חשבון אינפיניטסימלי 3

סיכומי הרצאותיו של אור שליט

נכתב ועוצב על ידי רן קירי בסמסטר חורף 2016-2017
תיקונים ועריכה ע"י אור שליט בסמסטר אביב 2017-2018
עדכונים נוספים ע"י אור שליט בסמסטר חורף 2019-2020

הקדמה

רשימות אלו הן חומר העזר הרשמי בקורס אינפי 3, 104295, כפי שניתן בסמסטר חורף תש״פ, בפקולטה למתמטיקה בטכניון.

בסמסטר חורף תשע״ו, כאשר לימדתי את הקורס ״אינפי 3 104282״, רן קירי, בהיותו סטודנט, הקליד את ההרצאות שלי. זהו המקור של רשימות אלו. בסמסטר אביב תשע״ז לימדתי את הקורס שוב, ובסמסטר חורף תש״פ לימדתי את הקורס במתכונת החדשה (ארבע שעות במקום שלוש) עם מספר הקורס החדש ״אינפי 3, 104295״. בכל פעם עדכנתי ותיקנתי, והוספתי חומר כפי שהתאפשר מהשעה הנוספת.

כעת הרשימות הללו יכולות להוות חומר עזר למרצים ולסטודנטים בקורס ״אינפי 3, 104295״. קצב הפרקים מתאים פחות או יותר לקצב ההתקדמות בכיתה. לא כל הדוגמאות שהבאתי בכיתה מופיעות כאן, ולהיפך.

ברצוני להודות לרן קירי, שללא היוזמה הראשונית שלו וההשקעה הרבה רשימות אלו לא היו באות לעולם.

בהכנת הרשימות נעזרתי בספרים ובחומרים הבאים:

1) "Multivariable Mathematics" מאת Shifrin.
2) ״Calculus on Manifolds״ מאת Spivak.
3) ״חשבון אינפיניטסימלי מתקדם א׳ ו-ב׳״ מאת לינדנשטראוס.
4) Lecture Notes באינפי 3 מאת רועי משולם, שניתן למצוא באתר הבית שלו.

כמו כן, לקראת הפעם הראשונה שלימדתי את הקורס התייעצתי עם אביב צנזור ורום פנחסי, מה שעזר לי לגבש את הגישה שלי לגבי הקורס.

אני רוצה להודות לסטודנטים שלמדו אצלי שקריאתם הראשונית הובילה לשיפורים ותיקונים. אני רוצה להודות לשירה איל, אדיר גיגי, בועז מואב, ורועי קליין. לצערי אינני זוכר את שמותיהם של כל מי שהעיר וכל מי שתיקנה, ואני רוצה להודות (ולהתנצל בפני) כל מי ששכחתי.

אני אודה לכל מי שמוצא טעויות לפנות אלי כדי שאוכל לתקן : oshalit@technion.ac.il

אור שליט, החמישי לפברואר, 2020.

פרק 1 – טופולוגיה של המרחב האוקלידי

הנושא של טופולוגיה של המרחב האוקלידי נידון במספר קורסים: אינפי 2, מבוא למרחבים מטריים וטופולוגיים, וכן באלגברה ב' / אלגברה לינארית ב' (באלגברה ב' דנים במרחב האוקלידי כמרחב מכפלה פנימית ומרחב נורמי). בקורס אינפי 3 צריך להבין היטב את התכונות הבסיסיות של המרחב האוקלידי כמרחב מטרי, ולכן הקורס מתחיל במבוא מהיר וממוקד בנושאים שחשובים במיוחד לחשבון דיפרנציאלי ואינטגרלי בכמה משתנים. החפיפה עם קורסים אחרים נלקחת בחשבון, וחלק מן ההוכחות לא ניתנות במלואן, בהתאם להחלטת המרצה.

<u>סקירה על הטופולוגיה של $\mathbb{R}^d$:</u>

1.1 הגדרה –

$$\mathbb{R}^d := \left\{ (x_1, \cdots, x_d) \middle| \begin{matrix} 1 \le i \le d \\ x_i \in \mathbb{R} \end{matrix} \right\}$$

1.2 הגדרה – הנורמה האוקלידית מוגדרת על ידי:

$$\|x\| = \|x\|_2 = \left(\sum_{i=1}^{d} |x_i|^2 \right)^{\frac{1}{2}}$$

1.3 הגדרה – המטריקה האוקלידית מוגדרת על ידי:

$$d(x,y) = d_2(x,y) = \|x-y\|$$

1.4 <u>סדרה</u> – סדרה של נקודות ב-$\mathbb{R}^d$ היא סדרה $(x^n)_{n=1}^{\infty} = (x^1, x^2, \dots)$, **כאשר** $x^n = (x_1^n, x_2^n, \cdots, x_d^n)$.

1.5 הגדרה (התכנסות) – סדרה $(x^n)_{n=1}^{\infty}$ **מתכנסת ל-$\boldsymbol{x} \in \mathbb{R}^d$** אם מתקיים:

$$\lim_{n\to\infty} \|x^n - x\| = 0$$

וכן מסמנים:

$$x^n \to x \quad , \quad \lim_{n\to\infty} x^n = x$$

<u>הערה:</u>

$x^n \to x$ אם ורק אם $\lim_{n\to\infty} x_i^n = x_i$ לכל $1 \le i \le d$.

<u>הערה:</u> כאשר $d = 1$, הנורמה האוקלידית היא פשוט הערך המוחלט, ואנו נקבל את המובן המוכר של התכנסות של נקודות ב- $\mathbb{R} = \mathbb{R}^1$. כאשר $d > 1$, ניתן לטפל בסדרות באופן דומה לסדרות של מספרים ממשיים, כאשר הנורמה מחליפה את הערך המוחלט. למשל, קל להראות (עם הוכחה דומה לזו שניתנה באינפי 1 לגבי גבולות של מספרים ממשיים) שהגבול, אם הוא קיים, הינו יחיד.

1.6 הגדרה (סדרת קושי) – סדרה $(x^n)_{n=1}^{\infty}$ נקראת **סדרת קושי** אם מתקיים:

$$\|x^n - x^m\| \xrightarrow{n,m\to\infty} 0$$

כלומר שלכל $\varepsilon > 0$ קיים $N \in \mathbb{N}$ כך שלכל $m, n > N$ מתקיים $\|x^n - x^m\| < \varepsilon$

1.6.1 משפט ($\mathbb{R}^d$ הינו מרחב שלם) – כל סדרת קושי ב-$\mathbb{R}^d$ מתכנסת.

<u>הוכחה:</u>

נניח כי $(x^n)_{n=1}^{\infty}$ הינה סדרת קושי, אזי לכל $1 \le i \le d$, מתקיים:

$$|x_i^n - x_i^m| = \sqrt{\left|x_i^n - x_i^m\right|^2} \le \sqrt{\sum_{j=1}^{d} \left|x_j^n - x_j^m\right|^2} = \|x^n - x^m\| \xrightarrow{m,n\to\infty} 0$$

ולכן $(x_i^n)_{n=1}^{\infty}$ הינה סדרת קושי ב-$\mathbb{R}$ ולכן מתכנסת ב-$\mathbb{R}$. נגדיר $x_i = \lim_{n\to\infty} x_i^n$ ונסמן:

$$x = (x_1, x_2, \cdots, x_d)$$

ונקבל כנדרש כי מתקיים: $\|x^n - x\| = \sqrt{\sum_{j=1}^{d} \left|x_j^n - x_j\right|^2} \leq \sqrt{d} \max_{1\leq j\leq d} \left|x_j^n - x_j\right|$ והביטוי הימני שואף לאפס, כלומר

$$\lim_{n\to\infty} x^n = x \qquad \text{מש"ל.}$$

1.7 הגדרה – כדור פתוח מוגדר על ידי:

$$B_r(x) = \{y \in \mathbb{R}^d \mid \|x - y\| < r\} \qquad \begin{matrix} r > 0 \\ x \in \mathbb{R}^d \end{matrix}$$

1.8 הגדרה – קבוצה $E \subseteq \mathbb{R}^d$ נקראת **קבוצה פתוחה** אם לכל $x \in E$ קיים $r > 0$ כך שמתקיים:

$$B_r(x) \subseteq E$$

דוגמאות: כדור פתוח, מלבן סגור לא, המרחב כולו, הקבוצה הריקה.

1.9 הגדרה – אם $x \in \mathbb{R}^d$, קבוצה $N \subseteq \mathbb{R}^d$ נקראת **סביבה** של x אם x מוכל בקבוצה פתוחה המוכלת ב N.

1.10 הגדרה – קבוצה $E \subseteq \mathbb{R}^d$ נקראת **קבוצה סגורה** אם מתקיים התנאי:

$$\forall (x^n)_{n=1}^{\infty} \subseteq E \quad x^n \to x \Longrightarrow x \in E$$

דוגמאות: כדור פתוח לא, מלבן סגור כן, וכמו כן המרחב כולו והקבוצה הפתוחה.

1.11 טענה – קבוצה $E \subseteq \mathbb{R}^d$ קבוצה סגורה אם ורק אם המשלימה שלה $E^c \coloneqq \mathbb{R}^d \setminus E$ קבוצה פתוחה.

<u>הוכחה:</u>

($\Longleftarrow$) נניח כי E^c אינה פתוחה, ונראה ש- E איננה סגורה. E^c אינה פתוחה, לכן קיים $x \notin E$ כך שלכל $r > 0$ מתקיים: $B_r(x) \cap E \neq \emptyset$. נבחר, אם כן, $x^n \in B_{\frac{1}{n}}(x) \cap E$, ונשים לב כי x^n מתכנסת ל-x ולכן E אינה סגורה.

($\Longrightarrow$) אם E לא סגורה, אזי יש נקודה $x \notin E$ שאליה מתכנסת סדרת נקודות מתוך E. נובע שהמשלים של E לא קבוצה פתוחה. (למה?)

מש"ל.

1.12 הגדרה (קומפקטיות) – קבוצה $E \subseteq \mathbb{R}^d$ נקראת **קומפקטית** אם לכל סדרה $(x^n)_{n=1}^{\infty} \subseteq E$ קיימת תת סדרה $(x^{n_k})_{k=1}^{\infty} \subseteq E$ כך ש-$x^{n_k} \to x \in E$.

1.13 הגדרה – קבוצה $E \subseteq \mathbb{R}^d$ נקראת **חסומה** אם קיים $R > 0$ כך שלכל $x \in E$ מתקיים $\|x\| < R$.

1.14 משפט – קבוצה $E \subseteq \mathbb{R}^d$ קומפקטית אם ורק אם היא סגורה וחסומה.

<u>הוכחה:</u>

($\Longleftarrow$) אם E לא חסומה יש $\|x^n\| > n$ לכל $n \in \mathbb{N}$. לסדרה זו אין תת-סדרה מתכנסת (כי כל תתי הסדרות מתבדרות לאינסוף). אם E לא סגורה, אז ישנה סדרה $(x^n)_{n=1}^{\infty}$ כך ש-$E \supseteq x^n \to x \notin E$, ולסדרה זו אין תת-סדרה שמתכנסת לנקודה ב- E.

($\Longrightarrow$) נשתמש בטענת עזר:

1.15 טענת עזר – לכל סדרה חסומה ב-$\mathbb{R}^d$ יש תת סדרה מתכנסת ב-$\mathbb{R}^d$.

הוכחת טענה זו דומה להוכחה מאינפי 1. אפשרות אחרת: באינדוקציה, באמצעות תתי סדרות מתכנסות של הרכיב הראשון (ממשפט בולצנו-ויירשטראס בממד אחד) שמשרה תת סדרה מתכנסת של הרכיב השני וכן הלאה לכל הרכיבים ולבסוף מתקבלת תת סדרה מתכנסת.

לשם השלמות נציג כאן את ההוכחה של טענת העזר באינדוקציה. המקרה $d=1$ ידוע מאינפי 1. כעת נניח ש - $d>1$ ושטענת העזר נכונה עבור ממדים קטנים מ- d, כלומר נניח שהוכחנו כבר שלכל $k<d$ מתקיים שלכל סדרה חסומה ב - $\mathbb{R}^k$ יש תת סדרה מתכנסת. תהא $\{x^n\}$ סדרה חסומה ב - $\mathbb{R}^d$. נוכל לרשום $x^n=(y^n,z^n)$, כאשר $y^n\in\mathbb{R}^{d-1}$ ו - $z^n\in\mathbb{R}$. כיון ש - $\|x^n\|^2=\|y^n\|^2+\|z^n\|^2$, נקבל שהסדרות $\{y^n\}$ ו - $\{z^n\}$ חסומות. מהנחת האינדוקציה, קיימת תת-סדרה מתכנסת $y^{n_k}\to y$. $\{z^{n_k}\}$ עדיין סדרה חסומה, לכן יש לה תת-סדרה מתכנסת $z^{n_{k_l}}\to z$. נקבל ש - $\{(y^{n_{k_l}},z^{n_{k_l}})\}$ תת-סדרה מתכנסת ל- (y,z). זה מוכיח את טענת העזר.

סיום הוכחת משפט 1.14 – נניח אם כן, כי E סגורה וחסומה. לכל סדרה $(x^n)_{n=1}^\infty\subseteq E$ קיימת תת סדרה מתכנסת $(x^{n_k})_{k=1}^\infty$ המתכנסת לאיבר $x\in\mathbb{R}^d$. מסגירות E $x\in E$ נקבל את הקומפקטיות כנדרש. מש"ל.

1.16 הערה – קומפקטיות שקולה לתכונה הבאה – בכל פעם שנתון אוסף $\{U_i\}_{i\in I}$ של קבוצות פתוחות כך ש-$E\subseteq$ $\bigcup_{i\in I}U_i$. אזי ניתן לבחור תת אוסף סופי $U_{i_1},\cdots,U_{i_n}$ כך שמתקיים:

$$E\subseteq\bigcup_{j=1}^{n}U_{i_j}$$

התנאי הזה נקרא **תנאי הכיסוי הסופי**, והוא אולי מוכר לחלקכם מהקורס אינפי 1 כתנאי המופיע במשפט היינה-בורל. בספרות מקובל להגדיר קבוצה **קומפקטית** כאשר היא מקיימת את תנאי הכיסוי הסופי, ואילו לקרוא לקבוצה המקיימת את התכונה בהגדרה 1.11 לעיל **קומפקטית סדרתית**. בתרגיל הבית אתם תוכיחו ששתי ההגדרות הן שקולות (בהקשר של תתי קבוצות של המרחב האוקלידי).

<u>פנים, סגור, ושפה</u>

תהי $E\subseteq\mathbb{R}^d$.

1.17 הגדרה – **הפנים** של E זו קבוצת כל הנקודות שיש להן סביבה שמוכלת ב- E הקבוצה

$$int(E)\subseteq\{x\in\mathbb{R}^d:\exists r>0.\,B_r(x)\subseteq E\}$$

הוכיחו: הפנים של כל קבוצה זו הקבוצה הפתוחה הגדולה ביותר המוכלת בה.

1.18 הגדרה – **הסגור** של E זו הקבוצה

$$\bar{E}\subseteq\{x\in\mathbb{R}^d:\exists(x_n)\subseteq E.\,x_n\to x\}$$

הוכיחו: הסגור של קבוצה זו הקבוצה הסגורה הקטנה ביותר המכילה אותה.

1.19 הגדרה – **השפה** של E זו הקבוצה

$$\partial E=\{x\in\mathbb{R}^d:\forall r>0.\,\exists y\in E.\,z\in E^c.\,y,z\in B_r(x)\}$$

1.20 תרגיל – $\bar{E}=E\cup\partial E$.

1.21 תרגיל – נגדיר את הכדור הסגור ברדיוס r סביב x להיות

$$\overline{B_r}(x)=\{y\in\mathbb{R}^d|\|x-y\|\le r\}$$

הראו ש - $\overline{B_r}(x) = \overline{B_r(x)}$ **,** כלומר שהסגור של הכדור הפתוח זה הכדור הסגור.

<u>פונקציות רציפות:</u>

1.22 הגדרה – פונקציה $f\colon \overset{\subseteq\mathbb{R}^d}{A} \to \overset{\subseteq\mathbb{R}^m}{B}$ נקראת **רציפה** בנקודה $x \in A$ אם מתקיים שלכל $\varepsilon > 0$ קיימת $\delta > 0$ כך שלכל $y \in A$, אם $\|x - y\| < \delta$ אזי מתקיים $\|f(x) - f(y)\| < \varepsilon$. נאמר ש- f **רציפה ב-** A אם היא רציפה בכל נקודה $x \in A$.

1.23 הגדרה – פונקציה $f\colon \overset{\subseteq\mathbb{R}^d}{A} \to \overset{\subseteq\mathbb{R}^n}{B}$ נקראת **רציפה במידה שווה** אם לכל $\varepsilon > 0$ קיימת $\delta > 0$ כך שלכל $x, y \in A$, אם $\|x - y\| < \delta$ אזי $\|f(x) - f(y)\| < \varepsilon$.

1.24 הגדרה (שקולה לרציפות) – לכל סדרה $(x^n)_{n=1}^{\infty} \subseteq A$ מתקיים – אם $x^n \to x$ אזי בהכרח:

$$f(x^n) \to f(x)$$

1.25 הגדרה (שקולה, אם A, B פתוחות) – לכל קבוצה פתוחה $U \subseteq B$, גם $f^{-1}(U) = \{x \in A | f(x) \in U\}$ הינה קבוצה פתוחה.

הערה: הטענה שהגדרות 1.15, 1.17 ו- 1.18 שקולות דורשת הוכחה (תרגיל).

1.26 משפט

1.27 – תהא $A \subseteq \mathbb{R}^d$ קבוצה קומפקטית, וכן תהא $f\colon A \to \mathbb{R}$ רציפה. אזי f חסומה ב-A ומקבלת שם מקסימום ומינימום מוחלטים.

<u>הוכחה:</u>

תהא $(x^n)_{n=1}^{\infty}$ סדרה כך ש-$f(x^n) \to \sup_{x \in A} f(x)$. ניתן לבנות סדרה כזו מהגדרת הסופרמום. אך נשים לב כי הסדרה $(x^n)_{n=1}^{\infty}$ <u>לא</u> בהכרח מתכנסת, אך היא מוכלת בקבוצה סגורה וחסומה, ולכן, נקבל מקומפקטיות הקבוצה כי קיימת לסדרה זו תת סדרה מתכנסת $x^{n_k} \to x_0 \in A$ עבורה מתקיים $f(x^{n_k}) \to \sup_{x \in A} f(x)$ וכן $f(x^{n_k}) \to f(x_0)$ מרציפות.

