n+k]$ ל־k קטעים

$$[n,n+1]\ ,\ [n+1,n+2]\ ,\ \ldots\ ,\ [n+k-1,n+k]$$

שכל אחד מהם באורך 1, ונקבל

$$\int_n^{n+k} f = \int_n^{n+1} f + \int_{n+1}^{n+2} f + \ldots + \int_{n+k-1}^{n+k} f$$

ולכן מהנחת המונוטוניות ולפי מה שראינו למעלה מתקיים

$$f(n+1)+f(n+2)+\ldots+f(n+k) \le \int_n^{n+k} f \le f(n)+f(n+1)+\ldots+f(n+k-1)$$

כך אנו מקבלים קשר בין אינטגרל של f על קטעים מהסוג $[n,n+k]$ לבין סכומים של הערכים של f בנקודות $n,n+1,\ldots$.

משפט 9.9.4 (מבחן האינטגרל להתכנסות טורים) תהי $f:[0,\infty)\to\mathbb{R}$ פונקציה מונוטונית. אז הטור $\sum_{k=1}^{\infty} f(k)$ מתכנס אמ"מ $\int_1^\infty f(x)dx$ קיים.

הוכחה נסמן $a_k=f(k)$, ונניח למשל ש־f יורדת. תחת הנחה זו קל לבדוק שאם f מקבלת ערכים שליליים אז הן הטור $\sum_{k=1}^{\infty} a_k$ והן האינטגרל $\int_1^\infty f$ מתבדרים ל־$-\infty$, ולכן נניח ש־ $f\ge 0$, ומכאן שגם $a_k\ge 0$.

הטור $\sum_{k=1}^{\infty} a_k$ הוא טור חיובי ולכן הוא מתכנס אמ"מ הוא טור חסום, דהיינו אם סדרת הסכומים החלקיים שלו חסומה. לפי הדיון הקודם מתקיים

$$\sum_{k=2}^{n} a_k \le \int_1^n f \le \sum_{k=1}^{n-1} a_k$$

ולכן סדרת הסכומים החלקיים של $\sum_{k=1}^{\infty} a_k$ חסומה אמ"מ הסדרה $(\int_1^n f)_{n=1}^{\infty}$ חסומה. מכיוון ש־ $f\ge 0$ הפונקציה $F(t)=\int_0^t f$ עולה, ולכן הסדרה $F(n)=\int_1^n f$ חסומה אמ"מ הפונקציה $F(t)$ חסומה, וזה שקול לקיום הגבול $\lim_{t\to\infty}F(t)$ (שוב, כי F עולה). קיום הגבול האחרון שקול לפי הגדרה לקיום האינטגרל $\int_1^\infty f$, כפי שרצינו. ■

הערות

1. כמובן שאין בנקודה 0 משהו מיוחד. באותו אופן מראים שאם f מונוטונית אז לכל n טבעי, $\int_n^\infty f$ קיים אמ"מ $\sum_{k=n}^{\infty} f(k)$ קיים.
2. הנחת המשפט לגבי המונוטוניות של f היא הכרחית. למשל, אם f היא הפונקציה שמקבלת את הערך 1 בשלמים ואת הערך $\frac{1}{n^2}$ בקטע $(n,n+1)$ אז הטור $\sum 1 = \sum_{n=1}^{\infty} f(n)$ מתבדר, אבל $\int_1^\infty f = \sum_{n=1}^{\infty}\frac{1}{n^2}$ מתכנס (מדוע השוויון נכון?). אפשר למצוא דוגמאות כאלה גם כאשר האינטגרנד רציף.

מבחן האינטגרל הוא כלי חזק מאד להוכחת התכנסות והתבדרות של אינטגרלים. למשל, יהי $\alpha \in \mathbb{R}$ ונתבונן בטור $\sum_{n=1}^{\infty} n^{\alpha}$. אנו יודעים משיקולים אלמנטריים שאם $\alpha \leq -2$ הטור מתכנס, ואם $\alpha \geq -1$ הוא מתבדר. המקרה $\alpha \in (-2,-1)$ הוכרע רק בשיטות מתוחכמות יותר כמו מבחן העיבוי. הנה דרך קלה: נשים לב שעבור $f(x)=x^{-\alpha}$ מתקיים $f(n)=n^{-\alpha}$, ומכיוון ש־f מונוטונית יורדת, התכנסות הטור $\sum_{n=1}^{\infty} n^{\alpha}$ שקולה להתכנסות האינטגרל $\int_{1}^{\infty} x^{\alpha}$, אבל ראינו כבר על־ידי חישוב ישיר שאינטגרל זה מתכנס אמ"מ $\alpha < -1$. לכן הוכחנו ש־ $\sum_{n=1}^{\infty} n^{\alpha}$ מתכנס אמ"מ $\alpha < -1$.

תרגילים

1. חשבו את השטחים הבאים:
   (א) השטח שמתחת לגרף e^x בקטע $[0,1]$.
   (ב) השטח שבין הגרפים של e^x ו־ $-x$ בקטע $[0,1]$.
   (ג) השטח שבין הגרפים x ו־x^2 בקטע שבין נקודות המפגש שלהם.
   (ד) השטח של הגזרה המוגדרת על ידי הגרף של x^2 והגרפים של $-\frac{x}{3}$ ו־ $x^4-\frac{1}{3}$.
2. חשבו את הנפחים הבאים:
   (א) נפח גוף הסיבוב המתקבל מ־ $f(x)=x^2-1$ בקטע $[-1,1]$.
   (ב) נפח גוף הסיבוב המתקבל מ־ $f(x)=|x|-1$ בקטע $[-1,1]$.
   (ג) נפח גוף הסיבוב המתקבל מהגזרה שבין הפונקציה $f(x)=x^3$ לבין $g(x)=x^2$ בקטע $[1,2]$.
   (ד) נפח גוף הסיבוב המתקבל מ־ $f(x)=e^x$ בקטע $[-1,1]$.
3. חשבו את האורכים של העקומות הבאות:
   (א) $f(x)=\sqrt{1-x^2}$ בקטע $[\arccos\theta,1]$ כאשר $-1 \leq \theta \leq 1$ קבוע.
   (ב) $f(x)=\sqrt{1+x^2}$ בקטע $[0,1]$.
   (ג) $f(x)=\ln x$ בקטע $[1,2]$.
4. יהיו $x,y:[a,b]\to\mathbb{R}$ פונקציות גזירות ברציפות. מצאו ביטוי לאורך של העקומה $\{(x(t),y(t)) : t\in[a,b]\}$ (הפעילו את אותו ההיגיון שהפעלנו קודם כדי לחשב את האורך של גרף. תוכלו גם להיעזר בדיון בעמוד 335 ובמשפט ערך הביניים של קושי, משפט 8.5.10).

   בעזרת בנוסחה שקבלתם חשבו את האורכים של המסילות הבאות:
   (א) $(x(t),y(t))=(\cos t,\sin t)$ בקטע $[0,2\pi]$.
   (ב) $(x(t),y(t))=(\cos^3 t,\sin^3 t)$בקטע $[0,1]$.
   (ג) $(x(t),y(t))=(e^t,e^t)$ בקטע $[1,2]$.

5. הראו שלגרף של פונקציה מונוטונית תמיד קיים אורך. היעזרו בכך כדי לתת הצדקה פורמלית לנוסחה שקיבלנו בשאלה 12 בעמוד 230.

6. בהנתן פונקציה ממשית f על $[a,b]$, נאמר שהאורך של הגרף שלה הוא אינסוף אם M יש $\delta > 0$ כך שלכל חלוקה P עם $\lambda(P) < \delta$ מתקיים $\ell(f,P) > M$.

(א) הראו שלפונקציה $f : [0,1] \to \mathbb{R}$, המוגדרת על ידי $f(0) = 0$ ועל ידי $f(x) = \sin\frac{1}{x}$ עבור $x \in (0,1]$, יש אורך אינסופי.

(ב) (*) הראו שהאורך של גרף של פונקציה בקטע הוא תמיד סופי או אורך אינסופי.

7. (*) בשאלה זו נראה שהעקומה הקצרה ביותר בין שתי נקודות היא קו ישר. הראו שאם $f : [0,1] \to \mathbb{R}$ היא פונקציה כך ש־ $f(0) = 0$ ו־ $f(1) = 1$, ואם ℓ הוא האורך שלה במובן של הגדרה 9.9.2, אז $\ell \geq \sqrt{2}$, וייתכן שוויון רק אם $f(x) = x$.

8. הכריעו בעזרת מבחן האינטגרל אילו מהטורים הבאים מתכנסים (בכל מקרה הצדיקו מדוע אפשר להשתמש במבחן זה):

(א) $\sum_{n=1}^{\infty} ne^{-n}$.

(ב) $\sum_{n=2}^{\infty} \frac{1}{n\ln^{\alpha} n}$ (α קבוע חיובי).

(ג) $\sum_{n=1}^{\infty} \ln(1+n^2)$.

9. הוכיחו שאם f יורדת ממש והאינטגרל $\int_0^{\infty} f$ קיים אז $\sum_{n=1}^{\infty} f(n) \neq \int_0^{\infty} f$, אבל שקיים מספר $a \in (0,1)$ כך ש־ $\sum_{n=1}^{\infty} f(n) = \int_a^{\infty} f$.

10. תהי $f : [0,\infty) \to \mathbb{R}$ רציפה. הוכיחו או הפריכו:

(א) אם $\int_0^{\infty} f(x)dx$ מתכנס אז $\sum_{n=0}^{\infty} f(n)$ מתכנס.

(ב) אם $\sum_{n=0}^{\infty} f(n)$ מתכנס אז $\int_0^{\infty} f(x)dx$ מתכנס.

(ג) אם $\int_0^{\infty} f(x)dx$ וגם $\sum_{n=0}^{\infty} f(n)$ מתכנסים אז $\sum_{n=0}^{\infty} f(n+\frac{1}{2})$ מתכנס.

11. הטכניקה שהשתמשנו בה בהוכחת מבחן האינטגרל מאפשרת גם להעריך סכומים סופיים על־ידי שימוש באינטגרל מתאים. בשאלה זו נוכיח שיש קבועים c', c'' כך ש־

$$c' \cdot (\frac{n}{e})^n \leq n! \leq c'' \cdot (\frac{n+1}{e})^{n+1}$$

(למעשה נוכל לחשב את c', c'' באופן מפורש). ראשית, הצדיקו את השוויון $\ln(n!) = \sum_{k=1}^{n} \ln k$. לאחר מכן העריכו את הסכום $\sum_{k=1}^{n} \ln k$ מלמעלה ומלמטה על־ידי אינטגרל מתאים, וחשבו את האינטגרלים. הסיקו את האי־שוויון המבוקש.

ההערכה שקיבלנו ל־$n!$ אינה הטובה ביותר האפשרית. שימו לב למשל שהיחס בין אגף ימין לאגף שמאל שואף לאינסוף. קיימת גרסה חזקה יותר של האי־שוויון, שתוכח בסעיף 9.11 להלן.

12. נגדיר $\gamma_n = \sum_{k=1}^{n} \frac{1}{k} - \ln n$. הראו ש־ $\gamma = \lim_{n\to\infty} \gamma_n$ קיים. מספר זה נקרא **קבוע אוילר**. נציין שמשערים שקבוע אוילר הוא אי־רציונלי (ואף טרנסצנדנטי), אך לא קיימת כיום הוכחה לכך.

9.10 אינטגרציה נומרית

המשפט היסודי נותן כלים לחישוב אינטגרלים בתנאי שיש לאינטגרנד פונקציה קדומה אלמנטרית. ראינו בסעיף 9.7 שיטות המאפשרות למצוא פונקציה קדומה כזו במקרים מסוימים, אך אם האינטגרנד מסובך, מציאת פונקציה קדומה כזו יכולה להיות קשה או אפילו בלתי אפשרית. במקרים כאלה עלינו להסתפק בפתרון מקורב. בסעיף הנוכחי נעסוק בקצרה בכמה שיטות קירוב כאלה, כאשר הדגש הוא על יעילות והערכת המרחק של הקירוב מהערך האמתי.[3]

דרך טבעית לקרב אינטגרל מסוים היא להשתמש בסכומי רימן. ליתר דיוק, אם רוצים לחשב את $\int_a^b f$ ניתן לבחור חלוקה P ונקודות ξ לחלוקה, ולחשב את $\sigma(f,P,\xi)$. מספר זה ניתן לחישוב במידה שאנו יודעים לחשב את המספרים $f(\xi_i)$. כעת נשאלת השאלה, כיצד לבחור את P,ξ באופן ש־$\sigma(f,P,\xi)$ יהיה קרוב ל־$\int_a^b f$? אנו יודעים שכל מידת קרבה ניתנת להשגה על ידי בחירת חלוקה עדינה מספיק, אולם בהינתן $\varepsilon > 0$ לא ברור כמה עדינה צריכה להיות החלוקה על מנת להבטיח ש־ $|\int_a^b f - \sigma(f,P,\xi)| < \varepsilon$. בסעיף זה נראה שבמקרים מסוימים ניתן להעריך את ההפרש.

ראינו (תרגיל (2) בעמוד 409) שאם f אינטגרבילית ב־$[a,b]$, אם P חלוקה של $[a,b]$ ו־ξ בחירת נקודות ל־P אז

$$|\int_a^b f - \sigma(f,P,\xi)| \le \omega(f,P)$$

לכן מטרתנו היא לבחור חלוקה P שהתנודה שלה קטנה.

נניח ש־f גזירה ב־$[a,b]$ וש־M חסם של הנגזרת f'. אם $P = \{x_0, x_1, \ldots, x_k\}$ חלוקה של $[a,b]$ אז לכל $x', x'' \in [x_{i-1}, x_i]$ שונים קיים לפי משפט לגרנג' מספר c בין x', x'' כך שמתקיים

$$|f(x') - f(x'')| = |f'(c)||x' - x''| \le M\Delta x_i$$

משמע

$$\omega_i(f,P) = \sup\{f(x') - f(x'')\ :\ x', x'' \in [x_{i-1}, x_i]\} \le M\Delta x_i$$

[3] כיום החישוב עצמו מבוצע לעתים קרובות על ידי מחשב, אך אין זה פתרון קסם. בעיות מסובכות מעמידות במבחן גם את המחשבים המהירים ביותר, וחשוב למצוא שיטות קירוב יעילות.

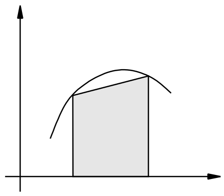
ולכן

$$\begin{aligned}\omega(f,P) &= \sum_{i=1}^{k}\Delta x_i\omega_i(f,P)\\ &\le \sum_{i=1}^{k}\lambda(P)\omega_i(f,P)\\ &\le \lambda(P)\sum_{i=1}^{k}M\Delta x_i\\ &= \lambda(P)M(b-a)\end{aligned}$$

מכאן קיבלנו

משפט 9.10.1 נניח ש־f גזירה ב־$[a,b]$ ו־ $|f'(x)| \le M$ לכל $x \in (a,b)$. יהי $\varepsilon > 0$. אז לכל חלוקה P של $[a,b]$ ולכל בחירת נקודות ξ עבור P מתקיים

$$|\sigma(f,P,\xi) - \int_a^b f| < \lambda(P)\cdot M\cdot(b-a)$$



איור 9.10.1 קירוב שטח מתחת לגרף על ידי טרפז

אפשר לקבל הערכה טובה יותר לאינטגרל $\int_a^b f$ אם במקום לקרב את השטח מתחת לגרף של f על ידי מלבנים, נקרב אותו על ידי טרפזים. השטח של טרפז שבסיסו באורך Δx ושגובה הניצבים שלו h_1 ו־h_2 הוא $\Delta x\cdot\frac{1}{2}(h_1+h_2)$. בהינתן חלוקה $P=\{x_0,x_1,\ldots,x_k\}$ של $[a,b]$ אפשר להתבונן בכל קטע חלוקה $[x_{i-1},x_i]$ על הטרפז שבסיסו הוא $[x_{i-1},x_i]$ וגבהי הצדדים שלו הם $f(x_{i-1}),f(x_i)$. אנו מצפים ש־$\int_{x_{i-1}}^{x_i} f$ יהיה קרוב לשטח $\Delta x_i\cdot\frac{1}{2}(f(x_{i-1})+f(x_i))$ של הטרפז, וכך מקבלים קירוב ל־$\int_a^b f$ מהצורה

$$\sum_{i=1}^{k}\frac{1}{2}\Delta x_i(f(x_{i-1})+f(x_i))$$

קירוב זה באמת טוב מאד אם הפונקציה גזירה פעמיים. על מנת להעריך את טיב הקירוב ניעזר בלמה הבאה:

למה 9.10.2 (נוסחת הטרפז) תהי $f:[a,a+h]\to\mathbb{R}$ פונקציה גזירה פעמיים. נסמן ב־T את שטח הטרפז שבסיסו $[a,a+h]$ וניצביו בגובה $f(a),f(a+h)$, כלומר $T=\frac{h}{2}(f(a)+f(a+h))$. אז יש נקודה $c\in(a,a+h)$ כך ש־

$$\int_a^{a+h} f = T - \frac{h^3}{12}f''(c)$$

הוכחה אנו רוצים להעריך את ההפרש $\int_a^{a+h} f - \frac{h}{2}(f(a)+f(a+h))$. נחשוב על ביטוי זה כפונקציה ב־h ונסמן אותו ב־

$$\varphi(t) = \int_a^{a+t} f - \frac{t}{2}(f(a)+f(a+t))$$

$\varphi(t)$ מוגדר לכל $0 \le t \le h$ ואנו רוצים להעריך את $\varphi(h)$. גזירה של φ לפי המשפט היסודי וכללי הגזירה נותן

$$\begin{aligned} \varphi'(t) &= f(a+t) - \frac{1}{2}(f(a)+f(a+t)) - \frac{t}{2}f'(a+t) \\ &= \frac{1}{2}f(a+t) - \frac{1}{2}f(a) - \frac{t}{2}f'(a+t) \end{aligned}$$

וגזירה פעם נוספת נותנת

$$\begin{aligned} \varphi''(t) &= \frac{1}{2}f'(a+t) - 0 - (\frac{1}{2}f'(a+t) + \frac{t}{2}f''(a+t)) \\ &= -\frac{t}{2}f''(a+t) \end{aligned}$$

כמו־כן הצבה מראה ש־ $\varphi(0) = \varphi'(0) = 0$.

כדי להעריך את $\varphi(h)$ די להעריך את ההפרש $\varphi(h) - \varphi(0)$. לשם כך נשווה את ההפרש להפרש אחר בעזרת משפט קושי. לפי משפט זה, קיים מספר $c \in (0,h)$ כך ש־

$$\frac{\varphi(h)-\varphi(0)}{h^3-0^3} = \frac{\varphi'(c)}{3c^2}$$

(בדקו שתנאי משפט קושי מתקיימים!) נשים לב שהביטוי מימין ניתן לכתיבה כ־

$$\frac{\varphi'(c)}{3c^2} = \frac{\varphi'(c)-\varphi'(0)}{3h^2-3\cdot 0^2}$$

כי $\varphi'(0) = 0$. לכן נוכל להפעיל את משפט קושי פעם נוספת, ולהסיק שקיים $d \in (0,c)$ כך ש־

$$\frac{\varphi'(c)-\varphi'(0)}{3c^2-2\cdot 0^2} = \frac{\varphi''(d)}{6d} = -\frac{1}{12}f''(a+d)$$

(בשוויון האחרון הצבנו את הביטוי שקיבלנו עבור φ''). בסיכום קיבלנו שקיים $d \in (0,h)$ כך ש־

$$\frac{\varphi(h)}{h^3} = -\frac{1}{12}f''(a+d)$$

והלמה נובעת על ידי העברת אגפים. ■

משפט 9.10.3 אם f גזירה פעמיים ב־$[a,b]$ ואם $|f''(x)| \leq M$ לכל $x \in (a,b)$ אז לכל חלוקה $P = \{x_0, x_1, \ldots, x_k\}$ של $[a,b]$ מתקיים

$$|\int_a^b f - \sum_{i=1}^k \frac{1}{2}\Delta x_i(f(x_{i-1}) + f(x_i))| \leq \frac{1}{12}M \cdot \lambda(P)^2 \cdot (b-a)$$

הערה מכיוון שפרמטר החלוקה מופיע באגף ימין בחזקה שנייה אומרים שהמשפט נותן קירוב מסדר שני של האינטגרל. משפט 9.10.1 הוא קירוב מסדר ראשון.

הוכחה מתכונות האינטגרל מתקיים

$$\int_a^b f = \sum_{i=1}^k \int_{x_{i-1}}^{x_i} f$$

ולכן

$$|\int_a^b f - \sum_{i=1}^k \frac{1}{2}\Delta x_i(f(x_{i-1}) + f(x_i))| \leq \sum_{i=1}^k |\int_{x_{i-1}}^{x_i} f - \frac{1}{2}\Delta x_i(f(x_{i-1}) + f(x_i))|$$

לפי אי־שוויון הטרפז לכל $i = 1, \ldots, k$ יש $c_i \in (x_{i-1}, x_i)$ כך ש־

$$|\int_{x_{i-1}}^{x_i} f - \frac{1}{2}\Delta x_i(f(x_{i-1}) + f(x_i))| = \frac{(\Delta x_i)^3}{12}|f''(c_i)| \leq \frac{\lambda(P)^2}{12} \cdot \Delta x_i \cdot M$$

ולכן

$$\begin{aligned} |\int_a^b f - \sum_{i=1}^k \frac{1}{2}\Delta x_i(f(x_{i-1}) + f(x_i))| &\leq \sum_{i=1}^k \frac{\lambda(P)^2}{12} \cdot \Delta x_i \cdot M \\ &= \frac{M \cdot \lambda(P)^2}{12} \sum_{i=1}^k \Delta x_i \\ &= \frac{1}{12}M \cdot \lambda(P)^2 \cdot (b-a) \end{aligned}$$

כנדרש. ■

תרגילים

העריכו את השטח מתחת לגרף של $\sqrt{1-x^2}$ בקטע $[-\frac{1}{2}, \frac{1}{2}]$ וחשבו את π בקירוב של שתי ספרות אחרי הנקודה.

9.11 נוסחת סטירלינג

המספר $n!$ והמקדמים הבינומיים $\frac{n!}{m!(n-m)!}$ מופיעים במקומות רבים במתמטיקה, ולעתים מתעורר צורך להעריך את גודלם בעזרת גודל שקל יותר לניתוח. בסעיף זה ניתן הערכה מדויקת של $n!$, הנקראת **נוסחת סטירלינג**[4].

יש כמה הערכות בסיסיות עבור $n!$ שאפשר לקבלן באמצעים אלמנטריים. קל לראות ש־ $\frac{a^n}{n!} \to 0$ לכל a קבוע, ומצד שני $\frac{n^n}{n!} \to \infty$. ממבחן המנה אפשר אף להסיק שהמנה $\frac{(n/a)^n}{n!}$ שואפת לאינסוף עבור $a < e$, ואילו עבור $a > e$ היא שואפת לאפס (הוכיחו!). לכן הסדרה $(\frac{n}{e})^n$ היא קירוב טוב יחסית ל־$n!$. הערכה טובה עוד יותר קיבלנו בתרגיל (11) בעמוד 454, שם ראינו שיש קבועים c', c'' כך ש־

$$c' \cdot (\frac{n}{e})^n \leq n! \leq c'' \cdot (\frac{n+1}{e})^{n+1}$$

זו עדיין לא הערכה מדויקת במיוחד: היחס בין החסם העליון והתחתון שואף לאינסוף.

במשפט הבא, לעומת זאת, אנו מקבלים הערכה מדויקת:

משפט 9.11.1 (נוסחת סטירלינג) $\sqrt{2\pi}(\frac{n}{e})^n \leq n! \leq \sqrt{2\pi n}(\frac{n}{e})^n \cdot e^{1/12n}$, ובפרט

$$\lim_{n\to\infty} \frac{\sqrt{2\pi}(\frac{n}{e})^n}{n!} = 1$$

הוכחה הרעיון דומה להוכחה של תרגיל (11) מסעיף 454 אבל ההערכה שנבצע עדינה יותר. נשתמש בשוויון

$$\ln n! = \sum_{k=1}^{n} \ln k = \sum_{k=2}^{n} \ln k$$

שנובע מתכונות הלוגריתם (השוויון האחרון כי $\ln 1 = 0$). נגדיר את ε_n להיות ההפרש בין $\ln k$ לאינטגרל $\int_{k-1}^{k} \ln$, דהיינו $\ln k = \int_{k-1}^{k} \ln x\, dx + \varepsilon_k$. אז

$$\begin{aligned}
\ln n! &= \sum_{k=2}^{n} (\varepsilon_k + \int_{k-1}^{k} \ln x\, dx) \\
&= \sum_{k=2}^{n} \varepsilon_k + \int_{1}^{n} \ln x\, dx \\
&= \sum_{k=2}^{n} \varepsilon_k + n \ln n - n
\end{aligned}$$

[4] James Stirling, 1692–1770.

נותר להעריך את ה־ε_k־ים. לפי כלל הטרפז 9.10.2 והעובדה ש־ $(\ln x)'' = -\frac{1}{x^2}$ אנו יודעים שיש $c_k \in (k-1,k)$ כך ש־

$$\begin{aligned}\int_{k-1}^{k} \ln x\,dx &= \frac{1}{2}(\ln(k-1)+\ln k) - (-\frac{1}{12}\cdot\frac{1}{c_k^2})\\ &= \ln k + \frac{1}{2}(\ln(k-1)-\ln k) + \frac{1}{12c_k^2}\end{aligned}$$

ולכן

$$\varepsilon_k = \frac{1}{2}(\ln k - \ln(k-1)) - \frac{1}{12c_k^2}$$

ומתקיים

$$\begin{aligned}\sum_{k=2}^{n}\varepsilon_k &= -\sum_{k=2}^{n}\frac{1}{12c_k^2} + \frac{1}{2}\sum_{k=2}^{n}(\ln k - \ln(k-1))\\ &= -\frac{1}{12}\sum_{k=2}^{n}\frac{1}{c_k^2} + \frac{1}{2}\ln n\end{aligned}$$

כי הסכום $\sum_{k=2}^{n}(\ln k - \ln(k-1))$ טלסקופי ו־ $\ln 1 = 0$. קיבלנו אפוא

$$\ln n! = n\ln n - n + \frac{1}{2}\ln n - \frac{1}{12}\sum_{k=2}^{n}\frac{1}{c_k^2}$$

נפעיל על שני האגפים את פונקציית האקספוננט ונקבל

$$n! = \sqrt{n}(\frac{n}{e})^n \cdot \exp(-\frac{1}{12}\sum_{k=2}^{n}\frac{1}{c_k^2})$$

נזכור כעת ש־ $c_k \geq k-1$. מכאן שהטור $\sum_{k=2}^{\infty}\frac{1}{c_k^2}$ נשלט על ידי הטור $\sum_{k=2}^{\infty}\frac{1}{(k-1)^2}$, ולכן מתכנס. נסמן $C = \sum_{k=2}^{\infty}\frac{1}{c_k^2}$. אז $\sum_{k=2}^{n}\frac{1}{c_k^2} \leq C$ ובפרט

$$\sum_{k=2}^{n}\frac{1}{c_k^2} = C - \sum_{k=n+1}^{\infty}\frac{1}{c_k^2} \geq 0$$

מאידך, מאחר ש־ $c_k \leq k$ מתקיים

$$\sum_{k=n+1}^{\infty}\frac{1}{c_k^2} \leq \sum_{k=n+1}^{\infty}\frac{1}{k^2} \leq \int_{n}^{\infty}\frac{1}{x^2dx} = \frac{1}{n}$$

(ממה נובע האי־שוויון בין הטור לאינטגרל?). קיבלנו ש־

$$C - \frac{1}{n} \leq \sum_{k=2}^{n}\frac{1}{c_k^2} \leq C$$

מכאן

$$\sqrt{n}(\frac{n}{e})^n e^{-C/12} \leq n! \leq \sqrt{n}(\frac{n}{e})^n e^{-(C-\frac{1}{n})/12}$$

ואם נסמן $B = e^{-\frac{1}{12}C}$ אז

$$B\sqrt{n}(\frac{n}{e})^n \leq n! \leq B\sqrt{n}(\frac{n}{e})^n e^{1/12n}$$

נותר רק להראות ש־ $B = \sqrt{2\pi}$. את זה תוכיחו בתרגילים (2) ו־(3) להלן. ■

תרגילים

1. עבור $0 \leq x \leq 1$ נסמן $H(x) = -x\ln x - (1-x)\ln(1-x)$, כאשר מפרשים $0\ln 0 = 0$. הוכיחו שלכל $\varepsilon > 0$ יש קבועים C', C'' כך ש־

$$C' e^{H(\varepsilon)} \leq \binom{n}{[\varepsilon n]} \leq C e^{H(\varepsilon)}$$

2. נוכיח בשאלה זו את **נוסחת וואלס**[5]. נסמן $J_m = \int_0^{\frac{\pi}{2}} \sin^m x dx$.

(א) הראו שמתקיים $J_{m+1} \leq J_m$.

(ב) חשבו את J_0 ו־ J_1 והוכיחו שלכל $m \geq 2$ מתקיים $J_m = \frac{m-1}{m} J_{m-2}$.

(ג) לכל $m > 0$ טבעי נסמן

$$m!! = \begin{cases} 2\cdot 4\cdot 6\cdot \ldots \cdot m, & \text{זוגי } m \\ 1\cdot 3\cdot 5\cdot \ldots\cdot m, & \text{אחרת} \end{cases}$$

הסיקו את הנוסחה

$$J_m = \begin{cases} \frac{\pi}{2}\frac{(m-1)!!}{m!!} & \text{זוגי } m \geq 2 \\ \frac{(m-1)!!}{m!!} & \text{אי־זוגי } m \geq 3 \end{cases}$$

ואת האי־שוויון

$$\frac{1}{2m+1}\left(\frac{(2m)!!}{(2m-1)!!}\right)^2 \leq \frac{\pi}{2} \leq \frac{1}{2m}\left(\frac{(2m)!!}{(2m-1)!!}\right)^2$$

[5] John Wallis, 1616־1703.

(ד) הוכיחו את נוסחת וואלס

$$\lim_{m\to\infty}\frac{1}{2m}\left(\frac{(2m)!!}{(2m-1)!!}\right)^2=\frac{\pi}{2}$$

3. בעזרת נוסחת וואלס הוכיחו שמתקיים

$$\sqrt{\pi}=\lim_{m\to\infty}\frac{2^{2m}}{\sqrt{m}\binom{2m}{m}}$$

היעזרו בנוסחת סטירלינג כדי להעריך את $\binom{2m}{m}$, והסיקו שהקבוע בנוסחת סטירלינג הוא $\sqrt{2\pi}$.

פרק 10

סדרות וטורי פונקציות

אחד ממושגי המפתח שפיתחנו הוא מושג ההתכנסות של סדרת מספרים. בפרק זה נפתח מושג דומה עבור סדרות של פונקציות. ליתר דיוק, נפתח כמה מושגים כאלה, שכן יש יותר מדרך טבעית אחת להגדיר התכנסות של סדרת פונקציות.

השאלה העיקרית שתעסיק אותנו היא אילו תכונות של פונקציות נשמרות תחת גבול, כלומר: אם סדרת פונקציות $(f_n)_{n=1}^{\infty}$ מתכנסת לפונקציה f אילו מהתכונות של f_n מועברות ל־f? למשל, האם גבול של פונקציות רציפות הוא רציף? האם האינטגרל של f שווה לגבול האינטגרלים של f_n? וכן הלאה. התשובות אינן תמיד חיוביות, ונאלץ למצוא תנאים שונים שיבטיחו את ההתנהגות המבוקשת.

המושגים החדשים יפתחו בפנינו אפשרויות חדשות להבנה של פונקציות. למשל, נוכל לנתח פונקציה מסובכת על ידי הצגתה כגבול של סדרת פונקציות פשוטות יותר (בתנאי, כמובן, שאנו יודעים למצוא סדרה כזו). כמו־כן נוכל לבנות דוגמאות חדשות לפונקציות על ידי תיאורן כגבול של סדרת פונקציות מתאימה. ניישם רעיונות אלה בפרק הזה ובפרק הבא.

10.1 התכנסות נקודתית של סדרות פונקציות

סדרת פונקציות היא סדרה $(f_n)_{n=1}^{\infty}$ שכל אחד מאיבריה הוא פונקציה. אם (f_n) סדרת פונקציות וכולן מוגדרות בנקודה x_0 אז על ידי הפעלת כל אחת מהפונקציות על x_0 אנו מקבלים סדרה חדשה $(f_n(x_0))_{n=1}^{\infty}$, שהיא סדרה של מספרים ממשיים (ולא של פונקציות!). ייתכן שסדרה זו מתכנסת וייתכן שלא, וכמובן שהתשובה יכולה להיות תלויה בנקודה x_0.

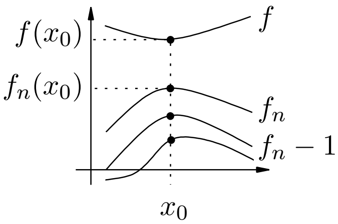
לדוגמה, לכל n נגדיר את $e_n : \mathbb{R} \to \mathbb{R}$ על ידי

$$e_n(x) = (1 + \frac{x}{n})^n$$

אז (e_n) היא סדרת פונקציות. בסעיף 5.11 ראינו שלכל $x \in \mathbb{R}$ מתקיים $e_n(x) \to e^x$ (נזכיר שוב שמהרגע שקבענו את המספר x הסדרה $(e_n(x))_{n=1}^{\infty}$ היא סדרה של מספרים, ולא סדרה של פונקציות). על כן טבעי לומר שהפונקציה $\exp$ היא הגבול של סדרת הפונקציות e_n.