קיבלנו כי x_0 הוא בדיוק המקסימום המוחלט שלה, ובאותה מכה קיבלנו גם שהפונקציה חסומה מלעיל. מש"ל.

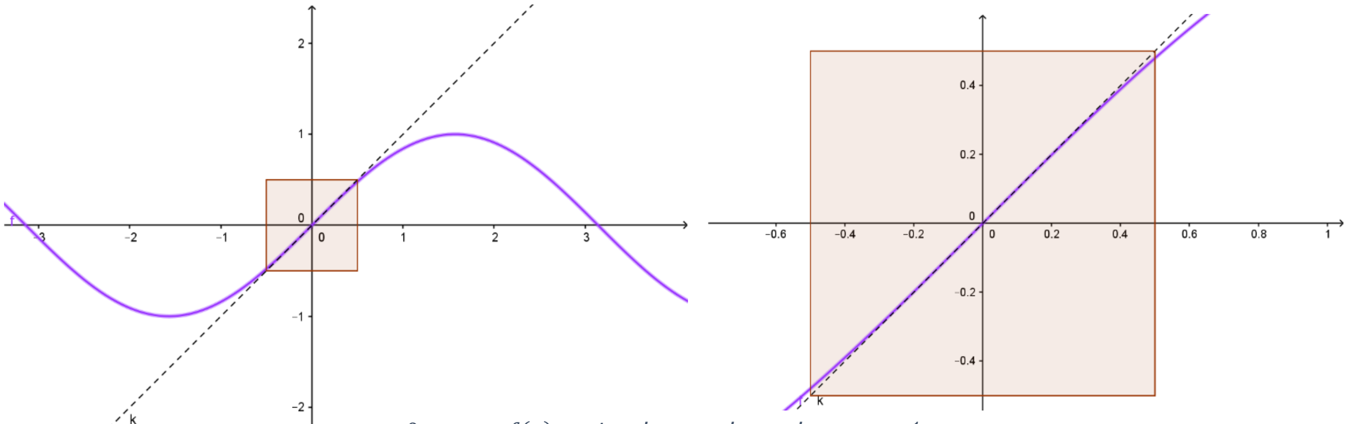
כתרגיל, הוכיחו את המשפטים החשובים הבאים:

1.28 משפט – פונקציה רציפה על קבוצה קומפקטית הינה רציפה במידה שווה.
1.29 משפט – תמונה של קבוצה קומפקטית על-ידי פונקציה רציפה הינה קבוצה קומפקטית.

הערה חשובה:

<u>לתרגול/תרגיל בית:</u> בתרגולים ובתרגילי הבית, ילמדו ויתרגלו גם המושגים של **סגור, שפה, פנים, קשירות, וקשירות מסילתית.** הוכחה שקבוצה פתוחה ב- $\mathbb{R}^n$ הינה קשירה אם ורק אם היא קשירה מסילתית, אם ורק אם היא קשירה פוליגונלית.

פרק 2 – גזירות



איור 1 – הדגמה לקירוב ליניארי של $f(x) = \sin x$ בסביבת $x = 0$

בהינתן $A \in \mathbb{R}^{m \times n}$ מוגדרת העתקה ליניארית $T_A: \mathbb{R}^n \to \mathbb{R}^m$ על ידי:

$$T_A(x) = Ax \quad x \in \mathbb{R}^n$$

A מכונה המטריצה המייצגת של T_A לפי הבסיס הסטנדרטי. אנחנו בדרך כלל מזהים בין T_A ל-A.

2.1 הגדרה – תהא $T: V \to W$ העתקה ליניארית בין מרחבי מכפלה פנימית. **נורמת האופרטורים** של T מוגדרת על ידי:

$$\|T\|_{op} = \|T\| \coloneqq \sup_{v \neq 0} \frac{\|Tv\|_W}{\|v\|_V}$$

2.1.1 הגדרה – T נקראת **חסומה** אם מתקיים $\|T\| < \infty$.

2.2 מסקנה – אם T חסומה, אזי מתקיים:

$$\forall v \in V \quad \|Tv\|_W \leq \|T\| \|v\|_V$$

2.3 טענה – תהא $T = T_A: \mathbb{R}^n \to \mathbb{R}^m$ העתקה ליניארית. אזי:

$$\|T_A\| \leq \left(\sum_{i,j} a_{ij}^2\right)^{\frac{1}{2}} < \infty$$

<u>הוכחה:</u>

על פי ההגדרה ואי-שוויון קושי-שוורץ מתקיים:

$$\|T_A x\|^2 = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j\right)^2 \leq \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}^2\right)\left(\sum_{j=1}^{n} x_j^2\right) = \left[\sum_{i,j} a_{ij}^2\right] \cdot \|x\|^2$$

2.4 מסקנה – לכל $x, y \in \mathbb{R}^n$ מתקיים:

$$\|T_A(x)-T_A(y)\| \le \|T_A\|\|x-y\|$$

כלומר בפרט T_A רציפה ואף רציפה ליפשיץ עם קבוע ליפשיץ $k=\|T_A\|$.

(**הגדרה:** פונקציה $f: A \to B$ נקראת **ליפשיצית** (או **רציפה ליפשיץ**) אם קיים k כך ש $\|f(x)-f(y)\| \le k\|x-y\|$ לכל $x,y \in A$. כל קבוע k כזה נקרא **קבוע ליפשיץ**).

2.5 הגדרה – פונקציה אפינית היא פונקציה מהצורה $T_{A,b}(x)=Ax+b$ עבור $\begin{matrix} A \in \mathbb{R}^{m\times n} \\ b \in \mathbb{R}^m \end{matrix}$.

<u>תזכורת:</u>

בהינתן $I \subset \mathbb{R}$ קטע פתוח, ו-$a \in I$, אזי $f: I \to \mathbb{R}$ נקראת **גזירה ב-a** אם קיים הגבול:

$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}$$

במידה והפונקציה אכן גזירה ב-a, מסמנים את הגבול בסימון $f'(a)$. יחד עם זאת, ניתן באופן שקול להגדיר את גזירות f ב-a, אם $f(a+h)-[f(a)+f'(a)h]=o(h)$. כאשר $o(h)$ מקיימת:

$$\lim_{h\to 0}\frac{o(h)}{h}=\lim_{h\to 0}\frac{f(a+h)-f(a)-f'(a)h}{h}=0$$

(הערה: דרך נכונה יותר להגדיר זאת תהיה על ידי קיומו של קבוע A כך ש-$f(a+h)-[f(a)+Ah]=o(h)$. אם קבוע כזה קיים, אז מסמנים אותו $A=f'(a)$).
במילים אחרות, אנו אומרים כי הפונקציה גזירה ב-a אם ניתן לקרב את f בסביבת a באמצעות פונקציה אפינית.

בעצם נרצה לומר, כי הישר $y(x)=f(a)+f'(a)(x-a)$ מהווה "קירוב טוב" ל-f בסביבת a. ניתן לראות זאת באיור 1 בתחילת העמוד הקודם בו ניתן לראות כיצד בסביבה קרובה של $x=0$, הישר $y=x$ מהווה קירוב טוב יותר ויותר של $f(x)=\sin x$ בסביבה זו (וככל שהסביבה תהא קטנה יותר הקירוב יהיה טוב יותר).

2.6 הגדרה – תהא $U \subseteq \mathbb{R}^n$ קבוצה פתוחה, ותהי $f: U \to \mathbb{R}^m$. בהנתן $a \in U$ נאמר כי f **גזירה ב-a** אם קיימת העתקה ליניארית $T: \mathbb{R}^n \to \mathbb{R}^m$ כך שמתקיים:

$$\lim_{h\to 0}\frac{f(a+h)-[f(a)+Th]}{\|h\|}=0$$

2.6.1 הערה – את הגבול עם הביטוי $h \to 0$ עבור h שהינו וקטור יש להבין באופן הבא: הכוונה היא שלכל $\varepsilon>0$ קיימת $\delta>0$ כך שלכל h עבורו $0<\|h\|<\delta$, מתקיים שהנורמה של המנה קטנה מ- ϵ. (כמובן נדרוש כי δ תהא קטנה די צורכנו על מנת ש-$B_\delta(a) \subseteq U$).

2.7 הגדרה (שקולה) – f גזירה ב-a אם קיימת העתקה ליניארית T ופונקציה ε המוגדרת בסביבת 0 כך שמתקיימת הזהות $f(a+h)=f(a)+Th+\varepsilon(h)$ וכן קיים הגבול:

$$\lim_{h\to 0}\frac{\varepsilon(h)}{\|h\|}=0 \qquad (*)$$

2.8 הגדרה – אם $\varepsilon: B_r(0) \to \mathbb{R}^m$, נכתוב $\varepsilon(h)=o(h)$ אם קיים הגבול (*) בהגדרה 2.7 (שימו לב שלא מדובר באמת בשוויון, זהו מקרה ברור של "התעללות בסימונים"! הסימן $o(h)$ אינו מציין פונקציה ספציפית שקוראים לה "או", ואילו $\varepsilon(h)$ הינה פונקציה ספציפית, והשוויון $\varepsilon(h)=o(h)$ פירושו פשוט שמתקיים הגבול (*)).

במקרה זה נוכל להגדיר בצורה נוספת גזירות על ידי:

$$f(a+h)-f(a)-Th=o(h)$$

2.9 טענה (יחידות הנגזרת) – אם f גזירה ב-a, אזי קיימת העתקה ליניארית יחידה T המקיימת את הגדרה הגזירות הנתונה בהגדרות 2.6,2.7.

<u>הוכחה:</u>

נניח כי אכן מתקיים:

$$f(a+h)=f(a)+T_ih+\varepsilon_i(h)$$

עבור $i=1,2$ (כלומר העתקות ליניאריות שונות המקיימות את הנ"ל על פי הגדרה 2.7). נרצה להראות כי:

$$\forall h\in\mathbb{R}^n \quad T_1h=T_2h$$

נשים לב כי:

$$\|T_1h-T_2h\| \overset{\text{אי שוויון המשולש}}{\leq} \|T_1h-[f(a+h)-f(a)]\|+\|[f(a+h)-f(a)]-T_2h\|=$$
$$\|\varepsilon_1(h)\|+\|\varepsilon_2(h)\|$$

עתה ניקח איבר במרחב $h\in\mathbb{R}^n$. לכל $t>0$ מתקיים :

$$\|T_1h-T_2h\|=\frac{t\|T_1h-T_2h\|}{t}=\frac{\|T_1(th)-T_2(th)\|}{t}\leq\frac{\|\varepsilon_1(th)\|}{t}+\frac{\|\varepsilon_2(th)\|}{t}$$

נשים לב כי שני הביטויים, באגף ימין, על פי הגדרה, מקיימים:

$$\lim_{t\to0}\frac{\|\varepsilon_i(th)\|}{t}=0 \quad i=1,2$$

ולכן גם האגף השמאלי חייב לשאוף ל-0. אך הביטוי בצד שמאל קבוע $\|T_1h-T_2h\|$ ולכן נסיק כי:

$$\|T_1h-T_2h\|=0\Longrightarrow T_1h-T_2h=0\rightarrow\boxed{T_1=T_2}$$ מש"ל.

2.10 הגדרה – אם f גזירה ב-a, נסמן ב-$f'(a)$, (או $Df(a)$) את ההעתקה הליניארית (היחידה) המקיימת:

$$f(a+h)=f(a)+f'(a)h+o(h)$$

$f'(a)$ נקראת **הנגזרת** (או **הדיפרנציאל**) של f ב-a.

ניתן גם לכתוב את המשוואה מהגדרה 2.10 באופן הבא:

$$f(x)=f(a)+f'(a)(x-a)+o(x-a)$$

2.11 הגדרה – אם f גזירה בכל נקודה ב-$U\subset\mathbb{R}^n$, נאמר כי f **גזירה ב-U.**

2.12 טענה – אם f פונקציה אפינית הנתונה ע"י $f(x)=Ax+b$, אזי f גזירה בכל $a\in\mathbb{R}^n$, ומתקיים $Df(a)=A$.

<u>הוכחה:</u> מתקיים $f(a+h)-[f(a)+Ah]=0$, ולכן מתקיים התנאי מהגדרה 2.6. מטענה 2.9 (יחידות הנגזרת), $Df(a)=A$. מש"ל.

מקרים פרטיים חשובים: f לינארית ו- f קבועה.

תרגיל – אם f גזירה ב-a אזי f רציפה ב-a.

2.13 טענה – בהינתן $f,g:\mathbb{R}^n\to\mathbb{R}^m$, אזי מתקיים:

א. $D(f+g)(a)=Df(a)+Dg(a)$

ב. אם $m=1$ אזי מתקיים $D(fg)(a)=g(a)Df(a)+f(a)Dg(a)$

ג. אם $m=1$ אזי מתקיים $D\left(\frac{f}{g}\right)=\frac{f(a)Dg(a)-g(a)Df(a)}{g(a)^2}$

הוכחת הטענה מושארת כתרגיל.

2.14 משפט (כלל השרשרת) – נניח כי f גזירה ב-a, ונניח כי g גזירה ב-$b:=f(a)$. אזי $g\circ f$ גזירה ב-a ומתקיים:

$$(g\circ f)'(a)=g'\big(f(a)\big)f'(a)$$

או בסימונים אחרים: $D(g\circ f)(a)=Dg(b)\circ Df(a)$.

<u>הוכחה:</u>

בהסתמכות על גזירות הפונקציות f,g, נכתוב:

$$\begin{aligned}f(a+h)&=f(a)+f'(a)h+\varepsilon_1(h)\\ g(b+k)&=g(b)+g'(b)k+\varepsilon_2(k)\end{aligned}\quad,\quad \frac{\varepsilon_i(v)}{\|v\|}\xrightarrow[\|v\|\to 0]{}0$$

אזי:

$$g\circ f(a+h)-g\circ f(a)=g\big(f(a+h)\big)-g\big(f(a)\big)=\circledast$$

בהגדרת $k=f(a+h)-f(a)=f'(a)h+\varepsilon_1(h)$ נקבל:

$$\circledast=g(b+k)-g(b)=g'(b)\big(f'(a)h+\varepsilon_1(h)\big)+\varepsilon_2(k)$$

$$=\underbrace{g'(b)f'(a)h}_{\text{רכיב ליניארי}}+g'(b)\varepsilon_1(h)+\varepsilon_2(k)$$

אך נשים לב כי:

$$\left\|\frac{g'(b)\varepsilon_1(h)}{\|h\|}\right\|\underset{\substack{g'(b)\\ \text{העתקה ליניארית}}}{=}\left\|g'(b)\left(\frac{\varepsilon_1(h)}{\|h\|}\right)\right\|\le\|g'(b)\|\cdot\left\|\frac{\varepsilon_1(h)}{\|h\|}\right\|\xrightarrow[h\to 0]{}0$$

ובנוסף:

$$k=f'(a)h+\varepsilon_1(h)\xrightarrow[h\to 0]{}0$$

$$\frac{\varepsilon_2(k)}{\|h\|}=\begin{cases}0 & k=0\\ \frac{\varepsilon_2(k)}{\|k\|}\cdot\frac{\|k\|}{\|h\|} & k\neq 0\end{cases}$$

במקרה שבו $k=0$ אנחנו נקבל את שרצינו. ונשים לב שעבור $k\neq 0$, ניתן לכתוב את הביטוי הימני במכפלה באופן הבא:

$$\frac{\varepsilon_2(k)}{\|k\|_{\neq 0}}\cdot\frac{\|f'(a)h+\varepsilon_1(h)\|}{\|h\|}\xrightarrow[h\to 0]{}0$$

ונקבל את הדרוש. מש"ל.

2.15 הגדרה – תהא $f\colon \underset{\subset\mathbb{R}^n}{U} \to \mathbb{R}$ ותהא $a \in U$ וכן יהא $0 \neq v \in \mathbb{R}^n$. אזי, אם קיים הגבול:

$$D_v f(a) = \lim_{t\to 0}\frac{f(a+tv)-f(a)}{t}$$

אזי נאמר כי f **גזירה בכיוון v**, ו - $D_v f(a)$ נקראת **הנגזרת הכיוונית/מכוונת של f בכיוון v.**

2.15.1 הגדרה – אם v (מהגדרה 2.14) הינו e_i, כלומר וקטור בסיס סטנדרטי, אזי $D_{e_i} f(a)$ נקראת **הנגזרת החלקית של f לפי המשתנה ה-i.** מסמנים:

$$D_{e_i} f(a) = D_i f(a) = \frac{\partial f}{\partial x_i}(a) = f_{x_i}(a)$$

2.16 טענה – תהא $f\colon \underset{\subset\mathbb{R}^n}{U} \to \mathbb{R}$ גזירה ב-a. אזי f גזירה בנקודה a בכל הכיוונים ובפרט קיימות כל הנגזרות החלקיות. הנגזרת בכיוון v נתונה על ידי:

$$D_v f(a) = Df(a)v$$

<u>הוכחה:</u>

בהינתן כי f גזירה ב-a, אזי קיימת העתקה ליניארית $Df(a)\colon \mathbb{R}^n \to \mathbb{R}$ כך שמתקיים:

$$f(a+h) = f(a) + Df(a)h + o(h)$$

ולכן יתקיים:

$$\frac{f(a+tv)-f(a)}{t} = \frac{Df(a)tv + o(tv)}{t} \quad \text{and} \quad \frac{o(tv)}{t}\xrightarrow[t\to 0]{} 0$$

וה-t מצטמצם בביטוי $\frac{Df(a)tv}{t}$ ולכן נקבל כי:

$$\circledast\ \boxed{\lim_{t\to 0}\frac{f(a+tv)-f(a)}{t} = Df(a)v}$$

במידה ו-f מקיימת את תנאי הטענה, מתקיים:

$$\frac{\partial f}{\partial x_j}(a) = Df(a)e_j$$

באופן כללי, המטריצה המייצגת של $Df(a)$ בבסיס הסטנדרטי היא:

$$\boxed{\text{הגרדיאנט} - \nabla f(a) = Df(a) := \left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_n}(a)\right)}$$

ולכן נקבל כי:

$$D_v f(a) = Df(a)v = \nabla f(a)\cdot v = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a)v_j$$

מש"ל.

הערה: עבור $\|v\| = 1$ נקבל כי הביטוי הנ"ל מקסימלי כאשר $v = \frac{\nabla f(a)}{\|\nabla f(a)\|}$. המשמעות היא שקצב הגידול של f הינו מקסימלי בכיוון של הגרדיאנט.

הערה: הנוסחה ⊛ שהופיעה בהוכחה לעיל הינה נכונה גם עבור $f: U \to \mathbb{R}^m$.

2.17 טענה – תהא $f: \underset{\subseteq \mathbb{R}^n}{U} \to \mathbb{R}$ גזירה ב-U. אם ל-f יש מקסימום או מינימום מקומי ב-$a \in U$, אז מתקיים:

$$\nabla f(a) = 0 \quad \begin{matrix} \frac{\partial f}{\partial x_j}(a) = 0 \\ \forall 1 \le j \le n \end{matrix}$$

<u>הוכחה:</u>

נשים לב כי אם $a \in U$ היא נקודה שבה מקסימום מקומי מתקבל, אזי נוכל להגדיר פונקציה חדשה, לכל $i = 1, \dots, n$, על ידי:

$$g_i(u) = f(a_1, a_2, \cdots, a_i + u, a_{i+1}, \cdots, a_n)$$

אך אנו יודעים כי עבור $u = 0$ מתקבל מקסימום מקומי, כלומר $g'(0) = 0$. אך מתקיים:

$$g'(0) = 0 = \lim_{u \to 0} \frac{f(a_1, a_2, \cdots, a_i + u, a_{i+1}, \cdots a_n) - f(a_1, a_2, \cdots, a_n)}{u} \overset{\text{לפי הגדרה}}{=} \frac{\partial f}{\partial x_i}(a)$$

כנדרש. מש"ל.

נחזור עתה לעיסוק במספר ממדים, קרי בפונקציות $f: \underset{\subset \mathbb{R}^n}{U} \to \mathbb{R}^m$. נכתוב $f = (f_1, f_2, \cdots, f_m)$ כאשר $f_i: \underset{\subset \mathbb{R}^n}{U} \to \mathbb{R}$ וכל f_i ניתנת על ידי $f_i = \pi_i \circ f$ כאשר:

$$\pi_i: \mathbb{R}^m \to \mathbb{R} \quad \pi_i(x_1, \cdots, x_m) = x_i$$

2.18 טענה – f גזירה ב-$a \in U$ אם ורק אם f_i גזירה ב-a לכל $i = 1, \cdots, m$ ומתקיים:

$$Df(a)v = (Df_1(a)v, \cdots, Df_m(a)v)$$

<u>הוכחה:</u>

אם f_i גזירה לכל i נגדיר:

$$Tv = (Df_1(a)v, \cdots, Df_m(a)v)^T$$

נבדוק כי אכן $f(a+h) - f(a) - Th = o(h)$:

$$\|f(a+h) - f(a) - Th\| \le \sum |f_i(a+h) - f_i(a) - Df_i(a)h| = \sum \epsilon_i(h) = o(h)$$

מצד שני, אם f גזירה ב-a אזי $f_i = \pi_i \circ f$ גזירה ב-a לפי כלל השרשרת ומתקיים:

$$Df_i(a) = D\pi_i\big(f(a)\big) \circ Df(a) = \pi_i \circ Df(a)$$

לכן התנאים שקולים.

לסיום אם $f\colon \underset{\subset\mathbb{R}^n}{U} \to \mathbb{R}^m$ כך ש-$f = (f_1, f_2, \cdots, f_m)$ גזירה ב-a, אזי ראינו ש $Df_i(a) = \nabla f_i(a) = \left(\frac{\partial f_i}{\partial x_1}, \ldots, \frac{\partial f_i}{\partial x_n}\right)$ זו השורה ה- i ב - $Df(a)$. בפרט כל הנגזרות החלקיות $\frac{\partial f_i}{\partial x_j}(a)$ קיימות והמטריצה המייצגת של $Df(a)$ לפי הבסיס הסטנדרטי היא:

$$\boxed{J_f(a) \coloneqq \left[\frac{\partial f_i}{\partial x_j}(a)\right]_{\substack{i=1,\cdots,m \\ j=1,\cdots,n}}}$$

מש"ל.

2.19 – הגדרה – המטריצה J_f מכונה **מטריצת היעקוביאן של f.**

נספח לפרק 2 – תכונות של נורמת האופרטורים וחבורת המטריצות ההפיכות

לשם פשטות בסימונים, אנו נזהה בין אופרטור לינארי על $\mathbb{R}^{n\times n}$ לבין המטריצה המייצגת אותו בבסיס הסטנדרטי.

תהי $A \in \mathbb{R}^{n\times n}$. הגדרנו לעיל את נורמת האופרטורים :

$$\|A\| = \|A\|_{op} = \sup\{\|Ax\| : \|x\| = 1\} = \sup\left\{\frac{\|Ax\|}{\|x\|} : x \neq 0\right\}$$

נורמת האופרטורים מקיימת לפי ההגדרה $\|Ax\| \leq \|A\|\|x\|$ לכל $x \in \mathbb{R}^n$, וכמו כן $\|A\|$ זה בדיוק הקבוע הקטן ביותר שמקיים את האי שוויון $\|Ax\| \leq \|A\|\|x\|$ לכל $x \in \mathbb{R}^n$. קל להוכיח ש-$A \mapsto \|A\|$ זו אכן נורמה, למשל, אם $A, B \in \mathbb{R}^{n\times n}$, אז

$$\|(A+B)x\| = \|Ax + Bx\| \leq \|Ax\| + \|Bx\| \leq \|A\|\|x\| + \|B\|\|x\| \leq (\|A\| + \|B\|)\|x\|$$

ולכן $\|A + B\| \leq \|A\| + \|B\|$. יתר תכונות הנורמה אף קלות יותר.

נורמת האופרטורים מקיימת גם את האי-שוויון החשוב

$$\|AB\| \leq \|A\|\|B\|$$

הנובע מיד מהחישוב: $\|ABx\| \leq \|A(Bx)\| \leq \|A\|\|Bx\| \leq \|A\|\|B\|\|x\|$.

תזכורת – מטריצה $A \in \mathbb{R}^{n\times n}$ נקראת **הפיכה** אם הגרעין שלה הוא טריוויאלי, כלומר אם $Ax = 0$ גורר $x = 0$. תנאי זה שקול לכך שההעתקה הלינארית $T_A\colon x \mapsto Ax$ היא העתקה חד-חד ערכית ועל של המרחב $\mathbb{R}^n$ על עצמו. זה גם שקול לכך ש - $\det A \neq 0$.

טענה – מטריצה $A \in \mathbb{R}^{n\times n}$ היא הפיכה אם ורק קיים $c > 0$ כך ש- $\|Ax\| \geq c\|x\|$ לכל x. למעשה, אם A הפיכה, אז הקבוע $c = \|A^{-1}\|^{-1}$ מקיים את אי השוויון הנ"ל, והוא הקבוע הגדול ביותר המקיים $\|Ax\| \geq c\|x\|$ לכל x.

<u>הוכחה:</u>

אם A הפיכה, אז עבור $\|A^{-1}\|$ מתקיים האי-שוויון הבא: $\|x\| = \|A^{-1}Ax\| \leq \|A^{-1}\|^{-1}\|Ax\|$, וכיון ש- $\|A^{-1}\| \neq 0$ אפשר לחלק בו ולקבל ש- $\|Ax\| \geq c\|x\|$ לכל x עבור $c = \|A^{-1}\|^{-1}$. זהו הקבוע הגדול ביותר המקיים אי שוויון זה, כי אם $\|Ax\| \geq r\|x\|$ לכל x, אז נרשום $y = Ax$, וכיון ש- A הפיכה, נסיק ש- $\|A^{-1}y\| \leq r^{-1}\|y\|$ לכל y. לכן $\|A^{-1}\| \leq r^{-1}$, או $r \leq \|A^{-1}\|^{-1}$.

בכיוון השני, אם קיים $c > 0$ כך ש- $\|Ax\| \geq c\|x\|$ לכל x אז נובע ש - $Ax = 0$ גורר $x = 0$, כלומר A הפיכה. מש"ל.

הערה: אם עבור אופרטור לינארי A על מרחב וקטורי כלשהו קיים $c > 0$ כך ש- $\|Ax\| \geq c\|x\|$ לכל x, נאמר ש- A **חסום מלמטה**.

טענה – אם $A, B \in \mathbb{R}^{n\times n}$, אם A הפיכה, ואם $\|A - B\| < \frac{\|A^{-1}\|^{-1}}{2}$, אזי B גם כן הפיכה, ומתקיים $\|B^{-1}\| \leq 2\|A^{-1}\|$.

<u>הוכחה:</u>

ע"פ הטענה הקודמת, מספיק לבדוק שהאופרטור ש- B מגדירה הינו חסום מלמטה. לכל x מתקיים

$$Bx = Ax + (B - A)x$$

ולכן

$$\|Bx\| \geq \|Ax\| - \|(B - A)x\| \geq \|A^{-1}\|^{-1}\|x\| - \frac{\|A^{-1}\|^{-1}}{2}\|x\| \geq \frac{\|A^{-1}\|^{-1}}{2}\|x\|$$

קיבלנו ש- B הפיכה וש- $\|B^{-1}\|^{-1} \geq \frac{\|A^{-1}\|^{-1}}{2}$, כנדרש. מש"ל.

הגדרה – נסמן ב - $GL(\mathbb{R}^n)$ (או GL_n) את אוסף המטריצות ההפיכות ב- $\mathbb{R}^{n\times n}$.

ברור ש- $GL(\mathbb{R}^n)$ הינה חבורה תחת כפל (מכפלת הפיכות הינה הפיכה, וההופכית של הפיכה היא הפיכה גם כן). משתמשים באותו סימון עבור אוסף ההעתקות הלינאריות על $\mathbb{R}^n$.

משפט – החבורה $GL(\mathbb{R}^n)$ הינה פתוחה ב - $\mathbb{R}^{n\times n}$, **והפעולות של כפל והופכי הינן רציפות.**

<u>הוכחה:</u>

לפי הטענה הקודמת, אם $A \in GL(\mathbb{R}^n)$ אז עבור $r = \frac{\|A^{-1}\|^{-1}}{2}$ מתקיים $B_r(A) \subseteq GL(\mathbb{R}^n)$, כלומר $GL(\mathbb{R}^n)$ הינה פתוחה. את הרציפות אפשר להוכיח לפי בחינה של הקואורדינטות של המטריצות, הרי ההופכית תלויות בצורה מפורשת באיברי המטריצה המקורית, לפי הנוסחה $A^{-1} = \frac{1}{\det A} adj(A)$, והאיבר הכללי של AB נתון ע"י הנוסחה הפשוטה $(AB)_{ij} = \sum_k A_{ik}B_{kj}$. אך מעדיפים לתת הוכחות לפי ההגדרה, מבלי להידרש לנוסחאות מפורשות אלו.

רציפות הכפל: שימו לב, צריך להוכיח שההתקה $m: GL(\mathbb{R}^n) \times GL(\mathbb{R}^n) \to GL(\mathbb{R}^n)$ הנתונה ע"י $m(A, B) = AB$ הינה רציפה. נניח ש- $A_k \to A$ ו- $B_k \to B$. אז

$$\|A_kB_k - AB\| = \|A_kB_k - A_kB + A_kB - AB\| \leq \|A_k\|\|B_k - B\| + \|A_k - A\|\|B\| \to 0$$

שימו לב שהשתמשנו בעובדה שסדרה מתכנסת הינה חסומה. לכן הכפל רציף.

רציפות ההופכי: נקבע $A \in GL_n$. יהי $\varepsilon > 0$. נגדיר $\delta = \min\left\{\frac{\|A^{-1}\|^{-1}}{2}, \frac{\|A^{-1}\|^{-2}}{2}\right\} \cdot \varepsilon$. כעת לכל B, אם $\|B - A\| < \delta$, אזי B הפיכה ומתקיים $\|B^{-1}\| \leq 2\|A^{-1}\|$.

לכן

$$\|A^{-1} - B^{-1}\| \leq \|A^{-1}(B - A)B^{-1}\| \leq 2\|A^{-1}\|^2\delta < \varepsilon$$

וזה מוכיח רציפות בנקודה A. מש"ל.

פרק 3

תזכורת:

העתקה $f:\mathbb{R}^n\to\mathbb{R}^m$ נקראת **גזירה** (או **דיפרנציאבילית**) אם ניתן לקרב אותה ליניארית (ליתר דיוק – אפינית). במקרה של פונקציה במשתנה אחד הקירוב נתון על ידי:

$$f(x)=f(a)+f'(a)(x-a)+o(x-a)$$

ובעבור f שהיא פונקציה במספר ממדים מתקיים:

$$f=\begin{pmatrix}f_1\\ \vdots\\ f_m\end{pmatrix}\qquad f'(a)=J_f(a)=\underbrace{\left[\frac{\partial f_i}{\partial x_j}(a)\right]}_{\text{יעקוביאן}}$$

כאשר $J_f(a)$ הינה מטריצה המכילה את כל הנגזרות החלקיות לפי כל המשתנים בנקודה $x=a$. במידה והן רציפות נקבל כי f דיפרנציאבילית בנקודה a :

3.1 משפט – תהא $f:\underset{\subseteq\mathbb{R}^n}{U}\to\mathbb{R}^m$ **מהצורה** $f=\begin{pmatrix}f_1\\ \vdots\\ f_m\end{pmatrix}$**. אזי, נניח כי לכל** i,j**, הנגזרת החלקית** $\frac{\partial f_i}{\partial x_j}$ **רציפות בסביבת** a **(כלומר, בכדור** $B_r(a)\subseteq U$ **כלשהו), אזי** f **גזירה (דיפרנציאבילית) ב-**a**.**

<u>הוכחה:</u>

לפי טענה משיעור קודם, מספיק להוכיח כי f_i גזירה לכל i[1]. לכן נניח כי $f:U\to\mathbb{R}$ ויהא $h\in\mathbb{R}^n$. ניתן לכתוב:

$$(h_1,h_2,\cdots,h_n)=h=\sum_{j=1}^{n}h_je_j$$

אזי נקבל כי:

$$f(a+h)-f(a)=\sum_{j=1}^{n}\left(f\left(a+\sum_{i=1}^{j}h_ie_i\right)-f\left(a+\sum_{i=1}^{j-1}h_ie_i\right)\right)^{2}$$

$$=\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}\left(a+\sum_{i=1}^{j-1}h_iej\right)h_j+o\left(h_j\right)=\nabla f(a)\cdot h+\sum_{j=1}^{n}\left(\frac{\partial f}{\partial x_j}\left(a+\sum_{i=1}^{j-1}h_ie_i\right)-\frac{\partial f}{\partial x_j}(a)\right)h_j+o(h)$$

נותר להראות כי:

---

[1] שכן בהרצאה הקודמת הראינו כי גזירות פונקציה מ-$\mathbb{R}^n$ ל-$\mathbb{R}$ מאפשרת לקבל קירוב ליניארי שלי, וכך, במידה ונוכיח זאת לכל הרכיבים, נוכל לקבל קירוב כנ"ל (כלומר דיפרנציאביליות) לכל הרכיבים ב-f, כלומר, היא תהא דיפרנציאבילית.

[2] נשים לב כי זהו טור טלסקופי מהצורה:

$$f(a)-f(a_1+h_1,a_2,\cdots,a_n)+f(a_1+h_1,a_2,\cdots,a_n)-f(a_1+h_1,a_2+h_2,a_3,\cdots,a_n)$$
$$+f(a_1+h_1,a_2+h_2,a_3,\cdots,a_n)\cdots-f(a_1+h_1,a_2+h_2,a_3+h_3,a_4,\cdots,a_n)\cdots$$

כך שמתקבל, ראשית, סכום טלסקופי שמותיר לנו את ההפרש המקורי שרצינו לחשב, אך עתה, נשים לב שכל צמד איברים בסכום זה הוא נגזרת חלקית לפי משתנה יחיד. היות וכל הנגזרות החלקיות קיימות, נוכל לבצע לכל צמד כזה קירוב ליניארי כמתואר בהמשך ההוכחה.

$$\sum_{j=1}^{n}\left(\frac{\partial f}{\partial x_j}\left(a+\sum_{i=1}^{j-1}h_i e_i\right)-\frac{\partial f}{\partial x_j}(a)\right)h_j = o(h)$$

אבל זה נובע מכך שהנגזרות החלקיות רציפות בסביבת a. לכן עבור $h \to 0$ אכן נקבל ש- f דיפרנציאבילית ב-a כנדרש.
מש"ל.

3.2 סימון – אם $\frac{\partial f_i}{\partial x_j}$ רציפות ב-U לכל i, j נרשום $f \in C^1$ או $f \in C^1(U)$ או $f \in C^1(U, \mathbb{R}^m)$ כאשר זהו סימן של קבוצת הפונקציות הגזירות ברציפות ב-U.

3.3 הערה – פונקציה $f: U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ הינה גזירה ברציפות אם ורק אם ההתקה $x \mapsto Df(x)$ היא רציפה מ- U לתוך מרחב ההעתקות הלינאריות $L\left(\mathbf{R}^n, \mathbf{R}^m\right)$ עם נורמת האופרטורים.

דוגמה – הפונקציה $A \mapsto A^{-1}$

נגדיר את הפונקציה $f: GL_n \to GL_n$ באופן הבא: $f(A) = A^{-1}$. ראינו בסוף של פרק 2 שזו פונקציה רציפה. נראה כעת שהפונקציה הזו גם גזירה.

הוכחה מעט "מלוכלכת", היא להיזכר שההפוך נתון על-ידי $A^{-1} = \frac{1}{\det A}adj(A)$. זוהי נוסחה ידועה מאלגברה לינארית, ומראה שכל רכיב במטריצה A^{-1} נתון כפונקציה מפורשת התלויה באופן C^1 באיברי המטריצה. למעשה: $(A^{-1})_{ij} = \frac{1}{\det A}(-1)^{i+j}M_{ji}$ כאשר M_{ij} זה הקו-פקטור ה-i,j כלומר הדטרמיננט של המטריצה המתקבלת מ- A על-ידי מחיקת השורה ה-i והעמודה ה- j. לכן הפונקציה $f: GL_n \to GL_n$ הינה ב - $C^1(GL_n)$, ולכן לפי משפט 3.1 נובע ש- f גזירה.
אני מזמין אתכם לנסות, כתרגיל, לחשב את $Df(A)$.

אנו רוצים להציג גם הוכחה לפי ההגדרה שהפונקציה f היא גזירה. מלבד ההיבט הפדגוגי, יתרון של הוכחה זו הוא שנקבל מיד ביטוי מפורש עבור הנגזרת.