באופן כללי יותר,

הגדרה 10.1.1 תהי $(f_n)_{n=1}^{\infty}$ סדרת פונקציות המוגדרות כולן בקבוצה $A \subseteq \mathbb{R}$ ותהי f פונקציה המוגדרת ב־A. נאמר ש־f הוא **הגבול הנקודתי** (pointwise limit) של הסדרה (f_n) ב־A, או ש־(f_n) **מתכנסת נקודתית** ל־f ב־A, אם לכל $x \in A$ הסדרה $(f_n(x))$ מתכנסת ל־$f(x)$. במקרה זה רושמים $f = \lim_{n\to\infty} f_n$ או $f_n \to f$.[1] קבוצת הנקודות x כך שסדרת המספרים $(f_n(x))_{n=1}^{\infty}$ מתכנסת (ובפרט מוגדרת לכל n גדול מספיק) נקראת **תחום ההתכנסות** של הסדרה (f_n).



איור 10.1.1 סדרת הערכים $f_n(x_0)$ של הפונקציות ב־x_0 מתכנסת ל־$f(x_0)$

בחרנו להשתמש במונח "התכנסות נקודתית" ולא במונח "התכנסות" כדי להבחין בין המושג שהגדרנו עתה לסוג התכנסות נוסף שאותו נגדיר בהמשך. שימו לב שבסימון $f = \lim_{n\to\infty} f_n$ לא מופיע העובדה שמדובר בהתכנסות נקודתית, וגם הקבוצה שבה מתרחשת ההתכנסות לא מצויינת. יש לציין זאת תמיד במפורש.

הערות

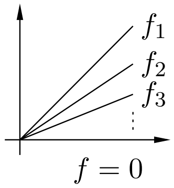
1. מההגדרה ברור שאם D הוא תחום ההתכנסות של סדרת פונקציות (f_n) אז הסדרה מתכנסת נקודתית ב־D. יתר על כן D היא הקבוצה המקסימלית שבה סדרת הפונקציות מתכנסת נקודתית, במובן זה שאם $f_n \to f$ נקודתית ב־A אז $A \subseteq D$.
2. אם הגבול הנקודתי של סדרת פונקציות קיים ב־A אז הוא יחיד במובן הבא: אם (f_n) מתכנסת נקודתית ל־f ול־g ב־A אז f, g מסכימות ב־A, כי לכל $x \in A$ מתקיים
$$f(x) = \lim_{n\to\infty} f_n(x) = g(x)$$
במילים אחרות, $f|_A = g|_A$.

דוגמאות

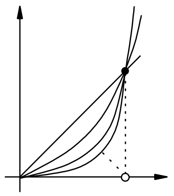
1. נסמן $e_n(x) = (1 + \frac{x}{n})^n$. אז $e_n \to \exp$ נקודתית בכל הישר $\mathbb{R}$.

---

[1] בספרים רבים מסמנים ב־ $f_n(x) \to f(x)$ את ההתכנסות הנקודתית $f_n \to f$ של פונקציות. אנו נקפיד להשתמש בסימון הראשון רק במובן של התכנסות סדרת המספרים $(f_n(x))_{n=1}^{\infty}$ למספר $f(x)$, ובסימון השני כדי לציין התכנסות פונקציות.



איור 10.1.2 סדרת הפונקציות $f_n = x/n$


איור 10.1.3 סדרת הפונקציות $f_n = x^n$

2. יהיו $f_n : \mathbb{R} \to \mathbb{R}$ נתונות על ידי $f_n(x) = \frac{x}{n}$. לכל $x \in \mathbb{R}$ מתקיים:

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{x}{n} = 0$$

לכן $f_n \to 0$ נקודתית ב־$\mathbb{R}$ (שימו לב שכאן 0 מציין את פונקציית האפס!), ותחום ההתכנסות של הסדרה הוא $\mathbb{R}$.

3. יהיו $g_n : [0,\infty) \to \mathbb{R}$ נתונות על ידי $g_n(x) = x^n$. לכל $x \in [0,\infty)$ מתקיים:

$$\lim_{n\to\infty} g_n(x) = \lim_{n\to\infty} x^n = \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1 \\ \infty & x > 1 \end{cases}$$

לכן g אינה מתכנסת נקודתית ב־$[0,\infty)$. תחום ההתכנסות שלה הוא $[0,1]$ והגבול שלה שם הוא הפונקציה

$$g(x) = \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1 \end{cases}$$

שימו לב שאפשר להרחיב את ההגדרה של g באופנים רבים מחוץ ל־$[0,1]$. כל ההרחבות האלה שונות כפונקציות אבל g_n מתכנסת נקודתית אל כולן ב־$[0,1]$. זה מדגיש את העובדה שהגבול הוא יחיד רק כשמצטמצמים לתחום ההתכנסות (עיינו הערה מלפני הדוגמאות).

4. תהי $(q_n)_{n=1}^{\infty}$ סדרה המכילה פעם אחת כל מספר הרציונלי בקטע $[0,1]$ (ראינו שיש סדרה כזאת בסעיף 5.12). לכל n נגדיר $h_n : [0,1] \to \mathbb{R}$ על ידי

$$h_n(x) = \begin{cases} 1 & x \in \{q_1, q_2, \ldots, q_n\} \\ 0 & x \notin \{q_1, q_2, \ldots, q_n\} \end{cases}$$

במילים אחרות, h_{n+1} שווה ל־h_n אלא שהנקודה q_n שינתה את ערכה מ־0 ל־1.

לכל $x \in [0,1]$ מתקיימת אפשרות אחת משתיים. אם x אי־רציונלי אז $h_n(x) = 0$ לכל n, ולכן $h_n(x) \to 0$. אחרת $x = q_n$ עבור n כלשהו, ואז $h_k(x) = 1$ לכל $k \geq n$, כלומר $h_n(x) \to 1$. קיבלנו אם כן

$$\lim_{n\to\infty} h_n(x) = \begin{cases} 1 & x \in \mathbb{Q} \cap [0,1] \\ 0 & x \notin \mathbb{Q} \cap [0,1] \end{cases}$$

כלומר $h_n \to D$ נקודתית ב־$[0,1]$, כאשר D פונקציית דירכלה.

5. אם $(a_n)_{n=1}^{\infty}$ היא סדרת מספרים נגדיר פונקציות קבועות $f_n \equiv a_n$ בכל הישר. אז (f_n) נקודתית ב־$\mathbb{R}$ אמ"מ (a_n) מתכנסת, ובמקרה זה הגבול היא הפונקציה הקבועה $f \equiv a$ כאשר $a = \lim a_n$.

6. כמובן שסדרת פונקציות אינה חייבת להתכנס נקודתית בשום קבוצה, דהיינו תחום ההתכנסות יכול להיות ריק. למשל, לפי הדוגמה הקודמת, אם (a_n) סדרת מספרים מתבדרת אז הפונקציות $f_n \equiv a_n$ מהוות סדרת פונקציות שאינה מתכנסת נקודתית בשום קבוצה.

מושג ההתכנסות הנקודתית מוגדר גם לטורי פונקציות:

הגדרה 10.1.2 תהי $(u_n)_{n=1}^{\infty}$ סדרת פונקציות המוגדרות בקבוצה $A \subseteq \mathbb{R}$ ותהי S פונקציה המוגדרת ב־A. נגדיר פונקציות S_N על ידי $S_N = \sum_{n=1}^{N} u_n$. נאמר ש**טור הפונקציות** $\sum_{n=1}^{\infty} u_n$ מתכנס נקודתית ל־S ב־A אם $S_N \to S$ נקודתית ב־A. במקרה זה נסמן $S = \sum_{n=1}^{\infty} u_n$, ונציין שההתכנסות היא נקודתית ומתרחשת בקבוצה A. **תחום ההתכנסות** של טור פונקציות הוא תחום ההתכנסות של סדרת הסכומים החלקיים.

כמו במקרה של טורי מספרים, נשתמש לפעמים בסימון $\sum_{k=1}^{\infty} u_k$ על מנת לציין את סדרת הפונקציות (u_k), גם במקרים שבהם הטור אינו מתכנס. לכן את המשפט "יהי $\sum_{k=1}^{\infty} u_k$ טור פונקציות" אין לפרש כהבטחה שהטור מתכנס אלא רק כהצגה של הסדרה. מההגדרה ברור שאם $\sum_{k=1}^{\infty} u_k$ מתכנס נקודתית ל־S ב־A, אז $S(x) = \sum_{k=1}^{\infty} u_k(x)$ לכל $x \in A$. כמו־כן הגבול S יחיד באותו מובן שגבול פונקציות הוא יחיד, דהיינו, אם גם T הוא גבול נקודתי של $\sum_{k=1}^{\infty} u_k$ ב־A אז $S|_A = T|_A$.

דוגמאות

1. נגדיר $u_n(x) = x^n$. $x \in (-1,1)$ אז

$$\begin{aligned}
\sum_{n=1}^{\infty} u_n(x) &= \lim_{N\to\infty} \sum_{k=1}^{N} u_k(x) \\
&= \lim_{N\to\infty} \left(1 + x + x^2 + \cdots x^N\right) \\
&= \lim_{N\to\infty} \frac{1 - x^{N+1}}{1-x} \\
&= \frac{1}{1-x}
\end{aligned}$$

ולכן $\sum_{1}^{\infty} u_n$ מתכנס נקודתית לפונקציה $\frac{1}{1-x}$ ב־$(-1,1)$.

דוגמה זו הייתה מקור לבלבול רב בתקופה שלפני הגדרת הגבול. הנוסחה $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$ נחשבה לנכונה, ואולם אם מתייחסים אליה כנכונה ללא סייג מתקבלות תוצאות משונות. למשל המתמטיקאי אוילר השתמש בנוסחה זו כדי להצדיק את אמונתו בשוויון $\ldots 1 - 1 + 1 - 1 + 1 - 1 = \frac{1}{2}$, שמתקבל מהנוסחה על ידי הצבת $x = -1$. לעומת זאת אם מציבים בה $x = 2$ מתקבל

השוויון $-1 = 1 + 2 + 4 + 8 + \ldots$, שבאגף שמאל שלו מופיע מספר שלילי, ובאגף ימין סכום של מספרים חיוביים. תופעות אלה היו מקור למבוכה ולחילוקי דעות בקרב הקהילה המתמטית, והיו בין המניעים לפיתוח תורה מדויקת יותר של הגבול.

2. כמו במקרה של סדרות וטורי מספרים, שאלות על טורי פונקציות ניתנות לתרגום לשאלות על סדרות של פונקציות, ולהפך. המעבר מטור לסדרה הוא בעזרת סדרת הסכומים החלקיים. בכיוון ההפוך, בהינתן סדרת פונקציות $(f_n)_{n=1}^{\infty}$ ניתן לקבל טור פונקציות $\sum_{n=1}^{\infty} u_n$ שהתכנסותו שקולה להתכנסות הסדרה על ידי

$$\begin{aligned} u_1 &= f_1 \\ u_n &= f_n - f_{n-1} \end{aligned}$$

לכל $n \geq 2$. אז $f_n \to f$ נקודתית ב־A אמ"מ $\sum_{n=0}^{\infty} u_n = f$ נקודתית ב־A.

המשפטים והמושגים שפיתחנו עבור סדרות וטורי מספרים נותנים משפטים ומושגים מקבילים עבור סדרות וטורי פונקציות. למשל,

הגדרה 10.1.3 נאמר שטור פונקציות $\sum_{n=1}^{\infty} u_n$ הוא טור **מתכנס בהחלט** (converges absolutely) בקבוצה A אם טור המספרים $\sum_{n=1}^{\infty} u_n(x)$ מתכנס בהחלט לכל $x \in A$.

טור מספרים מתכנס בהחלט הוא בפרט טור מתכנס, ולכן אם טור פונקציות $\sum_{n=1}^{\infty} u_n$ מתכנס בהחלט בקבוצה $A \subseteq \mathbb{R}$ אז הטור מתכנס ב־A נקודתית.

הנה דוגמה נוספת: נניח ש־$(f_n), (g_n)$ סדרות פונקציות ו־ $f_n \to f$ ו־ $g_n \to g$ נקודתית ב־A. מכאן נסיק ש־ $f_n + g_n \to f + g$ נקודתית על A, כי מהנתון לכל $x \in A$ מתקיים

$$(f_n + g_n)(x) = f_n(x) + g_n(x) \xrightarrow{n\to\infty} f(x) + g(x) = (f+g)(x)$$

וזו בדיוק ההגדרה של התכנסות נקודתית של $(f_n + g_n)$ ל־ $f + g$ ב־A.

תרגילים

1. נגדיר בקטע $[0, \infty)$ סדרות פונקציות (f_n) בעזרת הנוסחאות הבאות. מצאו את תחום ההתכנסות ואת פונקציית הגבול של הסדרות.

   (א) $x^n(1 - x^n)$.

   (ב) $\sin^n(x)$.

   (ג) $\frac{x+n}{n}$.

   (ד) $\frac{x}{x+n}$.

(ה) $\frac{nx}{1+nx}$.
(ו) $\frac{nx}{1+n^2x^2}$.
(ז) $n^2x^2e^{-nx}$.
(ח) $\frac{1}{n}\log(1+nx)$.
(ט) $\frac{x^n}{1+x^n}$.

2. הוכיחו:

(א) אם $f_n \to f$ נקודתית ב־A, B אז $f_n \to f$ נקודתית ב־$A \cup B$.
(ב) אם $f_n \to f$ נקודתית ב־A ו־ $B \subseteq A$ אז $f_n \to f$ נקודתית ב־B.
(ג) אם $\sum_{n=1}^{\infty} u_k$ מתכנס נקודתית ב־A אז $u_n \to 0$ נקודתית ב־A.

3. תהי (u_k) סדרת פונקציות. נסחו גרסה של משפט לייבניץ שתבטיח שהטור $\sum_{k=1}^{n}(-1)^k u_k$ מתכנס נקודתית בקבוצה A.

4. תהי f פונקציה המוגדרת בקבוצה $A \subseteq \mathbb{R}$, תהי (a_n) סדרת מספרים המקיימת $a_n \to a$, יהי $x_0 \in A$ קבוע ותהי $(x_n) \subseteq A$ סדרה כך ש־ $x_n \to x_0$. הוכיחו או הפריכו:

(א) לכל n נגדיר פונקציה f_n על ידי שינוי הערך של f בנקודה x_0, כך ש־ $f_n(x_0) = a_n$. האם הסדרה (f_n) מתכנסת נקודתית ב־A, ואם כן, מה הגבול?
(ב) לכל n נגדיר פונקציה $\widehat{f_n}$ על ידי שינוי f בנקודה x_n בלבד, כך ש־ $\widehat{f}_n(x_n) = a_n$. האם הסדרה $(\widehat{f}_n)$ מתכנסת נקודתית ב־A, ואם כן, מה הגבול?
(ג) לכל n נגדיר פונקציה $\overline{f}_n$ על ידי שינוי f בנקודות $x_1, \ldots, x_n$, ובכל הנקודות לפונקציה החדשה יהיה ערך a_n. האם הסדרה $(\overline{f}_n)$ מתכנסת נקודתית ב־A, ואם כן, מה הגבול?

5. הוכיחו או הפריכו:

(א) הגבול הנקודתי של פונקציות אי־שליליות הוא אי־שלילי.
(ב) הגבול הנקודתי של פונקציות חיוביות הוא חיובי.
(ג) הגבול הנקודתי של פונקציות עולות חלש הוא עולה חלש.
(ד) הגבול הנקודתי של פונקציות עולות חזק הוא עולה חזק.
(ה) הגבול הנקודתי של פונקציות חסומות הוא חסום.

6. תהי $f : \mathbb{R} \to \mathbb{R}$ ונגדיר $a_n(x) = f(x+n)$, $b_n(x) = f(nx)$, $c_n(x) = f(\frac{x}{n})$. מצאו תנאים שיבטיחו התכנסות נקודתית ב־$\mathbb{R}$ של כל אחת מהסדרות $(a_n), (b_n), (c_n)$.

7. תהי $(f_n)_{n=1}^{\infty}$ סדרת פונקציות ליפשיציות בעלות קבוע M, דהיינו לכל x', x'' בתחום ולכל n מתקיים $|f_n(x') - f_n(x'')| \leq M|x' - x''|$. נניח $f_n \to f$ נקודתית ב־$[a, b]$. הוכיחו שגם f ליפשיצית עם קבוע M.

8. תהי $f_0 : [0,1] \to \mathbb{R}$ פונקציה אינטגרבילית. נגדיר ברקורסיה סדרת פונקציות, $f_n(x) = \int_0^x f_{n-1}$. האם (f_n) מתכנסת נקודתית ב־$[0,1]$, ואם כן לאן?

9. (*) בשאלה זו נראה שעבור מושג ההתכנסות הנקודתית של סדרת פונקציות משפט בולצאנו־ויירשטראס אינו תקף. נגדיר באינדוקציה סדרת פונקציות $f_n : [0,1] \to \mathbb{R}$ באופן הבא:

$$f_0(x) = \begin{cases} 0 & x \in [0, \frac{1}{2}) \\ 1 & x \in [\frac{1}{2}, 1] \end{cases}$$

ובהנחה שהגדרנו את f_n נגדיר

$$f_{n+1}(x) = \begin{cases} f_n(2x) & x \in [0, \frac{1}{2}) \\ f_n(2x-1) & x \in [\frac{1}{2}, 1] \end{cases}$$

הוכיחו ש־(f_n) אינה מתכנסת נקודתית ב־$[0,1]$, ויתר על כן אין לה תת־סדרה המתכנסת נקודתית ב־$[0,1]$.

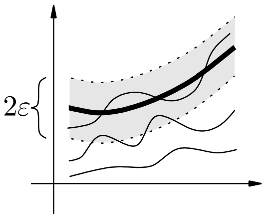
10.2 התכנסות במידה שווה

התכנסות נקודתית של סדרת פונקציות פירושה שבכל נקודה בנפרד התמונות מתכנסות. בסעיף זה נפתח מושג אחר של התכנסות של סדרת פונקציות, המבוסס על הרעיון שפונקציות הן קרובות זו לזו אם הן קרובות בו־זמנית בכל נקודה.

הגדרה 10.2.1 תהי $(f_n)_{n=1}^{\infty}$ סדרת פונקציות המוגדרות כולן בקבוצה $A \subseteq \mathbb{R}$ ותהי f פונקציה המוגדרת ב־A. נאמר ש־f היא **הגבול במידה שווה** (uniform limit) של הסדרה (f_n) ב־A, או בקיצור: הגבול במ"ש של (f_n) ב־A, אם לכל $\varepsilon > 0$ מתקיים שלכל n גדול מספיק המרחק של f_n מ־f בכל נקודה ב־A קטן יותר מ־ε. במקרה זה רושמים $f = \lim_{n\to\infty} f_n$ או $f_n \to f$ ומציינים שההתכנסות היא ב־A ושהיא במ"ש.

באופן מפורש יותר, $f_n \to f$ במ"ש ב־A אם לכל $\varepsilon > 0$ קיים N, כך שלכל $n \geq N$ ולכל $x \in A$ מתקיים $|f(x) - f_n(x)| < \varepsilon$. בסימנים,

$$f_n \to f \text{ במ"ש ב־ } A \iff \forall\varepsilon\, \exists N\, \forall n > N\, \forall x \in A\, |f_n(x) - f(x)| < \varepsilon$$



איור 10.2.1

מבחינה גאומטרית, התכנסות במ"ש של f_n ל־f פירושה שלכל פס ברוחב ε שנצייר סביב הגרף של f, החל ממקום מסוים בסדרה הגרפים של כל הפונקציות f_n נמצאים בתוך הפס, כפי שמתואר באיור 10.2.1.

כדאי להשוות בין ההגדרה החדשה להגדרת ההתכנסות הנקודתית:

$$f_n \to f \text{ נקודתית ב־} A \quad \Longleftrightarrow \quad \forall\varepsilon\, \forall x \in A\, \exists N\, \forall n > N\, |f_n(x) - f(x)| < \varepsilon$$

שימו לב להבדל הדק אך המכריע בין ההגדרות. במקרה של התכנסות נקודתית, N יכול להיות תלוי ב־x. זהו הביטוי הפורמלי לכך שבנקודות x שונות קצב ההתכנסות של סדרת המספרים $(f_n(x))$ ל־$f(x)$ יכול להיות שונה. לעומת זאת במקרה של התכנסות במ"ש N אינו תלוי ב־x, ואותו N טוב לכל הנקודות ב־A.

מכל מקום, אם $f_n \to f$ במ"ש ב־A אז לכל $\varepsilon > 0$ ולכל $x \in A$ יש N (שבמקרה מתאים גם לנקודות אחרות) כך שלכל $n > N$ מתקיים $|f(x) - f_n(x)| < \varepsilon$. מכאן אנו מסיקים את הגרירה הבסיסית בין התכנסות במ"ש להתכנסות נקודתית:

טענה 10.2.2 אם $f_n \to f$ במ"ש ב־A אז $f_n \to f$ נקודתית ב־A.

כפי שנראה בדוגמאות להלן, הגרירה ההפוכה אינה נכונה: התכנסות נקודתית אינה גוררת התכנסות במ"ש.

נציין שלטענה האחרונה משמעות מעשית: בבואנו להוכיח שסדרת פונקציות (f_n) מתכנסת במ"ש יש רק מועמד אחד להיות הגבול, והוא הגבול הנקודתי. לכן בבואנו למצוא את הגבול במ"ש של סדרה (f_n) בקבוצה כלשהי נמצא קודם כל את הגבול הנקודתי f של הסדרה. אם זה אינו קיים ממילא אין גבול במ"ש. אחר־כך נבדוק האם ההתכנסות ל־f ב־A היא גם במ"ש.

דוגמאות

1. יהיו $f_n : [0,1] \to \mathbb{R}$ נתונות על ידי $f_n(x) = \frac{x}{n}$. אז $f_n \to 0$ נקודתית ב־$[0,1]$. נראה כעת שההתכנסות היא גם במ"ש. יהי $\varepsilon > 0$. עלינו למצוא N כך שלכל $n > N$ ולכל $x \in [0,1]$ מתקיים

$$|f_n(x) - 0| = |\frac{x}{n}| < \varepsilon$$

אבל לכל $x \in [0,1]$ מתקיים $|\frac{x}{n}| \leq \frac{1}{n}$ ולכן לכל n די לבחור N כך שלכל $n > N$ מתקיים $n > \frac{1}{\varepsilon}$ (למשל $N = \lceil \frac{1}{\varepsilon} \rceil$).

2. יהיו $f_n(x) = \frac{x}{n}$ כמו בדוגמה הקודמת, אך כעת נחשוב עליהן כפונקציות המוגדרות בכל הישר. עדיין מתקיים ש־ $f_n \to 0$ נקודתית ב־$\mathbb{R}$, אבל כעת ההתכנסות אינה במידה שווה. כדי להוכיח זאת די שנראה שלכל $N > 0$ אפשר למצוא מספר $n > N$ ונקודה $x \in \mathbb{R}$ עבורם

$$|f_n(x) - 0| = |\frac{x}{n}| \geq 1$$

ואמנם בהינתן N יהיו $x = n = N + 1$. אז $|f_n(x) - 0| = |\frac{n+1}{n}| \geq 1$, כנדרש.

3. יהיו $g_n : [0,1] \to \mathbb{R}$ נתונות על ידי $g_n(x) = x^n$. ראינו כבר בדוגמה (3) בעמוד 465 ש־ $g_n \to g$ נקודתית ב־$[0,1]$, כאשר

$$g(x) = \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1 \end{cases}$$

נראה שההתכנסות אינה במ"ש. נשים לב שכל אחת מהפונקציות g_n רציפה, ומקבלת ב־1 את הערך 1. לכן לכל n יש סביבה שמאלית של 1 שבה g_n גדולה מ־$\frac{1}{2}$, ומאידך $g(x) = 0$. בפרט לכל n יש $x \in [0,1)$ כך ש־

$$|g_n(x) - g(x)| \geq \frac{1}{2}$$

וזה פוסל התכנסות במ"ש של g_n ל־g (באופן מפורש יותר, בהינתן n אפשר לבחור $x = \sqrt[n]{\frac{1}{2}}$ ואז $|g_n(x) - g(x)| = \frac{1}{2}$).

4. יהי $0 < r < 1$ ונראה שהפונקציות g_n מהדוגמה הקודמת מתכנסות לאפס במ"ש ב־$[0,r]$. נראה שלכל $0 < r < 1$ מתקיים $g_n \to 0$ במ"ש על $[0,r]$. ואמנם, יהי $\varepsilon > 0$. עלינו להראות שלכל n גדול מספיק מתקיים

$$|x^n| < \varepsilon$$

לכל $x \in [0,r]$. לכל x כזה מתקיים $x \leq r$ ולכן $|x^n| \leq r^n$. לכן אם נראה שקיים N כך שלכל $n > N$ מתקיים $r^n < \varepsilon$, נקבל שלכל $x \in [0,r]$ מתקיים $|x^n| \leq r^n < \varepsilon$. העובדה שיש N כזה היא מסקנה מכך ש־ $r^n \to 0$.

5. בסעיף 7.14 הוכחנו שאם f רציפה בקטע $[a,b]$ אז יש סדרה של פוליגונים f_n כך ש־ $f_n \to f$ במ"ש ב־$[a,b]$ (אם כי שם לא ניסחנו זאת כך!).

שימו לב לתופעה מעניינת: הפונקציות x^n מדוגמה (4) למעלה מתכנסות במ"ש לאפס בכל תת־קטע סגור של $[0,1)$ אך הן אינן מתכנסות במ"ש בקטע $[0,1)$ כולו, שהוא איחוד הקטעים הללו. מכך נובע שלא תמיד קיימת קבוצה גדולה ביותר שבה סדרת פונקציות מתכנסת במ"ש, וממילא אין סיכוי להגדיר "תחום התכנסות במ"ש" באופן מקביל להגדרה של תחום התכנסות נקודתית של סדרת פונקציות.

בדוגמאות הקודמות השתמשנו במובלע באפיון הבא של התכנסות במ"ש:

משפט 10.2.3 יהיו f, f_n פונקציות שכולן מוגדרות ב־A. אז $f_n \to f$ במ"ש ב־A אמ"מ קיימת סדרת מספרים (ε_n) השואפת ל־0 כך שלכל n מתקיים

$$\sup_{x \in A} |f(x) - f_n(x)| < \varepsilon_n$$

השקילות מידית מההגדרות (בדקו את הפרטים!). כדוגמה ננסח מחדש את ההוכחה בדוגמה (4) במונח המשפט. באותה דוגמה ראינו שאם $r > 0$ אז לכל n מתקיים $\sup_{x\in[0,r]} |x^n - 0| \le r^n$. כיוון ש־ $0 < r < 1$ הרי $r^n \to 0$ ולכן $x^n \to 0$ במ"ש ב־$[0,r]$.

נעבור לדון בהתכנסות במ"ש של טורים:

הגדרה 10.2.4 תהי $(u_k)_{k=1}^{\infty}$ סדרת פונקציות המוגדרות בקבוצה $A \subseteq \mathbb{R}$. יהיו $S_N : A \to \mathbb{R}$ הסכומים החלקיים $S_N = \sum_{k=1}^{N} u_k$ ותהי S פונקציה המוגדרת ב־A. נאמר שהטור $\sum_{k=1}^{\infty} u_k$ **מתכנס במידה שווה** ל־S ב־A אם $S_N \to S$ במ"ש ב־A.

באופן מפורש, $S = \sum_{k=1}^{\infty} u_k$ במ"ש ב־A אם לכל $\varepsilon > 0$ קיים N כך שלכל $n > N$ ולכל $x \in A$ מתקיים

$$|S(x) - \sum_{k=1}^{n} u_k(x)| < \varepsilon$$

כיוון שהתכנסות במ"ש ב־A של סדרת הסכומים החלקיים גוררת התכנסות נקודתית שלהם שם, הרי שאם $\sum_{k=1}^{\infty} u_k$ מתכנס במ"ש ל־S ב־A אז ההתכנסות היא גם נקודתית.

דוגמאות

1. יהיו $g_n(x) = x^n$. אם נסמן $S_N(x) = \sum_{k=0}^{N} g_k(x)$ אז מהנוסחה לסכום גאומטרי מתקיים $S_N(x) = \frac{1-x^{N+1}}{1-x}$. אם $|x| < 1$ אז $S_N(x) \to \frac{1}{1-x}$ ולכן $\sum_{k=1}^{\infty} g_k$ מתכנס נקודתית ל־$\frac{1}{1-x}$ ב־$(-1,1)$.

   ההתכנסות אינה במ"ש ב־$(-1,1)$. נבחר למשל $\varepsilon = 1$ (כל מספר אחר מתאים גם־כן). מכיוון ש־

   $$|S_N(x) - \frac{1}{1-x}| = |\frac{x^{N+1}}{1-x}|$$

   לכל N יש $x \in (-1,1)$ כך שביטוי זה גדול מ־ε. למשל, בהינתן N אפשר לבחור $x = \sqrt[N+1]{\frac{3}{4}}$, ואז $|S_N(x) - \frac{1}{1-x}| \ge 3$.

   מאידך, אם $0 < r < 1$ אז $\sum_{k=0}^{\infty} g_k$ מתכנס ל־$\frac{1}{1-x}$ במ"ש ב־$[-r,r]$. שכן לכל $x \in [-r,r]$ מתקיים

   $$|S_N(x) - \frac{1}{1-x}| = |\frac{x^N}{1-x}| \le \frac{r^N}{1-r}$$

   הסדרה מימין שואפת ל־0 ולכן לפי משפט 10.2.3, $S_N \to \frac{1}{1-x}$ במ"ש ב־$[-r,r]$, כפי שרצינו.

2. נביט בטור

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n+x^2}$$

לכל $x \in \mathbb{R}$ זהו טור לייבניץ, ולכן הטור מתכנס נקודתית. נסמן את פונקציית הגבול ב־S. ממשפט לייבניץ 6.4.6 אנו יודעים שאפשר לחסום את הזנב ה־N של טור לייבניץ על ידי ערכו המוחלט של המחובר ה־$N+1$. לכן לכל $x \in \mathbb{R}$ מתקיים

$$|S(x) - \sum_{n=1}^{N} \frac{(-1)^n}{n+x^2}| \leq \frac{1}{(N+1)+x^2} \leq \frac{1}{N+1}$$

לכל $\varepsilon > 0$ יש N כך שלכל $n > N$ מתקיים $\frac{1}{n+1} < \varepsilon$, ואז לכל $x \in \mathbb{R}$ מתקיים

$$|S(x) - \sum_{n=1}^{N} \frac{(-1)^n}{n+x^2}| < \varepsilon$$

מכאן שהטור מתכנס במ"ש ב־$\mathbb{R}$ (אפשר גם להפעיל את משפט 10.2.3).

3. כרגיל, אפשר להמיר בעיית התכנסות במ"ש של סדרת פונקציות בבעיית התכנסות במ"ש של טור פונקציות. אנו משאירים את הפרטים כתרגיל.

נסיים במספר כלים משוכללים יותר לבדיקת התכנסות. השוו את המשפט הבא עם תנאי קושי להתכנסות סדרות (משפט 5.9.5):

משפט 10.2.5 (תנאי קושי להתכנסות במ"ש של סדרות פונקציות) תהי $(f_n)_{n=1}^{\infty}$ סדרת פונקציות המוגדרות בקבוצה $A \subseteq \mathbb{R}$. אז הסדרה (f_n) מתכנסת במ"ש ב־A אמ"מ לכל $\varepsilon > 0$ קיים N כך שלכל $n,m > N$ ולכל $x \in A$ מתקיים $|f_n(x) - f_m(x)| < \varepsilon$.

הוכחה נראה שהתנאי הכרחי. נניח שהסדרה מתכנסת במידה שווה לפונקציית הגבול $f(x)$. יהי $\varepsilon > 0$. מהגדרת ההתכנסות במידה שווה קיים N טבעי כך שלכל $n > N$ ולכל $x \in A$ מתקיים $|f_n(x) - f(x)| \leq \frac{\varepsilon}{2}$. לכן אם $n,m > N$ ו־ $x \in A$ אז

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

ותנאי קושי מתקיים.

נראה שהתנאי מספיק. תחילה נוכיח שהסדרה מתכנסת נקודתית. נשים לב שמההנחה שסדרת הפונקציות מקיימת את תנאי קושי נובע שלכל $x_0 \in A$ סדרת המספרים $(f_n(x_0))_{n=1}^{\infty}$ מקיימת את תנאי קושי להתכנסות סדרות מספרים, ולכן מתכנסת. מכאן שהסדרה (f_n) מתכנסת נקודתית ב־A. נסמן את גבולה ב־f.