נחזור על החישוב מהסוף של פרק 2 (שם ראינו ש- f רציפה)

$$f(A+H)-f(A) = (A+H)^{-1}-A^{-1} = (A+H)^{-1}\big(A-(A+H)\big)A^{-1} = -(A+H)^{-1}HA^{-1}$$

מטרתנו להציג את הביטוי הנ"ל כפונקציה לינארית ב- H ועוד איבר שגיאה שהולך לאפס מהר יותר מ- H.

$$-(A+H)^{-1}HA^{-1} = -A^{-1}HA^{-1} + (A^{-1}-(A+H)^{-1})HA^{-1}$$

ההעתקה $H \mapsto A^{-1}HA^{-1}$ היא לינארית (ב- H !!!) ונותר להראות ש - $(A^{-1}-(A+H)^{-1})HA^{-1} = o(H)$. אבל ראינו כבר בסוף של פרק 2 ש -f רציפה, לכן

$$\left\|\frac{(A^{-1}-(A+H)^{-1})HA^{-1}}{\|H\|}\right\| \le \frac{\|f(A)-f(A+H)\|\cdot\|H\|\cdot\|A^{-1}\|}{\|H\|}$$
$$= \|f(A)-f(A+H)\|\|A^{-1}\| \xrightarrow{H\to 0} 0$$

לסיכום, קיבלנו ש- f גזירה וש- $Df(A)H = -A^{-1}HA^{-1}$. אנו מקבלים מהנוסחה הזו ש- f גם גזירה ברציפות.

מעניין להשוות את הנוסחה הזו עם הנוסחה של הנגזרת של הפונקציה הממשית $t \mapsto \frac{1}{t}$.

נספח לפרק 3 – שיטת הגרדיאנט למציאת מינימום ושיטת הריבועים המינימליים

תהא $f: U \subset \mathbb{R}^n \to \mathbb{R}$ פונקציה גזירה. נניח שאנו מעוניינים למצוא את המינימום (אול לפחות, מינימום לוקלי) של f. הגרדיאנט הוא כלי חשוב בפתרון הבעיה הזו.

שיטה ראשונה: היא להשתמש בעובדה שאם f מקבלת מינימום מקומי בנקודה $a \in U$ (כלומר קיים $r > 0$ כך שמתקיים $f(a) \leq f(x)$ לכל $x \in B_r(a)$) , אז $\nabla f(a) = 0$. כלומר אנו מקבלים n משוואות ב n נעלמים

$$\frac{\partial f}{\partial x_i}(a_1, \dots, a_n) = 0 \quad , \quad i = 1, \dots, n$$

לפעמים יש בידינו לפתור ממש את n המשוואות הללו, ולמצוא נקודה או נקודות שם הנגזרת מתאפסת, ונקודות אלו הן חשודות למינימום. אם מסיבה כלשהיא אנו יודעים שהמינימום חייב להתקבל, ויש רק נקודה אחת בה הגרדיאנט מתאפס, אזי אנו יודעים בביטחון שהמינימום מתקבל בנקודה זו.

דוגמה: נתבונן בפונקציה $f(x) = \|Ax - b\|^2$, כאשר $A \in \mathbb{R}^{m \times n}$ ו- $b \in \mathbb{R}^m$. כאשר הדרגה של A היא m, ניתן לפתור את מערכת המשוואות הלינאריות

$$\circledast \quad Ax = b$$

והמינימום הוא אפס. הבעיה מתחילה להיות מעניינת יותר כאשר $m > n$, ואז המשמעות היא שלא ניתן תמיד לפתור את המשוואה, אבל אנו מחפשים "פתרון מקורב" x, כלומר מחפשים x כזה שיפתור את המשוואה הלינארית $\circledast$ עם השגיאה הכי קטנה. בעיה זו נקראת "בעיית הריבועים המינימליים" (כי השגיאה אותה אנו מנסים להביא למינימום היא סכום של ריבועים: $\|Ax - b\|^2 = \sum_i \left(\sum_j a_{ij}x_j - b_i\right)^2$) או "רגרסיה לינארית", והיא חשובה מאוד כמעט בכל תחום בהנדסה, מדע או סטטיסטיקה, לרבות עיבוד תמונה ולמידת מכונה. במקרה זה, כאשר $m > n$, זו כבר לא בעיה טהורה באלגברה לינארית אלא בעיה אנליטית.

במקרה של הבעיה הפשוטה הזו, ניתן למצוא נוסחה סגורה עבור הפתרון. ראשית נחשב את הגרדיאנט. אפשר להשתמש בכלל השרשרת (ושיטה זו מופיעה ברשימות התרגולים של רן קירי), אבל כאן נראה שיטה אחרת. בחישוב הבא נתייחס לוקטורים ב- $\mathbb{R}^n$ כאל וקטורי עמודה, ופונקציונלים לינאריים על $\mathbb{R}^n$ ייצוגו כוקטורי שורה (המכפילים את וקטורי העמודה משמאל לפי כפל מטריצות רגיל). נשים לב ש –

$$f(x) = \langle Ax - b, Ax - b\rangle = \langle A^T Ax, x\rangle - 2\langle Ax, b\rangle + \langle b, b\rangle =: g(x) + h(x)$$

כאשר $g(x) = \langle A^T Ax, x\rangle = x^T A^T Ax$ ו- $h(x) = -2\langle Ax, b\rangle + \langle b, b\rangle$. מתקיים $\nabla f = \nabla g + \nabla h$.נשים לב ש - h הינה פונקציה אפינית, ולכן הנגזרת שלה היא החלק הלינארי $Dh(x): h \mapsto -2\langle Ah, b\rangle = -2b^T Ah$, כלומר

$$\nabla h(x) = -2b^T A$$

כעת נחשב את $Dg(x)$. עם קצת אלגברה:

$$g(x + h) - g(x) = \langle A^T A(x + h), x + h\rangle - \langle A^T Ax, x\rangle = 2\langle A^T Ax, h\rangle + \|Ah\|^2$$

אנו מזהים ש- $2\langle A^T Ax, h\rangle$ זה ביטוי לינארי ב-h, ואילו $\|Ah\|^2 \leq \|A\|^2\|h\|^2 = o(\|h\|)$. לכן אין ברירה אחרת: $Dg(x): h \mapsto 2\langle A^T Ax, h\rangle = 2x^T A^T Ah$, זאת אומרת ש - $\nabla g(a) = 2x^T A^T A$. לכן

$$\nabla f(x) = 2x^T A^T A - 2b^T A$$

כדי למצוא את נקודת המינימום של f , נוכל לפתור את המשוואה $\nabla f(x) = 0$ – ה x – בו הגרדיאנט מתאפס הוא נקודת מינימום מקומי, ואם יש רק x בו הגרדיאנט מתאפס, אזי זוהי נקודת מינימום גלובלית (לא נכנס כעת להסבר מדוע f חייבת לקבל מינימום, כיון שמטרתנו עכשיו היא להתמקד בשיטות מעשיות למציאת מינימום).

המשוואה

$$\nabla f(x) = 2x^T A^T A - 2b^T A = 0$$

שקולה למשוואה הלינארית

$$(*) \qquad A^T A x = A^T b$$

בדרך כלל בשימושים, כאשר $m > n$, מתקיים ש - $rankA = n$, ולכן המטריצה הריבועית $A^T A$ הינה גם כן מדרגה n, כלומר הפיכה (הוכחה: אם $rankA = n$ אזי מהנוסחה $\dim KerA + rankA = n$ מתקיים שההעתקה $x \mapsto Ax$ הינה חד-חד ערכית, לכן אם $A^T Av = 0$ אזי $\|Av\| = \langle A^T Av, v\rangle = 0$, כלומר $v = 0$, וזה מוכיח ש- $A^T A$ מטריצה הפיכה) .
במקרה הזה למשוואה $(*)$ יש פתרון יחיד והוא נתון על-ידי הנוסחה:

$$x = (A^T A)^{-1} A^T b$$

נציין שפועל את המשוואה $(*)$ פותרים לאו דווקא על-ידי חישוב המטריה ההופכית $(A^T A)^{-1}$, כיון שחישוב הופכי זו פעולה מורכבת מבחינה נומרית ויש אלגוריתמים לפתרון מערכת משוואות שאינם דורשים את חישוב ההופכית (כפי שפתרתם מערכות משוואות באלגברה לינארית בלי לחשב את ההופכית באופן מפורש).

<u>שיטה שנייה למציאת מינימום של פונקציה גזירה – שיטת הגרדיאנט (*Gradient descent*):</u>

בדוגמה לעיל מערכת המשוואות $\nabla f(x) = 0$ הייתה מערכת משוואות לינאריות, ולכן היה אפשר למצוא פתרון. בפועל, מציאת האפס של מערכת משוואות כללית הוא עניין מסובך ולפעמים אי אפשר למצוא פתרון "סגור". שיטת הגרדיאנט למציאת מינימום משתמשת בתכונה החשובה של הגרדיאנט, והיא שהכיוון של וקטור הגרדיאנט בכל נקודה הוא הכיוון בו הפונקציה גדלה בקצב הכי מהיר. "שיטת הגרדיאנט" זהו בעצם שם קוד לאוסף של שיטות הפועלות לפי העיקרון הבא. בהינתן פונקציה גזירה $f: U \subset \mathbb{R}^n \to \mathbb{R}$, נחפש את המינימום שלה לפי השלבים הבאים:

1. ננחש פתרון x לבעיית המינימיזציה. נחשב את הגרדיאנט $\nabla f(x)$ בנקודה x.
2. אם $\nabla f(x)$ שווה לאפס (או אם $\|\nabla f(x)\|$ קטן מאוד), אז מצאנו נקודת מינימום ו- x הינה הנקודה המבוקשת, וכאן האלגוריתם מסתיים.
3. אחרת, מבצעים $x \leftarrow x - c\nabla f(x)$, כלומר מחליפים את x ב - $x - c\nabla f(x)$. המשמעות היא שאנו מתקנים את הניחוש הקודם שלנו בכך שאנו מזיזים אותו בכיוון שבו הפונקציה f יורדת הכי מהר. הקבוע c הוא בדרך כלל קטן מ-1 , ומטרתו לדאוג שלא נלך יותר מדי בכיוון הגרדיאנט.
4. חוזרים לשלב 2.

השלבים הרשומים למעלה הם רק בסיס הרעיון, ובדרך כלל המימוש כולל כל מיני תוספות. למשל, צריך להגדיר מה זאת אומרת "$\|\nabla f(x)\|$ קטן מאוד", כלומר מגדירים איזשהו $\varepsilon > 0$, ובשלב 2 בודקים אם, $\|\nabla f(x)\| < \varepsilon$, אם כן מחזירים את x כפתרון, ואם לא אז ממשיכים. כמו כן, קובעים מספר מקסימלי של איטרציות, כדי שאם האלגוריתם לא מוצא מינימום הוא לא ימשיך לרוץ לנצח.

בעמוד הבא קוד *MATLAB* לדוגמה המדגים את שיטת הגרדיאנט עבור בעיית הריבועים המינימליים, ומשווה את הפתרון המקורב עם הפתרון המתקבל על-ידי הנוסחה הסגורה. מציין שעובר מטריצות ענקיות, לפעמים עדיף להשתמש בשיטת הגרדיאנט מאשר בפתרון סגור משיקולי סיבוכיות.

<u>תרגיל:</u> קראו את הקוד הבא והסבירו מהו התפקיד של לולאת ה- *while*.

קוד לדוגמה (מאטלב):

```
% Gradient descent demo

% Generate random data for the problem:
m = 15;
n = 4;
A = randn(m,n);
b = randn(m,1);

% Maximal number of iteration that we are going to run:
Niter = 100;

% Set a threshold value for when we think we reached:
thresh = 10^(-4);

% Random initial guess
x = randn(n,1);
err = norm(A*x-b)^2;

for i=1:Niter
    % The following line computes the gradient of the function ||A*x-b||^2
    grad = 2*A'*A*x - 2*A'*b;

    % Check if the gradient is very small, in which case stop:
    if norm(grad) < thresh
        break
    end
    c = 1;
    x_tmp = x;
    err_tmp = err;
    while (err <= err_tmp) & (c>2*eps)
        % The gradient descent step:
        x_tmp = x - c*grad;
        err_tmp = norm(A*x_tmp-b)^2;
        c = c/2;
    end
    err = err_tmp;
    x = x_tmp;
end

% Display the solution found by the algorithm:
disp('x = '), disp(x');
disp('||A*x - b||^2 = '), disp(err);

% Now find the minimum using the closed form formula found in the tutorial:
X = inv(A'*A)*A'*b;

% Dispaly the optimal solution:
disp('True solution X = '), disp(X');
disp('||A*X - b||^2 = '), disp(norm(A*X-b)^2);
```

פרק 4

נגזרות מסדר גבוה:

תהא $U \subseteq \mathbb{R}^n$ קבוצה פתוחה ותהא $f: U \to \mathbb{R}$ עבורה קיימות הנגזרות החלקיות:

$$D_i f = \frac{\partial f}{\partial x_i}$$

לכל $1 \le i \le n$ ובכל נקודה ב- U . יתכן שנגזרות אלו גזירות בעצמן, שכן הן פונקציות מהצורה $\frac{\partial f}{\partial x_i}: U \to \mathbb{R}$, ונוכל לכתוב:

$$D_i \frac{\partial f}{\partial x_j} = \frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

סימונים מקובלים נוספים הם:

$$\partial_i = \partial_{x_i} = D_i$$

ואם $f = f(x, y, \cdots)$ אזי בהנתן:

$$\frac{\partial f}{\partial x} = f_x$$

מסמנים:

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial y \partial x} = (f_x)_y = f_{xy}$$

ניתן להגדיר נגזרות מסדר גבוה יותר באופן דומה:

$$D_{i_1} D_{i_2} \cdots D_{i_k} = \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}$$

4.1 הגדרה –נסמן ב-$C^k(U)$ את כל הפונקציות $f: U \to \mathbb{R}$ המקיימות שכל הנגזרות החלקיות מסדר קטן או שווה ל-k קיימות והן כולן רציפות. הפונקציות ב - $C^k(U)$- נקראות **פונקציות גזירות k- פעמים ברציפות**.

4.2 משפט – תהא $U \subseteq \mathbb{R}^n$ פתוחה ותהא $f \in C^2(U)$. אזי, לכל i, j מתקיים:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}, \quad D_i D_j = D_j D_i$$

<u>הוכחה:</u>

מספיק להוכיח עבור $n = 2$. נרצה להראות, כי $f_{xy} = f_{yx}$. לשם כך, תהא $(a, b) \in U$ ונתבונן בביטוי:

$$G(h, k) = f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b)$$

<u>הערה</u> – מומלץ לסטודנטים לבדוק את הגבול:

$$\lim_{k \to 0} \lim_{h \to 0} \frac{G(h, k)}{h \cdot k}$$

ולראות כי ביטוי זה, הוא בדיוק $f_{xy}(a,b)$.

נגדיר עתה:

$$\varphi(t) = f(t, b+k) - f(t,b)$$

ונשים לב, כי:

$$G(h,k) = \varphi(a+h) - \varphi(a) \overset{\text{ערך הביניים}}{=} \varphi'(p)h = \big(f_x(p,b+k) - f_x(p,b)\big)h$$

$$\overset{\text{ערך הביניים}}{=} f_{xy}(p,q)kh$$

מצד שני, ניתן לבצע את התהליך בסדר הפוך ולקבל כי:

$$f_{xy}(p,q)kh = G(h,k) = f_{yx}(p',q')hk$$

עבור $h,k \to 0$ נקבל כי $p',p \to a$, וכן $q',q \to b$ ולכן מרציפות f_{xy}, f_{yx}, נקבל כי:

$$\boxed{f_{xy}(a,b) = \lim_{k,h\to 0} f_{xy}(p,q) = \lim_{k,h\to 0} f_{yx}(p',q') = f_{yx}(a,b)}$$

כנדרש. מש"ל.

<u>הערה:</u> שימו לב שבאמצעות ההוכחה שנתנו אפשר להוכיח את המשפט גם תחת תנאים מעט חלשים יותר: משפיק להניח שהנגזרות המעורבות f_{xy}, f_{yx} מוגדרות בסביבה של הנקודה, ו**שאחת מהן** רציפה בנקודה.

4.3 סימונים (Multi-Index Notation) – נסמן עבור $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{N}^n$:

$$|\alpha| = \alpha_1 + \cdots + \alpha_n \qquad \alpha! = \alpha_1! \cdot \cdots \cdot \alpha_n!$$

$$x = (x_1, \cdots, x_n) \in \mathbb{R}^n \quad x^\alpha = x_1^{\alpha_1} \cdot \cdots \cdot x_n^{\alpha_n} \in \mathbb{R} \quad \partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$$

<u>דוגמאות:</u>

$$(x_1 + x_2)^k = \sum_{|\alpha|=k} \frac{|\alpha|!}{\alpha!} x^\alpha = \sum_{\alpha_1+\alpha_2=k} \frac{k!}{\alpha_1!\,\alpha_2!} x^{\alpha_1} x^{\alpha_2}$$

$$(x_1 + \cdots + x_n)^k = \sum_{|\alpha|=k} \frac{|\alpha|!}{\alpha!} x^\alpha = \sum_{\alpha_1+\cdots+\alpha_n=k} \frac{k!}{\alpha_1!\,\alpha_2! \cdots \alpha_n!} x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

4.4 משפט (פיתוח טיילור) – תהא $U \subseteq \mathbb{R}^n$ פתוחה. תהא $f \in C^{k+1}(U)$. אזי לכל $a \in U$ ולכל h כך שמתקיים $(a+h) \in B_r(a) \subset U$ מתקיים:

$$f(a+h) = f(a) + \sum_{i=1}^n D_i f(a) h_i + \sum_{m=2}^k \sum_{|\alpha|=m} \frac{D^\alpha f(a)}{\alpha!} h^\alpha + R_k(f; a, h)$$

כאשר:

$$R_k(f;a,h)=\sum_{|\alpha|=k+1}\frac{D^\alpha f(a+ch)}{\alpha!}h^\alpha$$

עבור $0<c<1$ כלשהו. זה נקרא **פיתוח טיילור מסדר k (סביב a) ב-n משתנים**.

נוכל לכתוב את הביטוי הנ"ל מחדש על ידי:

$$f(x)=\underbrace{\sum_{|\alpha|\le k}\frac{D^\alpha f(a)}{\alpha!}(x-a)^\alpha}_{\substack{f\text{ פולינום טיילור של}\\ k \text{ מסביב } a \text{ מסדר}}}+R_k$$

4.5 הגדרה – פונקציה $p:\mathbb{R}^n\to\mathbb{R}$ מהצורה $p(x)=\sum_{|\alpha|\le k}c_\alpha x^\alpha$ נקראת **פולינום** (אם קיים מקדם $c_\alpha\neq 0$ עבור $|\alpha|=k$, נאמר ש - p הינו **פולינום מסדר** k). שימו לב שגם ביטוי מהצורה $\sum_{|\alpha|\le k}c_\alpha(x-a)^\alpha$ הוא פולינום מסדר k. הפולינום המופיע בפיתוח טיילור נקרא **פולינום טיילור**. הביטוי $R_k=R_k(f;a,h)$ נקרא **איבר השגיאה** בפיתוח טיילור (שימו לב: איבר השגיאה איננו פולינום). הביטוי ה"מפורש" שיקבלנו עבור R_k נקרא **השארית בצורת לגרנז'**.

פולינומים מהווים את אחד מאוספי הפונקציות הכי פשוטים שיש. אמנם לפתור משוואות פולינומיאליות מסדר גבוה אי-אפשר, אבל קל מאוד לחשב את הערך של פולינום בנקודה, ולנתח את ההתנהגות של הגרף של פולינום.

<u>תרגיל:</u> הוכיחו שאם $p(x)=\sum_{|\alpha|\le k}c_\alpha x^\alpha$ פולינום מסדר לכל היותר k המקיים $p(x)=o\left(\|x\|^k\right)$, אזי $p\equiv 0$.

פרק 5

5.1 משפט (פיתוח טיילור) – תהא $U \subseteq \mathbb{R}^n$ פתוחה ותהא $f \in C^{k+1}(U)$. אזי לכל $a \in U$ מתקיים:

$$f(a+h) = \sum_{|\alpha| \le k} \frac{D^\alpha f(a)}{\alpha!} h^\alpha + R_k(f;a,h)$$

לכל $h \in \mathbb{R}^n$ קטן מספיק, כאשר:

$$R_k(f;a,h) = \sum_{|\beta|=k+1} \frac{D^\beta f(a+ch)}{\beta!} h^\beta$$

עבור איזשהו $0 < c < 1$.

מקובל גם לרשום (בסימונים מעט שונים):

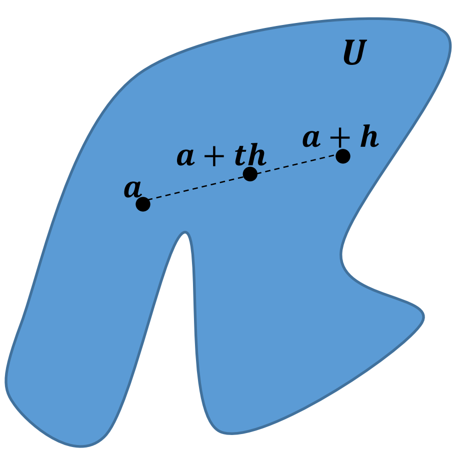
$$f(x) = \sum_{|\alpha| \le k} \frac{D^\alpha f(a)}{\alpha!} (x-a)^\alpha + R_k(f;a,x)$$

ו-

$$R_k(f;a,x) = \sum_{|\beta|=k+1} \frac{D^\beta f(u)}{\beta!} (x-a)^\beta$$

עבור נקודה u הנמצאת על הישר בין a ו- x .

<u>הערה:</u> באיור 2 המופיע כאן מדגים שאפשר להפעיל את משפט טיילור כל עוד הקו הישר בין a ו - $a+h$ מוכל ב- U.



איור 2

<u>הוכחת המשפט:</u>

נגדיר את הפונקציה $g(t) = f(a+th)$.

כך ש-$g: I \to \mathbb{R}$, וזאת כאשר I קטע המכיל את $[0,1]$ (למשל קטע פתוח מהצורה $(-\varepsilon, 1+\varepsilon)$ עבור $\varepsilon > 0$ קטן כרצוננו).

5.1.1 טענה – $g \in C^{k+1}(I)$ **ומתקיים לכל** $0 \le m \le k+1$:

$$\frac{d^m}{dt^m} g(t) = \sum_{|\alpha|=m} \frac{m!}{\alpha!} h^\alpha D^\alpha f(a+th)$$

<u>הוכחה:</u>

נראה זאת באינדוקציה –

בסיס האינדוקציה עבור $m=0$ טריוויאלי. עבור $m=1$ נראה כי מתקיים[3]:

[3] זאת בהתאם לכלל השרשרת עבור ההרכבה $f \circ \psi$ כאשר $\psi: I \mapsto \mathbb{R}^n$ המוגדרת על ידי $g = f \circ \psi$, כלומר $\psi(t) = a + th$.

$$\frac{d}{dt}f(a+th) = Df(a+ht)\cdot h = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a+th)h_i = \left(\sum_{i=1}^{n}h_iD_i\right)f(a+th)$$

את הביטוי $\left(\sum_{i=1}^{n}h_iD_i\right)f(a+th)$ אפשר להבין פשוט כקיצור לביטוי שמופיע משמאלו. אפשר גם לחשוב על $\sum_{i=1}^{n}h_iD_i$ כעל אופרטור לינארי שפועל על מרחב הפונקציות הגזירות m פעמים ברציפות למרחב הפונקציות הגזירות $m-1$ פעמים ברציפות. שימו לב שבעצם הוכחנו לכל פונקציה גזירה $\psi: U \to \mathbb{R}$, שמתקיים

$$(*)\quad \frac{d}{dt}\phi(a+th) = D\phi(a+ht)\cdot h = \sum_{i=1}^{n}\frac{\partial \phi}{\partial x_i}(a+th)h_i = \left(\sum_{i=1}^{n}h_iD_i\right)\phi(a+th)$$

כעת, נחשב (תוך שימוש בלינאריות של האופרטור $\frac{d}{dt}$) :

$$\frac{d^2}{dt^2}g(t) = \frac{d}{dt}\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a+th)h_i = \sum_{i=1}^{n}h_i\frac{d}{dt}\frac{\partial f}{\partial x_i}(a+th)$$

לכן אם נחליף את ϕ ב - $\frac{\partial f}{\partial x_i}(a+th)$ במשוואה $(*)$, נקבל

$$\frac{d^2}{dt^2}g(t) = \sum_{i=1}^{n}h_i\frac{d}{dt}\frac{\partial f}{\partial x_i}(a+th) = \sum_{i=1}^{n}h_i\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_j\partial x_i}(a+th)h_j = \sum_{j=1}^{n}h_j\left(\frac{\partial}{\partial x_j}\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}h_i\right)\Bigg|_{x=a+th}$$

נוכל לכתוב בקיצור:

$$\frac{d^2}{dt^2}g(t) = \left(\sum_{j=1}^{n}h_jD_j\right)\left(\sum_{i=1}^{n}h_iD_i\right)f(a+th) = \left[\left(\sum_{j=1}^{n}h_jD_j\right)\left(\sum_{i=1}^{n}h_iD_i\right)f\right]_{x=a+th}$$

כיון ש - $\frac{\partial^2 f}{\partial x_j\partial x_i} = \frac{\partial^2 f}{\partial x_i\partial x_j}$, מתקיים

$$\frac{d^2}{dt^2}g(t) = \left(\sum_{j=1}^{n}h_jD_j\right)^2 f(a+th) = \left(\sum_{|\alpha|=2}\frac{2!}{\alpha!}h^\alpha D^\alpha\right)f(a+th)$$

ובאופן כללי מתקיים[4]:

$$\frac{d^m}{dt^m}g(t) = \left(\sum_{i=1}^{n}h_iD_i\right)^m f(a+th) = \left(\sum_{|\alpha|=m}\frac{m!}{\alpha!}h^\alpha D^\alpha\right)f(a+th) = \sum_{|\alpha|=m}\frac{m!}{\alpha!}h^\alpha D^\alpha f(a+th)$$

עד כאן הוכחנו את הטענה. (ניתן היה גם להוכיח באינדוקציה בלי להגדיר את האופרטורים $\sum_{i=1}^{n}h_iD_i$. אם לא נוח לכם לעבוד עם אופרטורים, נסו לכתוב את ההוכחה הנ"ל בלעדיהם. שיהיה בכיף!).

---

[4] בהרכבה m פעמים של האופרטור נשים לב כי בכל שלב מדובר על אופרטור ממרחב הפונקציות הגזירות ברציפות m פעמים למרחב של פונקציות גזירות ברציפות $m-1$ פעמים, כלומר למרחב גדול יותר. לא מדובר באותו אופרטור. הסיבה לכתיב הנ"ל היא מכך שבגזירות הפונקציה ברציפות בכל שלב נסיק כי כל הנגזרות החלקיות מתחלפות בסדר ולכן ניתן לקבל את הנוסחה מקומבינטוריקה מתחשיב מולטינומים.

עתה, נשים לב, כי ממשפט טיילור עם שארית לגרנז' (בפיתוח סביב הראשית), מתקיים:

$$g(1)=\sum_{m\le k}\frac{g^{(m)}(0)}{m!}1^m+\frac{1}{(k+1)!}g^{(k+1)}(c)1^{m+1}$$

עבור נקודה $0<c<1$ כלשהי.

אם נציב ביטויים אלה שקיבלנו, נקבל כי מתקיים:

$$f(a+h)=g(1)=\sum_{m=0}^{k}\frac{1}{m!}\sum_{|\alpha|=m}\frac{m!}{\alpha!}D^\alpha h^\alpha f(a)+\frac{1}{(k+1)!}\sum_{|\alpha|=k+1}\frac{(k+1)!}{\alpha!}D^\alpha f(a+ch)h^\alpha$$

ולאחר צמצום נקבל את הדרוש. מש"ל.

<u>מסקנה מידית</u> – תהי $U\subseteq\mathbb{R}^n$ פתוחה, וכן תהא $f\in C^{k+1}(U)$, אזי השגיאה היא $o\left(\|h\|_2^k\right)$ בפיתוח טיילור מסדר k.

<u>הוכחה:</u> נניח כי בסביבה של a מתקיים $|D^\alpha f(x)|\le M$ כך ש-$|\alpha|=k+1$. אזי מתקיים:

$$|R_k(f;a,h)|\le\sum\frac{M}{\alpha!}|h^\alpha|=\frac{M}{(k+1)!}(|h_1|+\cdots+|h_n|)^{k+1}\le\frac{M}{(k+1)!}n^{\frac{k+1}{2}}\|h\|_2^{k+1}=o\left(\|h\|_2^k\right)$$

מש"ל.

למעשה, ניתן לקבל מסקנה חדה יותר:

5.2 מסקנה (משפט טיילור) – תהי $U\subseteq\mathbb{R}^n$ פתוחה, וכן תהא $f\in C^k(U)$, אזי מתקיים:

$$f(x)=\sum_{|\alpha|\le k}\frac{D^\alpha f(a)}{\alpha!}(x-\alpha)^\alpha+o(\|x-\alpha\|^k)$$

כמו כן, פולינום טיילור הוא הפולינום היחיד P ממעלה קטנה או שווה ל-k המקיים

$.f(x)-P(x)=o(\|x-a\|^k)$

מסקנה זו שונה מעט מה-"מסקנה המידית" ממשפט טיילור לעיל, שכן היא דורשת גזירות k פעמים בלבד ולאו דווקא גזירות $k+1$ פעמים.

<u>הוכחה:</u> נשתמש במשפט 5.1 עם $k-1$ במקום k , ונקבל

$$f(x)=\sum_{|\alpha|\le k-1}\frac{D^\alpha f(a)}{\alpha!}(x-a)^\alpha+\sum_{\alpha=k}\frac{D^\alpha f(a+ch)}{\alpha!}(x-a)^\alpha$$

$$=\sum_{|\alpha|\le k}\frac{D^\alpha f(a)}{\alpha!}(x-a)^\alpha+\underbrace{\sum_{|\alpha|=k}\frac{1}{\alpha!}[D^\alpha f(a+ch)-D^\alpha f(a)](x-\alpha)^\alpha}_{\star}$$

נרצה להוכיח כי הביטוי $\star$ אכן מתפקד כ-$o\left(\|x-a\|^k\right)$. כלומר, נרצה להראות כי מתקיים:

$$\frac{[D^\alpha f(a+ch)-D^\alpha f(a)](x-\alpha)^\alpha}{\|x-a\|^k}\xrightarrow{x\to a}0$$

היות והפונקציה f גזירה ברציפות k פעמים נקבל כי לכל α מתקיים:

$$D^\alpha f(a+ch) - D^\alpha f(a) \xrightarrow{h\to a} 0$$

וכן נשים לב שמתקיים:

$$\left|\frac{(x-a)^\alpha}{\|x-a\|^k}\right| = \frac{|x_1-a_1|^{\alpha_1}\cdots|x_n-a_n|^{\alpha_n}}{\|x-a\|^k} \le \frac{\|x-a\|^{\alpha_1}\cdots\|x-a\|^{\alpha_n}}{\|x-a\|^k} = 1$$

ולכן בפרט:

$$\left|\frac{[D^\alpha f(a+ch) - D^\alpha f(a)](x-\alpha)^\alpha}{\|x-a\|^k}\right| \le \left|\frac{[D^\alpha f(a+ch) - D^\alpha f(a)]\|x-a\|^k}{\|x-a\|^k}\right| \to 0$$

וכך נקבל את הדרוש.

את היחידות מוכיחים בנקל ע"י התרגיל שמופיע בסוף פרק 4. מש"ל.

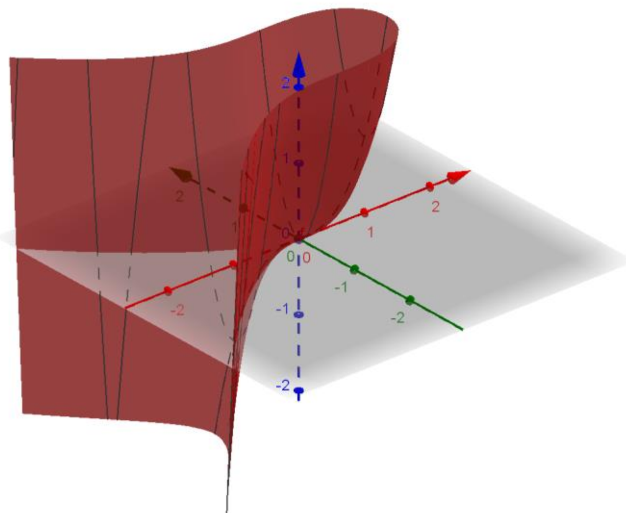
<u>שאלת העשרה:</u> משפט טיילור במשתנה אחד (משפט 5.2) נכון עבור פונקציות גזירות k פעמים בנקודה, ולא דורשים גזירות ברציפות. כמו כן, נוסחת טיילור עם שארית בצורת לגרנג' (משפט 5.1) במקרה החד-ממדי דורשת רק גזירות $k+1$ פעמים, ולא גזירות ברציפות $k+1$ פעמים. איפה השתמשנו בגזירות ברציפות? האם ניתן היה להימנע מכך?

פרק 6

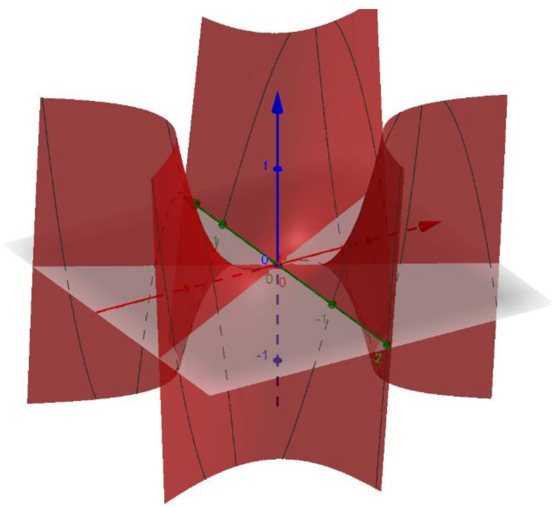
סיווג נקודות קריטיות:

6.1 הגדרה – נניח ש: $f: U \to \mathbb{R}$ ו-$U \subseteq \mathbb{R}^n$ פתוחה, $a \in U$.

1. בהינתן אם קיים $r > 0$ עבורו כך ש-$f(a) \leq f(x)$ לכל $x \in B_r(a)$ או לחילופין $f(a) \geq f(x)$, אזי a נקראת **נקודת מינימום / מקסימום מקומית.**
2. אם האי שוויון חזק, אזי נאמר כי זו **נקודת מינימום / מקסימום חזקה.**
3. אם f גזירה ב - a ומתקיים $\nabla f(a) = 0$, אזי a נקראת **נק' קריטית** של f.
4. אם a נק' קריטית של f , אזי a נקראת **נקודת אוכף** אם לכל $r > 0$ קיימים $x, y \in B_r(a)$ כך ש-$f(x) < f(a) < f(y)$.



איור 4 – הפונקציה $f(x,y) = x^3 + 5y^2$ בעלת נקודת אוכף ב-$(0,0)$.



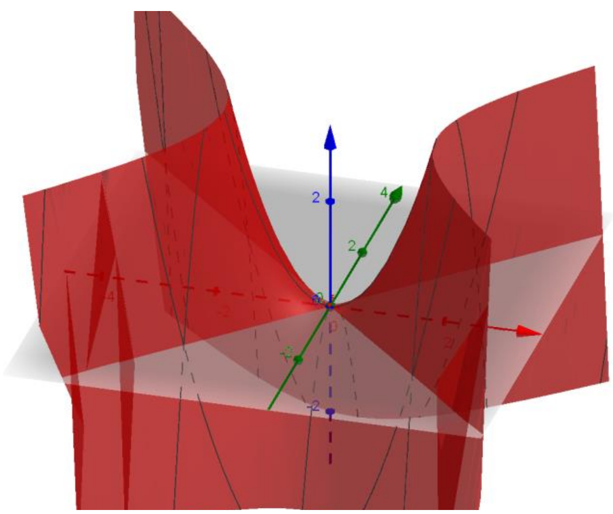
איור 2 – הפונקציה $f(x,y) = 3xy^2 - x^3$ כאשר $(0,0)$ הינה נקודת אוכף שלה 3 כיווני ירידה ו3 כיווני עליה ("**אוכף קופים**")

6.2 הגדרה – תהא $f \in C^2(U)$. אזי **ההסיאן** (Hessian) של f בנקודה a, הוא המטריצה:

$$H_f = H_f(a) = \left[\frac{\partial^2 f(a)}{\partial x_i \partial x_j}\right]_{i,j=1}^{n}$$

זו מטריצה סימטרית והיא מגדירה תבנית ריבועית על ידי:

$$\forall v \in \mathbb{R}^n$$
$$\langle H_f(a)v, v\rangle = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) v_i v_j$$



איור 3 – הפונקציה $f(x,y) = x^2 - y^2$ כאשר $(0,0)$ הינה נקודת אוכף שלה 2 כיווני ירידה ברורים ו-2 כיווני עליה

נשים לב, כי בעבור פונקציה כנ"ל, מתקיים:

$$f(a+h)=f(a)+\sum_{i=1}^{n}D_if(a)h_i+\sum_{|\alpha|=2}\frac{D^\alpha f(a)}{\alpha!}h^\alpha+o(\|h\|^2)$$

כאשר, כאמור, מתקיים:

$$|\alpha|=2\Rightarrow \alpha_1+\cdots+\alpha_n=2$$

כלומר בכל אחד מהמקרים האפשריים, או שאחד מהאינדקסים הללו הינו 2 והיתר 0, או שיש שני מקדמים שערכם 1 והיתר 0. לכן ניתן לרשום:

$$\sum_{|\alpha|=2}\frac{D^\alpha f(a)}{\alpha!}h^\alpha=\sum_{i=1}^{n}\frac{1}{2}\frac{\partial^2 f}{\partial x_i^2}(a)h_i^2+\sum_{1\le i<j\le n}\frac{\partial f}{\partial x_i\partial x_j}(a)h_ih_j$$

$$f(a+h)=f(a)+\langle\nabla f(a),h\rangle+\frac{1}{2}\langle H_f(a)h,h\rangle+o(\|h\|^2)$$

<u>תרגיל</u>: להראות את השוויון הנ"ל בצורה מדוקדקת.