נותר להוכיח שההתכנסות של (f_n) ל־f היא גם במ"ש ב־A. יהי $\varepsilon > 0$. מתנאי קושי קיים N כך שלכל $m,n > N$ ולכל $x \in A$ מתקיים $|f_n(x) - f_m(x)| < \frac{\varepsilon}{2}$. בהינתן אינדקס $n > N$ ונקודה $x \in A$, לכל בחירה של $m > N$ מתקיים

$$|f(x) - f_n(x)| \leq |f(x) - f_m(x)| + |f_m(x) - f_n(x)| < |f(x) - f_m(x)| + \frac{\varepsilon}{2}$$

שימו לב ש־m אינו מופיע כלל באגף שמאל. נשאיף את m לאינסוף בשני האגפים ונקבל

$$|f(x) - f_n(x)| \leq \lim_{m\to\infty} |f(x) - f_m(x)| + \frac{\varepsilon}{2} = 0 + \frac{\varepsilon}{2} < \varepsilon$$

כי $f_m(x) \to f(x)$ לפי הגדרת f. אבל $x \in A$ הייתה נקודה שרירותית, וקיבלנו שלכל $n > N$ ולכל $x \in A$ מתקיים $|f(x) - f_n(x)| \leq \varepsilon$. מכאן נובע ש־ $f_n \to f$ במ"ש ב־A. ■

לדוגמה נשתמש בתנאי קושי כדי להוכיח שסדרת הפונקציות $e_n(x) = (1 + \frac{x}{n})^n$ מתכנסת במ"ש ל־e^x בכל קטע $[a,b]$ (אנו כבר יודעים שהסדרה מתכנסת נקודתית ל־e^x בכל הישר). יהי $\varepsilon > 0$. עלינו להראות שיש N כך שלכל $m,n > N$ מתקיים $|e_n(x) - e_m(x)| < \varepsilon$ לכל $x \in [a,b]$.

מתרגיל (7) בעמוד 347 נובע שיש N_0 כך שלכל $m,n > N_0$, המקסימום של הפונקציה $|e_n - e_m|$ ב־$[a,b]$ מתקבל בקצוות הקטע. מכיוון שסדרת המספרים $(e_n(a))_{n=1}^{\infty}$ מתכנסת היא מקיימת את תנאי קושי, ולכן קיים N_1 כך שלכל $m,n > N_1$ מתקיים $|e_n(a) - e_m(a)| < \varepsilon$. כמו־כן הסדרה $(e_n(b))_{n=1}^{\infty}$ מתכנסת ולכן יש N_2 כך שלכל $m,n > N_2$ מתקיים $|e_n(b) - e_m(b)| < \varepsilon$. יהי אם כן $N = \max\{N_0, N_1, N_2\}$ ויהיו $m,n > N$. אז לכל $x \in [a,b]$,

$$\begin{aligned} |e_n(x) - e_m(x)| &\leq \max\{|e_n(a) - e_m(a)|\,,\, |e_n(b) - e_m(b)|\} \\ &< \max\{\varepsilon, \varepsilon\} \\ &= \varepsilon \end{aligned}$$

כנדרש.

משפט 10.2.6 (תנאי קושי להתכנסות במ"ש של טורי פונקציות) תהי $(u_k)_{k=1}^{\infty}$ סדרת פונקציות המוגדרות בקבוצה $A \subseteq \mathbb{R}$. אז הטור $\sum_{k=1}^{\infty} u_k$ מתכנס במידה שווה ב־A אמ"מ לכל $\varepsilon > 0$ קיים N כך שלכל $x \in A$ ולכל $m > n > N$ מתקיים $|\sum_{k=n}^{m} u_k(x)| < \varepsilon$.

המשפט נובע מיד מהפעלת תנאי קושי לסדרות פונקציות על סדרת הסכומים החלקיים של הטור.

תנאי קושי מאפשר לנסח עבור טורי פונקציות מבחן השוואה להתכנסות במ"ש:

משפט 10.2.7 (מבחן M של ויירשטראס) תהי $(u_n)_{n=1}^{\infty}$ סדרת פונקציות המוגדרות בקבוצה $A \subseteq \mathbb{R}$. נניח שקיים טור מתכנס של מספרים חיוביים $\sum_{n=0}^{\infty} M_n$ כך שלכל n ולכל $x \in A$ מתקיים

$$|u_n(x)| \leq M_n$$

אז $\sum_{n=1}^{\infty} u_n$ מתכנס במ"ש ב־A.

הוכחה יהי $\varepsilon > 0$. לפי תנאי קושי להתכנסות טורי מספרים קיים N כך שלכל $m > n > N$,

$$M_n + M_{n+1} + \cdots M_m < \varepsilon$$

עבור n, m כאלה מתקיים לכל $x \in A$

$$\begin{aligned}|u_n(x) + u_{n+1}(x) + \cdots + u_m(x)| &\leq |u_n(x)| + |u_{n+1}(x)| + \cdots + |u_m(x)| \\ &\leq M_n + M_{n+1} + \cdots M_m < \varepsilon\end{aligned}$$

ולפי תנאי קושי להתכנסות במ"ש של טורי פונקציות (משפט 10.2.6) הטור $\sum_{k=1}^{\infty} u_k$ מתכנס במידה שווה. ■

שימו לב שאם הנחות המשפט מתקיימות אז הן מתקיימות גם לגבי הטור $\sum_{k=1}^{\infty} |u_k|$ ולכן גם $\sum_{k=1}^{\infty} |u_k|$ מתכנס במידה שווה ב־A (ובפרט נקודתית), וקיבלנו שהטור $\sum_{k=1}^{\infty} u_k$ מתכנס בהחלט ב־A. לכן המשפט יכול לעזור רק בהכרעת ההתכנסות של טורי פונקציות שמתכנסים בהחלט.

דוגמאות

1. הטור $\sum_{n=1}^{\infty} \frac{sin(nx)}{n^2}$ מתכנס במידה שווה ב־$\mathbb{R}$ הואיל ולכל x מתקיים

$$|\frac{sin(nx)}{n^2}| \leq \frac{1}{n^2}$$

וטור החסמים $\sum_{n=1}^{\infty} \frac{1}{n^2}$ הוא טור מספרים חיובי מתכנס.

2. נביט שוב בטור $S(x) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n+x^2}$ מדוגמה (2) בעמוד 473. מתקיים

$$|\frac{(-1)^n}{n+x^2}| \leq \frac{1}{n}$$

לכל $x \in \mathbb{R}$, אבל הטור $\sum_{n=1}^{\infty} \frac{1}{n}$ מתבדר, ולכן לא נוכל להפעיל את משפט ויירשטראס. יתר על כן הטור $\sum_{n=1}^{\infty} \frac{1}{n}$ הוא טור החסמים הטוב ביותר שאפשר למצוא לאיברי הטור $\sum_{n=1}^{\infty} \frac{(-1)^n}{n+x^2}$, היות והפונקציה $|\frac{(-1)^n}{n+x^2}|$ חסומה בין $\frac{1}{n}$ ל־ $\frac{c}{n}$ עבור קבוע c התלוי ב־x (אבל לא ב־n). לכן אי אפשר להסיק ממשפט ויירשטראס שהטור מתכנס במ"ש ב־$\mathbb{R}$, אף שראינו שזה אכן המצב. אנו רואים שוב שמבחן M של ויירשטראס הוא תנאי מספיק, אך אינו תנאי הכרחי.

תרגילים

1. הוכיחו כי סדרות הפונקציות (f_n) הבאות מתכנסות במ"ש בקטע I הנתון, ומתכנסות נקודתית אבל לא במ"ש בקטע J. בכל הסעיפים $r > 0$.
   (א) $\sqrt[n]{\sin(x)}\ ,\ I=[r,\frac{\pi}{2}]\ ,\ J=(0,\frac{\pi}{2}]$.
   (ב) $\sin^n(x)\ ,\ I=[0,\frac{\pi}{2}-r]\ ,\ J=[0,\frac{\pi}{2})$
   (ג) $\frac{x+n}{n}\ ,\ I=[a,b]\ ,\ J=(-\infty,\infty)$.
   (ד) $\frac{x}{x+n}\ ,\ I=[0,b]\ ,\ J=[0,\infty)$.
   (ה) $\frac{nx}{1+n^2x^2}\ ,\ I=[r,\infty)\ ,\ J=(0,\infty)$.
   (ו) $\frac{nx}{1+nx}\ ,\ I=[r,\infty)\ ,\ J=(0,\infty)$.
   (ז) $n^2x^2e^{-nx}\ ,\ I=[r,\infty)\ ,\ J=(0,\infty)$.
   (ח) $\frac{1}{n}\log(1+nx)\ ,\ I=[0,r]\ ,\ J=[0,\infty)$.

2. הוכיחו או הפריכו התכנסות במידה שווה של הסדרות הבאות בתחומים הנתונים:
   (א) $\sqrt{n}e^{-nx}$ ב־$[0,1]$.
   (ב) $n\log(1+\frac{1}{nx})$ ב־$[1,4]$.
   (ג) $\frac{1}{n}\cos(n^2x)$ ב־$(-\infty,\infty)$.
   (ד) $\frac{\sin(nx)}{1+nx}$ ב־$(0,\infty)$ וב־(r,∞) עבור $r>0$.
   (ה) $\frac{x^n}{1+x^n}$ ב־$[0,1]$ וב־ $[0,1-\delta)$ עבור $\delta>0$.

3. הוכיחו או הפריכו:
   (א) אם $f_n \to f$ במ"ש ב־A ואם $B \subseteq A$ אז $f_n \to f$ במ"ש ב־B.
   (ב) אם $f_n \to f$ במ"ש בקבוצות $D_1,\ldots,D_n$ אז $f_n \to f$ במ"ש באיחוד $\bigcup_{k=1}^n D_k = D_1 \cup D_2 \cup \ldots \cup D_n$.
   (ג) אם $f_n \to f$ במ"ש בקבוצות $D_1, D_2, \ldots$ אז $f_n \to f$ במ"ש באיחוד $\bigcup_{k=1}^\infty D_k = D_1 \cup D_2 \cup \ldots$.
   (ד) אם $f_n \to f$ במ"ש ב־$[a,b)$ ונניח גם ש־ $f_n(b) \to f(b)$ אז $f_n \to f$ במ"ש ב־$[a,b]$.

4. יהיו f_n פונקציות חסומות ב־A ונניח ש־ $f_n \to f$ במ"ש ב־A. הוכיחו ש־f חסומה.

5. נניח שהסדרות $(f_n)_{n=1}^\infty$ ו־ $(g_n)_{n=1}^\infty$ מתכנסות במידה שווה בתחום משותף D.
   (א) הראו שהסדרה $(f_n+g_n)_{n=1}^\infty$ מתכנסת במ"ש ב־D.
   (ב) נניח שמספר M הוא **חסם משותף** של הסדרות, כלומר $|f_n(x)| \le M$ לכל $x \in D$, וגם $|g_n(x)| \le M$ לכל $x \in D$. הראו ש־$(f_n \cdot g_n)_{n=1}^\infty$ מתכנסת במ"ש ב־D.

(ג) הראו שללא ההנחה על חסימות משותפת של הסדרות המסקנה בסעיף הקודם אינה נכונה.

6. תהי f רציפה ב־$[0,1]$. תהי $F(x) = \int_0^x f$. לכל n ולכל $x \in [0,1]$ נגדיר $F_n(x) = \frac{1}{n}\sum_{k=1}^{n} f(\frac{k}{n}x)$ הוכיחו ש־ $F_n \to F$ במ"ש על $[a,b]$ (שימו לב ש־F_n היא סכום רימן).

7. תהי $f_0 : [-a,a] \to \mathbb{R}$ פונקציה רציפה כך ש־$f_0(0) = 0$, וכך שלכל $x \in [-a,a]$ פרט ל־0 מתקיים $|f_0(x)| < |x|$. נגדיר סדרת פונקציות באופן רקורסיבי על ידי $f_n = f_0 \circ f_{n-1}$. הראו שסדרת הפונקציות $(f_n)_{n=0}^{\infty}$ מתכנסת במידה שווה לפונקציית האפס בקטע $[-a,a]$.

8. תהי $f_0 : [0,1] \to \mathbb{R}$ פונקציה אינטגרבילית ונגדיר ברקורסיה סדרת פונקציות $f_n(x) = \int_0^x f_{n-1}$ (ראו תרגיל (8) מעמוד 469). הראו שהסדרה $(f_n)_{n=0}^{\infty}$ מתכנסת לפונקציית האפס במידה שווה בקטע $[0,1]$.

9. יהיו $f_n : [a,b] \to [c,d]$ פונקציות ותהי $g : [c,d] \to \mathbb{R}$ פונקציה רציפה. הראו שאם $f_n \to f$ במ"ש אז $g \circ f_n \to g \circ f$ במ"ש. הראו שאם g אינו רציפה המסקנה אינה נכונה, ושהמסקנה גם אינה נכונה אם מחליפים את הקטע $[c,d]$ בקטע פתוח.

10. הראו כי הטורים הבאים מתכנסים במידה שווה בתחום הנתון. $0 < \alpha$ בכל הסעיפים בהם הוא מופיע.

(א) $\sum(\frac{\log(x)}{x})^n, \quad [1,\infty)$

(ב) $\sum n^2 x^n, \quad [-\frac{1}{2},\frac{1}{2}]$

(ג) $\sum \frac{e^{nx}}{5^n}, \quad (-\infty, \log(5)-\alpha]$

(ד) $\sum \frac{x^n}{n^2}, \quad [-1,1]$

(ה) $\sum (x\log(x))^n, \quad (0,1]$

(ו) $\sum \frac{x^n}{n^n}, \quad [a,b]$

(ז) $\sum \frac{\sin(nx)}{n^2+1}, \quad (-\infty,\infty)$

(ח) $\sum \frac{1}{n^x}, \quad [1+\alpha,\infty)$

11. נגדיר $u_n : \mathbb{R} \to \mathbb{R}$ על ידי

$$u_n(x) = \begin{cases} 0 & x \neq n \\ \frac{1}{n} & x = n \end{cases}$$

הוכיחו שהטור $\sum_{n=1}^{\infty} u_n$ מתכנס במ"ש ב־$\mathbb{R}$ אבל אי אפשר להוכיח זאת בעזרת מבחן M של ויירשטראס.

10.3 גבולות במ"ש של פונקציות רציפות

במבוא לפרק שאלנו אילו תכונות של פונקציות עוברות מהאיברים של סדרת פונקציות אל פונקציית הגבול. ליתר דיוק, היינו רוצים שיתקיימו משפטים מהסוג

"אם כל פונקציה f_n מקיימת תכונה P, ואם $f_n \to f$, אז גם f מקיימת את P". קיימות כמה תכונות של פונקציות שנשמרות תחת גבולות נקודתיים, כמו אי־שליליות ומונוטוניות. אולם יש גם תכונות שאינן נשמרות, וביניהן תכונות חשובות כמו חסימות, רציפות, גזירות ואינטגרביליות. בסעיף זה ובבאים נראה שכשמדובר בגבולות במ"ש, המצב טוב יותר.

בסעיף הנוכחי נתמקד בשאלה האם בהינתן פונקציות רציפות f_n שמתכנסות ל־f נובע ש־f רציפה. התשובה תלויה בסוג ההתכנסות בה מדובר. גבול נקודתי של פונקציות רציפות אינו בהכרח רציף. דוגמה (3) מעמוד 471 היא דוגמה נגדית: עבור $g_n(x) = x^n$ ראינו ש־ $g_n \to g$ נקודתית ב־$[0,1]$ כאשר

$$g(x) = \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1 \end{cases}$$

כאן g אינה רציפה ב־$[0,1]$ אף שכל ה־g_n־ים רציפות שם.

לעומת זאת התכנסות במ"ש כן מבטיחה רציפות של פונקציית הגבול:

משפט 10.3.1 תהי $(f_n)_{n=1}^{\infty}$ סדרת פונקציות המוגדרות בקטע $I \subseteq \mathbb{R}$, ומתכנסות בו במידה שווה לפונקציה f. תהי $x_0 \in I$. אם כל ה־f_n־ים רציפות ב־x_0 אז גם f רציפה ב־x_0 (אם x_0 נקודת קצה של I אנו מפרשים את הרציפות במובן החד־צדדי). בפרט, אם כל ה־f_n רציפות ב־I אז f רציפה ב־I.

הוכחה יהי $x_0 \in I$. עלינו להראות שלכל $\varepsilon > 0$ קיים $\delta > 0$ כך שלכל $x \in I$, אם $|x - x_0| < \delta$ אז $|f(x) - f(x_0)| < \varepsilon$. הרעיון הוא לבחור פונקציה f_n שקרובה מאד ל־f בכל נקודה, ולהסיק שלכל $x \in I$ המרחק בין $f(x), f(x_0)$ שווה כמעט למרחק בין $f_n(x), f_n(x_0)$. אז אפשר להשתמש ברציפות של f_n כדי להסיק שאם x קרוב מספיק ל־x_0 אז $f_n(x), f_n(x_0)$ קרובים, ולכן $f(x), f(x_0)$ קרובים.

הנה הפרטים. בהינתן $\varepsilon > 0$ נבחר n כך שלכל $x \in I$ מתקיים $|f(x) - f_n(x)| < \frac{\varepsilon}{3}$. מכאן שלכל $x \in I$ מתקיים

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \\ &< \frac{\varepsilon}{3} + |f_n(x) - f_n(x_0)| + \frac{\varepsilon}{3} \end{aligned}$$

f_n רציפה, לכן קיים $\delta > 0$ כך שאם $x \in I$ מקיים $|x - x_0| < \delta$ אז מתקיים $|f_n(x) - f_n(x_0)| < \frac{\varepsilon}{3}$. לכן לכל x כזה מתקיים

$$|f(x) - f(x_0)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

כפי שרצינו. ■

הנה התוצאה המקבילה לטורי פונקציות (הוכיחו אותה בפירוט!):

מסקנה 10.3.2 אם $\sum_{k=1}^{\infty} u_k$ טור פונקציות המתכנס במ"ש ל־S בקטע I ואם $x_0 \in I$ נקודת רציפות של כל ה־u_k־ים אז S רציפה ב־x_0. בפרט אם כל ה־u_k־ים רציפות ב־I אז S רציפה ב־I.

דוגמאות

1. מאחר שהסדרה $e_n(x) = (1+\frac{x}{n})^n$ מתכנסת במ"ש ל־e^x בכל קטע סגור (עמוד 474) וכל אחת מה־e_n־ים רציפה (כי היא פולינום!), אנו מקבלים הוכחה חדשה לכך שהפונקציה e^x רציפה בכל קטע סגור, ולכן רציפה בכל הישר.
2. אנו מקבלים הוכחה חדשה שהפונקציות $g_n(x) = x^n$ אינן מתכנסות במ"ש ב־$[0,1]$, כי כל ה־g_n־ים רציפות אבל הגבול הנקודתי של הסדרה אינו רציף ב־$[0,1]$.

הקשר בין סוג ההתכנסות לרציפות הפונקציה הגבולית קיימת כמובן רק בכיוון אחד: רציפות הפונקציה הגבולית אינה מבטיחה שההתכנסות היא במ"ש. למשל, תהי $f(x) = e^x$. נתבונן בהזזות של f ימינה, הנתונות על ידי

$$f_n(x) = f(x-n) = e^{x-n}$$

אז $f_n(x) \to 0$ לכל $x \in \mathbb{R}$ ולכן $f_n \to 0$ נקודתית. הפונקציה הגבולית בוודאי רציפה אך ההתכנסות אינה במ"ש כי $f_n(n) = 1$ לכל n. בעזרת רעיונות דומים אפשר למצוא דוגמאות של פונקציות רציפות המתכנסות נקודתית בקטע סופי לפונקציה רציפה, אבל ההתכנסות אינה במ"ש.

למרות האמור לעיל, במקרים מיוחדים רציפות של פונקציית הגבול כן מעידה על התכנסות במ"ש. כדי לנסח את המשפט בעניין זה נזדקק להגדרה:

הגדרה 10.3.3 סדרת פונקציות $(f_n)_{n=1}^\infty$ המוגדרות בקבוצה $A \subseteq \mathbb{R}$ היא סדרה **עולה** אם לכל $x \in A$ סדרת המספרים $(f_n(x))_{n=1}^\infty$ היא סדרה עולה, כלומר

$$f_{n+1}(x) \geq f_n(x)$$

לכל $n \in \mathbb{N}$. באופן דומה מגדירים סדרת פונקציות יורדת. סדרת פונקציות נקראת מונוטונית אם היא עולה או יורדת.

שימו לב שמונוטוניות של סדרת פונקציות אין פירושה שהפונקציות בסדרה הן מונוטוניות, וגם אין כל גרירה בכיוון ההפוך.

משפט 10.3.4 (משפט דיני)[2] תהי $(f_n)_{n=1}^\infty$ סדרה מונוטונית של פונקציות רציפות בקטע הסגור $[a,b]$ המתכנסת נקודתית ב־$[a,b]$ לפונקציה רציפה f. אז ההתכנסות היא במידה שווה.

[2] Ulisse Dini, 1845-1918.

הוכחה נניח בלי הגבלת הכלליות ש־(f_n) סדרה עולה, ונתבונן בסדרת ההפרשים $R_n = f - f_n$. זו סדרה יורדת של פונקציות רציפות, והיא מתכנסת נקודתית לאפס ב־$[a,b]$. בפרט $R_n \geq 0$ לכל n. אם נראה ש־ $R_n \to 0$ במ"ש ב־$[a,b]$ ינבע $f_n \to f$ במ"ש ב־$[a,b]$, כנדרש.

נניח בשלילה כי $R_n \to 0$ נקודתית אך לא במ"ש ב־$[a,b]$. אז קיים $\varepsilon > 0$ כך שלכל N קיים $n > N$ ונקודה $x_n \in [a,b]$ כך ש־

$$|R_n(x_n)| = R_n(x_n) \geq \varepsilon$$

נבחר תת־סדרה (x_{n_k}) המתכנסת לנקודה $x_0 \in [a,b]$ (הדבר אפשרי לפי משפט בולצאנו־ויירשטראס). אז $x_{n_k} \to x_0$ ולכל k מתקיים $R_{n_k}(x_{n_k}) \geq \varepsilon$.

נקבע אינדקס $m \in \mathbb{N}$. כיוון שסדרת הפונקציות (R_n) יורדת, אם $n_k > m$ אז לכל $x \in [a,b]$ מתקיים

$$R_m(x) \geq R_{n_k}(x)$$

ובפרט עבור $x = x_{n_k}$:

$$R_m(x_{n_k}) \geq R_{n_k}(x_{n_k}) \geq \varepsilon$$

לפי ההנחה R_m היא פונקציה רציפה, ולכן באי־שוויון האחרון נוכל לעבור לגבול כאשר $k \to \infty$ ונקבל

$$R_m(x_0) \geq \varepsilon$$

אי־שוויון זה נכון לכל $m \in \mathbb{N}$, בסתירה להנחה ש־ $R_m \to 0$ נקודתית ב־$[a,b]$. ■

מסקנה 10.3.5 תהי $(f_n)_{n=1}^{\infty}$ סדרה מונוטונית של פונקציות רציפות בקטע סגור $[a,b]$. אז (f_n) מתכנסת במ"ש ב־$[a,b]$ לפונקציה רציפה אמ"מ היא מתכנסת שם נקודתית לפונקציה רציפה.

משפט דיני נכון גם עבור טורי פונקציות בניסוח הבא:

משפט 10.3.6 תהי $(u_k)_{k=1}^{\infty}$ סדרת פונקציות אי־שליליות ונניח ש־ $\sum_{k=1}^{\infty} u_k$ מתכנס נקודתית בקטע $[a,b]$ ל־S. נניח גם ש־S רציפה ב־$[a,b]$. אז הטור מתכנס ל־S במ"ש ב־$[a,b]$.

המשפט האחרון נובע בקלות ממשפט דיני לסדרות, ומושאר כתרגיל.

תרגילים

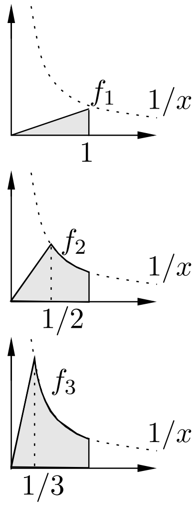
1. נניח שסדרת הפונקציות $(f_n)_{n=1}^{\infty}$ מתכנסת במידה שווה לפונקציה f בקטע $[a,b]$. תהי $x_0 \in [a,b]$ ונניח שהגבול $L_n = \lim_{x\to x_0} f_n(x)$ קיים לכל n. הוכיחו שהגבול $\lim_{n\to\infty} L_n$ קיים ושווה לגבול $\lim_{x\to x_0} f(x)$ (שימו לב שיש להוכיח גם את קיום הגבולות). נסחו טענה דומה לגבי טורי פונקציות.
2. תנו דוגמה המראה שהטענה מהשאלה הקודמת אינה נכונה אם מחליפים התכנסות במ"ש בהתכנסות נקודתית.
3. הוכיחו או הפריכו:
   (א) אם $f_n \to f$ במ"ש ואם הפונקציות f_n אינן רציפות אז f אינה רציפה.
   (ב) אם $f_n \to f$ נקודתית ב־A ואם f אינה רציפה אז ההתכנסות אינה במ"ש.
   (ג) אם $f_n \to f$ נקודתית ב־A ואם f_n, f כולן רציפות אז ההתכנסות במ"ש.
   (ד) אם $f_n \to f$ נקודתית ב־A ואם f רציפה אבל ה־f_n־ים אינן רציפות, אז ההתכנסות אינה במ"ש.
4. בסעיף 7.14 הוכחנו שפונקציה רציפה בקטע סגור היא גבול במידה שווה של פונקציות פוליגוניות. האם גם ההפך נכון?
5. (*) תהי $f : [a,b] \to \mathbb{R}$. הראו שאם f היא גבול במ"ש ב־$[a,b]$ של סדרה של פולינומים (P_n), ואם יש מספר d כך שכל הפולינומים ממעלה קטנה מ־d, אז f היא פולינום ממעלה שאינה עולה על d (רמז: הראו שאם מסמנים ב־ $a_{n,k}$ את המקדם של x^k ב־P_n אז לכל $0 \le k \le d$ הסדרה $(a_{n,k})_{n=1}^{\infty}$ חסומה. על ידי שימוש במשפט בולצאנו־ויירשטראס ומעבר לתת־סדרות הראו שאפשר להניח שלכל k הסדרה $(a_{n,k})_{n=1}^{\infty}$ מתכנסת למספר b_k. הראו ש־ $f = \sum_{k=0}^{d} b_k x^k$).

10.4 אינטגרציה איבר־איבר

כמו תכונת הרציפות, גם תכונות האינטגרביליות וערך האינטגרל אינן נשמרות תחת גבולות נקודתיים. כלומר: אם $f_n : [a,b] \to \mathbb{R}$ אינטגרביליות ואם $f_n \to f$ נקודתית ב־$[a,b]$ לא ניתן להסיק ש־f אינטגרבילית, וגם אם f אינטגרבילית ב־$[a,b]$ לא ניתן להסיק ש־ $\int_a^b f_n \to \int_a^b f$.

דוגמאות

1. התכנסות נקודתית אינה יכולה להבטיח שהפונקציה הגבולית אינטגרבילית. למשל, ישנן סדרות של פונקציות רציפות המתכנסות נקודתית לפונקציה לא



איור 10.4.1

חסומה, שבוודאי אינה אינטגרבילית. דוגמה כזו היא הסדרה

$$f_n(x) = \begin{cases} n^2x & 0 \le x \le \frac{1}{n} \\ \frac{1}{x} & \frac{1}{n} < x \le 1 \end{cases}$$

(ראו איור 1). הגבול הנקודתי של (f_n) הוא

$$f(x) = \begin{cases} 0 & x = 0 \\ \frac{1}{x} & 0 < x \le 1 \end{cases}$$

אשר אינה חסומה ובפרט אינה אינטגרבילית (אפילו כאינטגרל לא אמתי). יתר על כן

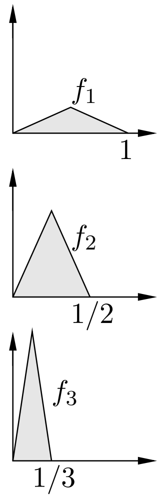
$$\int_0^1 f_n(x)dx = \int_0^{\frac{1}{n}} n^2xdx + \int_{\frac{1}{n}}^1 \frac{dx}{x} = \frac{1}{2} + \log(n)$$

וזו סדרה שלא קיים לה גבול סופי.

2. דוגמה קיצונית עוד יותר היא דוגמה (4) מעמוד 465. שם ראינו שיש פונקציות $r_n : [0,1] \to \mathbb{R}$ כך שכל r_n נבדלת מפונקציית האפס במספר סופי של נקודות, ולכן אינטגרבילית, אבל הגבול הנקודתי $\lim_{n\to\infty} r_n$ קיים ב־$[0,1]$ ושווה לפונקציית דירכלה, שאינה אינטגרבילית באף תת־קטע.

3. גם במקרה ש־ $s_n \to s$ נקודתית בקטע $[a,b]$ וכל הפונקציות אינטגרביליות ב־$[a,b]$ ערך האינטגרל אינו בהכרח נשמר בגבול, כלומר לא מתקיים בהכרח $\int_a^b s_n \to \int_a^b s$. למשל, יהיו $s_n : [0,1] \to \mathbb{R}$ מוגדרות על ידי

$$s_n(x) = \begin{cases} 4n^2x & x \in [0, \frac{1}{2n}] \\ 4n - 4n^2x & x \in (\frac{1}{2n}, \frac{1}{n}) \\ 0 & x \in [\frac{1}{n}, 1] \end{cases}$$



איור 10.4.2

(ראו איור 10.4). אלה פונקציות רציפות המתכנסות נקודתית לאפס. כדי לראות זאת, נשים לב ש־ $s_n(x) = 0$ לכל $\frac{1}{n} \le x \le 1$, ולכל $x > 0$ אם $n > \frac{1}{x}$ אז $s_n(x) = 0$ וממילא $\lim_{n\to\infty} s_n(x) = 0$. מאידך, $s_n(0) = 0$ לכל n ולכן $s_n(0) \to 0$.

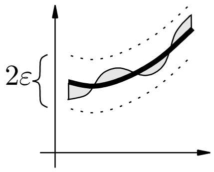
מצד שני, חישוב קל מראה ש־ $\int_0^1 s_n = 1$ לכל n. אם נסמן ב־s את פונקציית האפס אז קיבלנו ש־ $s_n \to s$ נקודתית ב־$[0,1]$ אבל $\int_0^1 s_n \not\to \int_0^1 s$.

אינטגרביליות ואינטגרלים נשמרים טוב יותר תחת גבולות במידה שווה. נדחה את הדיון באינטגרביליות של פונקציית הגבול להמשך ונראה תחילה שאם פונקציית הגבול אינטגרבילית אז ערך האינטגרל נשמר תחת גבולות במ"ש.

משפט 10.4.1 (רציפות האינטגרל) אם (f_n) היא סדרת פונקציות אינטגרביליות בקטע $[a,b]$ המתכנסת במ"ש ב־$[a,b]$ לפונקציה f שגם היא אינטגרבילית ב־$[a,b]$, אז

$$\lim_{n\to\infty}\int_a^b f_n = \int_a^b f$$

ובפרט הגבול באגף שמאל קיים.



איור 10.4.3 פונקציות הנבדלות בכל נקודה בפחות מ־ε. התחום המודגש מייצג את הפרש השטחים

הוכחה מבחינה גאומטרית המשפט ברור: עבור n־ים גדולים מספיק הגרף של f_n נמצא בתוך פס בעובי ε סביב הגרף של f, ולכן ההפרש בין השטחים שמתחת לגרפים אינו עולה על השטח של הפס, שהוא המכפלה של אורך הקטע $[a,b]$ עם גובה הפס (ראו איור 10.4.3).

באופן פורמלי, לכל $\varepsilon > 0$ קיים N כך שלכל $n > N$ ולכל $t \in [a,b]$ מתקיים

$$|f_n(t) - f(t)| < \varepsilon$$

ולכן לכל $n > N$ הפונקציה $|f_n - f|$ קטנה מ־ε בכל נקודה ולכן

$$|\int_a^b (f_n - f)| \le \int_a^b |f_n - f| \le (b-a)\cdot\varepsilon$$

מכאן נובע

$$\lim_{n\to\infty} |\int_a^b (f_n - f)| = 0$$

כלומר $\lim_{n\to\infty}(\int_a^b f_n - \int_a^b f) = 0$, והמשפט נובע. ■

דוגמאות

1. תהי $f_n(x) = \cos\frac{x}{n}$. אז $f_n \to 1$ במ"ש ב־$[0,1]$ (הוכיחו!). לפי המשפט היסודי מתקיים

$$\int_0^1 f_n = n(\sin\frac{1}{n} - 0) = n\sin\frac{1}{n}$$

וכמובן $\int_0^1 1 = 1$ ולכן ההתכנסות $\int_a^b f_n \to \int_a^b 1$ גוררת את השוויון

$$\lim_{n\to\infty} n\sin\frac{1}{n} = 1$$

2. יהיו $e_n : [a,b] \to \mathbb{R}$ נתונות על ידי $e_n(x) = (1+\frac{x}{n})^n$. ראינו בעמוד 474 ש־$e_n \to \exp$ במ"ש ב־$[a,b]$. קל לוודא ש־

$$\int (1+\frac{x}{n})^n dx = \frac{n}{n+1}(1+\frac{x}{n})^{n+1}$$

(בדקו!) מכאן ש־

$$\begin{aligned}
\int_a^b e^x dx &= \lim_{n\to\infty} \int_a^b e_n(x)dx \\
&= \lim_{n\to\infty} (\frac{n}{n+1}(1+\frac{a}{n})^{n+1} - \frac{n}{n+1}(1+\frac{b}{n})^{n+1}) \\
&= e^b - e^a
\end{aligned}$$

כך הצלחנו לחשב את האינטגרל של e^x ב־$[a,b]$ מבלי להשתמש בתכונות הדיפרנציאליות של האקספוננט.

הערה מהדוגמה האחרונה והמשפט היסודי אפשר לתת הוכחה חדשה לזהות $\exp' = \exp$.

3. אם נציב x^2 במקום x בפונקציות e_n נקבל ש־$e^{x^2} \to (1+\frac{x^2}{n})^n$ נקודתית וקל להראות שההתכנסות היא במ"ש בכל קטע סגור (הוכיחו שזה נובע מההתכנסות במ"ש של e_n בקטעים סגורים!). כפי שהזכרנו בסעיף 9.7, אין פונקציה קדומה אלמנטרית ל־e^{x^2}. לעומת זאת אנו יודעים כעת ש־

$$\int_0^1 e^{x^2} dx = \lim_{n\to\infty} \int_0^1 (1+\frac{x^2}{n})^n dx$$

ובעזרת משפט הבינום ניתן לפתוח סוגריים ולחשב את האינטגרל בגבול מימין. כך מקבלים ביטוי עבור $\int_0^1 e^{-x^2} dx$ כגבול של ביטויים חשבוניים פשוטים.

המשפט הקודם כלל את ההנחה שפונקציית הגבול אינטגרבילית. הנחה זו מיותרת, כפי שמראה המשפט הבא:

משפט 10.4.2 אם (f_n) היא סדרת פונקציות אינטגרביליות ב־$[a,b]$ המתכנסת במ"ש ב־$[a,b]$ לפונקציה f, אז f אינטגרבילית ב־$[a,b]$.

הערה אם (f_n) סדרה של פונקציות רציפות ואם $f_n \to f$ במ"ש ב־$[a,b]$ אז לפי משפט 10.3.1 גם f רציפה ב־$[a,b]$, ולכן אינטגרבילית שם. אבחנה זו מספיקה ברוב המקרים שניתקל בהם. המשפט הנוכחי מוסיף מידע רק כשמדובר בסדרת פונקציות שאינן רציפות.

הוכחה נשתמש בסימונים לסכום תחתון ועליון מסעיף 9.1. די שנראה שלכל $\varepsilon > 0$ יש חלוקה P של $[a,b]$ כך ש־ $\overline{s}(f,P) - \underline{s}(f,P) < \varepsilon$.

יהי $\varepsilon > 0$. קיים n כך ש־ $|f(x) - f_n(x)| < \varepsilon$ לכל $x \in [a,b]$, או במילים אחרות,

$$f_n(x) - \varepsilon \leq f(x) \leq f_n(x) + \varepsilon$$

לכל $x \in [a,b]$. לכן לכל חלוקה P של $[a,b]$ מתקיים

$$\underline{s}(f_n - \varepsilon, P) \leq \underline{s}(f,P) \leq \overline{s}(f,P) \leq \overline{s}(f_n + \varepsilon, P)$$

אבל

$$\begin{aligned} \underline{s}(f_n - \varepsilon, P) &= \underline{s}(f_n, P) - \varepsilon(b-a) \\ \overline{s}(f_n + \varepsilon, P) &= \overline{s}(f_n, P) + \varepsilon(b-a) \end{aligned}$$

(מדוע?) הצבת שוויונות אלה באי־שוויון הקודם נותן

$$\underline{s}(f_n, P) - \varepsilon(b-a) \leq \underline{s}(f,P) \leq \overline{s}(f,P) \leq \overline{s}(f_n, P) + \varepsilon(b-a)$$

ולכן

$$\overline{s}(f,P) - \underline{s}(f,P) \leq \overline{s}(f_n, P) - \underline{s}(f_n, P) + 2\varepsilon(b-a)$$

מכיוון ש־f_n אינטגרבילית, יש חלוקה P של $[a,b]$ כך ש־ $\overline{s}(f_n, P) - \underline{s}(f_n P) < \varepsilon$ ולאותה חלוקה מתקיים

$$\overline{s}(f,P) - \underline{s}(f,P) < 2\varepsilon(b-a) + \varepsilon = (1 + 2(b-a))\varepsilon$$

וזה מספיק. ■

הנה הניסוח של המשפטים האחרונים עבור טורי פונקציות:

משפט 10.4.3 (אינטגרציה איבר־איבר) אם (u_n) סדרת פונקציות אינטגרביליות ב־$[a,b]$ והטור $\sum_{n=1}^{\infty} u_n$ מתכנס במידה שווה ב־$[a,b]$ לפונקציה S, אז S אינטגרבילית ב־$[a,b]$ ומתקיים

$$\int_a^b S = \sum_{n=1}^{\infty} (\int_a^b u_n)$$

הוכחה נסמן $S_N = \sum_{n=1}^{N} u_n$. לפי ההנחה $S_N \to S$ במ"ש, וממילא לכל N מתקיים $\int_a^b S_N = \sum_{n=1}^{N} \int_a^b u_n$, לכן

$$\int_a^b S = \lim_{n\to\infty} \int_a^b S_N = \lim_{n\to\infty} \sum_{n=1}^{N} \int_a^b u_n = \sum_{n=1}^{\infty} \int_a^b u_n$$

כנדרש. ■

אפשר לרשום את מסקנת המשפט כך:

$$\int_a^b (u_1 + u_2 + \cdots) = \int_a^b u_1 + \int_a^b u_2 + \cdots$$

השוויון המקביל לסכומים סופיים נובע מתכונת הלינאריות של האינטגרל. המשפט נחוץ כדי לטפל בסכומים אינסופיים. שימו לב שבמהלך הוכחת המשפט השתמשנו בגרסה לסכומים סופיים.

תרגילים

1. נגדיר בקטע $[-1,1]$:

$$f_n(x) = \frac{n^2x^2}{1+n^2x^2}$$

(א) חשבו את הפונקציה הגבולית בקטע: $f(x) = \lim_{n\to\infty} f_n(x)$

(ב) הוכיחו כי התכנסות הסדרה $f_n(x)$ **אינה** מתכנסת במידה שווה בקטע, ולמרות זאת

$$\lim_{n\to\infty} \int_{-1}^1 f_n(x)dx = \int_{-1}^1 f(x)dx$$

(ג) הוכיחו אותן טענות עבור סדרת הפונקציות הבאה בקטע $[0,1]$:

$$f_n(x) = nx(1-x)^n$$

2. האם קיימת סדרה (f_n) של פונקציות אינטגרביליות ב־$[0,1]$ אשר מתכנסת נקודתית ב־$[0,1]$ אבל הגבול $\lim_{n\to\infty} \int_0^1 f_n(x)dx$ אינו קיים גם במובן הרחב?

3. תהי $R : [0,1] \to \mathbb{R}$ פונקציית רימן, המוגדרת על ידי

$$R(x) = \begin{cases} 0 & x \notin \mathbb{Q} \\ \frac{1}{q} & x = \frac{p}{q}, \end{cases}$$

הוכיחו כי $R(x)$ אינטגרבילית וחשבו את האינטגרל $\int_0^1 R(x)dx$ על ידי הצגת R כגבול של פונקציות פשוטות יותר. תוכלו להיעזר בדוגמה (4) מעמוד 465.

4. יהיו $f_n : \mathbb{R} \to \mathbb{R}$ פונקציות אינטגרביליות במובן הרחב ונניח ש־ $f_n \to f$ במ"ש ב־$\mathbb{R}$. האם נובע ש־ $\int_{-\infty}^{\infty} f = \lim_{n\to\infty} \int_{-\infty}^{\infty} f_n$?

5. יהיו $f_n : [a,b] \to \mathbb{R}$ פונקציות אינטגרביליות ונניח $f_n \to f$ נקודתית ב־$[a,b]$. נניח עוד שיש חסם M כך ש־ $|f_n| \leq M$ לכל n, ושלכל $\delta > 0$ הסדרה (f_n) מתכנסת במ"ש ל־f בקטע $[a+\delta, b-\delta]$. הוכיחו ש־ $\int_a^b f_n \to \int_a^b f$.

6. נגדיר סדרות של מספרים

$$\begin{aligned} c_n &= (-1)^{n+1} \prod_{k=1}^{n-1} \frac{n}{2n-1} \\ a_n &= c_n \cdot \frac{1}{(2n-1)2^n} \end{aligned}$$

הוכיחו על ידי אינטגרציה בחלקים שלכל n מתקיים

$$\int_0^1 \frac{1}{1+x^2} dx = \sum_{k=1}^{n-1} a_k + \int_0^1 \frac{c_n x^{2(n-1)}}{(1+x^2)^n}$$

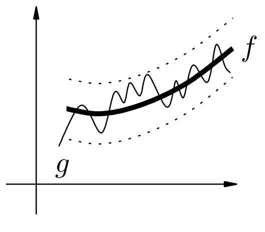
הראו ש־ $\lim_{n\to\infty} \int_0^1 \frac{c_n x^{2(n-1)}}{(1+x^2)^n} = 0$ (תוכלו להיעזר בתרגיל הקודם) והסיקו ש־

$$\frac{\pi}{2} = \sum_{n=1}^{\infty} a_n$$

היעזרו בטור זה כדי לחשב את $\frac{\pi}{2}$ ברמת דיוק של שלוש ספרות אחרי הנקודה העשרונית.

10.5 גזירה איבר־איבר

בוודאי כבר לא תתפלאו שגזירה איבר־איבר אינה נשמרת תחת גבולות נקודתיים. אם $g_n \to g$ סדרת פונקציות גזירות המתכנסות נקודתית בסביבה של נקודה x_0 לפונקציה שאינה רציפה ב־x_0, אז g בוודאי גם אינה גזירה ב־x_0. למשל עבור $g_n(x) = x^n$ כל ה־g_n־ים גזירות משמאל ב־1 אך פונקציית הגבול אינה רציפה משמאל ב־1 וממילא אינה גזירה משמאל ב־1.



איור 10.5.1 הפונקציה g קרובה ל־f בכל נקודה אך בעלת שיפועים גדולים בהרבה

מפתיע יותר לגלות שגזירות וערך הנגזרת אינם נשמרים גם תחת התכנסות במ"ש. הסיבה הגאומטרית היא שגם פונקציות שהגרפים שלהן קרובים מאד זה לזה בכל נקודה אינן בהכרח בעלות שיפועים דומים: אפילו אם ε קטן ואם הגרף של פונקציה g מוכל בפס בעובי ε סביב לגרף של פונקציה f, ייתכן שהשיפוע של g גדול בהרבה מהשיפוע של f בנקודות רבות. התנהגות כזו מודגמת באיור 10.5.1

דוגמאות

1. תהי $f_n : [-1,1] \to \mathbb{R}$ נתונה על ידי

$$f_n(x) = |x|^{1+\frac{1}{n}}$$

אפשר לבדוק ישירות ש־f_n גזירה בכל נקודה (הנקודה הבעייתית היחידה היא $x = 0$. ודאו ש־f_n גזירה שם!). קל גם לוודא ש־$f_n(x) \to |x|$ נקודתית ב־$[-1,1]$. נראה שההתכנסות היא במ"ש. ואמנם, לכל n מתקיים

$$|f_n(x) - |x|| = (1 - |x|^{1/n})|x| \leq \begin{cases} 1 - |1/n|^{1/n} & 1/n \leq |x| \leq 1 \\ 1/n & 0 \leq |x| < 1/n \end{cases}$$

ולכן

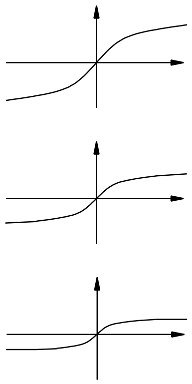
$$\sup_{x\in[-1,1]} |f_n(x) - |x|| \leq \max\{\frac{1}{n}, 1 - (\frac{1}{n})^{1/n}\}$$

אגף ימין הוא סדרה של מספרים ששואפת ל־0 כאשר n שואף לאינסוף, וקיבלנו ש־f_n מתכנסת במ"ש ב־$[-1,1]$ לפונקציית הערך המוחלט, שאינה גזירה באפס.

2. גם במקרים שבהם $f_n \to f$ במ"ש וכל הפונקציות f_n וגם f גזירות בנקודה x_0, לא מתקיים בהכרח ש־$f_n'(x_0) \to f'(x_0)$. למשל יהיו f_n נתונות על ידי

$$f_n(x) = \frac{1}{n}\arctan(nx)$$

(שימו לב לאופן שבו הסדרה התקבלה: כיווצנו פי n לאורך ציר ה־x את הפונקציה arctan, ואז כיווצנו פי n בכיוון y). ראו איור 10.5.2. מכיוון ש־arctan חסומה בין -1 ל־1, ברור ש־$-\frac{1}{n} \leq f_n \leq \frac{1}{n}$ ולכן $f_n \to 0$ במ"ש ב־$\mathbb{R}$. גזירה מראה ש־$f_n'(0) = 1$ לכל n. אבל הנגזרת של פונקציית האפס היא 0 בכל נקודה. כלומר, אף שהנגזרת של פונקציית הגבול קיימת, וגם גבול הנגזרות של f_n קיים, השניים אינם שווים.



איור 10.5.2 לפונקציות שיפוע זהה ב־0 אבל הן מתכנסות במידה שווה לפונקציית האפס

אם כך, לא ניתן לנסח משפט על גזירות באותה כלליות כמו במשפטי הרציפות והאינטגרציה שהוכחנו בסעיפים הקודמים. כדי לנסח משפט על הנגזרת של גבול של פונקציות גזירות (f_n), נוסיף דרישה לגבי התכנסות סדרת פונקציות הנגזרת $(f_n')_{n=1}^{\infty}$. ניתן כאן שתי גרסאות למשפט על גזירות הפונקציה הגבולית. הראשונה היא המשפט הבא, שהוכחתו מבוססת על משפט האינטגרציה איבר־איבר:

משפט 10.5.1 (גזירה איבר־איבר, תחת הנחת רציפות הנגזרות) תהיה $(f_n)_{n=1}^{\infty}$ סדרת פונקציות גזירות ברציפות בקטע I, ונניח ש־$f_n \to f$ במ"ש ב־I ו־$f_n' \to g$ במ"ש שם. אז f גזירה ומתקיים $f' = g$ (גזירות בקצוות מפורשת כגזירות חד־צדדית).

הוכחה תהי $c \in I$ נקודה שרירותית. נשים לב שלכל $x \in I$ מתקיים

$$\begin{aligned} f(x) &= \lim_{n\to\infty} f_n(x) \\ &= \lim_{n\to\infty} (f_n(c) + \int_c^x f_n') \\ &= \lim_{n\to\infty} f_n(c) + \lim_{n\to\infty} \int_c^x f_n' \\ &= f(c) + \int_c^x g \end{aligned}$$

(השוויון הראשון כי $f_n \to f$ נקודתית, ובפרט ב־x, השוויון השני כי לפי המשפט היסודי $f_n(x) = f_n(c) + \int_c^x f_n'$, השלישי מאריתמטיקה של גבולות. ברביעי השתמשנו שוב בעובדה ש־ $f_n \to f$ ובפרט ב־c, במשפט 10.4.1 על אינטגרציה איבר־איבר, ובנתון ש־ $f_n' \to g$ במ"ש ב־I).

מאחר ש־(f_n') רציפות ומתכנסות במ"ש ל־g ב־I, לפי משפט 10.3.2 g רציפה ב־I. לכן מהמשפט היסודי ומהשוויון $f(x) = f(c) + \int_c^x g$ נובע ש־f גזירה ומקיימת $f' = g$, כנדרש. ■

אם תנאי המשפט מתקיימים לסדרת פונקציות (f_n), אפשר לרשום את מסקנת המשפט כ־

$$\frac{d}{dx} \lim_{n\to\infty} f_n = \lim_{n\to\infty} \frac{d}{dx} f_n$$

עקב כך אומרים לפעמים שפעולת הגזירה מתחלפת (מבחינת סדר פעולות) עם פעולת הגבול.

לדוגמה, ראינו ש־ $e_n(x) = (1 + \frac{x}{n})^n \to e^x$ במ"ש בכל קטע סגור. כמו־כן חישוב קל מראה

$$e_n'(x) = (1 + \frac{x}{n})^{n-1}$$

וזו סדרת פונקציות רציפות. קל להסיק שגם היא מתכנסת ל־e^x במ"ש בכל קטע חסום. לכן לפי המשפט,

$$\frac{d}{dx} e^x = \frac{d}{dx} (\lim_{n\to\infty} (1 + \frac{x}{n})^n) = \lim_{n\to\infty} \frac{d}{dx} (1 + \frac{x}{n})^n = \lim_{n\to\infty} (1 + \frac{x}{n})^{n-1} = e^x$$

בהוכחה של המשפט האחרון ההנחה בדבר רציפות הנגזרות אפשרה להפעיל את המשפט היסודי על f_n', g. אפשר להיפטר מהנחה זו:

משפט 10.5.2 (גזירה איבר־איבר, המקרה הכללי) תהי $(f_n)_{n=1}^{\infty}$ סדרת פונקציות גזירות בקטע I. נניח ש־ $f_n \to f$ במ"ש ב־I וגם $f_n' \to g$ במ"ש שם. אז f גזירה ב־I ומתקיים $f' = g$ (גזירות בקצוות מפורשת כגזירות חד־צדדית).

הוכחה נקבע $x_0 \in I$ ונוכיח ש־ $f'(x_0) = g(x_0)$. תהי $F : I \to \mathbb{R}$ הפונקציה

$$F(x) = \begin{cases} g(x_0) & x = x_0 \\ \frac{f(x)-f(x_0)}{x-x_0} & x \neq x_0 \end{cases}$$

די להראות ש־F רציפה ב־x_0, שכן מכאן נובע ש־f גזירה ומקיימת $f'(x_0) = g(x_0)$. נגדיר $F_n : I \to \mathbb{R}$ באופן דומה:

$$F_n(x) = \begin{cases} f_n'(x_0) & x = x_0 \\ \frac{f_n(x)-f_n(x_0)}{x-x_0} & x \neq x_0 \end{cases}$$

לפי ההנחה ש־f_n גזירות נובע ש־F_n רציפות ב־x_0 (ובכל I, אך אין לנו צורך בכך). אם נראה ש־ $F_n \to F$ במ"ש ב־I ינבע ש־F רציפה ב־x_0, כפי שאנו רוצים.

נשים לב שלפי ההנחות $F_n \to F$ נקודתית ב־I. לכן די להראות שהסדרה (F_n) מתכנסת במ"ש ב־I, וממילא ינבע שהגבול הוא F. נפעיל לכן את תנאי קושי ונראה שבהינתן $\varepsilon > 0$ יש N כך שלכל $m, n > N$ ולכל $x \in I$ מתקיים $|F_n(x) - F_m(x)| < \varepsilon$.

יהי אם כן $\varepsilon > 0$. בעזרת משפט לגרנג', לכל m, n אנו מקבלים

$$\begin{aligned} F_n(x) - F_m(x) &= \frac{f_n(x) - f_n(x_0) - f_m(x) + f_m(x_0)}{x - x_0} \\ &= \frac{(f_n - f_m)(x) - (f_n - f_m)(x_0)}{x - x_0} \\ &= (f_n - f_m)'(y) \end{aligned}$$

כאשר y נקודה בין x_0 ל־x. אבל $(f_n - f_m)'(y) = f_n'(y) - f_m'(y)$, והסדרה $(f_k')_{k=1}^{\infty}$ מקיימת את תנאי קושי לסדרות פונקציות כי היא מתכנסת במ"ש ב־I. לכן יש N כך שלכל $m, n > N$ מתקיים $|f_n'(y) - f_m'(y)| < \varepsilon$, כלומר לכל $m, n > N$ ולכל $x \in I$ מתקיים $|F_n(x) - F_m(x)| < \varepsilon$, כפי שרצינו. ■

אפשר להחליש מעט יותר את ההנחות בשני המשפטים האחרונים. בשניהם הנחנו התכנסות במ"ש הן של הסדרה (f_n) והן של סדרת הנגזרות (f_n'). שני התנאים יחד נחוצים להוכחה, אבל מסתבר שאת ההתכנסות במ"ש של (f_n) אפשר להסיק מהנחה חלשה יותר:

למה 10.5.3 נניח ש־f_n גזירות בקטע חסום I, שסדרת הנגזרות (f_n') מתכנסת במ"ש ב־I ושיש נקודה $x_0 \in I$ כך שהסדרה $(f_n(x_0))_{n=1}^{\infty}$ מתכנסת. אז (f_n) מתכנסת במ"ש.

הוכחה נגדיר את F_n כמו בהוכחת משפט 10.5.2 ונשים לב ששם הוכחנו ש־(F_n) היא סדרת קושי ב־I מבלי להשתמש בהנחה ש־(f_n) מתכנסת במ"ש. לכן (F_n) היא

סדרת קושי גם תחת ההנחות הנוכחיות. יהי $\varepsilon > 0$. אז יש N כך שלכל $m, n > N$ מתקיים $|F_n(x) - F_m(x)| < \varepsilon$ לכל $x \in I$, כלומר

$$|\frac{f_n(x) - f_n(x_0) - f_m(x) + f_m(x_0)}{x - x_0}| = |F_n(x) - F_m(x)| < \varepsilon$$

ולאחר העברת אגפים אנו מקבלים שלכל $m, n > N$ ולכל $x \in I$ מתקיים

$$|(f_n(x) - f_n(x_0)) - (f_m(x) - f_m(x_0)| \leq |x - x_0|\varepsilon \leq |I|\varepsilon$$

כאשר $|I|$ הוא אורך הקטע I. אנו מסיקים שהפונקציות $g_n = f_n - f_n(x_0)$ מהוות סדרת קושי ב־I. אבל $(f_n(x_0))_{n=1}^{\infty}$ היא סדרת מספרים מתכנסת, ומכאן קל להסיק ש־(f_n) סדרת קושי (השלימו את הפרטים!). ■

מסקנה 10.5.4 אם (f'_n) מתכנסות במ"ש בקטע I אז (f_n) מתכנסות במ"ש ב־I אמ"מ (f_n) מתכנסות בנקודה אחת.

כצפוי, גם עבור טור שמחובריו פונקציות גזירות די בהתכנסות במ"ש כדי להבטיח גזירות של פונקציית הגבול. הפעלת המשפט על סדרת הסכומים החלקיים של טור פונקציות נותנת:

משפט 10.5.5 תהי $(u_k)_{k=1}^{\infty}$ סדרת פונקציות גזירות בקטע חסום I. נניח שטור הנגזרות $\sum_{k=1}^{\infty} u'_k$ מתכנס במ"ש ב־I לפונקציה g, ושקיימת נקודה $x_0 \in I$ כך ש־ $\sum_{k=1}^{\infty} u_k(x_0)$ מתכנס. אז הטור $\sum_{k=1}^{\infty} u_k$ מתכנס במידה שווה ב־I ומתקיים $(\sum_{k=1}^{\infty} u_k)' = \sum_{n=1}^{\infty} u'_k$.

הוכחה יהיו $S_N = \sum_{k=1}^{N} u_k$. מכללי הגזירה לסכומים סופיים, $S'_N = \sum_{k=1}^{N} u'_k$ ולכן מהנתון, (S'_N) מתכנסת במ"ש. כמו־כן $(S_N(x_0))$ היא סדרת מספרים מתכנסת ולכן לפי למה 10.5.3 הסדרה (S_N) מתכנסת במ"ש ב־I. לכן ממשפט 10.5.2 על גזירה איבר־איבר אנו מקבלים שב־I מתקיים

$$(\lim S_N)' = \lim S'_N$$

או במילים אחרות, $\sum_{k=1}^{\infty} u_k$ גזירה ב־I ומתקיים

$$(\sum_{k=1}^{\infty} u_k)' = \sum_{k=1}^{\infty} u'_k$$

כנדרש. ■

תוצאת המשפט מקבילה לנוסחה $(\sum_{k=1}^{N} u_k)' = \sum_{k=1}^{N} u'_k$ אשר נובעת מהלינאריות של הנגזרת לסכומים סופיים. מסקנת המשפט מכונה גזירה איבר־איבר כי אפשר לרשום זאת כך:

$$\frac{d}{dx}(u_1(x) + u_2(x) + \cdots) = \frac{d}{dx}u_1(x) + \frac{d}{dx}u_2(x) + \cdots$$

כלומר, כדי לגזור את הסכום האינסופי יש לגזור כל איבר בסכום בנפרד ולסכם את התוצאות.

תרגילים

1. יהיו $f, g : [0,1] \to R$ פונקציות גזירות ונניח ש־ $|f(x) - g(x)| < \frac{1}{100}$ בכל נקודה. האם ייתכן ש־ $|g'(x) - f'(x)| > \frac{1}{10}$ בכל נקודה $x \in [0,1]$?
2. נתבונן בטור הפונקציות

$$\sum_{n=1}^{\infty} \arctan\left(\frac{x}{n^2}\right)$$

(א) מהו תחום התכנסותו של הטור?
(ב) הראו שהטור אינו מתכנס במ"ש בתחום התכנסותו.
(ג) נמקו למה בכל זאת מותר לגזור את הטור איבר־איבר.

3. נגדיר סדרת פונקציות בקטע $[0,1]$:

$$f_n(x) = \frac{\sin(nx)}{n}$$

הוכיחו ש־ $f_n \to 0$ במ"ש ב־$[0,1]$ בעוד שסדרת הנגזרות $(f_n'(x))_{n=1}^{\infty}$ אינה מתכנסת נקודתית ל־0 שם.

4. נגדיר

$$f_n(x) = \frac{n^2x^2}{n + n^2x^2}$$

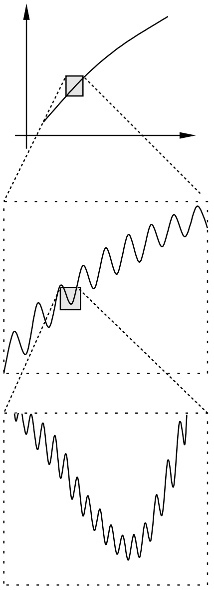
הוכיחו ש־ $f_n \to 0$ במ"ש ב־$\mathbb{R}$ אבל $f_n'(0) \not\to 0$.

10.6 פונקציה רציפה שאינה גזירה באף נקודה

הקשר בין תכונות הרציפות והגזירות לא נבחנה לעומק עד שלהי המאה התשע עשרה. על אף שהיה ידוע לכל שקיימות פונקציות רציפות עם נקודות אי־גזירות, הרי שבמידה ופונקציות לא גזירות התעוררו התייחסו אליהן כחריגות. ממילא חלק גדול מהבעיות שעניינו את המתמטיקאים באו מהפיזיקה, ושם מניחים שכל הפונקציות גזירות. במידה שהשאלה התעוררה, הדעה הרווחת הייתה שפונקציה רציפה חייבת להיות גזירה ברוב הנקודות, גם אם לא בכולן. ואולם בשנת 1861 הראה ויירשטראס[3] שקיימות פונקציות רציפות שאינן גזירות באף נקודה.

[3]למעשה בולצאנו מצא דוגמה כזאת קודם לכן, אך התגלית לא זכתה לתשומת הלב הראויה. נציין גם שבאותה תקופה רימן גילה פונקציה רציפה שאינה גזירה בקבוצה צפופה של נקודות. פגשנו דוגמה כזו בתרגיל (7) בעמוד 414.

לא נוכל לצייר תמונה מדויקת של פונקציה כזו אך באופן גס המראה שלה הוא כמו באיור 10.6.1: ככל שמתקרבים אליה מגלים תנודות בעלות שיפוע גדול, המצטברות לכדי אי־גזירות, אך אינן גורמות לאי־רציפות כי תרומתן הכוללת לפונקציה קטנה. הקושי בבניית פונקציה כזו הוא באיזון בין שתי התכונות: אם התנודות קטנות מדי הפונקציה תהיה גזירה, אך אם התנודות חזקות מדי נקבל פונקציה לא רציפה, כפי שקורה במקרה של פונקציית דרכלה.


איור 10.6.1 פונקציה עם תנודות בכל קנה מידה

משפט 10.6.1 (פונקציית ויירשטראס) תהי

$$w(x)=\sum_{k=1}^{\infty}\frac{\cos(\pi\cdot 100^k x)}{10^k}$$

אז w רציפה ב־$\mathbb{R}$ ואינה גזירה באף נקודה.

הוכחה אנו נוכיח משפט מעט יותר כללי: נראה שהפונקציה

$$w(x)=\sum_{k=1}^{\infty}\frac{\cos\pi b^k x}{a^k}$$

רציפה אך לא גזירה בשום מקום, בתנאי ש־ $a>5$ ו־ $b>a+\frac{\pi a(a-1)}{a-5}$. המקרה $a=10$ ו־ $b=100$ מתקבל כמקרה פרטי של טענה זו.

הטענה למעלה כוללת שני מרכיבים: האחד בדבר רציפות w, והשני בדבר אי־גזירותה. הרציפות נובעת בקלות מהכלים שפיתחנו בפרק זה: אם נסמן

$$u_k(x)=a^{-k}\cos(\pi b^k x)$$

אז u_k רציפה ב־$\mathbb{R}$ עבור כל k. מכיוון ש־ $|u_k(x)|\le a^{-k}$ לכל $x\in\mathbb{R}$ והטור המספרי $\sum_{k=1}^{\infty}a^{-k}$ מתכנס, לפי מבחן ויירשטראס (משפט 10.2.7) הטור $\sum_{k=1}^{\infty}u_k$ מתכנס במ"ש ב־$\mathbb{R}$, ומכאן ש־w היא גבול במ"ש של פונקציות רציפות ולכן רציפה בעצמה.

החלק המורכב יותר בהוכחת המשפט הוא ההוכחה ש־w אינה גזירה באף נקודה. יהי $x_0\in\mathbb{R}$. כדי להראות ש־w אינה גזירה ב־x_0 נראה שיש סדרה של נקודות $x_k\to x_0$ כך

$$\lim_{k\to\infty}\left|\frac{w(x_k)-w(x_0)}{x_k-x_0}\right|=\infty$$

מכאן נובע שהגבול

$$\lim_{y\to x_0}\frac{w(y)-w(x_0)}{y-x_0}$$

אינו קיים (אלא אולי במובן הרחב), ולכן w אינה גזירה ב־x_0.

הרעיון בבחירת x_k הוא להשתמש בתנודתיות של המחובר u_k. ליתר דיוק, נבחר את x_k כך שהשיפוע של u_k בין x_0 ל־x_k הוא התורם העיקרי לשיפוע של w בין הנקודות.

לכל N נסמן את הסכום החלקי ה־N והזנב ה־N של טור הפונקציות $\sum_{k=1}^{\infty} u_k$ על ידי

$$\begin{aligned} S_N &= \sum_{k=1}^{N} u_k \\ R_N &= \sum_{k=N+1}^{\infty} u_k \end{aligned}$$

אז לכל k מתקיים

$$w = S_{k-1} + u_k + R_k$$

לכל y השיפוע של המיתר בגרף של w בין x ל־y מקיים את החסם

$$\begin{aligned} &|\tfrac{w(y)-w(x_0)}{y-x_0}| \\ &\quad = |\tfrac{(S_{k-1}+u_k+R_k)(y)-(S_{k-1}+u_k+R_k)(x_0)}{y-x_0}| \\ &\quad \geq |\tfrac{u_k(y)-u_k(x_0)}{y-x_0}| - |\tfrac{R_k(y)-R_k(x_0)}{y-x_0}| - |\tfrac{S_{k-1}(y)-S_{k-1}(x_0)}{y-x_0}| \end{aligned}$$

שלושת המחוברים בשורה האחרונה מייצגים את השיפועים של המיתרים בגרפים של הפונקציות u_k, R_k, S_{k-1}, בהתאמה.