עתה, נניח לרגע ש- $n=2$. המטריצה $H_f(a)$ סימטרית, ולכן היא לכסינה אורתוגונלית, וכך נקבל כי היא דומה למטריצה מהצורה $\begin{bmatrix}\lambda & 0\\ 0 & \mu\end{bmatrix}$. נניח כי a נקודה קריטית. אם נעבור למערכת קואורדינטות בה הבסיס הסטנדרטי מורכב מוקטורים עצמיים של ההסיאן, אז נשים לב כי:

$$f\big(a+(x,y)\big)-f(a)\approx\frac{1}{2}\langle H_f(a)(x,y),(x,y)\rangle=\lambda x^2+\mu y^2$$

ניתן לראות כי מיחסי λ,μ מבחינת חיוביות / שליליות, נוכל לזהות נקודות מקסימום ומינימום. כלומר:

6.3 הגדרה – תהא $A\in M_n(\mathbb{R})$ מטריצה סימטרית. אזי:

א. A נקראת **(מוגדרת) אי שלילית** אם לכל $v\in\mathbb{R}^n$ מתקיים $\langle Av,v\rangle\ge 0$, ומסמנים $A\ge 0$.
ב. A נקראת **(מוגדרת) חיובית** אם אי השוויון של א' חזק לכל $0\neq v\in\mathbb{R}^n$. מסמנים $A>0$.
ג. A נקראת **(מוגדרת) אי חיובית** אם לכל $v\in\mathbb{R}^n$ מתקיים $\langle Av,v\rangle\le 0$ ומסמנים $A\le 0$.
ד. A נקראת **(מוגדרת) שלילית** אם אי השוויון של ב' חזק לכל $0\neq v\in\mathbb{R}^n$. מסמנים $A<0$.
ה. A נקראת **לא מוגדרת** אם קיימים $u,v\in\mathbb{R}^n$ עבורם $\langle Au,u\rangle<0<\langle Av,v\rangle$.

6.4 טענה – תהא $U\subseteq\mathbb{R}^n$ פתוחה ותהא $f\in C^2(U)$. אזי בהנתן $a\in U$ נקודה קריטית (כלומר נק' a בה מתקיים: $\nabla f(a)=0$)

א. אם $H_f(a)>0$ אזי a נקודת מינימום חזקה.
ב. אם $H_f(a)<0$ אזי a נקודת מקסימום חזקה.
ג. אם $H_f(a)$ לא מוגדרת אזי a נקודת אוכף.
ד. בשאר המקרים, כל האפשרויות יכולות לקרות (מקס' חזק, מקס חלש, אוכף וכו').

<u>הערה</u> – סעיפים א'-ג' לא מכסים כמובן את כל המקרים. במקרים שבהם $H_f(a)\le 0$ או $H_f(a)\ge 0$ ייתכן שהנקודה תהא מקסימום או מינימום חלש או נקודת אוכף, כפי שמצוין בסעיף ד'.

<u>הוכחה:</u>

תהא a נקודה קריטית, קרי נקודה שבה $\nabla f(a) = 0$. אזי מתקיים:

$$f(a+h) = f(a) + \frac{1}{2}\langle H_f(a)h, h\rangle + \varepsilon(h) \quad , \quad \frac{\varepsilon(h)}{\|h\|^2} \xrightarrow{h\to 0} 0$$

נניח כי $H_f(a) > 0$. אזי נקבל כי:

$$0 < c \coloneqq \min_{\|v\|=1}\langle H_f(a)v, v\rangle = \min_{\|v\|\neq 0}\frac{\langle H_f(a)v, v\rangle}{\|v\|^2}$$

אזי לכל h מתקיים: $\langle H_f(a)h, h\rangle \geq c\|h\|^2$. יהא $r > 0$ כך ש-$B_r(a) \subseteq U$, המקיים שעבור $0 < \|h\| < r$:

$$\frac{|\varepsilon(h)|}{\|h\|^2} < \frac{1}{2}c$$

ולכן $|\varepsilon(h)| < \frac{c}{2}\|h\|^2$. עבור h כזה מתקיים:

$$f(a+h) = f(a) + \frac{1}{2}\langle H_f(a)h, h\rangle + \varepsilon(h) \geq f(a) + \frac{c}{2}\|h\|^2 - |\varepsilon(h)| > f(a)$$

ולכן a נקודת מינימום חזקה, כנדרש. באופן דומה מראים שאם $H_f(a) < 0$, אזי a היא נקודת מקסימום חזקה.

עתה, נניח כי $H_f(a)$ אינה מוגדרת (במובן חיוביות/אי חיובית). יהיו $u, v \in \mathbb{R}^n$ וקטורי יחידה המקיימים

$$b := \langle H_f(a)u, u\rangle < 0$$
$$c := \langle H_f(a)v, v\rangle > 0$$

נבחר $r > 0$ כך שלכל h מתקיים:

$$0 < \|h\| < r \Longrightarrow \left|\frac{\varepsilon(h)}{\|h\|^2}\right| < \frac{1}{2}\min\{|b|, c\}$$

וכך, עבור כל $0 < t < r$ יתקיים:

$$f(a+tu) = f(a) + \frac{1}{2}\langle H_f(a)tu, tu\rangle + \varepsilon(tu) \leq f(a) + \frac{b}{2}t^2 + |\varepsilon(tu)| < f(a)$$

וזאת משום שכאמור, דרשנו $|\varepsilon(tu)| < \frac{1}{2}|b|t^2$. באופן דומה, עבור $f(a+tv) > f(a)$ ניתן להראות בצורה זהה.

מש"ל.

<u>לתרגול/תרגיל בית:</u> תרגול סיווג נק' קריטיות, חקירת אוכפים עם מספר שונה של "כיווני עליה וירידה", נק' קריטיות בהן ההסיאן מתאפס.

<u>פרק 7</u>

7.1 הגדרה - $C \subset \mathbb{R}^n$ נקראת **קמורה** אם מתקיים:

$$\forall x, y \in C \quad \forall t \in [0,1] \quad tx + (1-t)y = y + t(x-y) \in C$$

7.2 טענה – תהא $U \subseteq \mathbb{R}^n$ קבוצה קמורה ופתוחה ונניח $f: U \to \mathbb{R}^m$ גזירה. נניח כי מתקיים $\sup\limits_{x \in U} \|f'(x)\| \le M \le \infty$, אזי לכל $x, y \in U$ מתקיים:

$$\|f(x) - f(y)\| \le M\|x - y\|$$

<u>הוכחה:</u>

נסמן $v = \frac{f(x)-f(y)}{\|f(x)-f(y)\|}$ (ניתן לעשות זאת שכן אם המכנה מתאפס נקבל $f(x) = f(y)$ וסיימנו). אזי מתקיים:

$$\|f(x) - f(y)\| = \langle v, f(x) - f(y) \rangle$$

נגדיר, עתה:

$$g: \mathbb{R}^m \to \mathbb{R} \quad g(z) = \langle v, z \rangle$$

ונגדיר:

$$h: [0,1] \to U \quad h(t) = y + t(x-y)$$

נשים לב, כי מתקיים:

$$g' = v \quad h' = (x-y)$$

עתה נגדיר:

$$\varphi = g \circ f \circ h$$

פונקציה זו גזירה כהרכבה של גזירות בתחומים המתאימים, ולכן לפי משפט לגרנז', נקבל כי קיימת נקודה $c \in (0,1)$ כך שמתקיים:

$$\varphi'(c) = \frac{\varphi(1) - \varphi(0)}{1 - 0} = \varphi(1) - \varphi(0) = \langle v, f(x) \rangle - \langle v, f(y) \rangle = \|f(x) - f(y)\|$$

מצד שני, מתקיים:

$$|\varphi'(c)| = \left\| g'\big(f\big(y + c(x-y)\big)\big) f'\big(y + c(x-y)\big) h'(c) \right\| \le \|v\| \cdot M \cdot \|x - y\|$$

כיון ש- $\|v\| = 1$ נקבל כי:

$$\|f(x) - f(y)\| \le M\|x - y\|$$

מ.ש.ל.

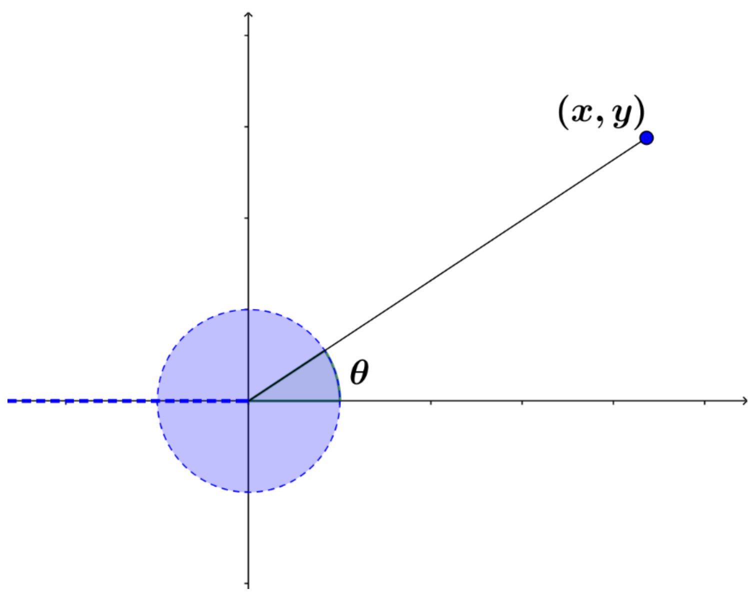
<u>תרגיל:</u> הראו שהמשפט נשאר נכון אם משתמשים בנורמה $\|\cdot\|_\infty$ במקום $\|\cdot\|_2$. שימו לב שאז גם צריך לחשב את הנורמה האופרטורית של הנגזרת ביחס לנורמה $\|\cdot\|_\infty$, כלומר מגדירים $\|T\|_{\infty,op} = \sup\limits_{x \neq 0} \frac{\|Tx\|_\infty}{\|x\|_\infty}$ ואז ההנחה במשפט הופכת להיות $\sup\limits_{x \in U} \|f'(x)\|_{\infty,op} \le M \le \infty$. (הדרכה: במקום הוקטור v בהוכחה, מצאו וקטור v שמקיים $\|f(x) - f(y)\|_\infty = \langle v, f(x) - f(y) \rangle$)

7.3 מסקנה – אם $f \in C^1$ אז f ליפשיצית בכל תת קבוצת קומפקטית של תחום הגדרתה.

הוכחה: תרגיל.

<u>דוגמה:</u>

$$D = \left\{ (x,y) \in \mathbb{R}^2 \middle| \begin{matrix} x > 0 \text{ או } y \neq 0 \\ \text{וגם} \\ x^2 + y^2 > 1 \end{matrix} \right\}$$



איור 3 – תיאור של התחום D (הקו המקווקו והתחום הצבוע לא כלולים)

ונגדיר את הפונקציה הבאה:

$$f(x) = \begin{cases} \arctan\left(\frac{y}{x}\right) & x > 0 \\ \frac{\pi}{2} - \arctan\left(\frac{x}{y}\right) & y > 0 \\ -\frac{\pi}{2} - \arctan\left(\frac{x}{y}\right) & y < 0 \end{cases}$$

ונשים לב כי מתקיים עבור $x > 0$:

$$\|\nabla f(x,y)\|^2 = \left\| \left(\frac{\frac{-y}{x^2}}{1 + \left(\frac{y}{x}\right)^2}, \frac{\frac{1}{x}}{1 + \left(\frac{y}{x}\right)^2} \right) \right\|^2 = \frac{y^2}{(x^2+y^2)^2} + \frac{x^2}{(x^2+y^2)^2} = \frac{1}{\|(x,y)\|^2}$$

כלומר מתקיים:

$$\sup_D \|f'\| = 1$$

נראה כי:

$$\|f(x) - f(y)\| \nleq \sup_D \|f'\| \, \|x - y\|$$

על ידי בחירת:

$$x = \left(0, \frac{3}{2}\right) \quad y = \left(0, -\frac{3}{2}\right)$$

ונקבל כי:

$$\|f(x) - f(y)\| = \pi \nleq 3 = \|x - y\|$$

ניתן להראות גם שבקבוצה החסומה $D \cap B_R$ (עבור R כלשהו גדול), הפונקציה f כלל איננה ליפשיצית (צריך להסתכל על זוגות של נקודות מהצורה $(-2, \pm\varepsilon)$.)

כעת אנו מגיעים למשפט החשוב ביותר בפרק זה, למעשה אחד החשובים בקורס: **משפט הפונקציה ההפוכה**.

7.4 משפט הפונקציה ההפוכה – תהא $U \subseteq \mathbb{R}^n$ פתוחה, $f \in C^1(U, \mathbb{R}^n)$ וכן תהא $a \in U$. אם $f'(a)$ הפיכה, אזי קיימת קבוצה פתוחה $a \in V \subseteq U$, וקבוצה פתוחה $f(a) \in W$ כך ש- $f: V \to W$ **היא חד-חד ערכית ועל** ו- $f^{-1}: W \to V$ **גם כן גזירה, ומתקיים:**

$$(f^{-1})'(f(x)) = [f'(x)]^{-1} \quad \text{או} \quad (f^{-1})'(y) = \left[f'\left(f^{-1}(y)\right)\right]^{-1}$$

__דוגמה:__

נתבונן בפונקציה המוגדרת על ידי:

$$f(x,y) = (e^x \cos y\,, e^x \sin y)$$

נשים לב כי:

$$J_f(x,y) = \begin{bmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{bmatrix}$$

מתקיים:

$$\det(J_f(x,y)) = e^{2x}(\cos^2 y + \sin^2 y) = e^{2x} \neq 0$$

הפונקציה f איננה הפיכה, כי איננה חד-חד ערכית. אך לפי המשפט, סביב כל נקודה בתחום הגדרתי קיימת סביבה כלשהי בה היא חד-חד ערכית, וניתן להגדיר לה פונקציה הפוכה על התמונה של סביבה זו. **שלא כמו במקרה החד מימדי, אי אפשר להסיק על הפיכות גלובלית באמצעות תנאי על הנגזרת בלבד.**

ניתן לחשוב על f כעל ההעתקה $z \mapsto e^z$ במישור המרוכב $\mathbb{C} \simeq \mathbb{R}^2$, ולפונקציה זו יש הפיך $w \mapsto \log w$ (צריך להגדיר ענף מתאים), אבל הפיך זה אינו הפיך גלובלי.

__**הוכחת המשפט:**__

__**שלב I – מציאת כדור סגור $a \in \overline{B_r(a)} =: \bar{B} \subseteq U$ בו f חד-חד ערכית:**__

ההעתקה $x \mapsto f'(x)$ רציפה ביחס לנורמת האופרטורים. אכן, ההעתקה:

$$U \ni x \mapsto \left[\frac{\partial f_i}{\partial x_j}(x)\right]_{i,j=1}^{n}$$

הינה העתקה רציפה, מהנחת גזירות ברציפות. העתקה זו מקיימת (לפי טענה 2.3 בפרק מס' 2):

$$\|f'(x) - f'(y)\|_{op} \leq \left(\sum_{i,j}\left(\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(y)\right)^2\right)^{\frac{1}{2}}$$

ולכן, ההעתקה $x \mapsto f'(x)$ אכן רציפה ביחס לנורמת האופרטורים.

נסמן $A = f'(a)$. בעצם ההנחה שלנו היא ש- A הפיכה, כלומר $\det A \neq 0$. כיון ש- $f \in C^1$, קיים $R > 0$ כך שמתקיים:

$$(*) \quad \forall x \in B_R(a) \subseteq U. \quad \|f'(x) - f'(a)\|_{op} < \frac{\|A^{-1}\|^{-1}}{2}$$

נבחר איזשהו $0 < r < R$ ונגדיר $\bar{B} := \overline{B_r(a)}$. מתקיים גם:

$$\forall x \in \bar{B}. \quad \|f'(x) - f'(a)\|_{op} < \frac{\|A^{-1}\|^{-1}}{2}$$

בשלב זה, אנו רוצים להוכיח שבסביבה $\bar{B}$ הפונקציה f הינה חד-חד ערכית. לשם כך, נראה שקיים קבוע חיובי C כך שמתקיים לכל $x, y \in \bar{B} := \overline{B_r(a)}$:

$$\|f(x) - f(y)\| \geq C\|x - y\|$$

נסביר קודם כל את הרעיון. היות ו-f גזירה, ישנו קירוב ליניארי מהצורה:

$$f(x) \approx f(a) + A(x - a)$$

כלומר, ניתן לומר כי:

$$f(x) - f(y) \approx f(a) + A(x - a) - f(a) - A(y - a)$$

אנו יודעים כי A חסומה מלמטה: $\|A(x - y)\| \geq \|A^{-1}\|^{-1}\|x - y\|$, ולכן, ננסה לחסום את $\|f(x) - f(y)\|$ מלמטה בעזרת A :

$$\begin{aligned}\|f(x) - f(y)\| &= \|A(x - y) - (f(y) + Ay - f(x) - Ax)\| \\ &\geq \|A(x - y)\| - \|f(x) - Ax - (f(y) - Ay)\| \overset{?}{\geq} \|A^{-1}\|^{-1}\|x - y\| - ????\end{aligned}$$

כדי למצוא חסם לביטוי $\|f(x) - Ax - (f(y) - Ay)\|$, נגדיר פונקציית עזר:

$$g(x) := f(x) - Ax$$

מתקיים:

$$g'(x) = f'(x) - A$$

ולכן, לפי טענה 7.2, מתקיים לכל $x, y \in \bar{B}$:

$$\|f(x) - Ax + (f(y) - Ay)\| = \|g(x) - g(y)\| \leq \sup_{z \in B_R(a)} \|g'(z)\| \cdot \|x - y\| \leq \frac{\|A^{-1}\|^{-1}}{2}\|x - y\|$$

כיון שלפי $(*)$ בחרנו את R קטן מספיק כך שמתקיים אי השוויון $\sup_{z \in B_R(a)} \|g'(z)\| \leq \frac{1}{2}\|A^{-1}\|^{-1}$. ולכן אם נחזור אל הביטוי שניסינו להעריך לפני הגדרת פונקציית העזר, נקבל לכל $x, y \in \bar{B}$:

$$\begin{aligned}\|f(x) - f(y)\| &= \big\|A(x - y) - \big(Ax - f(x) - Ay + f(y)\big)\big\| \\ &\geq \|A(x - y)\| - \|f(x) - Ax - (f(y) - Ay)\| \geq \frac{1}{2}\|A^{-1}\|^{-1}\|x - y\|\end{aligned}$$

כלומר, לכל $x, y \in \bar{B}$ מתקיים $\|f(x) - f(y)\| \geq \frac{\|A^{-1}\|^{-1}}{2}\|x - y\|$ ולכן בפרט $f: \bar{B} \to f(\bar{B})$ חד-חד ערכית.