אנו רוצים לבחור נקודה x_k שמרחקה מ־x_0 קטן (ליתר דיוק, שואף לאפס כאשר $k \to \infty$), וכך שהמחובר הראשון $|\frac{u_k(x_k)-u_k(x_0)}{y-x_0}|$ בביטוי שקיבלנו יהיה גדול בהרבה משני המחוסרים. לשם כך ניעזר בעובדה הבאה, שהוכחתה קלה ומושארת כתרגיל: לפונקציה $\cos \pi x$ יש התכונה שלכל x, קיים y כך ש־ $\frac{1}{2} \leq |y-x| \leq 1$ ו־ $|\cos \pi y - \cos \pi x| \geq 1$ (הוכיחו טענה זו!). נובע מכך שעבור $u_k(x) = a^{-k} \cos(\pi b^k x)$ קיימת נקודה x_k שמקיימת $\frac{1}{2} b^{-k} \leq |x_k - x_0| \leq b^{-k}$ ו־ $|u_k(x_k) - u_k(x_0)| \geq a^{-k}$, ובפרט

$$|\frac{u_k(x_k) - u_k(x_0)}{y - x_0}| \geq \frac{b^k}{a^k}$$

אם נציב בחירה זו של x_k בהערכה שלנו לשיפוע המיתר ב־w נקבל

$$|\frac{w(x_k) - w(x_0)}{x_k - x_0}| \geq \frac{b^k}{a^k} - |\frac{R_k(x_k) - R_k(x_0)}{x_k - x_0}| - |\frac{S_{k-1}(x_k) - S_{k-1}(x_0)}{x_k - x_0}|$$

כדי להעריך את המחוסר הראשון, נשים לב שאפשר לחסום את הגודל של הזנב R_k. לכל z מתקיים

$$|R_k(z)| \leq \sum_{m=k+1}^{\infty} |u_m(z)| \leq \sum_{m=k+1}^{\infty} a^{-m} = a^{-(k+1)} \frac{a}{a-1}$$

לכן

$$|\frac{R_k(x_k) - R_k(x)}{x_k - x}| \leq \frac{2a^{-(k+1)}}{|x_k - x|} \cdot \frac{a}{a-1} \leq \frac{4b^k}{a^k(a-1)}$$

כשהאי־שוויון הימני נובע שוב מההערכה $|x_k - x_0| \geq \frac{1}{2}b^{-k}$. אם נציב זאת בהערכה שלנו לשיפוע המיתר ב־w נקבל

$$\begin{aligned} |\frac{w(x_k) - w(x_0)}{x_k - x_0}| &\geq \frac{b^k}{a^k} - \frac{4b^k}{a^k(a-1)} - |\frac{S_{k-1}(x_k) - S_{k-1}(x_0)}{x_k - x_0}| \\ &\geq \frac{(a-5)b^k}{(a-1)a^k} - |\frac{S_{k-1}(x_k) - S_{k-1}(x_0)}{x_k - x_0}| \end{aligned}$$

כדי להעריך את המחוסר השני אי אפשר להשתמש באותו נימוק שבו השתמשנו עבור המחוסר הראשון, כי באופן טיפוסי S_{k-1} הוא גדול מאד ביחס ל־u_k. במקום זאת נשתמש בעובדה שכל המיתרים ב־S_{k-1} הם בעלי שיפוע קטן יחסית. ניעזר במשפט לגרנג', לפיו יש c בין x_0 ל־x_k כך ש־

$$|\frac{S_{k-1}(x_k) - S_{k-1}(x_0)}{x_k - x_0}| = S'_{k-1}(c)$$

לכן אם נקבל חסם מלמעלה על הנגזרת של S_{k-1} נקבל הערכה למחוסר האחרון. ובכן, נחשב:

$$\begin{aligned} |S'_{k-1}(c)| &= |\sum_{m=1}^{k-1} u'_m(c)| \\ &\leq \sum_{m=1}^{k-1} |u'_m(c)| \\ &= \sum_{m=1}^{k-1} \pi \frac{b^m}{a^m} |\cos \pi b^m c| \\ &\leq \pi \sum_{m=1}^{k-1} (\frac{b}{a})^m \\ &= \pi \frac{b}{a} \cdot \frac{(b/a)^{k-1} - 1}{b/a - 1} \\ &\leq \pi \cdot \frac{b^k}{a^k} \cdot \frac{a}{b-a} \end{aligned}$$

נציב זאת בהערכה שלנו עבור שיפוע המיתר ונקבל

$$|\frac{w(x_k)-w(x_0)}{x_k-x_0}| \geq \frac{(a-5)b^k}{(a-1)a^k} - \frac{\pi b^k a}{a^k(b-a)} \geq (\frac{a-5}{a-1} - \frac{\pi a}{b-a})\frac{b^k}{a^k}$$

לסיכום, ההערכה $|x_k - x_0| \leq b^{-k}$ גוררת ש־ $x_k \to x_0$. מצד שני החשבון לעיל מראה ש־

$$|\frac{w(x_k)-w(x_0)}{x_k-x_0}| \geq (\frac{a-5}{a-1} - \frac{\pi a}{b-a})(\frac{b}{a})^k$$

ואם $b > a > 0$ וגם $\frac{a-5}{a-1} - \frac{\pi a}{b-a} > 0$ אז אגף ימין שואף לאינסוף כאשר k שואף לאינסוף. זה יקרה אם

$$b > a + \frac{\pi a(a-1)}{a-5}$$

ובפרט אם $a = 10, b = 100$. כפי שרצינו. ∎

המשפט האחרון הוא הוכחה ניצחת שגזירות אינה נשמרת תחת גבולות במ"ש אפילו של פונקציות יפות מאד. כל אחת מהפונקציות u_k מהמשפט גזירה אינסוף פעמים אך אין אפילו נקודה אחת שבה פונקציית הגבול גזירה!

בניגוד למה שחשבו בני דורו של ויירשטראס, פונקציות מעין אלה קיימות בטבע, או לפחות בתיאור המתמטי של הטבע. למשל תנועה של חלקיק בתמיסה היא בקירוב כזו, שכן החלקיק מתנגש כל העת בחלקיקים אחרים ומשנה כיוון עם כל התנגשות. כמו־כן אפשר להראות שבמובן מסוים רוב הפונקציות הרציפות אינן גזירות באף נקודה. על הנושא האחרון אפשר לקרוא בספר [9].

פרק 11

פולינומי טיילור־מקלורן וטורי חזקות

בפרק זה נתעניין בקירוב של פונקציות ממשיות על ידי פולינומים והצגתן כטורי חזקות, שאפשר לראות בהם פולינומים מוכללים. אנו נראה שלפונקציות רבות יש קירובים והצגות כאלה.

בחלק הראשון של הפרק נכליל את מושג המשיק: כפי שהישר המשיק הוא הישר המקרב בצורה הטובה ביותר את פונקציה בקרבת נקודת ההשקה, לכל n נחפש עתה את הפולינום ממעלה n המקרב את הפונקציה בצורה הטובה ביותר (במובן שיוגדר להלן). הפולינום המתקבל הוא פולינום טיילור־מקלורן.

לאחר מכן נדון בטורי חזקות, שהם טורי פונקציות עם מחוברים מהצורה ax^n. טורי חזקות מתנהגים בצורה דומה מאד לפולינומים, ויש להם גם קשר הדוק עם פולינומי טיילור־מקלורן. בעזרת השיטות שנפתח נקבל נוסחאות שיאפשרו לחשב בכל רמת דיוק רצויה את הפונקציה המעריכית, הלוגריתם, הפונקציות הטריגונומטריות ועוד. נראה גם מספר שימושים.

11.1 פולינומים

לפני שניתן הגדרות חדשות נחזור בקצרה על התכונות של פולינומים. כזכור, פולינום הוא פונקציה שניתן לרשום אותה בצורה $p(x) = \sum_{k=0}^{d} a_k x^k$, כש־$a_k$ מספרים ממשיים (זכרו ש־ $x^0 = 1$ לכל x, כך שהמחובר הראשון בסכום הוא הקבוע a_0). אם $a_d \neq 0$ אז d נקרא מעלת הפולינום.

אם $p(x)=\sum_{k=0}^{d}a_kx^k$ אז כללי הגזירה מראים ש־

$$\begin{aligned} p'(x) &= \sum_{k=1}^{d} k a_k x^{k-1} \\ p''(x) &= \sum_{k=2}^{d} k(k-1) a_k x^{k-2} \end{aligned}$$

ובאופן כללי לכל $0\le n\le d$ מתקיים

$$p^{(n)}(x)=\sum_{k=n}^{d} k(k-1)\cdot\ldots\cdot(k-n+1)a_kx^{k-n}$$

מאידך עבור $n>d$ מתקיים $p^{(n)}\equiv 0$.

הנוסחה עבור נגזרות הפולינום גוררת שהערכים של הנגזרות בנקודה קובעים את מקדמי הפולינום. שכן אם נציב 0 בנוסחה עבור $p^{(n)}$ נקבל

$$p^{(n)}(0)=n!\cdot a_n$$

כלומר, סדרה המקדמים $a_0,a_1,\ldots,a_d$ שווה לסדרה $\frac{1}{0!}p(0),\frac{1}{1!}p'(0),\ldots,\frac{1}{d!}p^{(d)}(0)$ (כזכור $0!=1$ לפי הגדרה). אגב, זו הוכחה שאם פונקציה היא פולינום אז יש לה הצגה יחידה מהסוג $\sum_{n=0}^{d}a_nx^n$ עם $a_d\neq 0$, כי המקדמים נקבעים על ידי הנגזרות של הפונקציה.

את הלמה הבאה הוכחנו בדוגמה (3) בעמוד 418.

למה 11.1.1 אם p הוא פולינום אז כל פונקציה קדומה של p ב־$\mathbb{R}$ היא פולינום.

מסקנה 11.1.2 נניח ש־f היא פונקציה גזירה d פעמים וש־ $f^{(d+1)}\equiv 0$. אז f היא פולינום והיא נתונה על ידי הנוסחה $f(x)=\sum_{k=0}^{d}\frac{f^{(k)}(0)}{n!}x^k$.

הוכחה הפונקציה $f^{(d+1)}$ היא פולינום האפס, ועל ידי הפעלה חוזרת של הלמה אנו מסיקים ש־$f^{(d)}$ היא פולינום, $f^{(d-1)}$ היא פולינום, וכן הלאה, עד שמקבלים ש־ $f=f^{(0)}$ היא פולינום. נרשום $f(x)=\sum_{k=0}^{d}a_kx^k$. אז ראינו שהמקדמים a_k נתונים על ידי הנוסחה $f^{(k)}(0)=k!a_k$, ומכאן הנוסחה עבור f. ■

תרגילים

1. יהיו $f,g:[a,b]\to\mathbb{R}$ גזירות n פעמים ונניח ש־ $f^{(n)}=g^{(n)}$. הוכיחו ש־ f,g נבדלות בפולינום. האם זה נכון גם אם התחום של f,g אינו קטע?
2. תהי $f:[a,b]\to\mathbb{R}$ גזירה n פעמים ויהי $c\in(a,b)$. הראו ש־f נקבעת על ידי הפונקציה $f^{(n)}$ והערכים $f(c),f'(c),\ldots,f^{(n)}(c)$.
3. הוכיחו שאם $a_0,\ldots,a_n$ מספרים כך ש־ $a_0+\frac{a_1}{2}+\ldots+\frac{a_n}{n+1}=0$ אז יש $x\in[0,1]$ כך $\sum_{k=0}^{n}a_kx^k=0$.

11.2 פולינומי טיילור־מקלורן

נתבונן בידידנו הוותיק החלקיק, אשר נע בקו ישר ומיקומו בזמן t הוא $f(t)$. נניח שמדדנו את מיקומו ומהירותו בזמן 0 ואנו מתבקשים לנחש את מיקומו בזמן $t > 0$. בהיעדר מידע נוסף, הניחוש הטבעי הוא שהחלקיק נע במהירות קבועה, ואז ננחש ש־f נתונה על ידי הנוסחה $\ell(x) = f(0) + f'(0)t$. לא נצפה שקירוב זה יהיה טוב עבור t־ים גדולים אך כאשר t קרוב ל־0 אנו יודעים שזהו הפולינום ממעלה אחד המקרב בצורה הטובה ביותר את f: ℓ הוא הישר המשיק, וכפי שראינו בפרק על נגזרות הוא הישר היחיד שמקיים $f(t) = f(0) + f'(0)t + o(t)$.

היינו רוצים לקבל הערכה דומה במקרה שיש בידינו מידע רב יותר. תהי f פונקציה ונניח שנתונים לנו הערכים שלה ושל n הנגזרות הראשונות שלה ב־0, דהיינו $f(0), f'(0), \ldots, f^{(n)}(0)$. אנו מחפשים פונקציה p אשר מקרבת את f בצורה טובה לפחות עבור t קרוב ל־0. מתקבל על הדעת שערכי הנגזרות של p יסכימו עם הנתונים שבידינו, כלומר אנו מחפשים פונקציה p שמקיימת

$$p(0) = f(0) \quad , \quad p'(0) = f'(0) \quad , \; \ldots \; , \quad p^{(n)}(0) = f^{(n)}(0)$$

כמו־כן, בהיעדר מידע נוסף נניח ש־ $p^{(n+1)} \equiv 0$ (כך נהגנו כשידענו רק את הנגזרת הראשונה: אז בחרנו קירוב שהוא ישר, דהיינו שהנגזרת השנייה שלו מתאפסת זהותית).

לפי הדיון בסעיף הקודם, ישנה רק פונקציה p אחת שעומדת בדרישות אלה. היא פולינום כי $p^{(n+1)} \equiv 0$, והמקדמים שלה נקבעים על ידי הנגזרות של f ב־0 לפי הנוסחה $p^{(n)}(0) = n! \cdot a_n$ שפיתחנו בסעיף הקודם. לכן נגדיר:

הגדרה 11.2.1 תהי f פונקציה גזירה n פעמים ב־0. **פולינום מקלורן**[1] (MacLaurin polynomial) מסדר n של f הוא הפולינום P_n הנתון על־ידי הנוסחה

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$$

כפי שציינו בדיון למעלה, P_n נבחרה כך ש־n הנגזרות הראשונות שלה יתלכדו עם הנגזרות המתאימות של f, והנגזרת ה־$n+1$ של P_n היא זהותית 0. בפרט, P_0 היא הפונקציה הקבועה שערכה $f(0)$, ו־P_1 היא הפונקציה $P_1(x) = f(0) + f'(0) \cdot x$ אשר מתארת את המשיק ב־0 לגרף של f.

[1] Colin MacLaurin, 1698־1746.

הערות

1. f אינה מופיעה במפורש בסימון של פולינום מקלורן P_n שלה. כאשר נדבר על מספר פונקציות שונות נשתמש בסימונים שונים, למשל אם g היא פונקציה בדרך־כלל נציין ב־Q_n את פולינום טיילור ה־n של g. סימנים אלה יוגדרו כאשר יתעורר הצורך.
2. פולינום מקלורן ה־n של f הוא פולינום ממעלה שאינה עולה על n אך ייתכן שמעלתו קטנה יותר. כפי שרואים מהנוסחה של P_n, זה יקרה בדיוק כאשר $f^{(n)}(0)=0$.

דוגמאות

1. יהי $p(x)=\sum_{k=0}^{d} a_k x^k$ פולינום ויהי P_n פולינום מקלורן מסדר n של p, כאשר $n \le d$. מאחר ש־ $p^{(k)}(0)=k!a_k$ הרי

$$P_n(x)=\sum_{k=0}^{n}\frac{k!a_k}{k!}x^k=\sum_{k=0}^{n}a_kx^k$$

כלומר הקירוב מסדר n מתקבל מהפולינום p על ידי מחיקת כל המחוברים מסדר גדול מ־n. מאידך עבור $n \ge d$ מתקיים $P_n=p$. זהו אכן קירוב מצוין!

2. תהי $f(x)=e^x$. לכל n מתקיים $f^{(n)}(x)=e^x$ ולכן $f^{(n)}(0)=1$ לכל n. מכאן ש־

$$P_n(x)=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots+\frac{x^n}{n!}=\sum_{k=0}^{n}\frac{x^k}{k!}$$

3. נתבונן בפונקציה $\sin(x)$. לפי כללי הגזירה שלה מתקיים $\sin'=\cos$, לכן $\sin''=-\sin$. $\sin^{(3)}=-\cos$ ו־ $\sin^{(4)}=\sin$, וחזרנו לפונקציה המקורית. מכאן נובע

$$\sin^{(n)}(x) = \begin{cases} \sin(x) & n=4k \\ \cos(x) & n=4k+1 \\ -\sin(x) & n=4k+2 \\ -\cos(x) & n=4k+3 \end{cases}$$

ועל ידי הצבת $x=0$ מקבלים

$$\sin^{(n)}(0)=\begin{cases} 0 & n=2k \\ (-1)^k & n=2k+1 \end{cases}$$

לכן לכל n,

$$\begin{aligned}
P_n(x) &= 0+\frac{1}{1!}\cdot x+\frac{0}{2!}x^2-\frac{1}{3!}x^3+\frac{0}{4!}x^4+\ldots+\frac{(-1)^m}{m!}x^{2m+1}\\
&= \frac{x}{1!}-\frac{x^3}{x!}+\frac{x^5}{5!}-\frac{x^7}{6!}+\ldots+\frac{(-1)^m}{m!}x^{2m+1}
\end{aligned}$$

כאשר $m=[\frac{n-1}{2}]$ הוא המספר הטבעי הגדול ביותר כך ש־ $2m+1\le n$. על כן נוכל לרשום את P_n כך:

$$P_n(x)=\sum_{k=0}^{[(n-1)/2]}\frac{(-1)^k}{(2k+1)!}x^{2k+1}$$

נוסחה דומה קיימת לפולינום מקלורן Q_n מסדר n של cos:

$$Q_n(x)=\sum_{k=0}^{[n/2]}\frac{(-1)^k}{(2k)!}x^{2k}$$

הוכיחו זאת!

4. תהי $f(x)=\ln(1+x)$. אז

$$\begin{aligned}
f'(x) &= \frac{1}{1+x}\\
f''(x) &= -\frac{1}{(1+x)^2}\\
\vdots &\quad \vdots\\
f^{(n)}(x) &= (-1)^{n-1}(n-1)!\frac{1}{(1+x)^n}
\end{aligned}$$

(תנו הוכחה מפורטת לנוסחה זו!) לכן $f^{(n)}(0)=(-1)^{n+1}(n-1)!$, ומקבלים

$$P_n(x)=\sum_{k=1}^{n}\frac{(-1)^{k+1}}{k}x^k$$

(שימו לב שהסכום מתחיל מ־1 כי $f(0)=\ln(1+0)=0$).

ניתן להכליל מעט את מה שעשינו עד כה אם נחליף את הנקודה 0 בנקודה שרירותית x_0. כעת אנו מתבוננים בפונקציה f הגזירה n פעמים ב־x_0 ושואלים מהו הניחוש הטבעי ל־f בהינתן הנגזרות $f(x_0),f'(x_0),\ldots,f^{(n)}(x_0)$. התשובה היא

הגדרה 11.2.2 תהי f פונקציה גזירה n פעמים בנקודה x_0. **פולינום טיילור**[2] (Taylor polynomial) מסדר n של f סביב x_0 הוא הפולינום

$$P_n(x)=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

[2] Brook Taylor, 1685־1731.

הערות

1. לעתים אומרים ש**מפתחים** (develop) את f לפולינום טיילור סביב x_0.
2. פולינום מקלורן הוא מקרה פרטי של פולינום טיילור, כפי שרואים אם מציבים $x_0 = 0$ בנוסחה של פולינום טיילור. השימוש במונח "פולינום מקלורן" משמש לכן בעיקר להדגיש שנקודת הפיתוח היא 0 ולא נקודה אחרת, ולפעמים נקרא לפולינום מקלורן גם בשם פולינום טיילור.
3. כעת הסימון P_n מסתיר מידע: הוא אינו מראה מה הפונקציה ואינו מראה מהי הנקודה x_0. כשיהיה צורך בכך נשתמש בסימון מפורש יותר.

קל לבדוק על ידי גזירה שכאשר P_n הוא פולינום טיילור של f סביב x_0 מתקיים $P_n^{(k)}(x_0) = f^{(k)}(x_0)$ לכל $k \leq n$ (בדקו בעצמכם!). להלן נסיק זאת גם בדרך אחרת.

תהי f פונקציה גזירה n פעמים ב־x_0, ונגדיר $\widetilde{f}(x) = f(x + x_0)$, כך ש־$\widetilde{f}$ גזירה n פעמים ב־0. נשאלת השאלה, מה הקשר בין פולינום טיילור P_n של f ב־x_0 לפולינום מקלורן $\widetilde{P}_n$ של $\widetilde{f}$. התשובה היא שמתקיים אותו קשר כמו בין הפונקציות: $\widetilde{P}_n(x) = P_n(x + x_0)$. כדי לראות זאת, נשים לב שמכלל השרשרת נובע ש־$\widetilde{f}$ גזירה n פעמים ב־0 והנגזרות הן $\widetilde{f}^{(k)}(0) = f^{(k)}(x_0)$. לכן פולינום מקלורן $\widetilde{P}_n$ מסדר n של $\widetilde{f}$ הוא

$$\widetilde{P}_n(x) = \sum_{k=0}^{n} \frac{\widetilde{f}^{(k)}(0)}{k!} x^k = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} x^k$$

ולכן

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = \widetilde{P}_n(x - x_0)$$

כלומר $\widetilde{P}_n(x) = P_n(x + x_0)$.

תרגילים

1. חשבו את פולינום מקלורן מסדר n של הפונקציות הבאות:
   (א) $(1+x)^\alpha$ עבור $\alpha \in \mathbb{R}$ (הבחינו בין המקרה $\alpha \in \mathbb{N}$ ו־ $\alpha \notin \mathbb{N}$).
   (ב) $\ln(1-x)$
   (ג) $\frac{1}{2}(e^x + e^{-x})$
   (ד) $\frac{1}{2}(e^x - e^{-x})$
2. בהינתן פולינום מקלורן $P_n(x) = \sum_{k=0}^n a_k x^k$ של f, חשבו את פולינום מקלורן מסדר k של $f(2x)$ לכל $k \leq n$.

3. כתבו נוסחה לפולינום p ממעלה שלוש שמקיים $p(2) = 0$, $p'(2) = 1$, $p''(2) = 1$, $p^{(3)}(2) = 1$.

4. יהי $p(x) = x^5 + 4x^4 + 7x^3 + 2x^2 + 3x + 1$. מצאו את פיתוח טיילור שלו סביב הנקודה 1, כלומר את המקדמים a_k כך ש־ $p(x) = \sum_{k=0}^{5} a_k (x-1)^k$.

11.3 תכונות הקירוב של פולינום טיילור

הגדרנו את פולינום מקלורן בתקווה שיהיה קירוב טוב לפונקציה שהתחלנו ממנה. מתעוררות בהקשר זה שתי שאלות טבעיות. ראשית, עבור n קבוע, אפשר לשאול עד כמה הפולינום P_n מקרב את f. שנית, עבור נקודה קבועה x אפשר לשאול האם סדרת הקירובים $P_n(x)$ מתקרבת לערך $f(x)$ כשמגדילים את n, ובאופן כללי יותר האם סדרת הפונקציות (P_n) מתכנסת ל־f נקודתית או במ"ש בקבוצה כלשהי. בסעיף הנוכחי נבחן את השאלה הראשונה ובשנייה נדון בסעיפים הבאים.

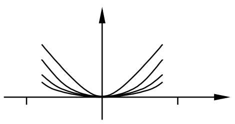
סימון 11.3.1 אם f גזירה n פעמים ב־0 אז **שארית טיילור** (remainder) מסדר n של f היא הפונקציה

$$R_n = f - P_n$$

ככל ש־$R_n(x)$ קטן יותר בערכו המוחלט נוכל לומר ש־P_n היא קירוב טוב יותר ל־f בנקודה x.

עבור n קבוע אין סיבה לצפות ש־R_n תהיה קטנה בכל נקודה. למשל, אם $f = \sin$ אז $P_1(x) = x$ והשארית R_1 אינה אפילו חסומה מלמעלה (הוכיחו זאת!). מצד שני, בקרבת 0 השארית R_1 תמיד קטנה, שהרי R_1 אינה אלה הישר המשיק ל־f ב־0, ומתקיים $R_1(x) = f(x) - P_1(x) = o(x)$. מסתבר ש־$R_n$ מקיימת תכונות דומות גם עבור $n \geq 2$.

הגדרה 11.3.2 עבור n טבעי נאמר שפונקציה α היא **אפסית מסדר** n אם מתקיים $\alpha(x) = x^n \cdot \alpha^*(x)$ עבור פונקציה α^* שהיא רציפה ב־0 ו־ $\alpha^*(0) = 0$. במקרה זה גם נאמר ש־α היא $o(x^n)$ (קרי: α היא o**־קטן** של x^n).


איור 11.3.1 הפונקציות x^n נעשות יותר ויותר שטוחות בסביבת 0

מושג זה מכליל את מושג הפונקציה האפסית (הגדרה 8.2.1) והסימון $o(x^n)$ מכליל את הסימון $o(x)$, שהוא מקרה פרטי כאשר $n = 1$. הכללים לשימוש בסימון $o(x^n)$ דומים לכללים לגבי $o(x)$: עיינו בדיון בעמוד 310.

אם כן, α היא אפסית מסדר n אמ"מ בסביבת 0 היא שואפת ל־0 מהר יותר מ־x^n (הפונקציה x^n עצמה איננה אפסית מסדר n). מבחינה גאומטרית פירוש הדבר שקרוב לאפס פונקציה שהיא אפסית מסדר n היא שטוחה יותר מהפונקציה x^n.

שימו לב שכשמגדילים את n, הפונקציה x^n נעשית יותר ויותר שטוחה, כפי שרואים באיור 11.3.1.

שימו לב שאם $m < n$ ואם $f(x)$ אפסית מסדר n, אז היא אפסית גם מסדר m. ההיפך כמובן אינו נכון, כפי שמדגימה הפונקציה x^m.

ההוכחה של הלמה הבאה דומה להוכחה של למה 8.2.2, ואנו משאירים אותה לכם:

למה 11.3.3 פונקציה α היא $o(x^n)$ אמ"מ α רציפה ב־0 ומקיימת $\lim_{x\to 0} \frac{\alpha(x)}{x^n} = 0$.

גם התכונות הבאות של פונקציות אפסיות קלות להוכחה ומושארות כתרגיל (השוו אותם עם למה 8.2.3).

למה 11.3.4 (אריתמטיקה של פונקציות אפסיות מסדר n):

1. $o(x^m) + o(x^n) = o(x^{\min\{m,n\}})$
2. אם f פונקציה חסומה בסביבת 0 אז $f(x)o(x^n) = o(x^n)$.
3. $o(x^m) \cdot o(x^n) = o(x^{m+n})$.
4. אם $\alpha(x) = o(x^m)$ ואם $\beta(x) = o(x^n)$ אז $\beta \circ \alpha(x) = o(x^{mn})$.

זכרו שקוראים שוויונות כאלה משמאל לימין. למשל, הסעיף הראשון אומר שהסכום של פונקציה אפסית מסדר m ופונקציה אפסית מסדר n הוא פונקציה אפסית מסדר $\min\{m,n\}$.

כפי שהישר המשיק הוא הפולינום היחיד ממעלה אחת שההפרש בינו לבין הפונקציה הוא אפסי מסדר אחת, נראה עתה שפולינום מקלורן מסדר n הוא הפולינום היחיד ממעלה n שההפרש בינו לבין הפונקציה אפסי מסדר n. נוכיח תחילה שלפולינום מקלורן באמת מקיים תכונה זו:

משפט 11.3.5 תהי f גזירה n פעמים ב־0 ו־R_n שארית טיילור מסדר n של f. אז $R_n(x) = o(x^n)$, או באופן שקול, $f(x) = P_n(x) + o(x^n)$.

הוכחה לפי הלמה די להראות ש־

$$\lim_{x\to 0} \frac{R_n(x)}{x^n} = 0$$

מנה זו היא מנה של פונקציות אשר מתאפסות ב־0, ולכן ננסה להפעיל את כלל לופיטל. גם אחרי גזירה של המונה והמכנה פעם אחת מסתבר שהמונה והמכנה מתאפסים ב־0, אך כשמפעילים את כלל לופיטל $n-1$ פעמים מקבלים את התוצאה הרצויה.

ניגש לפרטים. תחילה נגזור את R_n: לכל $k = 1, \ldots, n$ מתקיים

$$R_n^{(k)}(x) = f^{(k)}(x) - P_n^{(k)}(x)$$

בפרט $R_n^{(k)}$ רציפה בסביבת 0 לכל $k < n$ (הרי f גזירה n פעמים ב־0 ולכן $f^{(k)}$ רציפה בסביבת 0). לפי ההגדרה של פולינום טיילור,

$$R_n^{(k)}(0) = f^{(k)}(0) - P^{(k)}(0) = 0$$

מצד שני, הנגזרת ה־k של x^n היא $n(n-1)\cdot \ldots \cdot (n-k+1)x^{n-k}$, ופונקציה זו מתאפסת רק ב־0. לכן נוכל להפעיל את כלל לופיטל $n-1$ פעמים ונקבל

$$\begin{aligned}\lim_{x\to 0}\frac{R_n(x)}{x^n} &= \lim_{x\to 0}\frac{R_n'(x)}{nx^{n-1}}\\ &= \lim_{x\to 0}\frac{R_n''(x)}{n(n-1)x^{n-2}}\\ &\vdots\\ &= \lim_{x\to 0}\frac{R_n^{(n-1)}(x)}{n!x}\end{aligned}$$

בתנאי שהגבול האחרון קיים. גזירה $n-1$ פעמים של P_n נותנת

$$P_n^{(n-1)}(x) = f^{(n-1)}(0) + f^{(n)}(0)\cdot x$$

(בדקו!) ולכן

$$\begin{aligned}\lim_{x\to 0}\frac{R_n^{(n-1)}(x)}{n!x} &= \frac{1}{n!}\lim_{x\to 0}\frac{f^{(n-1)}(x) - P_n^{(n-1)}(x)}{x}\\ &= \frac{1}{n!}\lim_{x\to 0}\frac{f^{(n-1)}(x) - (f^{(n-1)}(0) - f^{(n)}(0)x)}{x}\\ &= \frac{1}{n!}\lim_{x\to 0}(\frac{f^{(n-1)}(x) - f^{(n-1)}(0)}{x-0} - f^{(n)}(0))\end{aligned}$$

הגבול האחרון קיים ושווה ל־0 מהגדרת הנגזרת, כי $f^{(n)} = (f^{(n-1)})'$. בכך סיימנו את ההוכחה. ■

המטרה הבאה שלנו היא להוכיח שמבין כל הפולינומים ממעלה n, פולינום טיילור P_n של f הוא היחיד שנבדל מ־f בפונקציה אפסית.

למה 11.3.6 אם $p \neq q$ פולינומים ממעלה לכל היותר n, אז $|p(x) - q(x)| \neq o(x^n)$.

הוכחה נסמן $p(x) = \sum_{k=0}^n a_k x^k$ ו־ $q(x) = \sum_{k=0}^n b_k x^k$. מההנחה $p \neq q$ נובע שיש m מינימלי, $m \leq n$, כך ש־ $a_m \neq b_m$, ואז

$$p(x) - q(x) = \sum_{k=m}^{n}(a_k - b_k)x^k$$

לכן

$$\frac{p(x)-q(x)}{x^n}=\sum_{k=m}^{n}(a_k-b_k)x^{k-n}=x^{m-n}\sum_{k=m}^{n}(a_k-b_k)x^{k-m}$$

הסכום באגף ימין הוא פולינום שאינו מתאפס ב־0 (כי $a_m \neq b_m$), ואילו הגורם x^{m-n} איננו חסום בסביבת 0 כאשר $m>n$, והוא שווה זהותית ל־1 כאשר $m=n$. לכן הגבול $\lim_{x\to 0}\frac{|p-q|}{x^n}$ הוא ∞ במקרה ש־ $m<n$ והוא שווה ל־$|a_n-b_n|$ במקרה ש־ $m=n$, ובכל מקרה איננו 0. ■

יחד עם המשפט הקודם נקבל את האפיון שהזכרנו עבור P_n.

משפט 11.3.7 אם f גזירה n פעמים ב־0 אז פולינום מקלורן P_n מסדר n הוא הפולינום היחיד ממעלה n המקיים $f(x)=P_n(x)+o(x^n)$

הוכחה ראינו קודם ש־ $f(x)-P_n(x)=o(x^n)$. יהי Q פולינום ממעלה n המקיים $f(x)-Q(x)=o(x^n)$. מאחר ש־

$$P_n-Q=(P_n-f)+(f-Q)$$

ושני המחוברים מימין הם $o(x^n)$, הרי ש־ P_n-Q גם כן אפסי מסדר $o(x^n)$ (הוכיחו!), ומהלמה נובע ש־ $Q=P_n$. ■

מהדיון למעלה ניתן לקבל הערכה דומה עבור מידת הקירוב של f על ידי פולינום טיילור סביב נקודה x_0 כלשהי. נניח ש־f גזירה n פעמים ב־x_0 ויהי P_n פולינום טיילור מסדר n של f ב־x_0. נגדיר $\widetilde{f}(x)=f(x+x_0)$ ו־$\widetilde{P}_n$ פולינום מקלורן של $\widetilde{f}$. כפי שראינו בדיון בעמוד 502, $P_n(x)=\widetilde{P}_n(x-x_0)$. לכן

$$\lim_{x\to x_0}\frac{f(x)-P_n(x)}{(x-x_0)^n}=\lim_{x\to x_0}\frac{\widetilde{f}(x-x_0)-\widetilde{P}_n(x-x_0)}{(x-x_0)^n}=\lim_{t\to 0}\frac{\widetilde{f}(t)-\widetilde{P}_n(t)}{t^n}=0$$

כאשר השוויון הימני נובע ממשפט 11.3.5. יתר על כן אם Q פולינום ממעלה n שמקיים

$$\lim_{x\to x_0}\frac{f(x)-Q(x)}{(x-x_0)^n}=0$$

נוכל להגדיר $\widetilde{Q}(x)=Q(x+x_0)$ ובדיקה מראה ש־ $\widetilde{f}(x)-\widetilde{Q}(x)=o(x^n)$ (בדקו!). מכאן ש־ $\widetilde{Q}=\widetilde{P}_n$, ולכן $Q=P_n$. אם כן הוכחנו את המשפט הבא:

משפט 11.3.8 תהי f גזירה n פעמים ב־x_0. אם נסמן ב־P_n את פולינום טיילור מסדר n של f סביב x_0, אז P_n הוא הפולינום היחיד ממעלה n שמקיים

$$\lim_{x\to x_0}\frac{f(x)-P_n(x)}{(x-x_0)^n}=0$$

דוגמאות

1. נחשב את $\lim_{x\to 0}\frac{x-\sin x}{x^2}$. אפשר לעשות זאת בעזרת כלל לופיטל (יש להפעילו פעמיים), אך הנה דרך מידית יותר שלא מעורב בה חישוב. מהנוסחה עבור פולינום מקלורן מסדר 2 של $\sin$ אנו יודעים ש־

$$\sin x = 0 + x + 0x^2 + o(x^3) = x + o(x^3)$$

ולכן

$$\frac{x-\sin x}{x^2} = \frac{x-(x+o(x^3))}{x^2} = \frac{o(x^3)}{x^2}$$

מהביטוי הימני ברור שהגבול של המנה ב־0 הוא 0 (למה?).