<u>שלב II – מציאת סביבות פתוחות $a \in V \subseteq B$ ו - $f(a) \in W \subseteq f(B)$ כך ש $f: V \to W$ חד-חד ערכית ועל:</u>

השלב הזה הוא קצת טריקי ושווה לעצור לרגע ולחשוב למה צריך אותו בכלל. המשפט שאנו מנסים להוכיח כולל את העובדה שניתן למצוא סביבות **פתוחות** של a ו - $f(a)$ כך ש- f העתקה הפיכה ביניהן. הכדור הפתוח B זו סביבה פתוחה של a, והצמצום של f ל- B היא חד-חד ערכית, וכמובן שהיא **על** התמונה שלה $f(B)$. הבעיה היא, שאיננו יודעים ש - $f(B)$

פתוחה. למעשה, הקבוצה $f(B)$ היא אכן פתוחה – אנו נקבל את זה כתוצאה מהמשפט! אבל לעת עתה, כדי להיות בטוחים שאנו לוקחים סביבה פתוחה כלשהיא של a לסביבה פתוחה של $f(a)$, עלינו להקטין עוד יותר את הסביבה של a עליה אנו מדברים, ולהשתמש בשיקולים הבאים.

נסמן $\partial B = \{y \mid \|y-a\| = r\}$ (השפה של הכדור $\bar{B}$). נסמן קבוצה חדשה על ידי $K = f(\partial B)$. הראינו כי f חד-חד ערכית ולכן נסיק כי $f(a) \notin K$ (שכן אחרת היינו מקבלים איבר נוסף ב - $\bar{B}$ שמקבל אותו ערך כמו a בסתירה לחד-חד ערכיות). מכאן שמתקיים $\delta \coloneqq \min\{\|f(x)-f(a)\| : x \in \partial B\} > 0$, הרי מדובר במינימום של פונקציה רציפה וחיובית על קבוצה קומפקטית (הפונקציה הרציפה היא $x \mapsto \|f(x)-f(a)\|$ המוגדרת על הקבוצה הקומפקטית ∂B.
אנו יודעים שמינימום על קבוצה קומפקטית מתקבל לפי משפט 1.27). במילים אחרות, לכל $x \in \partial B$ מתקיים:

$$\|f(x)-f(a)\| \geq \delta$$

כעת נגדיר:

$$W = B_{\frac{\delta}{2}}\big(f(a)\big)$$

כלומר ראינו שתמונת השפה ∂B בהכרח לא מועתקת ל-W על-ידי f.

מטרתנו כעת להראות, כי לכל $y \in W$ קיים $x \in B$ כך שמתקיים $f(x) = y$. אחרי שנשיג מטרה זו, נוכל לסמן:

$$V = f^{-1}(W) \cap B$$

והסביבות V, W תקיימנה את מסקנת המשפט.

נקבע $y \in W$ ונגדיר:

$$\varphi : \bar{B} \to \mathbb{R} \quad \varphi(x) = \|f(x)-y\|^2$$

אנו יודעים כי φ רציפה ב-$\bar{B}$ ולכן מקבלת בו מינימום.

אם $x \in \partial B$ אז מתקיים:

$$\|f(x)-y\| \geq \|f(a)-f(x)\| - \|f(a)-y\| > \delta - \frac{\delta}{2} = \frac{\delta}{2} > 0$$

מכאן נובע $\left(\frac{\delta}{2}\right)^2 < \varphi(x)$ לכל $x \in \partial B$. מצד שני, $\varphi(a) < \left(\frac{\delta}{2}\right)^2$, כך שהמינימום ב-$\bar{B}$ לא מתקבל על השפה, ולכן הוא מתקבל בכדור הפתוח B. דרך נוספת לכתוב את φ היא:

$$\varphi(x) = \sum_{i=1}^{n} (f_i(x) - y_i)^2$$

אם המינימום מתקבל בנקודה פנימית $z \in B$ אזי בהכרח מתקיים:

$$\nabla\varphi(x) = 0$$

לכן:

$$\forall 1 \leq j \leq n \quad \frac{\partial \varphi}{\partial x_j}(x) = \sum_{i=1}^{n} 2\big(y_i - f_i(x)\big)\frac{\partial f_i}{\partial x_j}(x) = 0$$

כלומר:

$$\nabla\varphi(x) = 2\big(y - f(x)\big)\cdot J_f(x) = 0$$

אבל $x \in \bar{B}$, לכן (*) מתחילת ההוכחה יחד עם טענה מהנספח לפרק 2 גוררים ש לכן $f'(x)$ הפיכה. לכן המטריצה המייצגת שלה $J_f(x)$ היא הפיכה. מכאן ש

$$2\big(y - f(x)\big)\cdot J_f(x) = 0 \Longrightarrow y - f(x) = 0$$

כלומר:

$$y = f(x)$$

כנדרש.

עד כאן מצאנו, אם כן, V, W פתוחות כך שהצמצום $f: V \to W$ הינה חד-חד ערכית ועל. בכך השלמנו חלק גדול מהוכחת משפט הפונקציה ההפוכה. נותר לנו להוכיח ש- $f^{-1}: W \to V$ היא גזירה, וכן להראות את הנוסחה לנגזרת.

<u>שלב III –הוכחה ש - f^{-1} ליפשיצית על W:</u>

עבור $f^{-1}: W \to V$ ראינו כי:

$$(**)\qquad \forall x, y \in V \subseteq \bar{B} \quad \|f(x) - f(y)\| \geq \frac{\|A^{-1}\|^{-1}}{2}\|x - y\|$$

ומכאן שמתקיים (אם ניקח $x = f^{-1}(u), y = f^{-1}(w)$):

$$\forall u, w \in W \quad \|f^{-1}(u) - f^{-1}(w)\| \leq 2\|A^{-1}\|\|u - w\|$$

כלומר f^{-1} ליפשיצית (ובפרט, רציפה).

<u>שלב IV – הוכחה ש - f^{-1} גזירה ב $f(a)$ וחישוב הנגזרת:</u>

עתה, יהיו $a, x \in V$ ונסמן $b = f(a), w = f(x) \in W$. אנו כעת נוכיח גזירות, וכן את הנוסחה עבור הנגשרת, לפי ההגדרה. עלינו להראות

$$f^{-1}(w) - \Big[f^{-1}(b) + \big(f'(a)\big)^{-1}(w - b)\Big] = o(w - b)$$

מצד אחד, מתקיים:

$$f^{-1}(w) - f^{-1}(b) = f^{-1}\big(f(x)\big) - f^{-1}\big(f(a)\big) = x - a$$

אך מצד שני ניתן לכתוב:

$$\big(f'(a)\big)^{-1}(w - b) = \big(f'(a)\big)^{-1}\big(f'(a)(x - a) + \varepsilon(x - a)\big)$$

כאשר השתמשנו בקירוב הליניארי של f :

$$w - b = f(x) - f(a) = f'(a)(x - a) + \varepsilon(x - a)$$

נשתמש במשוואות לעיל ונקבל:

$$\begin{aligned} f^{-1}(w) - \Big[f^{-1}(b) + \big(f'(a)\big)^{-1}(w - b)\Big] &= x - a - (x - a) - \big(f'(a)\big)^{-1}\mathcal{E}(x - a) \\ &= -\big(f'(a)\big)^{-1}\mathcal{E}(x - a) \end{aligned}$$

נותר להראות שמתקיים:

$$\lim_{\|w-b\|\to 0}\frac{\left(f'(a)\right)^{-1}\varepsilon(x-a)}{\|w-b\|}=0$$

נזכור כי $\left(f'(a)\right)^{-1}$ רציפה וליניארית וממילא חסומה. לכן מספיק להראות

$$\lim_{\|w-b\|\to 0}\frac{\varepsilon(x-a)}{\|w-b\|}=0$$

אבל לפי אי השוויון (**) למעלה

$$\|w-b\|=\|f(x)-f(a)\|\geq\frac{\|A^{-1}\|^{-1}}{2}\|x-a\|$$

ולכן

$$\frac{\|\varepsilon(x-a)\|}{\|w-b\|}=\frac{\|\varepsilon(x-a)\|}{\|x-a\|}\frac{\|x-a\|}{\|w-b\|}\leq 2\|A^{-1}\|\frac{\|\varepsilon(x-a)\|}{\|x-a\|}\to 0$$

כאשר נזכור כי מותר לחלק ב-$\|x-a\|$, כיון ש- f וגם f^{-1} הפיכות (בפרט חד-חד ערכיות) בסביבה הזאת ולכן לא יתכן $\|x-a\|=0$ כאשר $\|w-b\|\to 0\neq 0$. כמו כן, נשתמש בעובדה שאם $w\to b$ אז מתקיים שיש גם התכנסות $x=f^{-1}(w)\to a=f^{-1}(b)$, כיון ש- f^{-1} רציפה.

<u>שלב V – סיום ההוכחה: גזירות ברציפות ב – W:</u>

לסיום, נשים לב שהוכחנו ש – f^{-1} גזירה בנקודה $b=f(a)$, אבל אותו שיקול נכון בכל נקודה $w=f(x)\in W$, כיון ש - $f'(x)$ הפיכה בכל $x\in V$, ו- V הינה קבוצה פתוחה. לכן f^{-1} גזירה ב – W, ומתקיים

$$(f^{-1})'(w)=\left[f'\left(f^{-1}(w)\right)\right]^{-1}$$

מהנוסחה הזאת נובע ש- f^{-1} גזירה ברציפות. הרי אגף ימין מתקבל מהרכבת הפונקציות הגזירות f^{-1}, f', וכן ההעתקה $L\mapsto L^{-1}$, שהיא רציפה כפי שראינו בסוף פרק 2. מש"ל.

<u>נספח לפרק 7 – שיטת ניוטון-רפסון לפתרון משוואות</u>

נניח כי נתונות n-פונקציות, נסמנן $f_1,\cdots,f_n:\mathbb{R}^n\to\mathbb{R}$ אשר כולן גזירות ברציפות בתחום, ונניח כי ברצוננו לפתור מערכת משוואות מהצורה:

$$(*)\quad\begin{cases}f_1(x_1,x_2,\cdots,x_n)=0\\ \quad\vdots\\ f_n(x_1,x_2,\cdots,x_n)=0\end{cases}$$

דוגמה (GPS מישורי): נניח שיש שתי אנטנות במישור בנקודות שהמיקום שלהן ידוע $a=(a_1,a_2)$ ו- $b=(b_1,b_2)$, ונניח שאנו רוצים לחשב את מיקומו של מכשיר קליטה נייד כלשהו הנע על המישור ונמצא בנקודה $x=(x_1,x_2)$. כמו כן נניח שיש לנו דרך לחשב את המרחקים r_a ו - r_b של הנקודה x מהנקודות a ו – b, בהתאמה (למשל על-ידי שימוש בשעון מדויק מאוד ובדיקת הזמן בו מגיע אות מהאנטנות למכשיר הנייד). אנו נוכל למצוא את המיקום של המכשיר הנייד אם נצליח לפתור את מערכת המשוואות הבאה:

$$f_1(x) = (x_1 - a_1)^2 + (x_2 - a_2)^2 - r_a^2 = 0$$
$$f_2(x) = (x_1 - b_1)^2 + (x_2 - b_2)^2 - r_b^2 = 0$$

זוהי מערכת של שתי משוואות **לא** לינאריות בשני נעלמים, ולא ברור איך לגשת לבעיה הזו. זוהי כמובן רק דוגמה לדרך בה מערכת משוואות כזו יכולה לצוץ, ולמעשה הרעיון די קרוב לאיך ש- GPS עובד. אנחנו נפתח שיטה כללית לפתרון מקורב של מערכות משוואות, ונחזור מדי פעם לדוגמה הזו כדי להמחיש את העניינים.

השיטה שנציג למציאת פתרון למערכת משוואות זו היא שיטה איטרטיבית. מתחילים בניחוש מועמד לפתרון $x^1 \in \mathbb{R}^n$. בעצם הרעיון שיש לנו ניחוש התחלתי לפתרון הוא מציאותי מאוד. למשל, בדוגמה של ה- GPS המישורי, אנו יכולים להניח שבכל כמה שניות אנו מחשבים את המיקום מחדש, וכיון שבכמה שניות לא יכולנו לזוז הרבה, נוכל להשתמש במיקום הקודם שחושב בתור ניחוש התחלתי לחישוב הנוכחי של המיקום.

אם כן, מתחילים בניחוש מועמד לפתרון $x^1 \in \mathbb{R}^n$. נגדיר קירוב לפתרון x^k בצורה רקורסיבית. לכל $k \geq 1$ נגדיר:

$$(**) \quad x^{k+1} = x^k - \left[f'\left(x^k\right)\right]^{-1} f\left(x^k\right)$$

(שימו לב: אם $n = 1$, נקבל:

$$\boxed{x^{k+1} = x^k - \frac{f\left(x^k\right)}{f'\left(x^k\right)}}$$

וקיבלנו את שיטת ניוטון-רפסון לפתרון איטרטיבי של משוואות במשתנה אחד, שאולי ראיתם בקורס אחר, למשל באינפי 1.)

משפט: נניח שקיים פתרון $\bar{x}$ עבור מערכת המשוואות $(*)$, כך ש- $f'(\bar{x})$ הפיכה. אזי קיים $r > 0$ כך שלכל ניחוש ראשוני $x^1 \in B_r(\bar{x})$, הסדרה המתקבלת $\left(x^k\right)_{k=1}^{\infty}$ על ידי האיטרציות $(**)$ מתכנסת ל-$\bar{x}$.

הסבר קצר: במילים אחרות, המשפט אומר שאם באמת קיים פתרון $\bar{x}$ המקיים את התנאים במשפט, ואם נגדיר סדרה באמצעות האיטרציות $(**)$, אזי הסדרה המתקבלת תתכנס לפתרון האמיתי $\bar{x}$ בתנאי שהניחוש הראשוני שלנו היה מספיק קרוב לפתרון של מערכת המשוואות. בחיים האמיתיים, מפעילים אלגוריתם כמו שיטת ניוטון רפסון עם איזשהו ניחוש התחלתי סביר (מבלי לנסות לבדוק האם הוא מתאים ל - r התיאורטי שקיים), ופשוט מריצים אותו עד ש - $\left\|x^k - x^{k+1}\right\|$ מספיק קטן, ואז פשוט לוקחים את x^{k+1} בתור קירוב נומרי לפתרון. אם - $\left\|x^k - x^{k+1}\right\|$ לא נהיה קטן לשום k –, אז בשלב כלשהו מתייאשים ומנסים לפתור את הבעיה בדרך אחרת.

אנחנו מניחים שני דברים על פתרון המערכת: ראשית אנו מניחים שקיים פתרון. מובן שאם לא קיים פתרון לא נוכל למצוא אותו, בשום שיטה! התנאי השני שאנו מניחים הוא שהפתרון הוא כזה , שהנגזרת של הפונקציה מחושבת בפתרון היא הפיכה. שימו לב שאם התנאי השני לא מתקיים, אז לכל סדרה x^k המתכנסת לפתרון, יתקיים שסדרת הנגזרות $\left[f'\left(x^k\right)\right]^{-1}$ תהיה בעלת נורמה שואפת לאינסוף (מדוע?) ולכן אי אפשר לצפות מהשיטה לעבוד במקרה זה.

בדוגמה של ה- GPS המישורי, אנו יכולים לחשב את מטריצת היעקוביאן של f (המייצגת את הנגזרת בבסיס הסטנדרטי), והיא:

$$J_f(x) = \begin{pmatrix} 2(x_1 - a_1) & 2(x_2 - a_2) \\ 2(x_1 - b_1) & 2(x_2 - b_2) \end{pmatrix}$$

אפשר להראות (תרגיל) שהשורות של המטריצה הזאת תלויות לינארית אם ורק אם x נמצא על הישר המחבר בין a ו - b. וזה מצב שכמעט ולא קורה. בכל המיקומים האחרים (כלומר בכל המישור למעט בקו הישר הזה), שיטת ניוטון רפסון תיתן לנו סדרה שמתכנסת לפתרון המדויק, בתנאי שניחוש ההתחלתי שלנו הוא מספיק טוב.

מוטיבציה: לפני ההוכחה של המשפט, ננסה לתת מוטיבציה עבור ההגדרה $(**)$ של סדרת הקירובים לפתרון. נניח שקיים פתרון $\bar{x}$ המקיים את התנאים במשפט. ונניח ש-x^1 הניחוש הראשון שלנו ל - $\bar{x}$. נניח גם שהתמזל מזלנו, וש- x^1

באמת קרוב לפתרון האמיתי. אז ננסה לפתור את המשוואה $f(\bar{x})=0$, על-ידי החלפתה במשוואה אפינית והזנחת השגיאה: במקום $f(\bar{x})=0$ נתבונן ב-

$$(\&) \quad f(\bar{x}) \approx f(x^1)+Df(x^1)(\bar{x}-x^1)=0$$

זה נותן את המשוואה

$$[f'(x^1)]^{-1}f(x^1)=x^1-\bar{x}$$

שפתרונה הוא

$$\bar{x}=x^1-[f'(x^1)]^{-1}f(x^1)$$

כלומר אם במקום $\approx$ היה שוויון ממש במשוואה $(\&)$, אז האיטרציה $(**)$ הייתה נותנת לנו את הפתרון $\bar{x}$. כיון שאין לנו שוויון ממש במשוואה $(\&)$, האיטרציה לא נותנת את הפתרון האמיתי של המשוואה והשגיאה אינה אפס. אבל בגלל שהשגיאה היחסית קטנה מאוד ביחס ל- $\|\bar{x}-x^1\|$ אפשר לקוות ש - x^2 קרוב יותר לפתרון האמיתי מאשר x^1.