2. הנה הוכחה חלופית של משפט 8.6.3, המתבססת על פולינומי טיילור. נניח ש־ $n-1$ הנגזרות הראשונות של f ב־x_0 מתאפסות, ו־ $f^{(n)}(x_0)\neq 0$. אז פולינום טיילור מסדר n של f סביב x_0 הוא

$$P_n(x) = f(x_0) + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

קיימת פונקציה α^* רציפה ומתאפסת ב־0 כך ש־

$$\begin{aligned} f(x) &= f(x_0) + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + (x-x_0)^n\cdot\alpha^*(x-x_0) \\ &= f(x_0) + (\frac{f^{(n)}(x_0)}{n!} + \alpha^*(x-x_0))(x-x_0)^n \end{aligned}$$

קיום נקודות קיצון ועליה וירידה ב־x_0 עבור הפונקציה $(x-x_0)^n$ תלויה ב־n כמו בניסוח משפט 8.6.3. לכן אילו בסכום לעיל היה מופיע קבוע במקום הסכום $\frac{f^{(n)}(x_0)}{n!} + \alpha^*(x-x_0)$ אז התנהגות f בסביבת x_0 הייתה נקבעת על ידי הפונקציה $(x-x_0)^n$, כלומר, היא הייתה תלויה ב־n באופן שאנו רוצים. במקרה שלפנינו אמנם $\frac{f^{(n)}(x_0)}{n!} + \alpha^*(x-x_0)$ אינו קבוע, אלא תלוי ב־x, אבל זו פונקציה רציפה ב־x_0 שאינה מתאפסת שם, ואפשר להפוך את הנימוק הקודם למדויק.

נוכיח בפירוט את המקרה ש־n זוגי ו־ $f^{(n)}(x_0) > 0$ ואת יתר המקרים נשאיר כתרגיל. יהי $0 < \varepsilon < \frac{1}{n!}f^{(n)}(x_0)$. מרציפות של α^* יש סביבה V של x_0 כך שלכל $x\in V$ מתקיים $f^{(n)}(x_0) - \alpha^*(x-x_0) > \varepsilon$, ולכן באותה סביבה מתקיים

$$f(x) > f(x_0) + \varepsilon(x-x_0)^n$$

אבל n זוגי ולכן המחובר $\varepsilon(x-x_0)^n$ חיובי ממש לכל $x\neq x_0$ באותה סביבה, ואנו מסיקים ש־x_0 היא מינימום מקומי של f.

משפט 11.3.8 לעיל מאפשר להוכיח כמה משפטי אריתמטיקה לפולינומי טיילור. למעשה מדובר בהכללות של משפטי האריתמטיקה לנגזרות שהוכחנו בפרק 8.

למשל, יהיו P_n, Q_n פולינומי מקלורן מסדר n של f, g בהתאמה. אז $S_n = P_n + Q_n$ הוא פולינום ממעלה n ומתקיים

$$(f+g) - S_n = (f - P_n) + (g - Q_n) = o(x^n) + o(x^n) = o(x^n)$$

ולכן S_n הוא פולינום טיילור של $f+g$.

יהיו P_n, Q_n פולינומי מקלורן מסדר n של הפונקציות f, g בהתאמה, ויהי S_n פולינום מקלורן מסדר n של הפונקציה $f \cdot g$. אנו רוצים לבטא את S_n בעזרת P_n, Q_n. ניחוש ראשון הוא ש־S_n שווה למכפלה $P_n \cdot Q_n$, אך באופן כללי זהו פולינום ממעלה $2n$ וממילא לא יכול להיות שווה ל־S_n, שהוא פולינום ממעלה n לכל היותר. ננסה אם כן למחוק מ־$P_n \cdot Q_n$ את כל החזקות של x שמעלתן גדולה מ־n ונראה שהתוצאה היא S_n.

נשתמש בסימון הבא: עבור פולינום $r(x) = \sum_{k=0}^{d} a_k x^k$ ו־ $m \le d$ נסמן ב־$[r]_m$ את הפולינום

$$[r]_m(x) = \sum_{k=0}^{m} a_k x^k$$

המתקבל מ־r על ידי מחיקת המחוברים $a_{m+1}x^{m+1}, a_{m+2}x^{m+2}, \ldots, a_d x^d$.

למה 11.3.9 $r(x) - [r]_m(x) = o(x^{m+1})$.

הוכחה קל לראות שחלק מהמחוברים בהפרש $r - [r]_m$ מבטלים זה את זה:

$$\begin{aligned} r(x) - [r]_m(x) &= \sum_{k=0}^{d} a_k x^k - \sum_{k=0}^{m} a_k x^k \\ &= x^m \cdot \sum_{k=m+1}^{d} a_k x^{k-m} \end{aligned}$$

השורה האחרונה היא מכפלה של x^m עם פולינום ללא מקדם חופשי, וקיבלנו ש־$r(x) - [r]_m(x) = o(x^m)$. ■

נשוב לבעיה שלנו. נגדיר $S_n = [P_n \cdot Q_n]_n$, ונראה ש־S_n הוא פולינום מקלורן מסדר n של fg. לפי משפט 11.3.7 די לנו להראות ש־

$$f(x)g(x) - S_n(x) = o(x^n)$$

ואמנם

$$\begin{aligned} f(x)g(x) &= (P_n(x) + o(x^n))(Q_n(x) + o(x^n) \\ &= P_n(x)Q_n(x) + P_n(x)o(x^n) + Q_n(x)o(x^n) + o(x^n)o(x^n) \\ &= P_n(x)Q_n(x) + o(x^n) \end{aligned}$$

(המעבר בין שתי השורות האחרונות הוא לפי למה 11.3.4). קיבלנו אפוא

$$\begin{aligned} f(x)g(x) &= P_n(x)Q_n(x) + o(x^n) \\ &= S_n(x) + (P_n(x)Q_n(x) - S_n(x)) + o(x^n) \end{aligned}$$

אבל לפי הלמה האחרונה מתקיים $P_n(x)Q_n(x) - S_n(x) = o(x^n)$, וקיבלנו את המבוקש.

תרגילים

1. חשבו את פולינום טיילור מסדר 10 של $e^x \sin x$ ושל $e^{\sin x} - 1$.
2. מצאו את פולינום טיילור מסדר n של e^{x^2} (אל תנסו לגזור. במקום זאת מצאו את פולינום מקלורן Q_n של e^x ונחשו את פולינום טיילור של e^{x^2} על ידי התבוננות ב־$Q_n(x^2)$. השתמשו במשפט היחידות כדי להוכיח שזו התוצאה הנכונה.
3. היעזרו בפולינום טיילור כדי לחשב את הגבולות הבאים:
   (א) $\lim_{x\to 0} \frac{\cos x + x^2/2}{x^4}$.
   (ב) $\lim_{x\to 0} \frac{\cos x \cdot \sin x - x}{(\sin x)^3}$.
   (ג) $\lim_{x\to 0} \frac{(e^x-1)^2}{\ln(1+x)+x^2/2}$.
   (ד) $\lim_{x\to 0} \frac{\sin^2(e^x-1)\cos x}{\cos x - 1}$.
4. הוכיחו את למה 11.3.4 על כללי התחשיב של פונקציות אפסיות.
5. השלימו את הוכחת המקרים הנותרים בדוגמה (2) בעמוד 507.
6. אם R_n הוא שארית טיילור של פונקציה f ביחס לפיתוח סביב x_0, מתי מתקיים ש־ $R_n(x) = o(x^{n+1})$?
7. הסיקו את כלל המכפלה לנגזרות מכלל המכפלה לפולינומי טיילור.
8. יהיו f, g פונקציות גזירות n פעמים ב־x_0 ויהיו P_n, Q_n פולינומי טיילור מסדר n שלהן שם. האם $[P_nQ_n]_n$ הוא פולינום טיילור מסדר n של fg סביב x_0? אם כן הוכיחו, אם לא תנו דוגמה נגדית ומצאו את הנוסחה הנכונה.
9. תהי f גזירה n פעמים סביב x_0 ובעלת פולינום טיילור P_n שם, ותהי g גזירה n פעמים סביב $y_0 = f(x_0)$ ובעלת פולינום טיילור Q_n שם. נניח ש־ $g(y_0) = 0$. הראו שפולינום טיילור מסדר n של $g \circ f$ ב־x_0 נתון על ידי $[Q_n(P_n)]_n$. האם זה נכון ללא ההנחה ש־ $g(y_0) = 0$?

11.4 הערכת השארית של פולינום טיילור בנקודה

בסעיף זה נתעניין בשאלה מתי ובאיזו מידה הערכים של פולינום טיילור P_n בנקודה קבועה x מקרבים את הערך של f ב־x, ובפרט האם $\lim_{n\to\infty} P_n(x) = f(x)$ (או באופן שקול $\lim_{n\to\infty} R_n(x) = 0$). כדי לענות על השאלה נצטרך למצוא חסמים עבור $R_n(x)$. במקרים חשובים רבים השארית אכן שואפת ל־0 כאשר $n\to\infty$, אך לא תמיד.

משפט 11.4.1 (צורת לגרנג' לשארית טיילור) תהי f גזירה n פעמים בקטע הסגור I שקצותיו x_0, x, ונניח ש־$f^{(n)}$ גזירה בפנים של I ורציפה ב־I (כולל בקצוות). אם P_n, R_n הם פולינום טיילור ושארית טיילור מסדר n סביב x_0 אז קיימת נקודה c בפנים של I כך ש־

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}$$

הוכחה כזכור $R_n = f - P_n$, וכן $P_n^{(k)}(x_0) = f^{(k)}(x_0)$ לכל $0 \le k \le n$. בפרט $R_n^{(k)}(x_0) = 0$ לכל $0 \le k \le n$. לכן על ידי הפעלת משפט ערך הביניים של קושי על המנה $\frac{R_n}{(x-x_0)^{n+1}}$ מקבלים שיש נקודה c_1 בין x_0 ל־x כך ש־

$$\frac{R_n(x)}{(x-x_0)^{n+1}} = \frac{R_n(x) - R_n(x_0)}{(x-x_0)^{n+1} - (x_0-x_0)^{n+1}} = \frac{R_n^{(1)}(c_1)}{(n+1)(c_1-x_0)^n}$$

הפעלת אותו שיקול על המנה $\frac{R_n^{(1)}(c_1)}{(c_1-x_0)^n}$ נותנת נקודה c_2 בין x_0 ל־c_1 כך ש־

$$\begin{aligned}\frac{R_n^{(1)}(c_1)}{(n+1)(c_1-x_0)^{n+1}} &= \frac{R_n^{(1)}(c_1) - R_n^{(1)}(x_0)}{(n+1)(c_1-x_0)^n - (n+1)(x_0-x_0)^n} \\ &= \frac{R_n^{(2)}(c_2)}{(n+1)n(c_2-x_0)^{n-1}}\end{aligned}$$

נחזור על אותו שיקול $n+1$ פעמים ונסיק שיש נקודה $c = c_{n+1}$ בין x_0 ל־x כך ש־

$$\frac{R_n(x)}{(x-x_0)^{n+1}} = \frac{R^{(n+1)}(c)}{(n+1)!}$$

והמשפט מתקבל על ידי העברת אגפים. ■

שימו לב לדמיון הרב בין צורת לגרנג' של השארית למחוברים של פולינום טיילור, שצורתם $\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$. בעזרת צורת לגרנג' אפשר לרשום

$$f(x) = P_n(x) + R_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}$$

לכאורה כתבנו את f כפולינום, אך לא כך: זה אינו פולינום כי הנקודה c אינה קבועה, אלא תלויה ב־x (וכמובן גם ב־n).

דוגמאות

1. ראינו שאם $f(x) = e^x$ ואם P_n הוא פולינום טיילור של f סביב 0, אז $P_n(x) = \sum_{k=0}^{n} \frac{1}{k!} x^k$. כמו־כן לכל נקודה $x \in \mathbb{R}$ ולכל k טבעי מתקיים $f^{(k)}(x) = e^x$. לכן שארית טיילור $R_n(x)$ מקיימת, לפי צורת לגרנג',

$$R_n(x) = \frac{e^c}{(n+1)!} x^{n+1}$$

כאשר $0 < c < x$. נסמן $d = \max\{1, e^x\}$, כך ש־ $e^c \leq d$ ו־d אינו תלוי ב־n. מכאן ש־

$$|R_n(x)| \leq d \cdot \frac{x^{n+1}}{n!}$$

לכן $\lim_{n\to\infty} R_n(x) = 0$, או באופן שקול,

$$\lim_{n\to\infty} P_n(x) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k} x^k = e^x$$

הדבר נכון לכל x, וקיבלנו את הנוסחה

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

הערכת השארית נותנת לנו גם אפשרות לחשב את e (או e^x) בכל רמת דיוק שנרצה. למשל נחשב את e בדיוק של אלפית. אז $P_n(1)$ היא הערכה ל־ $e = e^1$ ברמת דיוק שלא נופלת מ־$R_n(1)$. אם נזכור ש־$2 \leq e \leq 3$ אנו מקבלים

$$\frac{2}{(n+1)!} \leq |R_n(1)| \leq \frac{3}{(n+1)!}$$

לכן על מנת ש־ $R_n(1) < 0.001$ די לבחור את n כך שיקיים $(n+1)! > 3000$. בדיקה מראה ש־ $n = 6$ מספיק, ואנו מקבלים את ההערכה

$$\begin{aligned} e &= P_6(1) \\ &= 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} + \frac{1}{720} \\ &\approx 2.718 \end{aligned}$$

2. תהי $f(x)=\sin x$. חישבנו את פולינום מקלורן של $\sin$ בדוגמה (3) בעמוד 500. כמו־כן $|\sin^{(k)}(x)|\leq 1$ לכל $x\in\mathbb{R}$ ולכל $k\in\mathbb{N}$ ולכן מצורת לגרנג' לשארית

$$|R_n(x)|=\frac{|\sin^{(n+1)}(c)|}{(n+1)!}|x|^{n+1}\leq\frac{|x|^{n+1}}{(n+1)!}$$

ובפרט ל־x קבוע מתקיים $\lim_{n\to\infty}R_n(x)=0$. קיבלנו שלכל $x\in\mathbb{R}$ מתקיים

$$\begin{aligned}\sin x &= \lim_{n\to\infty}P_n(x)\\ &= \lim_{n\to\infty}\sum_{k=0}^{[(n-1)/2]}\frac{(-1)^k}{(2k+1)!}x^{2k+1}\\ &= =\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}x^{2k+1}\end{aligned}$$

באופן דומה מראים ש־

$$\cos x=\sum_{k=0}^{\infty}\frac{(-1)^n}{(2k)!}x^{2k}$$

לכל $x\in\mathbb{R}$.

ישנן גם הערכות אחרות לשארית טיילור. למשל,

משפט 11.4.2 (צורת קושי לשארית טיילור) תהי f גזירה $n+1$ פעמים בקטע הסגור I שקצותיו x_0,x. אם P_n,R_n הם פולינום טיילור ושארית טיילור מסדר n של f סביב x_0 אז קיימת נקודה c בפנים של I כך ש־

$$R_n(x)=\frac{f^{(n+1)}(c)}{n!}(x-c)^n\cdot(x-x_0)$$

או באופן שקול, קיים מספר $\theta\in(0,1)$ כך ש־

$$R_n(x)=\frac{f^{(n+1)}\left(x_0+\theta\left(x_1-x_0\right)\right)}{n!}(1-\theta)^n(x-x_0)^{n+1}$$

הוכחה בהוכחה של משפט 11.4.1 הפעלנו את משפט ערך הביניים של קושי על המנה $\frac{R_n(x)}{(x-x_0)^{n+1}}$. השיטה הבאה מבוססת על אותו רעיון, אלא שבמקום להתייחס ל־$R_n(x)$ כפונקציה של x, נתייחס אל x כנקודה קבועה ונחשוב על R_n כפונקציה של הנקודה x_0 סביבה פיתחנו את f לפולינום טיילור. כדי להדגיש את התלות בשני הפרמטרים נשתמש בסימון כללי יותר לפולינום טיילור ולשארית: נרשום $\widehat{P}_n(s,t)$

עבור הערך בנקודה t של פולינום טיילור מסדר n של f, כשמפתחים אותו סביב s. כלומר:

$$\widehat{P}_n(s,t) = \sum_{k=0}^{n} \frac{f^{(k)}(s)}{k!}(t-s)^k$$

באופן דומה נסמן $\widehat{R}_n(s,t) = f(t) - \widehat{P}_n(s,t)$. נקבע כעת את x ונגדיר

$$\varphi(s) = \widehat{R}_n(s,x) = f(x) - (\sum_{k=0}^{n} \frac{f^{(k)}(s)}{k!}(x-s)^k)$$

כלומר, $\varphi(s)$ הוא מידת השגיאה בנקודה x בין f לבין פולינום טיילור מסדר n של f אשר פותח סביב s. כאשר $s = x$ השגיאה היא 0, כלומר $\varphi(x) = 0$ (זה ברור מהדיון, אך אפשר בקלות לבדוק זאת מהנוסחה עבור $\widehat{R}_n(x,x)$). לכן אם מניחים ש־φ גזירה אפשר להעריך את $\varphi(x_0)$ באופן הבא: לפי משפט לגרנג' קיימת נקודה c בין x_0 ל־x כך ש־

$$\varphi(x_0) - \varphi(x) = \varphi'(c)(x_0 - x)$$

אבל $\varphi(x) = 0$, ואילו $R_n(x_0) = \widehat{R}_n(x_0,x) = \varphi(x_0)$, וקיבלנו

$$R_n(x) = \varphi'(c)(x_0 - x)$$

לא נותר אלא לגזור ולראות איזה מין הערכה נותנת הנוסחה הזאת. לפי הנחות המשפט φ גזירה בקטע הפתוח שקצותיו x_0, x, ולכן לכל $k = 1, \ldots, n$ ולכל s בין x_0 ל־x מתקיים

$$\frac{d}{ds}(\frac{1}{k!}f^{(k)}(s)(x-s)^k) = \frac{1}{k!}f^{(k+1)}(s)(x-s)^k - \frac{1}{(k-1)!}f^{(k)}(s)(x-s)^{k-1}$$

ואילו עבור $k = 0$ מקבלים

$$\begin{aligned} \frac{d}{ds}(\frac{1}{k!}f^{(k)}(s)(x-s)^k) &= \frac{d}{ds}(\frac{1}{0!}f(s)\cdot 1) \\ &= f'(s) \end{aligned}$$

לכן בחישוב של φ' מופיע סכום טלסקופי, ואנו מקבלים

$$\begin{aligned} \varphi'(s) = &= \frac{d}{ds}(\sum_{k=0}^{n} \frac{f^{(k)}(s)}{k!}(x-s)^k) \\ &= f'(s) + \sum_{k=1}^{n}(\frac{1}{k!}f^{(k+1)}(s)(x-s)^k - \frac{1}{(k-1)!}f^{(k)}(s)(x-s)^{k-1}) \\ &= \frac{f^{(n+1)}(s)}{n!}(x-s)^n \end{aligned}$$

נציב את הנוסחה $\varphi'(s) = \frac{1}{n!} f^{(n+1)}(s)(x-s)^n$ בשוויון $R_n(x) = \varphi'(c)(x-x_0)$ ונקבל

$$R_n(x) = \frac{f^{(n+1)}(c)}{n!}(x-c)^n \cdot (x-x_0)$$

שהוא את הביטוי הראשון שרצינו להוכיח. כדי לקבל את הביטוי השני נחלק ונכפיל את הראשון ב־ $(x-x_0)^n$ ונקבל

$$R_n(x) = \frac{f^{(n+1)}(c)}{n!} \cdot (\frac{x-c}{x-x_0})^n (x-x_0)^{n+1}$$

נסמן $1-\theta = \frac{x-c}{x-x_0}$, כך שמתקיים $\theta \in (0,1)$ ו־ $c = x_0 + \theta(x-x_0)$. נציב שוויונות אלה בביטוי שקיבלנו עבור $R_n(x)$ ונקבל את הביטוי השני שרצינו. ■

שימו לב שהמספרים c, θ במשפט תלויים ב־x וב־n, אף שהדבר אינו מופיע במפורש בסימון.

אפשר להסיק את נוסחת לגרנג' באותה שיטה כמו בהוכחה האחרונה. נסמן $\psi(s) = (x_1 - s)^{n+1}$, ונשים לב ש־$\psi'$ קיים ואינו מתאפס בין x_0 ל־x_1. לכן לפי משפט קושי קיימת נקודה c בין x_0 ל־x_1 כך ש־

$$\frac{\varphi(x_0) - \varphi(x_1)}{\psi(x_0) - \psi(x_1)} = \frac{\varphi'(c)}{\psi'(c)}$$

אם נציב $\varphi(x_1) = 0$ ואת הביטויים השונים עבור φ', ψ, ψ' ונעביר אגפים, נקבל את צורת לגרנג' (בדקו!).

דוגמאות

1. תהי $f(x) = \ln(1+x)$. ראינו כבר ש־

$$f^{(n)}(x) = (-1)^{n-1}(n-1)!\frac{1}{(1+x)^n}$$

במקרה זה הערכת השארית בצורת לגרנג' אינה טובה מספיק (נסו להשתמש בה. מה קורה?) ולכן נשתמש בצורת קושי לשארית. קיים $\theta \in (0,1)$ כך ש־

$$\begin{aligned} R_n(x) &= (-1)^{n-1}\frac{(n-1)!}{n!} \cdot \frac{1}{(1+\theta x)^n}(1-\theta)^n x^{n+1} \\ &= \frac{(-1)^{n-1}}{n} \cdot (x\frac{1-\theta}{1+\theta x})^n \cdot x \end{aligned}$$

עבור $x \in (-1,1)$ מתקיים $|x \cdot \frac{1-\theta}{1+\theta x}| = |x| \cdot \frac{1-\theta}{1+\theta x} < 1$ ולכן עבור $x \in (-1,1)$ קבוע מתקיים $|R_n(x)| \leq \frac{1}{n}$, ומכאן ש־ $\lim_{n\to\infty} R_n(x) = 0$, כלומר

$$\ln(1+x) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{(-1)^k}{k} x^k = \sum_{k=0}^{\infty} \frac{(-1)^k}{k} x^k$$

אפשר להראות שהנוסחה נכונה גם כאשר $x = 1$. אנו משאירים את הפרטים כתרגיל.

2. הנה דוגמה לפונקציה שפולינום מקלורן שלה אינו קירוב טוב שלה. תהי $f(x) = e^{-1/x^2}$ עבור $x \neq 0$ ו־ $f(0) = 0$. אפשר להראות ש־f גזירה מכל סדר ומתקיים $f^{(k)}(0) = 0$ לכל k (ראו תרגיל (8) בעמוד 347), ולכן פולינום מקלורן הוא פולינום האפס לכל n. מצד שני, $f(x) \neq 0$ לכל $x \neq 0$ ולכן הסדרה $(P_n(x))_{n=1}^{\infty}$ של פולינומי מקלורן אינה מתכנסת ל־$f(x)$ לאף $x \neq 0$.

תרגילים

1. תהי f פונקציה גזירה אינסוף פעמים ב־0 ונניח שיש קבוע M כך שלכל n ולכל $t \in [0,x]$ מתקיים $|f^{(n)}(t)| \leq M$. הוכיחו שמתקיים $\lim_{n\to\infty} P_n(x) = f(x)$, כאשר P_n הם פולינומי מקלורן של f.

2. חשבו בדיוק של $\frac{1}{100}$ את:

   (א) e^2.

   (ב) $\sin 1$.

   (ג) $\ln 2$ (כאן מספיק לחשב בדיוק של $\frac{1}{10}$).

3. הוכיחו שהשארית R_n של פולינום מקלורן P_n של f מקיימת

$$R_n(x) = \frac{(-1)^n}{n!} \int_0^x f^{(n+1)}(t)(x-t)^n dt$$

(רמז: בצעו ל־$\int f^{(n+1)}(t)(x-t)^n dt$ אינטגרציה בחלקים $n+1$ פעמים). נוסחה זו נקראת **הצורה האינטגרלית לשארית**.

4. הראו ש־e אי־רציונלי (רמז: $e = P_n(1) + R_n(1)$ כאשר P_n, R_n פולינום ושארית מקלורן של הפונקציה e^x. הניחו בשלילה ש־ $e = \frac{r}{s}$ רציונלי והיעזרו בהערכת השארית של לגרנג').

11.5 טורי חזקות ונוסחת קושי־הדמר

הגדרה 11.5.1 טור חזקות (power series) הוא טור פונקציות מהצורה $\sum_{n=0}^{\infty} a_n x^n$, כאשר a_n מספרים ממשיים, הנקראים **המקדמים** (coefficients) של הטור. נאמר

שפונקציה ממשית f ניתנת לייצוג כטור חזקות בקבוצה $A \subseteq \operatorname{dom} f$ אם קיים טור חזקות $\sum_{n=0}^{\infty} a_n x^n$ שמתכנס ב־A ומקיים $f(x) = \sum_{n=0}^{\infty} a_n x^n$ בכל $x \in A$.

באופן כללי, **טור חזקות סביב** x_0 הוא טור פונקציות מהצורה $\sum_{n=0}^{\infty} a_n (x - x_0)^n$. נאמר ש־$f$ ניתנת לייצוג בקבוצה $A \subseteq \operatorname{dom} f$ כטור חזקות סביב x_0 אם קיים טור חזקות סביב x_0 התקיים $f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ לכל $x \in A$.

טורי חזקות הם הכללה טבעית של מושג הפולינום. ההבדל כמובן הוא שמספר המחוברים אינו סופי, וכפי שנראה מבחינת תורת הגזירה והאינטגרציה שלהם זהו הבדל טכני בלבד. ליתר דיוק, אפשר בדרך כלל להתייחס לטורי חזקות בדיוק כפי שמתייחסים לפולינומים, אלא שההצדקה של הפעולות השונות מורכבת יותר.

בסעיף הנוכחי נתייחס בעיקר לשאלת ההתכנסות של טורי חזקות, ונתייחס רק לטורי חזקות סביב אפס, שכתיבתם פשוטה יותר. המקרה של טורי חזקות כלליים יותר נובע ממקרה זה.

דוגמאות

1. כל פולינום הוא טור חזקות: אם $p(x) = \sum_{n=0}^{d} a_n x^n$ נגדיר $a_n = 0$ לכל $n > d$, ואז $p(x) = \sum_{n=0}^{\infty} a_n x^n$. בדוגמה זו שאלת ההתכנסות היא טריביאלית, והטור מתכנס בכל הישר (אפילו במ"ש).

2. נתבונן בטור החזקות $\sum_{n=0}^{\infty} x^n$. כפי שכבר ראינו בדוגמה 1 בעמוד 466, טור זה מתכנס נקודתית ב־$(-1,1)$ ל־ $f(x) = \frac{1}{1-x}$ וההתכנסות היא במ"ש בכל תת־קטע סגור של $(-1,1)$. הטור אינו מתכנס באף נקודה x המקיימת $|x| \geq 1$.

3. תהי f פונקציה גזירה אינסוף פעמים ב־0, ויהיו $a_n = \frac{f^{(n)}(0)}{n!}$. אנו מקבלים כך טור חזקות $\sum_{k=0}^{\infty} a_k x^k$. שימו לב שהסכומים החלקיים של טור זה הם בדיוק פולינומי מקלורן P_n של הפונקציה. טור החזקות הזה נקרא **טור טיילור** (או לפעמים טור טיילור־מקלורן) של f. השאלה האם טור טיילור של f מתכנס אליו בקבוצה כלשהי שקול לשאלה האם סדרת פולינומי מקלורן שלה מתכנסים אליה באותה הקבוצה.

4. בסעיף הקודם ראינו שהפונקציות $e^x, \ln(1+x), \sin x, \cos x$ הן הגבול הנקודתי (בקבוצה מתאימה) של פולינומי מקלורן שלהן. מכאן שפונקציות אלה ניתנות לייצוג כטור חזקות באותן קבוצות. למשל, פולינום מקלורן P_n של e^x סביב 0 הוא $P_n(x) = \sum_{k=0}^{n} \frac{1}{k!} x^k$. ראינו ש־ $\lim_{n\to\infty} P_n(x) = f(x)$ לכל $x \in \mathbb{R}$ ולכן לכל x מתקיים

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$$

באופן דומה ראינו שלכל x מתקיים

$$\begin{aligned}\sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \\ \cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}\end{aligned}$$

ושלכל $x \in (-1,1)$,

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k$$

הבעיה הבסיסית שנתעניין בה בסעיף זה היא אפיון תחום ההתכנסות של טור חזקות. בעוד שעבור טור פונקציות כללי תחום ההתכנסות יכול להיות קבוצת נקודות שרירותית, עבור טורי חזקות המצב פשוט יותר. שימו לב שטור חזקות סביב 0 תמיד מתכנס בנקודה 0.

הגדרה 11.5.2 יהי $\sum_{n=0}^{\infty} a_n x^n$ טור חזקות. מספר ממשי R נקרא **רדיוס ההתכנסות** (radius of convergence) של הטור אם הטור מתכנס עבור $|x| < R$ ומתבדר עבור $R < |x|$. אם הטור מתכנס לכל x אז אומרים שרדיוס ההתכנסות הוא ∞.

מההגדרה ברור ש־R כזה, אם הוא קיים, הוא יחיד. קיומו מובטח על ידי המשפט הבא.

משפט 11.5.3 (נוסחת קושי־הדמר)[3] יהי $\sum_{n=0}^{\infty} a_n x^n$ טור חזקות. יהי

$$R = \frac{1}{\limsup \sqrt[n]{|a_n|}}$$

(אם המכנה הוא ∞ נגדיר $R = 0$, ואם המכנה הוא 0 נגדיר $R = \infty$). אז הטור מתכנס אם $|x| < R$, ומתבדר אם $R < |x|$ (דהיינו, R הוא רדיוס ההתכנסות של הטור).

הוכחה נוכיח את המשפט למקרה $R \neq 0, \infty$. עלינו להראות שהטור $\sum_{n=0}^{\infty} a_n x^n$ מתכנס עבור $|x| < R$ ומתבדר עבור $|R| < x$.

בכיוון אחד, די שנראה שאם $|x| < R$ אז טור המספרים $\sum_{n=0}^{\infty} a_n x^n$ מתכנס בהחלט, כי אז הוא בוודאי מתכנס. נשים לב ש־

$$\limsup_{n\to\infty} \sqrt[n]{|a_n x^n|} = |x| \cdot \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{|x|}{R} < 1$$

[3] Jacques Hadamard, 1865-1963.

ולכן, לפי מבחן השורש (משפט 6.3.5), הטור $\sum |a_n x^n|$ מתכנס, כפי שרצינו.

בכיוון השני, נניח ש־ $|x| > R$, כלומר: $\limsup \sqrt[n]{|a_n|} > \frac{1}{|x|}$. נובע מכאן שלאינסוף n־ים מתקיים $\sqrt[n]{|a_n|} > 1/|x|$. לכן לאינסוף n־ים מתקיים $|a_n x^n| \geq 1$, ואיברי הטור $\sum_{n=1}^{\infty} a_n x^n$ אינם שואפים לאפס וממילא הטור אינו מתכנס.

הוכחת המקרים $R = 0, \infty$ דומה, אנו משאירים את הפרטים כתרגיל. ■

מסקנה 11.5.4 טור חזקות מתכנס נקודתית בקטע שמרכזו 0, או בכל הישר (ייתכן שהקטע הוא הקטע המנוון $[0, 0]$).

במקרים מסוימים קל יותר לחשב רדיוס התכנסות של טור חזקות בעזרת המבחן הבא:

מסקנה 11.5.5 יהי $\sum_{n=1}^{\infty} a_n x^n$ טור חזקות ונניח ש־$a_n \neq 0$ לכל n. נגדיר

$$R = (\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|})^{-1}$$

בהנחה שהגבול קיים במובן הרחב (את המקרה שהגבול שווה ל־0 או־∞ מפרשים כמו במשפט הקודם). אז הטור מתכנס עבור $|x| < R$ ומתבדר עבור $|x| > R$.

הוכחה בפרק על סדרות ראינו שאם a_n סדרת מספרים חיוביים ממש ואם הגבול $L = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}$ קיים במובן הרחב אז $\lim_{n\to\infty} \sqrt[n]{a_n} = L$ (ראו למה 5.4.12). כעת המסקנה נובעת ממשפט קושי־הדמר. ■

דוגמאות

1. נתבונן בטור $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ ונסמן $a_n = \frac{1}{n!}$. כאן קשה ליישם את מבחן קושי־הדמר אבל אפשר להשתמש במסקנה 11.5.5: מתקיים

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{n!}{(n+1)!} = \lim_{n\to\infty} \frac{1}{n+1} = 0$$

ולכן $R = \infty$, כלומר הטור מתכנס בכל הישר. יכולנו להסיק זאת גם מכך שאנו כבר יודעים שהטור מתכנס בכל נקודה ל־e^x.

2. יהי $\sum_{n=1}^{\infty} n^n x^n$. כאן מקדמי הטור הם $a_n = n^n$, ומתקיים $\sqrt[n]{a_n} = n \to \infty$. לכן $R = 0$, ולפי משפט קושי־הדמר הטור אינו מתכנס באף $x \neq 0$.

3. נתבונן שוב בטור $\sum_{n=0}^{\infty} x^n$. כאן $a_n = 1$ לכל n, ולפי נוסחת קושי־הדמר $R = 1$ (ואמנם ראינו קודם שהטור מתכנס עבור $|x| < 1$ ואינו מתכנס עבור $|x| > 1$). כאשר $x = \pm 1$, האיבר הכללי בטור אינו שואף לאפס, ולכן הטור אינו מתכנס שם. לפיכך תחום ההתכנסות הוא הקטע הפתוח $(-1, 1)$.

4. נתבונן בטור $\sum_{n=1}^{\infty}\frac{1}{n}x^n$. מקדמי הטור הם $\frac{1}{n}$ ומכיוון ש־ $\sqrt[n]{\frac{1}{n}}\to 1$, רדיוס ההתכנסות של הטור הוא 1. באשר לקצוות הקטע, כאשר $x=1$ הטור המתקבל הוא הטור ההרמוני $\sum_{n=1}^{\infty}\frac{1}{n}$, ולכן אינו מתכנס. מאידך, כאשר $x=-1$ הטור המתקבל הוא טור לייבניץ $\sum_{n=1}^{\infty}\frac{(-1)^n}{n}$, ולכן מתכנס. מכאן שתחום ההתכנסות הוא הקטע החצי־פתוח $[-1,1)$.

5. נתבונן בטור $\sum_{n=1}^{\infty}\frac{1}{n^2}x^n$. כעת $a_n=\frac{1}{n^2}$, ומנוסחת קושי הדמר מקבלים $R=1$ (בדקו!). בדוגמה הנוכחית, הטור המתקבל כאשר $x=1$ הוא $\sum_{n=1}^{\infty}\frac{1}{n^2}$, והטור המתקבל כאשר $x=-1$ הוא $\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}$. שני הטורים האלה מתכנסים, ולכן תחום ההתכנסות הוא הקטע הסגור $[-1,1]$.

הדוגמאות האחרונות מראות שאם מספר ממשי R הוא רדיוס ההתכנסות של טור חזקות אז משפט קושי־הדמר אינו נותן מידע על התכנסות או אי־התכנסות הטור בנקודות $\pm R$, וייתכן שתחום ההתכנסות הוא כל אחד מהקטעים $[-R,R]$,$[-R,R)$, $(-R,R]$, $(-R,R)$.

ענינו בצורה מלאה על השאלה כיצד נראה תחום ההתכנסות של טור חזקות. נעבור לשאלה באילו קבוצות טור חזקות מתכנס במידה שווה. עבור טור פונקציות כללי אין קשר פשוט בין תחום ההתכנסות לבין אותן קבוצות בהן הטור מתכנס במ"ש, למעט העובדה הטריביאלית שקבוצה שבה יש התכנסות במ"ש מוכלת בתחום ההתכנסות הנקודתית. לעומת זאת, עבור טורי חזקות אפשר לומר מעט יותר.

למה 11.5.6 אם R הוא רדיוס ההתכנסות של טור חזקות אז הטור מתכנס בהחלט בקטע $(-R,R)$.

הוכחה יהי $\sum_{n=1}^{\infty}a_nx^n$ טור חזקות עם רדיוס התכנסות R. המקדמים של טור החזקות $\sum_{n=0}^{\infty}|a_n|x^n$ שווים בערכם המוחלט למקדמים של הטור המקורי, ולכן לפי נוסחת קושי־הדמר רדיוס ההתכנסות שלו גם כן שווה ל־R. עבור נקודה $x_0\in(-R,R)$ מתקיים $|x_0|<R$ ולכן לפי האמור הטור $\sum_{n=1}^{\infty}|a_n||x_0|^n$ מתכנס; וזה שקול להתכנסות בהחלט של $\sum_{n=1}^{\infty}a_nx_0^n$, כפי שרצינו. ■

טענה 11.5.7 יהי $\sum_{n=0}^{\infty}a_nx^n$ טור חזקות בעל רדיוס התכנסות $R>0$. אז הטור מתכנס במ"ש בכל תת־קטע סגור של $(-R,R)$.

הוכחה די להראות שלכל $0<r<R$ הטור מתכנס במ"ש ב־$[-r,r]$, כי כל תת־קטע סגור של $(-R,R)$ מוכל בקטע $[-r,r]$ כזה (הבהירו לעצמכם שאתם מבינים למה זה מספיק!).

נשתמש במבחן ויירשטראס (משפט 10.2.7): עלינו למצוא מספרים M_n כך ש־ $|a_nx^n|\le M_n$ לכל $x\in[-r,r]$, וכך שהטור $\sum_{n=0}^{\infty}M_n$ מתכנס. מכאן נסיק שטור החזקות מתכנס במ"ש בתחום $[-r,r]$.

מתכונות המונוטוניות של הפונקציות x^n נובע שלכל $x\in[-r,r]$ מתקיים

$$|a_nx^n|=|a_n|\cdot|x|^n\le|a_n|r^n$$

אם נראה שטור המספרים $\sum_{n=0}^{\infty}|a_n|r^n$ מתכנס נוכל לבחור $M_n=|a_n|r^n$ ולהסיק ממשפט ויירשטראס שהטור מתכנס במ"ש בקטע $[-r,r]$. התכנסות הטור $\sum|a_n|r^n$ נובעת מהלמה הקודמת, כי לפי ההנחה $0<r<R$. ■

שאלה שלא ענינו עליה עדיין היא מתי ההתכנסות של טור חזקות היא במ"ש בכל תחום ההתכנסות שלו. נשוב לשאלה זו בסעיף 11.8.

אפשר לנסח את כל התוצאות הקודמות עבור טורי חזקות סביב נקודה x_0 כללית, אלא שאז תחום ההתכנסות הוא קטע שמרכזו x_0. ליתר דיוק, בהינתן הטור $\sum_{n=0}^{\infty}a_n(x-x_0)^n$, רדיוס ההתכנסות R מוגדר כמו בנוסחת קושי הדמר בתור $R=(\limsup\sqrt[n]{|a_n|})^{-1}$. אם $R=\infty$ הטור מתכנס בכל הישר ואם $R<\infty$ הטור מתכנס נקודתית בקטע (x_0-R,x_0+R) ואינו מתכנס באף נקודה x שמקיימת $|x-x_0|>R$. הטור מתכנס בהחלט בקטע (x_0-R,x_0+R) ומתכנס במ"ש בכל תת־קטע סגור שלו. ההוכחות מבוססות על הטענות הקודמות, ומושארות כתרגיל.

תרגילים

1. חשבו את רדיוס ההתכנסות ואת תחום ההתכנסות של טורי החזקות הבאים:
   (א) $\sum_{n=1}^{\infty}\frac{(n!)^2}{(2n)!}x^n$.
   (ב) $\sum_{n=1}^{\infty}\sqrt[n^2]{7^{3n}+5^{3n}}x^n$.
   (ג) $\sum_{n=1}^{\infty}\frac{n!}{n^n}x^n$.
   (ד) $\sum_{n=1}^{\infty}\frac{(-1)^n}{n^\alpha}x^n$ ($\alpha>0$ קבוע).
   (ה) $\sum_{n=1}^{\infty}\left(\sum_{k=1}^{n}\frac{1}{k}\right)x^n$.
   (ו) $\sum_{n=0}^{\infty}\frac{1}{n^2}x^{2^n}$.
   (ז) $\sum_{n=1}^{\infty}\frac{1}{n}x^{n^2}$.
2. השלימו את הוכחת משפט קושי־הדמר למקרה של רדיוס התכנסות 0 או ∞.
3. בעזרת הטור של $\ln(1+x)$ מצאו טור חזקות לפונקציה $\ln(r+x)$ עבור r קבוע, ומצאו את רדיוס ההתכנסות שלו.
4. הראו שטורי החזקות המייצגים את $e^x,\sin x,\cos x$ הם בעלי רדיוס התכנסות ∞, כלומר שהטורים מתכנסים נקודתית בכל הישר, אך אינם מתכנסים במ"ש בכל הישר.
5. אם לטור החזקות $\sum a_kx^k$ יש רדיוס התכנסות R, מה רדיוס ההתכנסות של הטור $\sum\frac{a_k}{2^k}x^k$?
6. נתון כי הטור $\sum_{n=0}^{\infty}a_nx^n$ הוא בעל רדיוס התכנסות R_1 ושהטור $\sum_{n=0}^{\infty}b_nx^n$ הוא בעל רדיוס התכנסות R_2. מה ניתן לומר על רדיוס ההתכנסות של $\sum_{n=0}^{\infty}(a_n+b_n)x^n$?
7. תנו דוגמה לטור חזקות שתחום התכנסותו הוא $[-2,2)$.

8. יהיו P_n פולינומים ונניח ש־ $P_n \to f$ במ"ש בכל הישר. הראו ש־f פולינום (רמז: הראו שמספיק להוכיח זאת תחת ההנחה ש־ $P_n(0) = 0$. הראו שתחת הנחה זו הסדרה (P_n) קבועה החל ממקום מסוים ו־$f = P_n$ לכל n גדול מספיק).
9. וודאו את הטענות אודות טורי חזקות מהסוג $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ המופיעות בפסקה האחרונה של הסעיף.

11.6 רציפות, גזירה ואינטגרציה של טורי חזקות

בעקבות העובדה שטור חזקות מתכנס במ"ש בכל קטע סגור בתחום ההתכנסות שלו, אפשר להפעיל את הכלים שפיתחנו על גבולות במ"ש בפרק 10 ולהוכיח רציפות וגזירות של טורי חזקות. כפי שנראה, כללי הגזירה והאינטגרציה של טורי חזקות דומים מאד לכללים עבור פולינומים.

משפט 11.6.1 (רציפות של טור חזקות) יהי $\sum_{n=0}^{\infty} a_n x^n$ טור חזקות בעל רדיוס התכנסות $R > 0$. אז הפונקציה $f : (-R, R) \to \mathbb{R}$ המוגדרת על־ידי

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

היא פונקציה רציפה ב־$(-R, R)$.

הוכחה לפי מסקנה 10.3.2 גבול במ"ש של טור שאיבריו רציפים הוא פונקציה רציפה. המחוברים בטור חזקות רציפים בכל הישר ולכן אילו הטור היה מתכנס במ"ש ב־$(-R, R)$ הטענה הייתה נובעת. אולם מאחר שהטור שלנו אינו בהכרח מתכנס במ"ש ב־$(-R, R)$ לא נוכל להפעיל את המשפט ישירות.

מאידך, אנו יודעים שטור החזקות מתכנס במ"ש בכל קטע $[-r, r] \subseteq (-R, R)$, ולכן הפונקציה f רציפה בכל קטע $[-r, r]$. כיוון שרציפות בנקודה היא תכונה אינפיניטסימלית, הרי שכדי להוכיח ש־f רציפה בנקודה $x_0 \in (-R, R)$, כל שעלינו לעשות הוא להראות ש־x_0 שייך לפנים של קטע מהסוג $[-r, r]$ אשר מוכל ב־$(-R, R)$. אבל זה ברור: יש רק לבחור r המקיים $|x_0| < r < R$. ■

נעבור לדון באינטגרל ובנגזרות של טור חזקות. נזדקק ללמה הבאה המקלה על החישוב של רדיוסי התכנסות:

למה 11.6.2 תהי (a_n) סדרה אי־שלילית. אז לכל $k \in \mathbb{Z}$ מתקיים

$$\limsup \sqrt[n+k]{a_n} = \limsup \sqrt[n]{a_n}$$

הוכחה תחום ההתכנסות של טורי החזקות $\sum_{n=k}^{\infty} a_n x^n$ ו־ $\sum_{n=k}^{\infty} a_n x^{n+k}$ זהה, כי בכל נקודה x הם מתקבלים זה מזה על ידי כפל במספר x^k. לכן רדיוסי ההתכנסות שלהם שווים, ולפי נוסחת קושי־הדמר

$$(\limsup \sqrt[n+k]{a_n})^{-1} = (\limsup \sqrt[n]{a_n})^{-1}$$

כפי שרצינו להראות. ■

משפט 11.6.3 (אינטגרציה איבר־איבר של טורי חזקות) יהי $f(x) = \sum_{n=0}^{\infty} a_n x^n$ טור חזקות בעל רדיוס התכנסות $R > 0$. אז לכל $x \in (-R, R)$ מתקיים

$$\int_0^x f(t)dt = \sum_{n=0}^{\infty} \int_0^x a_n t^n dt = \sum_{n=1}^{\infty} \frac{a_{n-1}}{n} x^n$$

(התכנסות הטור באגף ימין היא חלק מהטענה). בפרט, כל פונקציה קדומה של f ב־$(-R, R)$ גם־כן ניתנת להצגה כטור חזקות בעל רדיוס התכנסות R.

הוכחה יהי $x \in (-R, R)$. נניח למשל ש־ $x > 0$. אז בקטע $[0, x]$ טור החזקות המגדיר את f מתכנס במ"ש (טענה 11.5.7) ומאחר שכל איברי הטור הם פונקציות רציפות ב־$[0, x]$ מותר, לפי משפט 10.4.3, לבצע אינטגרציה איבר־איבר ומקבלים

$$\int_0^x f(t)dt = \sum_{n=0}^{\infty} \int_0^x a_n t^n dt$$

(התכנסות הטור באגף ימין היא אחת המסקנות של המשפט 10.4.3). חישוב פשוט בעזרת המשפט היסודי מראה ש־ $\int_0^x a_n t^n dt = \frac{a_n}{n+1} t^{n+1}$ (הוכיחו!) ולכן

$$\sum_{n=0}^{\infty} \int_0^x a_n t^n dt = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} = \sum_{n=1}^{\infty} \frac{a_{n-1}}{n} x^n$$

קיבלנו שלכל $x \in (-R, R)$ מתקיים $\int_0^x f = \sum_{n=1}^{\infty} \frac{a_{n-1}}{n} x^n$, דהיינו $\int_0^x f$ ניתנת לייצוג כטור חזקות ב־$(-R, R)$. לפי המשפט היסודי זו פונקציה קדומה של f, וכל פונקציה קדומה אחרת נבדלת ממנה בקבוע בקטע $(-R, R)$. לבסוף, נסמן ב־R' את רדיוס ההתכנסות של $\sum_{n=1}^{\infty} \frac{a_{n-1}}{n} x^n$. לפי נוסחת קושי־הדמר ולמה 11.6.2,

$$\begin{aligned} R' &= (\limsup \sqrt[n]{|a_{n-1}|/n})^{-1} \\ &= \lim_{n\to\infty} \sqrt[n]{n} \cdot (\limsup \sqrt[n]{|a_{n-1}|})^{-1} \\ &= R \end{aligned}$$

כאשר השוויון האחרון נובע מהלמה ומכך ש־ $\sqrt[n]{n} \to 1$. ■

הוכחה דומה מאד מאפשרת לגזור טור חזקות:

משפט 11.6.4 (גזירה איבר־איבר של טור חזקות) יהי $f(x) = \sum_{n=0}^{\infty} a_n x^n$ טור חזקות בעל רדיוס התכנסות $R > 0$. אז f גזירה בכל נקודה $x \in (-R, R)$ ומתקיים

$$f'(x) = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n$$

בפרט פונקציית הנגזרת של f בקטע $(-R, R)$ היא גם־כן טור חזקות בעל רדיוס התכנסות R.

הוכחה יהי $b_n = (n+1)a_{n+1}$. אז $\frac{d}{dx} a_{n+1} x^{n+1} = b_n x^n$, וגם $\frac{d}{dx} a_0 x^0 = 0$. לכן הטור $\sum_{n=0}^{\infty} b_n x^n$ הוא טור הנגזרות של $\sum_{n=0}^{\infty} a_n x^n$. רדיוס ההתכנסות R' של $\sum_{n=0}^{\infty} b_n x^n$ הוא

$$\begin{aligned} R' &= (\limsup \sqrt[n]{(n+1)\,|a_{n+1}|})^{-1} \\ &= (\lim_{n\to\infty} \sqrt[n]{n+1})^{-1} \cdot (\limsup \sqrt[n]{|a_{n+1}|})^{-1} \\ &= R \end{aligned}$$

יהי $x_0 \in (-R, R)$. די שנראה ש־f גזירה ב־x_0 ומקיימת $f'(x_0) = \sum_{n=0}^{\infty} b_n x_0^n$. נבחר r כך ש־ $|x_0| < r < R$. אז $x_0 \in (-r, r)$ והטורים $\sum a_n x^n$ ו־ $\sum b_n x^n$ מתכנסים במ"ש ב־$[-r, r]$. לכן לפי משפט 10.5.1 הפונקציה f גזירה ב־x_0 ומקיימת

$$f'(x_0) = \sum_{k=0}^{\infty} b_n x_0^n$$

כנדרש. ■

מסקנה 11.6.5 יהי $f(x) = \sum_{n=0}^{\infty} a_n x^n$ טור חזקות בעל רדיוס התכנסות $R > 0$. אז $f(x)$ היא פונקציה גזירה אינסוף פעמים בקטע הפתוח $(-R, R)$. יתר על כן, הנגזרות מכל הסדרים ניתנות לייצוג כטור חזקות בקטע $(-R, R)$, והנגזרת ה־k־ית נתונה במפורש על־ידי

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)\cdots(n-k+1) a_n x^{n-k}$$

הוכחה זו הפעלה חוזרת של משפט 11.6.4. ההוכחה היא באינדוקציה ומושארת כתרגיל. ■

דוגמאות

1. נגזור איבר־איבר את הפונקציה $f(x) = \sum_{n=0}^{\infty} \frac{1}{n!}x^n$, ונקבל

$$\begin{aligned} f'(x) &= \sum_{n=1}^{\infty} \frac{1}{n!}(n \cdot x^{n-1}) \\ &= \sum_{n=1}^{\infty} \frac{1}{(n-1)!}x^{n-1} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}x^n \\ &= f(x) \end{aligned}$$

אין זה מפתיע כי ראינו כבר ש־ $f(x) = e^x$.

2. נחפש ביטוי עבור $\int_0^x e^{-t^2}dt$. כפי שציינו בעבר, לאינטגרל הזה אין ביטוי אלמנטרי, אולם יש לו הצגה כטור חזקות. מאחר ש־ $e^t = \sum_{n=0}^{\infty} \frac{1}{n!}t^n$ לכל t אותו השוויון נכון כשמציבים t^2 במקום t:

$$e^{t^2} = \sum_{n=0}^{\infty} \frac{1}{n!}t^{2n}$$

רדיוס ההתכנסות של טור זה הוא ∞ (בדקו!) ולכן לכל x מתקיים

$$\int_0^x e^{-t^2}dt = \int_0^x \sum_{n=0}^{\infty} \frac{1}{n!}t^{2n}dt = \sum_{n=0}^{\infty} \frac{1}{(2n+1)n!}x^{2n+1}$$

3. נחפש נוסחה לטור $f(x) = \sum_{n=1}^{\infty} nx^n$. בדיקה מעלה שרדיוס ההתכנסות הוא 1. נשים לב ש־ $nx^n = x \cdot nx^{n-1}$, ולכן

$$f(x) = \sum_{n=1}^{\infty} x \cdot nx^{n-1} = x \cdot \sum_{n=1}^{\infty} nx^{n-1}$$

אבל הטור באגף ימין הוא גם בעל רדיוס התכנסות 1, ומתקיים

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{d}{dx}(\sum_{n=1}^{\infty} x^n) = \frac{d}{dx}(\frac{1}{1-x}) = \frac{1}{(1-x)^2}$$

(השוויון הראשון לפי משפט 11.6.4), ולכן

$$f(x) = \frac{x}{(1-x)^2}$$

מסקנה חשובה מהמשפט על גזירה איבר־איבר היא שטור החזקות המייצג פונקציה נקבע ביחידות על ידי הפונקציה:

משפט 11.6.6 תהי f פונקציה ונניח שיש ל־f ייצוג כטור חזקות $f(x)=\sum_{n=0}^{\infty} a_n x^n$ בסביבה כלשהי של 0. אז מקדמי הטור נתונים על ידי הנוסחה $a_n = \frac{1}{n!} f^{(n)}(0)$. בפרט, אם יש ל־$f$ ייצוג כטור חזקות בסביבה של 0 אז ההצגה יחידה.

הוכחה אם ל־f יש ייצוג כטור חזקות בסביבה של 0 אז היא גזירה אינסוף פעמים ב־0. נציב $x=0$ בנוסחה עבור $f^{(k)}$ במסקנה 11.6.5. כל המחוברים שבהם מופיע הגורם 0^{n-k} עבור $n-k>1$ מתאפסים, ולכן התרומה היחידה לטור באה מהמחובר הראשון, כלומר

$$f^{(k)}(0) = n(n-1)\dots(n-n+1)a_n$$

והנוסחה ל־a_n נובעת על ידי העברת אגפים. כיוון שמקדמי הטור שמייצג את f נקבעים על ידי הנגזרות של f, לכל טור אחר המייצג את f יש את אותם המקדמים, ומכאן הטענה לגבי היחידות. ■

המקדמים של טור חזקות נתונים במשפט האחרון על ידי אותה הנוסחה שהגדירה את המקדמים של פולינומי מקלורן של הפונקציה. ואמנם, הכלים שפיתחנו על טורי חזקות מראים שפולינומי מקלורן מתכנסים לפונקציה בסביבה של 0 בדיוק כאשר הפונקציה ניתנת להצגה כטור חזקות:

מסקנה 11.6.7 תהי f פונקציה גזירה אינסוף פעמים באפס. אז f ניתנת לייצוג כטור חזקות בקטע $(-r,r)$ אמ"מ f היא הגבול הנקודתי של פולינומי מקלורן שלה בקטע $(-r,r)$.

הוכחה נסמן $a_n = \frac{f^{(n)}(0)}{n!}$ ויהיו $P_n(x)=\sum_{k=0}^{n} a_k x^k$ פולינומי מקלורן של f. אם f היא גבול נקודתי של P_n בקטע כלשהו אז היא מקיימת $f(x)=\sum_{n=0}^{\infty} a_n x^n$ לכל x באותו קטע, וממילא מיוצגת כטור חזקות.

להפך, אם יש לה ייצוג כטור חזקות בסביבה של 0 אז הטור נתון על־ידי $\sum_{n=0}^{\infty} a_n x^n$ לאותם מקדמים a_n מהפסקה הקודמת, והסכומים החלקיים של טור זה הם P_n. לכן אם f ניתנת לייצוג כטור חזקות בקטע כלשהו אז f היא גבול נקודתי של P_n באותו קטע. ■

לפי מסקנה 11.6.5, גזירות אינסוף פעמים בקטע היא תנאי הכרחי לייצוג של פונקציה כטור חזקות שם. אבל תנאי זה אינו מספיק. כבר ראינו דוגמה לפונקציה אשר גזירה אינסוף פעמים בכל הישר אך שפולינומי מקלורן שלה אינם מתכנסים אליה באף קטע סביב 0 (דוגמה (2) בעמוד 515). לכן לפי המסקנה האחרונה הפונקציה אינה ניתנת לייצוג כטור חזקות באף קטע שמכיל את 0.

תרגילים

1. תנו הוכחה ישירה ללמה 11.6.2.

2. תהי $f(x) = \sum_{n=0}^{\infty} a_n x^n$. הראו ש־$f$ זוגית אמ"מ $a_n = 0$ לכל n אי־זוגי, וש־f אי־זוגית אמ"מ $a_n = 0$ לכל n זוגי.

3. נאמר ש־F פונקציה קדומה מסדר n של f אם $F^{(n)} = f$. נניח ש־f ניתנת לייצוג כטור חזקות בקטע $(-R, R)$. האם יש לה פונקציה קדומה מכל סדר?

4. טורי חזקות שימשו בעבר בפתרון **משוואות דיפרנציאליות** (differential equations). אלו הן משוואות בהן הנעלם הוא פונקציה. למשל, נניח שאנו מחפשים פונקציה f המוגדרת וגזירה בכל הישר ומקיימת $f' = f$. בהנחה שאפשר להציג את f כטור חזקות $f(x) = \sum_{n=0}^{\infty} a_n x^n$ בכל הישר, אנו מקבלים את השוויון

$$\sum_{n=0}^{\infty} a_n x^n = f(x) = f'(x) = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n$$

(אגף ימין התקבל על ידי גזירה איבר־איבר). מכיוון שהטורים בשני האגפים שווים הרי שהמקדם של x^n באגף ימין שווה למקדם שלו באגף שמאל (זה נובע ממשפט 11.6.6) ולכן קיבלנו $a_n = (n+1)a_{n+1}$. מכאן קל להסיק באינדוקציה ש־ $a_n = \frac{1}{n!} a_0$, ולכן

$$f(x) = \sum_{n=0}^{\infty} \frac{a_0}{n!} x^n = a_0 \sum_{n=0}^{\infty} \frac{1}{n!} x^n = a_0 e^x$$

אם נוסיף את התנאי ש־ $f(0) = 1$ נקבל ש־ $f(x) = e^x$. זו הוכחה חדשה לכך שהפונקציה המעריכית היא הפונקציה היחידה המקיימת $f' = f$ (ראו גם דוגמה (1) בעמוד 334).

(א) נמקו מדוע במקרה למעלה מוצדק להניח שיש לפונקציה פיתוח כטור חזקות.

(ב) פתרו את המשוואות הדיפרנציאליות הבאות ונסו להציג את הפתרונות בעזרת פונקציות אלמנטריות:

i. $f = f''$.

ii. $f'' = -f$.

iii. $f'(x) = x f(x)$.

5. יהי $f(x) = \sum_{n=0}^{\infty} a_n x^n$ טור חזקות בעל רדיוס התכנסות $R > 0$. הראו שלכל $x_0 \in (-R, R)$ יש ־f טור חזקות מהצורה $\sum_{n=0}^{\infty} b_n (x - x_0)^n$ בעל רדיוס התכנסות $R - |x_0|$.

6. יהי $f(x) = \sum_{n=0}^{\infty} a_n x^n$ טור חזקות בעל רדיוס התכנסות $R > 0$.

(א) נניח שיש סדרה $x_n \to 0$ כך ש־ $f(x_n) = 0$ לכל n. הראו ש־ $f \equiv 0$ (כלומר $a_n = 0$ לכל n).

(ב) הסיקו משאלה (5) לעיל שאם $x_0 \in (-R, R)$ ואם $f(x_0) = 0$ אז יש סביבה של x_0 שבה f אינה מתאפסת פרט מאשר בנקודה x_0 עצמה.

(ג) תהי g פונקציה חיובית בסביבה של 0, ונגדיר $f(x) = g(x)\sin(\frac{1}{x})$ אם $x \neq 0$ ו־ $f(0) = 0$. הראו ש־f אינה ניתנת להצגה כטור חזקות באף סביבה של 0.

הערה שאלה זו מצביעה על דמיון נוסף בין פולינומים לבין טורי חזקות. לפולינום יש מספר סופי של נקודות התאפסות. לטור חזקות יכול להיות אינסוף נקודות התאפסות (למשל $\sin$), אבל משאלה זו נובע שקבוצת נקודות ההתאפסות היא דיסקרטית, כלומר, שבכל קטע סגור $[a, b]$ יש מספר סופי של אפסים (לגבי קטעים פתוחים או חצי־פתוחים ייתכן שיש סדרת אפסים המתכנסת לנקודת קצה – מצאו דוגמא!).

11.7 טורי חזקות של הפונקציות האלמנטריות

את טורי החזקות המייצגים את $e^x, \ln(1+x), \sin x, \cos x$ מצאנו על ידי חישוב פולינומי טיילור שלהן והערכת השארית. ישנן גם שיטות עקיפות לחשב טורי חזקות של פונקציות. בסעיף זה נכיר כמה שיטות כאלה ונמצא טורי חזקות עבור הפונקציות האלמנטריות הבסיסיות.

הלוגריתם

השיטה הראשונה שנשתמש בה מנצלת את האפשרות לבצע אינטגרציה איבר־איבר. שיטה זו שימושית כשרוצים למצוא טור חזקות עבור פונקציה שאת הנגזרת שלה אנו כבר יודעים לייצג כטור חזקות.

לדוגמה, נשתמש בשיטה זו כדי לחשב בדרך חדשה את טור החזקות המייצג את $f(x) = \ln(1+x)$. גזירה מראה ש־ $f'(x) = \frac{1}{1+x}$. אנו יודעים שלכל $x \in (-1, 1)$ מתקיים $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$, ומאחר ש־ $x \in (-1, 1)$ אמ"מ $-x \in (-1, 1)$ הרי שלכל $x \in (-1, 1)$ מתקיים

$$f'(x) = \frac{1}{1+x} = \frac{1}{1-(-x)} = \sum_{n=0}^{\infty}(-x)^n = \sum_{n=0}^{\infty}(-1)^n x^n$$

תחום ההתכנסות של הטור החדש הוא גם־כן $(-1,1)$. לכן לכל $x\in(-1,1)$ נוכל לבצע אינטגרציה איבר־איבר לשני האגפים מ־0 עד x, ומהמשפט היסודי מקבלים

$$\begin{aligned}\ln(1+x) &= \ln 1+\int_0^x \frac{1}{1+x}dx\\ &= \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}x^n\end{aligned}$$

זה כמובן אותו טור עבור $\ln(1+x)$ שמצאנו בעזרת פולינומי טיילור בסעיף 11.2.

נוסחה זו מאפשרת לחשב את $\ln(1+x)$ לכל $x\in(-1,1)$, דהיינו את $\ln x$ לכל $x\in(0,2)$. על מנת לחשב את $\ln y$ עבור $y>0$ כללי ניעזר בתכסיס הבא. נשים לב שניתן לכתוב כל מספר ממשי $y>0$ בתור $y=\frac{1+x}{1-x}$ כאשר $x\in(-1,1)$ הוא המספר $x=\frac{y-1}{y+1}$. לכן

$$\ln y=\ln\frac{1+x}{1-x}=\ln(1+x)-\ln(1-x)$$

אבל ראינו קודם ש־ $\ln(1+x)=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}x^n$, ואם נציב בטור זה $-x$ במקום x (זה מותר כי $x\in(-1,1)$ אמ"מ $-x\in(-1,1)$) יוצא ש־

$$\ln(1-x)=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}(-x)^n=-\sum_{n=1}^{\infty}\frac{1}{n}x^n$$

ולכן

$$\begin{aligned}\ln\frac{1+x}{1-x} &= \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}x^n-\sum_{n=1}^{\infty}\frac{1}{n}x^n\\ &= -\sum_{n=0}^{\infty}\frac{2}{2n+1}x^{2n+1}\end{aligned}$$

לכן

$$\begin{aligned}\ln y &= -\sum_{n=0}^{\infty}\frac{2}{2n+1}(\frac{y-1}{y+1})^{2n+1}\\ &= \sum_{n=0}^{\infty}\frac{2}{2n+1}(\frac{1-y}{1+y})^{2n+1}\end{aligned}$$

הפונקציה arctan

תהי $f(x) = \arctan x$. כמו במקרה של הלוגריתם, ההגדרה של arctan מעט סתומה, אך נגזרתה בעלת הצורה הפשוטה $f'(x) = \frac{1}{1+x^2}$. בשוויון

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$$

נציב $x = (-t^2)$. מאחר ש־ $-t^2 \in (-1,1)$ באשר $t \in (-1,1)$ אנו מקבלים שלכל $t \in (-1,1)$ מתקיים

$$\frac{1}{1+t^2} = \sum_{n=0}^{\infty} (-t^2)^n = \sum_{n=0}^{\infty} (-1)^n t^{2n}$$

עבור $x \in (-1,1)$ נבצע אינטגרציה מ־0 ועד x, ומהמשפט היסודי נקבל

$$\begin{aligned}
\arctan(x) &= \arctan 0 + \int_0^x \sum_{n=0}^{\infty} (-1)^n t^{2n} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \\
&= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots
\end{aligned}$$

לכל $x \in (-1,1)$.

נוסחת הבינום עבור למעריך שלם שלילי

קודם לכן השתמשנו באינטגרציה איבר־איבר כדי לקבל טור חזקות ל־f מטור חזקות של f'. נלך עתה בכיוון ההפוך ונמצא טור חזקות ל־f על ידי גזירה של טור חזקות של פונקציה קדומה של f.

תהי $f(x) = \frac{1}{1-x}$. לכל $x \in (-1,1)$ מתקיים $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$. על ידי גזירה איבר־איבר נקבל שעבור $x \in (-1,1)$ מתקיים

$$\frac{1}{(1-x)^2} = (\frac{1}{1-x})' = \sum_{n=0}^{\infty} (n+1) x^n$$

גזירה נוספת מראה שלכל $x \in (-1,1)$ מתקיים

$$\frac{2}{(1-x)^3} = \sum_{n=0}^{\infty}(n+1)(n+2)x^n$$

או באופן שקול,

$$\frac{1}{(1-x)^3} = \sum_{n=0}^{\infty}\binom{n+2}{2}x^n$$

אם נמשיך לגזור, נקבל לכל $p \in \mathbb{N}$ טור חזקות המייצג בקטע $(-1,1)$ את הפונקציה $(1-x)^{-p}$. באינדוקציה ניתן להוכיח שלכל $x \in (-1,1)$ ולכל $p \in \mathbb{N}$ מתקיים

$$\frac{1}{(1-x)^{p+1}} = \sum_{n=0}^{\infty}\binom{n+p}{p}x^n$$

נוסחת הבינום, מקרה כללי

נוסחת הבינום עבור $p \in \mathbb{N}$ היא $(1+x)^p = \sum_{n=0}^{p}\binom{p}{n}x^n$, כשהמקדמים הבינומיים מוגדרים על ידי $\binom{p}{n} = \frac{p\cdot(p-1)\cdots(p-n+1)}{n!}$. אם מציבים בנוסחה זו $n > p$ טבעי מתקיים $\binom{p}{n} = 0$, כי אז המכפלה במונה כוללת גורם אפס. לכן אפשר לרשום את נוסחת הבינום כטור חזקות

$$(1+x)^p = \sum_{n=0}^{\infty}\binom{p}{n}x^n$$

היות וכל המחוברים מהמקום ה־$p+1$ ואילך מתאפסים, יש למעשה מספר סופי של מחוברים בסכום.

על מנת להכליל נוסחה זו למקרה שהמעריך אינו שלם נזדקק להכללה של המקדמים הבינומיים.

הגדרה 11.7.1 עבור $\alpha \in \mathbb{R}$ המקדם הבינומי המוכלל α מעל n מוגדר לפי אותה נוסחה כמו מקדם בינומי רגיל:

$$\binom{\alpha}{n} = \frac{\alpha\cdot(\alpha-1)\cdots(\alpha-n+1)}{n!}$$

מטרתנו בסעיף זה היא להוכיח את נוסחת הבינום המוכללת: לכל α ממשי ולכל $x \in (-1,1)$ מתקיים

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$$

שימו לב שאם α אינו שלם אז $\binom{\alpha}{n} \neq 0$ לכל $n \in \mathbb{N}$, ואז זהו טור אינסופי. כדאי לבדוק שהנוסחה הזו מתלכדת עם הנוסחה מהסעיף הקודם במקרה ש־n שלם שלילי וכשמציבים $-x$ במקום x (שהרי שניהם מתיימרים לתאר את הטור $(1-x)^n$).

יהי $\alpha \in \mathbb{R}$. על מנת להוכיח את השוויון $(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$ נאפיין את הפונקציה $(1+x)^\alpha$ בעזרת שני תנאים, ונראה שהפונקציה הנתונה על ידי טור החזקות האמור המקיימת אותם.

נסמן $f(x) = (1+x)^\alpha$. ברור ש־ $f(0) = 1$, וגזירה מראה שלכל $x > 0$ מתקיים

$$f'(x) = \alpha(1+x)^{\alpha-1} = \frac{\alpha f(x)}{1+x}$$

שתי תכונות אלו מאפיינות את הפונקציה f:

משפט 11.7.2 אם g היא פונקציה המקיימת $g(0) = 1$ ו־ $(1+x)g'(x) = \alpha g(x)$ לכל $x \in (-1,1)$ אז $g(x) = (1+x)^\alpha$ לכל $x \in (-1,1)$.

הוכחה נסמן

$$\varphi(x) = \frac{g(x)}{(1+x)^\alpha}$$

φ מוגדרת וגזירה בקטע $(-1,1)$. הטענה שאנו מנסים להוכיח שקולה לכך ש־ $\varphi \equiv 1$ בקטע $(-1,1)$. מהנתון $g(0) = 1$ אנו רואים ש־ $\varphi(0) = 1/1 = 1$. לכן די להראות ש־φ קבועה, ולשם כך די להראות שהנגזרת מתאפסת. נחשב ונמצא שאכן

$$\varphi'(x) = \frac{g'(x)(1+x)^\alpha - g(x)\alpha(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} = 0$$

כי לאחר הצבת הנתון $(1+x)g'(x) = \alpha g(x)$ המונה מתאפס. ■

אם כן, נותר להראות שהפונקציה $g(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$ מקיימת את שני תנאים. התנאי הראשון מידי, שכן $g(0) = \binom{\alpha}{0} = 1$. כדי לאמת את התכונה השנייה עלינו תחילה להראות ש־g מוגדרת ב־$(-1,1)$. ואמנם, המנות העוקבות של מקדמי הטור מקיימות

$$\left|\frac{\binom{\alpha}{n+1}}{\binom{\alpha}{n}}\right| = \left|\frac{\alpha - n}{n+1}\right| \xrightarrow{n\to\infty} 1$$

וממסקנה 11.5.5 נובע שרדיוס ההתכנסות הוא 1, ומכאן ש־g גזירה ב־$(-1,1)$. כעת עלינו להראות שמתקיים

$$(1+x)g'(x)=\alpha g(x)$$

נגזור את g איבר־איבר ונקבל

$$g'(x)=\sum_{n=1}^{\infty} n\binom{\alpha}{n}x^{n-1}$$

נשים לב כי

$$n\binom{\alpha}{n}=n\frac{\alpha\cdot(\alpha-1)\cdots(\alpha-n+1)}{n!}=\alpha\frac{(\alpha-1)\cdots(\alpha-n+1)}{(n-1)!}=\alpha\binom{\alpha-1}{n-1}$$

ולכן

$$g'(x)=\sum_{n=1}^{\infty}\alpha\binom{\alpha-1}{n-1}x^{n-1}=\alpha\sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^{n}$$

בנוסף, מבדיקה ישירה עולה כי

$$\binom{\alpha-1}{n-1}+\binom{\alpha-1}{n-1}=\binom{\alpha}{n}$$

(זו הכללה של נוסחת הנסיגה של המקדמים הבינומיים הרגילים. הוכיחו אותה!) ולכן

$$\begin{aligned}(1+x)g'(x) &= \alpha(1+x)\sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^n\\ &= \alpha(\sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^n+\sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^{n+1})\\ &= \alpha(1+\sum_{n=1}^{\infty}\left[\binom{\alpha-1}{n}+\binom{\alpha-1}{n-1}\right]x^n)\\ &= \alpha(1+\sum_{n=1}^{\infty}\binom{\alpha}{n}x^n)\\ &= \alpha\sum_{n=0}^{\infty}\binom{\alpha}{n}x^n\\ &= \alpha g(x)\end{aligned}$$

כפי שרצינו.

תרגילים

1. הציגו את $\frac{1}{(1-x)^2}$ כטור חזקות.
2. היעזרו בטור של $\arctan$ כדי לקבל טור מספרים השווה ל־$\frac{\pi}{6}$.
3. היעזרו בנוסחת הבינום כדי למצוא טור חזקות עבור הפונקציה $\frac{1}{\sqrt{1-x^2}}$, ובצעו אינטגרציה איבר־איבר לקבלת טור עבור $\arcsin$.
4. איך אפשר להשתמש בטור של $\sqrt{1+x}$ כדי לחשב את $\sqrt{3}$?
5. בפרק 7 לא הגדרנו את הפונקציות הטריגונומטריות או את π, אלא הנחנו את קיומם (ראו עמוד 227). עם הכלים שעומדים לרשותנו נוכל כעת לתת להם הגדרה פורמלית לחלוטין, ואף להוכיח יחידות. נגדיר את הפונקציות $\sin x, \cos x$ בעזרת טורי חזקות:

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \qquad , \qquad \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$$

בסעיפים הבאים נראה שיש מספר חיובי π כך שהשלשה $\sin, \cos, \pi$ מקיימת את התכונות בעמוד 227. רוב העבודה כבר נעשתה, ואנו מביאים כאן הפניות מתאימות.

(א) הראו שהטורים המגדירים את $\sin, \cos$ מתכנסים בכל הישר (ראו תרגיל (4) בעמוד 520), ולכן הפונקציות רציפות וגזירות בכל הישר.

(ב) $\sin$ פונקציה אי־זוגית ו־$\cos$ פונקציה זוגית (ראו תרגיל (2) בעמוד 526).

(ג) $\sin^2 + \cos^2 \equiv 1$ (ראו תרגיל (12) מעמוד 340).

(ד) לכל $x, y \in \mathbb{R}$ מתקיים $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$ וגם $\sin(x+y) = \sin(x)\cos(y) + \sin(y)\cos(x)$ (ראו תרגיל (4) בעמוד 206).

(ה) הראו של־$\cos$ קיימת נקודת התאפסות חיובית מינימלית x_0 (רמז: שימו לב ש־ $\cos 0 = 1$. השתמשו בעובדה שהטור עבור $\cos 2$ הוא טור לייבניץ כדי להסיק ש־ $\cos 2 < 0$ (היעזרו במשפט 6.4.6), והסיקו ממשפט ערך הביניים ומתרגיל (5) בעמוד 274, שיש נקודת התאפסות מינימלית $x_0 > 0$ של $\cos$. נסמן $\pi = 2x_0$.

(ו) הוכיחו ש־ $\sin(\frac{\pi}{2}) = 1$, והראו שלכל $x \in (0, \frac{\pi}{2})$ מתקיימים האי־שוויונות $0 < \cos(s), \sin(x) < 1$ ו־ $\sin(x) \le x \le \frac{\sin(x)}{\cos(x)}$ (רמז: חקירת פונקציות).

אם תעברו על הטענות תגלו שהוכחנו כמעט את כל התכונות שהשלשה $\sin, \cos \pi$ אמורה לקיים. יתר התכונות נובעות מהן, כפי שראינו בתרגיל (7) בעמוד 230. אנו מסיקים שקיימת שלשה $\sin, \cos, \pi$ המקיימת את התכונות של הפונקציות הטריגונומטריות כפי שהוצגו בעמוד 227. מאידך ראינו שאותה רשימת תכונות גוררת ש־$\sin, \cos$ נתונות על ידי טורי החזקות למעלה, וש־$\frac{\pi}{2}$ היא נקודת ההתאפסות המינימלית. לכן השלשה $\sin, \cos, \pi$ היא יחידה.

ישנם גם דרכים אחרות להגדיר את הפונקציות הטריגונומטריות. ראו למשל [5],[6].

11.8 משפט אבל ושימושיו

טור חזקות בעל רדיוס התכנסות R אמנם מתכנס במ"ש בכל קטע $[-r,r]$ עבור $0 \leq r < R$, אך ראינו שאין הוא חייב להתכנס בנקודות $\pm R$, וייתכן שההתכנסות בקטע $(-R,R)$ אינה במ"ש. בסעיף הנוכחי נחקור את הקשר בין ההתכנסות של טור חזקות בנקודות $\pm R$ לבין התכנסות במ"ש שלו בקטעים הכוללים את $\pm R$.

הטענה הכללית הבאה מקשרת בין התכנסות במ"ש בקטע פתוח להתכנסות בקצוות:

למה 11.8.1 יהי $\sum_{k=0}^{\infty} f_k(x)$ טור פונקציות שבו כל f_k רציפה ב־$[a,b]$, ונניח שהטור מתכנס במידה שווה בקטע $[a,b)$. אזי טור המספרים $\sum_{k=0}^{\infty} f_k(b)$ מתכנס, וטור הפונקציות מתכנס במ"ש בכל $[a,b]$.

הוכחה נראה שטור המספרים $\sum_{k=0}^{\infty} f_k(b)$ מקיים את תנאי קושי. יהי $\varepsilon > 0$. לפי תנאי קושי להתכנסות במ"ש של טורי פונקציות, קיים N כך שלכל $m > n > N$ ולכל $x \in [a,b)$ מתקיים:

$$|\sum_{k=n}^{m} f_k(x)| < \varepsilon$$

זהו סכום סופי של פונקציות רציפות ב־$[a,b]$, ולכן גם הסכום הוא פונקציה רציפה ב־$[a,b]$. הואיל והאי־שוויון נכון לכל $a \leq x < b$, מרציפות הפונקציות f_k משמאל ב־b נובע האי־שוויון

$$|\sum_{k=n}^{m} f_k(b)| \leq \varepsilon$$

לכן הטור $\sum_{k=n}^{m} f_k(b)$ מקיים את תנאי קושי ולכן מתכנס.

כעת טור הפונקציות $\sum_{k=n}^{m} f_k(x)$ מתכנס במ"ש ב־$[a,b)$ ומתכנס בנקודה b, ולכן מתכנס במ"ש ב־$[a,b]$ (זה מקרה פרטי של העובדה שאם סדרת פונקציות מתכנסת במ"ש בשתי קבוצות אז היא מתכנסת במ"ש באיחוד שלהן. הוכיחו!). ■

מסקנה 11.8.2 אם טור חזקות בעל רדיוס התכנסות R אינו מתכנס ב־R אז הטור אינו מתכנס במ"ש ב־$[0,R)$, ואם הוא אינו מתכנס ב־$-R$ אז הטור אינו מתכנס במ"ש ב־$(-R,0]$.

הוכחה המסקנה נובעת מידית מהטענה הקודמת, כי בטור חזקות כל מחובר הוא פונקציה רציפה בכל הישר. ■

גם הכיוון ההפוך של המסקנה נכון, אך ההוכחה מעט יותר קשה.

משפט 11.8.3 (משפט אבל) יהי $\sum_{n=0}^{\infty} a_n x^n$ טור חזקות בעל רדיוס התכנסות $0 < R$, ונניח ש־R שייך לתחום ההתכנסות. אז הטור מתכנס במ"ש בקטע $[0,R]$.

הוכחה נשתמש בקריטריון קושי להתכנסות במ"ש של טורים: נקבע $0 < \varepsilon$ ונחפש $N \in \mathbb{N}$ שיבטיח שלכל $m > n > N$ ולכל $x \in [0, R]$ מתקיים

$$|\sum_{k=n}^{m} a_k x^k| < \varepsilon$$

ע"פ הנתון בדבר התכנסות טור החזקות ב־ $x = R$, קיים $N \in \mathbb{N}$ כך שלכל $m \geq n > N$:

$$|\sum_{k=n}^{m} a_k R^k| < \varepsilon$$

יהיו $n > N$ ו־ $k \in \mathbb{N}$, ויהי $0 \leq x \leq R$ כלשהו. אזי:

$$|\sum_{i=n}^{n+k} a_i x^i| = |\sum_{i=1}^{k} a_{n+i} x^{n+i}| = |\sum_{i=1}^{k} a_{n+i} R^{n+i} (\frac{x}{R})^{n+i}|$$

ניזכר בנוסחת הסכימה של אבל (למה 6.4.8), הקובעת שלכל שתי סדרות $s_1, \ldots, s_k$ ו־ $t_1, \ldots, t_k$ מתקיים

$$\sum_{i=1}^{k} s_i t_i = S_k t_k + \sum_{i=1}^{k-1} S_i (t_i - t_{i-1})$$

כאשר $S_i = s_1 + \ldots + s_i$. עבור $s_i = a_{n+i} R^{n+i}$ ו־ $t_i = (\frac{x}{R})^{n+i}$ מקבלים

$$\begin{aligned} |\sum_{i=1}^{k} a_{n+i} R^{n+i} (\frac{x}{R})^{n+i}| &= |S_k (\frac{x}{R})^{n+k} + \sum_{i=1}^{k-1} S_i ((\frac{x}{R})^{n+i} - (\frac{x}{R})^{n+i-1})| \\ &\leq |S_k| |\frac{x}{R}|^{n+k} + \sum_{i=1}^{k-1} |S_i| \cdot |(\frac{x}{R})^{n+i-1} - (\frac{x}{R})^{n+i}| \end{aligned}$$

מכיוון שלכל i מתקיים

$$|S_i| = |\sum_{j=1}^{i} s_j| = |\sum_{j=1}^{i} a_{n+i} R^{n+i}| < \varepsilon$$

לפי בחירת N, הרי

$$|\sum_{i=1}^{k} a_{n+i} R^{n+i} (\frac{x}{R})^{n+i}| \leq \varepsilon |\frac{x}{R}|^{n+k} + \varepsilon \sum_{i=1}^{k-1} |(\frac{x}{R})^{n+i-1} - (\frac{x}{R})^{n+i}|$$

אבל $0 \le \frac{x}{R} \le 1$ ולכן ההפרש בכל אחד מהמחוברים מימין חיובי, ולכן אפשר להתעלם מסימן הערך המוחלט ומקבלים סכום טלסקופי:

$$\begin{aligned} |\sum_{i=1}^{k} a_{n+i}R^{n+i}(\frac{x}{R})^{n+i}| &\le \varepsilon + \varepsilon\sum_{i=1}^{k-1}((\frac{x}{R})^{n+i-1} - (\frac{x}{R})^{n+i}) \\ &= \varepsilon + \varepsilon((\frac{x}{R})^{n+1} - (\frac{x}{R})^{n+k}) \\ &\le 2\varepsilon \end{aligned}$$

ותנאי קושי מתקיים. ■

מסקנה 11.8.4 אם $\sum_{n=0}^{\infty} a_n x^n$ טור חזקות בעל רדיוס התכנסות R והוא מתכנס ב־$-R$ אז הטור מתכנס במ"ש בקטע $[-R, 0]$.

מסקנה 11.8.5 טור חזקות מתכנס במ"ש בכל תת־קטע סגור של תחום ההתכנסות הנקודתית שלו.

הוכחת המסקנות מושארת כתרגיל. מסקנה נוספת הראויה לתשומת לב בפני עצמה היא:

מסקנה 11.8.6 (משפט הרציפות והאינטגרציה של אבל) תהי $f(x) = \sum_{n=0}^{\infty} a_n x^n$ ויהי R רדיוס ההתכנסות של הטור. אם הטור מתכנס ב־R אז f רציפה משמאל ב־R, ואם הטור מתכנס ב־$-R$ אז f רציפה מימין ב־$-R$. בכל מקרה, אם הטור מתכנס ב־x אז $\int_0^x f(t)dt = \sum_{n=0}^{\infty} \frac{a_n}{n+1}x^{n+1}$.

הוכחה אם הטור מתכנס ב־R אז הטור מתכנס במ"ש ב־$[0, R]$ ולכן רציף משמאל ב־R כגבול במ"ש של פונקציות רציפות. כך גם אם הטור מתכנס ב־$-R$. המסקנה בעניין האינטגרל כבר הוכחה עבור המקרה $|x| < R$, ובמקרה $x = \pm R$ היא נובעת מאותו שיקול כמו בהוכחה ההיא, תוך שימוש בהתכנסות במ"ש של הטור בקטע שקצותיו 0 ו־x. ■

ראינו כבר שאם $\sum_{n=0}^{\infty} a_n x^n$ טור חזקות בעל רדיוס התכנסות R ואם R' הוא רדיוס ההתכנסות של טור הנגזרות אז $R = R'$. המסקנה האחרונה נותנת מידע נוסף: תחום ההתכנסות של הטור המקורי מכיל את תחום ההתכנסות של טור הנגזרות, דהיינו גזירה יכולה להשמיט (אך לא להוסיף כאלה) נקודות התכנסות. כיוון שרדיוסי ההתכנסות שווים נקודות אלה יכולות להיות רק נקודות קצה של $(-R, R)$. ואמנם, ייתכן שנקודות הקצה הושמטו: נתבונן למשל בטור $\sum_{n=1}^{\infty} \frac{1}{n}x^n$. תחום התכנסותו הוא $[-1, 1)$. נגזור ונקבל את הטור $\sum_{n=0}^{\infty} x^n$ שמתכנס בתחום $(-1, 1)$. נקודת הקצה 1 הושמטה מתחום ההתכנסות.

מסקנה 11.8.6 מבטיחה במקרים מסוימים רציפות חד־צדדית של טור חזקות ברדיוס ההתכנסות, ולתוצאה זו יש יישומים חשובים. לדוגמה, נראה ש־

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

נסמן ב־$f:(-1,1]\to\mathbb{R}$ את הפונקציה שערכה ב־x נתון ע"י הטור $\sum_{n=1}^\infty \frac{(-1)^n}{n}x^n$; ראינו שרהוא אמנם מוגדר (כלומר, שהטור מתכנס) בכל נקודה בתחום הנתון. לפי מסקנה 11.8.6, f רציפה בתחום כולו ובפרט f רציפה משמאל ב־1. מאידך, אנו יודעים שבקטע $(-1,1)$ מתקיים $f(x)=\ln(1+x)$, וגם הפונקציה $\ln(1+x)$ רציפה משמאל ב־1. לכן

$$\ln 2=\lim_{x\to 1-}\ln(1+x)=\lim_{x\to 1-}f(x)=f(1)=\sum_{n=1}^\infty\frac{(-1)^{n+1}}{n}$$

כפי שרצינו להוכיח.

הטענה ההפוכה למשפט הרציפות של אבל אינה נכונה. למשל, רדיוס ההתכנסות של טור החזקות $1-x+x^2-x^3+\cdots$ הוא 1, והיות והוא טור הנדסי קל לרשום ביטוי מפורש לסכומו: $S(x)=\frac{1}{1+x}$. זו פונקציה רציפה ב־$x=1$ (משמע, קיים הגבול באגף שמאל במשוואה האחרונה), ובכל זאת הטור אינו מתכנס ב־ $x=1$.

תרגילים

1. הוכיחו את מסקנה 11.8.4 (אין צורך לחזור על הוכחת משפט אבל, אלא אפשר להתבסס עליו).
2. הראו ש־
$$\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots$$
(רמז: היעזרו בטור של arctan).
3. השתמשו בנוסחה ל־arcsin שמצאתם בתרגיל (3) בעמוד 533 כדי לקבל ביטוי נוסף ל־π.

11.9 פונקציות יוצרות וסדרות רקורסיה

ראינו בסעיפים הקודמים שעבור טורי חזקות עם רדיוס התכנסות חיובי יש התאמה מלאה בין הפונקציה שהטור מתאר וסדרת המקדמים שלו. עד כה השתמשנו בהתאמה הזו בעיקר ככלי לחישוב או לתיאור של פונקציות נתונות: בדרך כלל יצאנו מפונקציה נתונה וחיפשנו את הטור המתאר אותה. כעת נתעניין בתהליך ההפוך: בהינתן סדרת מספרים נתבונן בטור החזקות שמקדמיו נתונים על ידי הסדרה, ונשאל איזו מין פונקציה זו ומה אפשר ללמוד מכך על הסדרה המקורית.

הגדרה 11.9.1 תהי $(a_n)_{n=0}^\infty$ סדרה של מספרים. **הפונקציה היוצרת** (gneerating function) של (a_n) היא $g(x)=\sum_{n=0}^\infty a_nx^n$.

אם הסדרה (a_n) גדלה מהר מדי הפונקציה היוצרת שלה תתכנס רק ב־0. לא נתעניין במקרה זה, אלא רק במקרה שרדיוס ההתכנסות חיובי. אז הפונקציה היוצרת וסדרת המקדמים שלה מכילים בדיוק את אותו מידע, כי כל אחת קובעת את השנייה. אפשר לומר שפונקציה יוצרת אינה אלא קידוד שונה של הסדרה. אלא שאת הפונקציה אפשר גם לנתח בעזרת שיטות אנליטיות ואלגבריות, והמידע המופק בצורה זו יכול לשפוך אור על התנהגות סדרת המקדמים.

נציג כאן רק שימוש אחד של פונקציות יוצרות, ונציג אותו בעזרת דוגמה. תהי $(a_n)_{n=0}^{\infty}$ הסדרה המוגדרת ברקורסיה על ידי $a_0 = a_1 = 1$ ו־ $a_n = a_{n-1} + a_{n-2}$ ל־ $n \geq 2$. זו סדרת פיבונאצ'י שאיבריה הראשונים הם $1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$. אנו מעוניינים למצוא נוסחה עבור האיבר ה־n בסדרה. אם תנסו למצוא נוסחה כזו בשיטות אלמנטריות תגלו שזו אינה משימה טריביאלית.

קל להראות באינדוקציה ש־ $0 \leq a_n \leq 2^n$, ולכן הטור $\sum_{n=0}^{\infty} a_n x^n$ מתכנס לפחות בקטע $(-\frac{1}{2}, \frac{1}{2})$. נרשום את הפונקציה היוצרת f של (a_n) באופן הבא:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = 1 + x + \sum_{n=2}^{\infty} a_n x^n$$

נציב את נוסחת הרקורסיה של a_n בטור באגף ימין, ונקבל

$$\begin{aligned} f(x) &= 1 + x + \sum_{n=2}^{\infty} (a_{n-1} + a_{n-2}) x^n \\ &= 1 + (x + \sum_{n=1}^{\infty} a_n x^{n+1}) + \sum_{n=0}^{\infty} a_n x^{n+2} \\ &= 1 + \sum_{n=0}^{\infty} a_n x^n \cdot x + \sum_{n=0}^{\infty} a_n x^n \cdot x^2 \\ &= 1 + x f(x) + x^2 f(x) \end{aligned}$$

חילוץ f מהשוויון $f(x) = 1 + x f(x) + x^2 f(x)$ שהתקבל נותן נוסחה אלגברית עבור f:

$$f(x) = \frac{1}{1 - x - x^2}$$

נרצה כעת לכתוב את הביטוי שקיבלנו בצורה פשוטה יותר. קל לוודא שלפולינום $1 - x - x^2$ יש שני שורשים שונים, שנסמנם ב־a, b, ושאפשר לרשום אותו בצורה $1 - x - x^2 = (x - a)(x - b)$. לכן לפי משפט 9.7.11 יש קבועים A, B כך ש־

$$f(x) = \frac{A}{x - a} + \frac{B}{x - b}$$

אבל נשים לב שהפונקציות $\frac{A}{x-a}, \frac{B}{x-b}$ ניתנות לייצוג כטורי חזקות:

$$\frac{A}{x - a} = -\frac{A}{a} \cdot \frac{1}{1 - (x/a)} = -\frac{A}{a} \sum_{n=0}^{\infty} (\frac{x}{a})^n$$

וכמובן גם $\frac{B}{x-b} = -\frac{B}{b}\cdot\sum_{n=0}^{\infty}(\frac{x}{b})^n$. יתר על כן שני הטורים האלה מתכנסים בסביבה לא מנוונת של 0. לכן בסביבה לא מנוונת של 0 מתקיים

$$f(x) = -\sum_{n=0}^{\infty}(\frac{A}{a}(\frac{x}{a})^n + \frac{B}{b}(\frac{x}{b})^n) = -\sum_{n=0}^{\infty}(\frac{A}{a^{n+1}} + \frac{B}{b^{n+1}})x^n$$

השוויון הזה נכון לכל x בסביבה לא טריביאלית של 0 ולכן מקדמי הטור מימין שווים למקדמים של f, כלומר

$$a_n = -(\frac{A}{a^{n+1}} + \frac{B}{b^{n+1}})$$

מכאן אנו כבר מפיקים מידע לא טריביאלי: a_n גדל בקצב מעריכי. לאחר חישוב הקבועים a, b, A, B מקבלים נוסחה מדויקת עבור a_n:

$$a_n = \frac{1}{\sqrt{5}}((\frac{1+\sqrt{5}}{2})^{n+1} + (\frac{1-\sqrt{5}}{2})^{n+1})$$

את החישוב, שאינו אלא אלגברה פשוטה, אנו משאירים כתרגיל.

אפשר לפתח את שיטת הפונקציות היוצרות גם מנקודות מבט אחרות (כלומר לא דרך החשבון האינפיניטסימלי). תוכלו לקרוא על כך בספר [11].

תרגילים

1. מצאו את האיבר הכללי של הסדרה הנתונה ברקורסיה על ידי $a_0 = 1$, $a_1 = 1$, $a_{n+1} = 2a_n - a_{n-1}$.
2. נאמר שסדרה $(a_n)_{n=0}^{\infty}$ של מספרים מוגדרת בעזרת **תנאי רקורסיה לינארי הומוגני** (homogeneous linear recurrence relation), או בקיצור – בעזרת רקורסיה לינארית, אם יש k וקבועים $c_1, c_2, \ldots, c_k$ כך שלכל $n \geq k$ מתקיים

$$a_n = c_1a_{n-1} + c_2a_{n-2} + \ldots + c_ka_{n-k}$$

(א) תהי (a_n) סדרת רקורסיה לינארית. הראו שיש מספר ממשי M כך ש־ $|a_n| < M^n$, והסיקו שלפונקציה היוצרת של הסדרה יש רדיוס התכנסות חיובי. הראו עוד שזו פונקציה רציונלית והמכנה שלה הוא פולינום ממעלה k לכל היותר.

(ב) נסמן ב־ $f(x) = \frac{p(x)}{q(x)}$ את הפונקציה היוצרת מהסעיף הקודם. נניח שאפשר להציג את המכנה q כמכפלה של גורמים לינאריים,

$$q(x) = c\cdot(1-\alpha_1x)(1-\alpha_2x)\cdot\ldots\cdot(1-\alpha_kx)$$

הוכיחו שיש קבועים $c_1, \ldots, c_k$ כך ש־ $a_n = c_1\alpha_1^n + c_2\alpha_2^n + \ldots + c_k\alpha_k^n$.

ביבליוגרפיה

קיימים מאות ספרים בחשבון אינפיניטסימלי, וגם בשפה העברית ישנו מבחר מסוים. הספרים נבדלים בסגנון הכתיבה, בקהל היעד, בהיקף החומר ובאופן שבו הוא מוצג.

הספרים הבאים קרובים בגישתם לספר שלנו:

[1] **חשבון אינפיניטסימלי**, דוד מיזלר, הוצאת אקדמון, 1968.

הספר הקלאסי בעברית בנושא החשבון האינפיניטסימלי. הוא תמציתי למדי אך מרחיב בכמה נושאים שלא עסקנו בהם, כמו פונקציות רבות משתנים.

[2] **חדו"א 1** ו־2, בן־ציון קון וסמי זעפרני, הוצאת בק, 1992.

ספרים אלה אינם מכסים את כל החומר שכיסינו, אך הם מכסים את עיקרו. תמצאו כאן דוגמאות ותרגילים רבים.

[3] **Real analysis and foundations**, Steven G. Krantz, CRC Press, 1991.

ספר המציג את החומר באופן מעט יותר מופשט. הוא כולל בין השאר בניה מסודרת של מערכות המספרים.

אפשר להיעזר גם בספרים הבאים, אם כי נציין שבניגוד אלינו, הם מציגים את תורת הפונקציות לפני תורת הסדרות והטורים.

[4] **חשבון אינפיניטסימלי** (חלקים 1 ו־2), האוניברסיטה הפתוחה, 1983־1978.

סדרת ספרים שנועדו ללימוד עצמי וכוללים תרגילים רבים.

[5] **Calculus** (3rd. ed.), M. Spivak, Publish or Perish inc., 1994.

ספר מקיף מאד הכתוב בצורה ברורה ומעניינת. הוא כולל, בין השאר, הוכחות לאי־רציונליות של π, מבוא לתורת הפונקציות המרוכבות, והוכחה של המשפט היסודי של האלגברה. יש בו גם ביבליוגרפיה מקיפה ומצוינת.

[6] **Calculus** (vol. 1,2), Tom M. Apostol, Blaisdell Publishing Company, 1961.

שני כרכים בהירים ומעניינים המכסים נושאים בסיסיים ומתקדמים כאחד.

הספרים הבאים מיועדים בעיקר לתלמידי המדעים. הם אינם שמים דגש על הוכחות אלא מתרכזים בשימושים של התורה.

[7] **חשבון דיפרנציאלי ואינטגרלי** (חלקים א' ו־ב'), הווארד אנטון, האוניברסיטה הפתוחה, 1997.

[8] **Calculus**, Lynn Loomis, Addison-Wesley Publishing Company, 1974.

בספרים הבאים תמצאו חומר מתקדם יותר.

[9] **חשבון אינפיניטסימלי מתקדם**, י. לינדנשטראוס, הוצאת אקדמון, 1986.

ספר מקיף מאד המתחיל בנקודה בה הפסקנו.

[10] **Introduction to Analysis**, Maxwell Rosenlicht, Dover, 1986.

אף שרוזנליכט מתחיל את הספר מאותה נקודה שהתחלנו אנו, הוא נוקט בגישה כללית ומתקדמת יותר, ומטפל בנושאים מתקדמים.

במהלך הספר הזכרנו נושאים שונים שאינם שייכים לחשבון האינפיניטסימלי. תוכלו לקרוא עליהם בספרים הבאים:

[11] **מתמטיקה בדידה**, נ. ליניאל ומ. פרנס, נ. בן־צבי מפעלי דפוס בע"מ, 2001.

כאן תמצאו דיון קצר בתורת הקבוצות, בלוגיקה, ובפונקציות יוצרות.

[12] **תורת הקבוצות**, שמואל ברגר, האוניברסיטה הפתוחה, 1997.

[13] **Naive set theory**, Paul R. Halmos, Springer-Verlag, 1974.

על אף שמו, זהו ספר רציני בתורת הקבוצות, מהיסודות האקסיומטיים שלו ועד תורת העוצמות.

[14] **לוגיקה מתמטית**, עזריאל לוי, הוצאת אקדמון, 1997.

[15] **Logic for mathematicians**, Alan G. Hamilton, Cambridge University, Press, 1978.

[16] **אלגברה א'**, ש. עמיצור, הוצאת אקדמון, 1969.

כאן תמצאו מבוא לתורת השדות ונושאים אחרים באלגברה.

[17] **Galois Theory**, 3rd ed., Ian Stewart, Chapman&Hall, 2004.

ספר מתקדם יותר בתורת השדות. אפשר למצוא בו הוכחה לטרנסצנדנטיות של e ו־π.

[18] **A History of Mathematics**, C. B. Boyer, Princeton University Press, 1968.

האלף־בית היווני

A,α אלפא (alpha)
B,β בתא (beta)
Γ,γ גמא (gamma)
Δ,δ דלתא (delta)
E,ε אפסילון (epsilon)
Z,ζ זתא (zeta)
E,η אתא (eta)
Θ,θ ת'תא (theta)
I,ι איותא (iota)
K,κ כפה (kappa)
Λ,λ למדא (lambda)
M,μ מיו (mu)
N,ν נו (nu)
Ξ,ξ קסי (xi)
O,o אומיקרון (omicron)
Π,π פאי (pi)
R,ρ רו (rho)
Σ,σ סיגמא (sigma)
T,τ תאו (tau)
Υ,υ אופסילון (upsilon)
Φ,ϕ,φ פי (phi)
X,χ חי (chi)
Ψ,ψ פסי (psi)
Ω,ω אומגה (omega)

רשימת סמלים

מפתח

ספר זה נועד ללוות קורס אוניברסיטאי בחשבון אינפיניטסימלי. הוא כולל מבוא בנושאים מתמטיים כלליים ופרקים על הגדרת המספרים ותכונת החסם העליון, גבולות של סדרות, טורי מספרים, גבולות ורציפות של פונקציות, הנגזרת והאינטגרל, סדרות וטורי פונקציות, פולינומי טיילור וטורי חזקות.
הדיון הפורמלי מלווה בהסברים, באיורים, בדוגמאות ובתרגילים בכל הרמות.

ISBN 978-965-350-150-8
9 789653 501508
0 007500 315184
דאנאקוד 75-31518