דרך אחרת ל"הצדיק" את האינטואיציה מאחורי האיטרציה $(**)$, היא לפתח את הקירוב הלינארי של f^{-1} סביב הנקודה $f(x^1)$, תוך שימוש ב - $f(\bar{x})=0$. בעזרת משפט הפונקציה ההפוכה, אנו מחשבים:

$$\bar{x}=f^{-1}(0)\approx f^{-1}\big(f(x^1)\big)+(f^{-1})'(x^1)\big(0-f(x^1)\big)=x^1-[f(x^1)]^{-1}f(x^1)$$

ואגף ימין הוא בדיוק כמו במשוואה $(**)$ עבור $k=1$.

כמובן, כל מה שהופיע למעלה תחת "מוטיבציה" איננו הוכחה, אלא רק אמור לתת תחושה מדוע יש טעם להגדיר את האיטרציות באופן הזה. וכעת סוף סוף ניגש להוכחת המשפט.

הוכחת המשפט: מתקיים $[f'(\bar{x})]^{-1}f'(\bar{x})=I$ וכן

$$f(x)=f(\bar{x})+f'(\bar{x})(x-\bar{x})+\varepsilon(x-\bar{x})$$

כאשר $\varepsilon(x-\bar{x})=o(x-\bar{x})$ בסביבת $\bar{x}$.

נבחר $r>0$ כך שלכל $x\in B_r(\bar{x})$ מתקיימים התנאים הבאים:

$$\|I-[f'(x)]^{-1}f'(\bar{x})\|<\frac{1}{4}$$

$$\|\varepsilon(x-\bar{x})\|\le\frac{1}{4}\|[f'(x)]^{-1}\|^{-1}\|x-\bar{x}\|$$

לכל k אנו נגדיר את האיבר x^k בסדרת הקירובים לפתרון כמו במשוואה $(**)$. כעת לכל k מתקיים (תוך שימוש בהגדרת הסדרה, בעובדה ש- $f(\bar{x})=0$, ובקירוב לינארי של f סביב הנקודה $\bar{x}$):

$$\begin{aligned}\|x^{k+1}-\bar{x}\| &= \left\|x^k-\left[f'\left(x^k\right)\right]^{-1}f\left(x^k\right)-\bar{x}\right\| \\ &\le \left\|x^k-\bar{x}-\left[f'\left(x^k\right)\right]^{-1}\left(f(\bar{x})+f'(\bar{x})\left(x^k-\bar{x}\right)\right)\right\|+\left\|\left[f'\left(x^k\right)\right]^{-1}\varepsilon\left(x^k-\bar{x}\right)\right\| \\ &= \left\|\left(I-\left[f'\left(x^k\right)\right]^{-1}f'(\bar{x})\right)\left(x^k-\bar{x}\right)\right\|+\left\|\left[f'\left(x^k\right)\right]^{-1}\varepsilon\left(x^k-\bar{x}\right)\right\|\end{aligned}$$

לכן, אם $x^k\in B_r(\bar{x})$ אז נקבל

$$\|x^{k+1}-\bar{x}\|\le\frac{1}{4}\|x^k-\bar{x}\|+\frac{1}{4}\|x^k-\bar{x}\|=\frac{1}{2}\|x^k-\bar{x}\|$$

מכאן נקבל באינדוקציה שאם $x^1\in B_r(\bar{x})$ אז גם $x^k\in B_r(\bar{x})$ לכל k, וכמו כן נקבל

$$\|x^{k+1} - \bar{x}\| \leq \frac{1}{2}\|x^k - \bar{x}\| \leq \cdots \leq \frac{1}{2^k}\|x^1 - \bar{x}\| \to 0$$

וזה מה שהיה צריך להוכיח! מש"ל.

קוד לדוגמה (פייתון):

```
# Newton-Raphson method demo
import numpy as np
from numpy.linalg import norm
# Location of two fixed points, "antenas"
a = np.array([10, 13])
b = np.array([-7,0])
# Pick a random location of some point, a "portable device":
x = 50*np.random.randn(2)
print('True location = ', x)
# The true distance from the antena's to the device:
r = np.array([norm(a-x)**2, norm(b-x)**2])
# Some intial guess of the location:
x_k = x + 10*np.random.randn(2)
print('Initial guess = ', x_k)

# The functions needed in the Newton-Raphson method
def f(P,A,B,r): # Calculates the distance squared of a point p from points A,
 B, minus the true known distance
    return np.array([norm(A-P)**2, norm(B-P)**2]) - r

def Df(P,A,B): # Derivative of f
    return 2*np.array([[P[0] - A[0], P[1] - A[1]], [P[0] - B[0], P[1] - B[1]]
])

# The following lines are a few iterations of the Newton-Raphson method
N = 10
for k in range(N):
    Dfinv = np.linalg.inv(Df(x_k,a,b))
    x_k = x_k - Dfinv @ f(x_k,a,b,r)

print('Calculated location = ', x_k)
```

פרק 8

תזכורת – משפט הפונקציה ההפוכה:

בהינתן $\underset{\text{פתוחה}}{U} \subseteq \mathbb{R}^n$ $a \in$ ובהינתן פונקציה $f \in C^1(U, \mathbb{R}^n)$ כך ש-$f'(a)$ הפיכה – אזי קיימת סביבות $a \in V \subseteq U$ וכן $f(a) \in W \subseteq f(U)$ כך ש-$f: V \to W$ חד-חד ערכית ועל. כמו כן, $f^{-1}: W \to V$ גזירה ברציפות ומתקיים:

$$\boxed{(f^{-1})'(y) = \left[f'\left(f^{-1}(y)\right)\right]^{-1}}$$

אנו יודעים כי f^{-1} גזירה ברציפות, שכן הנוסחה שמצאנו מראה כי הנגזרת של f^{-1} היא הרכבת f^{-1} שהיא רציפה, על f' שהיא רציפה, ועל הרכבה זו מרכיבים את ההעתקה שלוקחת מטריצה הפיכה להופכי שלה. זו העתקה רציפה לפי הנוספח של פרק 2, ולמעשה היא גזירה ברציפות לפי הדוגמה בפרק 3 (זה נובע גם מהקשר בין המטריצה $adj\ A$ לבין המטריצות A, A^{-1}).

8.1 הגדרה – פונקציה $f \in C^1\left(\underset{\subseteq \mathbb{R}^n}{U}, \mathbb{R}^m\right)$ נקראת **רגולרית ב-** לכל a ($a \in U$) אם $\operatorname{rank} J_f(a) = m$.
הפונקציה נקראת **רגולרית ב-**U אם לכל $a \in U$ היא רגולרית ב - a.

8.2 משפט (משפט ההעתקה הפתוחה) – תהא $f \in C^1(U, \mathbb{R}^m)$ רגולרית ב-U. אזי, f העתקה פתוחה – כלומר, לכל $V \subseteq U$ פתוחה, $f(V)$ גם היא פתוחה.

<u>הוכחה:</u>

המקרה הפשוט יותר, שבו $m = n$, מתקבל מידית ממשפט הפונקציה ההפוכה. אכן, לכל $a \in U$ קיימת סביבה פתוחה $V_a \subset U$ של a, וקיימת סביבה W_a של $f(a)$ כך ש - $W_a = f(V_a) \subseteq f(U)$. נסיק ש

$$f(U) = \bigcup_{a \in U} f(V_a)$$

איחוד של פתוחות, ולכן $f(U)$ קבוצה פתוחה. בהינתן קבוצה פתוחה $V \subseteq U$ אז נחזור על אותו שיקול בדיוק, כאשר חושבים על הצמצום של f ל-V כפונקציה $f: V \to \mathbb{R}^n$.

נעסוק עתה במקרה שבו $m < n$. נשים לב כי $\operatorname{rank} J_f(a) = m$ כלומר במטריצה $J_f(a)$ יש m עמודות בלתי תלויות ליניארית. בלי הגבלת הכלליות, יהיו אלה m העמודות הראשונות. נסמן:

$$\frac{\partial(f_1, \cdots, f_m)}{\partial(x_1, \cdots, x_m)} = \left[\frac{\partial f_i}{\partial x_j}\right]_{(i,j)=1}^{m}$$

נגדיר עתה:

$$F: U \to \mathbb{R}^n$$

על ידי:

$$F(x) = (f(x), x_{m+1}, \cdots, x_n)$$

כלומר, נשלים את הקואורדינטות כך שההעתקה תהיה בעלת n ממדים. לכן מתקיים:

$$J_F = \begin{bmatrix} J_f & \star \\ 0 & I_{(n-m)\times(n-m)} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial(f_1, \cdots, f_m)}{\partial(x_1, \cdots, x_m)} & \star \\ 0 & I \end{bmatrix}$$

נשים לב כי F מקיימת את תנאי משפט הפונקציה ההפוכה בכל נקודה, ולכן $F(V)$ פתוחה ב-$\mathbb{R}^n$ לכל $V \subseteq U$ פתוחה. מכאן נובע שגם $f(V) = \pi\big(F(V)\big)$ פתוחה, כאשר:

$$\pi: \mathbb{R}^n \to \mathbb{R}^m \quad , \quad (x_1, \cdots, x_n) \mapsto (x_1, \cdots, x_m)$$

הינה העתקה פתוחה. לשלמות הטיעון נראה כי העתקה זו, π, היא העתקה פתוחה.

תהא $W \subseteq \mathbb{R}^n$ פתוחה, ונרצה להראות כי $\pi(W)$ קבוצה פתוחה. לשם נוחות, נשתמש בסימון הבא:

$$(a,b) \in \mathbb{R}^n \quad \begin{matrix} a \in \mathbb{R}^m \\ b \in \mathbb{R}^{n-m} \end{matrix}$$

כל $a \in \pi(W)$ היא נקודה המתקבלת $(a,b) \in W$. אם W פתוחה, נסיק כי קיים $r > 0$ כך $B_r\big((a,b)\big) \subseteq W$. נראה כי $B_r(a) \subseteq \pi(W)$. יהא אם כן $a+h$, כך ש-$\|h\| < r$. נגדיר $(a+h,b) \in B_r(a,b) \subseteq W$ ולכן:

$$a + h \in \pi(W)$$

כלומר $B_r(a) \subseteq W$ לכל a ולכן $\pi(W)$ פתוחה כלומר π העתקה פתוחה כנדרש. מש"ל.

<u>מינימיזציה תחת אילוצים</u>

תהא $f \in C^1(U, \mathbb{R})$ כך ש-$U \subseteq \mathbb{R}^n$. נניח שאנו מחפשים מינימום של f על U. כלומר, אנחנו מחפשים $\bar{x}$ כך שמתקיים $f(\bar{x}) \leq f(x)$ לכל $x \in U$. אנו יודעים שאם U פתוחה, אז תנאי הכרחי עבור ש- $\bar{x}$ יהיה נקודת מינימום מקומי (ראו טענה 2.17) הוא שמתקיים:

$$\nabla f(x) = 0$$

במקרים רבים ניתן לפתור את המשוואה הנ"ל באופן אנליטי או נומרי, וכך למצוא מועמדים למינימום.

לעתים, נרצה לבצע מינימיזציה לא בכל המרחב הנתון U, אלא רק בתת קבוצה של המרחב. כאשר תת הקבוצה מוגדרת כקבוצת האפסים של פונקציה גזירה ברציפות ורגולרית ניתן להשתמש בכלים שפיתחנו בקורס זה בכדי למצוא את המינימום.

8.3 משפט – נניח ש $g \in C^1(U, \mathbb{R}^m)$ רגולרית כאשר $U \subset \mathbb{R}^n$ קבוצה פתוחה, $M = \{x \in U | g(x) = 0\}$ ותהי $f \in C^1(U, \mathbb{R})$. אם $a \in M$ ומתקיים $f(a) \leq f(x)$ לכל $x \in M$ (קרי, a מינימום של f על M), אזי מתקיים:

$$\nabla f(a) \in \operatorname{span}\{\nabla g_1(a), \nabla g_2(a), \cdots, \nabla g_m(a)\}$$

באופן מעשי – כדי למצוא מינימום, פותרים את המשוואות:

$$\begin{cases} \nabla f(a) = \sum_{i=1}^{m} \lambda_i \nabla g_i(a) \\ g(a) = 0 \end{cases}$$

שימו לב: המשוואה הראשונה נותנת n משוואות ב-$n+m$ נעלמים, (הנעלמים הם $(a_1, \cdots, a_n), \lambda_1, \cdots, \lambda_m$) והמשוואה השנייה היא בעצם m משוואות ב-n-נעלמים. בסה"כ יש $n+m$ משוואות ב-$n+m$ נעלמים, ובאופן טיפוסי למערכת כזו יש מספר סופי של פתרונות.

8.4 הגדרה – במשפט 8.3, הנעלמים λ_i לכל $1 \leq i \leq m$ נקראים **כופלי לגרנז'.**

לפני ההוכחה, אנו נתבונן בדוגמה.

<u>דוגמה:</u>

בדוגמה זו אנו נפתור מקרה מיוחד של הבעיה הבאה: בהינתן תת-מרחב ונקודה במרחב, מהי הנקודה בתת-המרחב שהינה קרובה ביותר לנקודה הנתונה?

נתון מישור ב-$\mathbb{R}^3$ המתואר על ידי המשוואה:

$$g(x,y,z) = ax + by + cz - d = 0$$

כאשר נניח $(a,b,c) \neq (0,0,0)$.

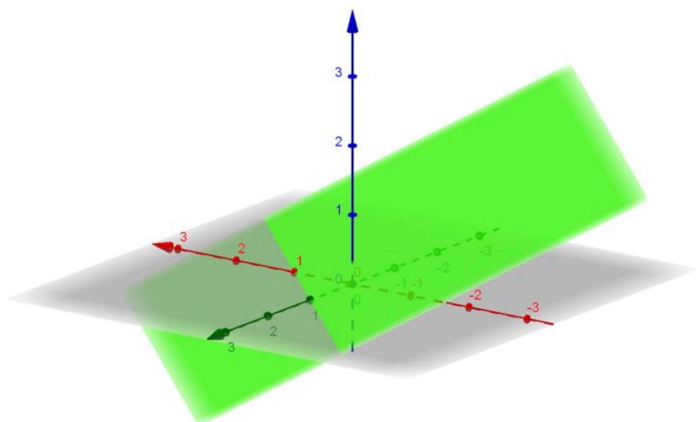
נניח שנתונה פונקציה:

$$f(x,y,z) = \|(x,y,z)\| = \sqrt{x^2 + y^2 + z^2}$$

אנו רוצים למצוא את הנקודה על המישור בה f מקבלת ערך מינימלי. זוהי בעצם הנקודה על המישור שהינה הכי קרובה לראשית הצירים.

לשם כך נחליף את f בפונקציה:

$$f(x,y,z) = x^2 + y^2 + z^2$$



איור 4 – דוגמה של מישור $ax + by + cz = d$ (בירוק) במרחב $\mathbb{R}^3$

ונשתמש במשוואות שהוגדרו במשפט 8.3:

$$(1) \quad \nabla f(x,y,z) = \lambda \nabla g(x,y,z)$$

$$(2) \quad g(x,y,z) = 0$$

נחשב את הגרדיאנטים:

$$\nabla f = (2x, 2y, 2z)$$
$$\nabla g = (a,b,c)$$

נשים לב שההנחה $(a,b,c) \neq (0,0,0)$ אומרת ש-g רגולרית בכל המרחב. נקבל ממשוואה (1) כי מתקיים:

$$(x,y,z) = \frac{\lambda}{2}(a,b,c)$$

נציב במשוואה (2) ונקבל כי:

$$g\left(\frac{\lambda}{2}(x,y,z)\right) = \frac{\lambda}{2}(a^2 + b^2 + c^2) - d = 0$$

ונקבל כי:

$$\lambda = \frac{2d}{a^2 + b^2 + c^2} \rightarrow \boxed{(x,y,z) = \frac{d(a,b,c)}{a^2 + b^2 + c^2}}$$

והמרחק המינימלי הוא $\|(x,y,z)\| = \frac{d}{\|(a,b,c)\|}$. כפי שיכולנו לצפות, קיבלנו בדיוק את הנוסחה (שאולי חלקכם מכירים) למרחק בין המישור לבין הראשית, כנדרש.

<u>הערה:</u> משפט 8.3 בעצם נותן רק תנאי הכרחי למינימום, הוא לא מבטיח שקיים מינימום, והוא לא מוכיח שהנקודות החשודות שמוצאים על-ידי שיטה זו הם אכן נקודות מינימום (הם אולי נקודות מקסימום, והם אולי לא מינימום ולא מקסימום). כדי לראות שקיים מינימום, או להראות שהנקודה שמצאנו היא מינימום, צריך להפעיל שיקול נוסף. בדוגמה

לעיל השיקול הנוסף הושמט, כדי שנוכל להתמקד בשימוש בכופלי לגרנז'. אבל כעת בואו ונתעכב על ההצדקה מדוע הנקודה שמצאנו היא אכן המינימום של f על המישור.

תהי (x_0, y_0, z_0) נקודה כלשהי על המישור (כיון ש - $(a, b, c) \neq (0,0,0)$ אז ידוע שקיימים פתרונות למשוואה $g(x, y, z) = 0$) נסמן $R = 2\|x_0, y_0, z_0\| > 0$ ונגדיר $U = B_R(0)$. נגדיר

$$W = \{(x, y, z) \in U : g(x, y, z) = 0\}$$

$$M = W \cap U$$

אז כיון ש - $(x_0, y_0, z_0) \in M$ ו- $f(x_0, y_0, z_0) < R^2$, ברור שהמינימום של f על המישור, אם הוא מתקבל בכלל, מתקבל ב- M. אבל בקבוצה הקומפקטית $W \cap \overline{B_R(0)}$ הפונקציה הרציפה חייבת לקבל מינימום, והמינימום קטן או שווה ל- $f(x_0, y_0, z_0) = \|x_0, y_0, z_0\|^2$, ולכן המינימום מתקבל בקבוצה $M = W \cap U$. כעת קל לבדוק ש f , g, U ו- M מקיימות את תנאי משפט 8.3, לכן בנקודת המינימום יתקיימו בהכרח המשוואות (1) ו - (2) . כיון שמצאנו רק נקודה אחת בכל W בה מתקיימות המשוואות, הנקודה שמצאנו היא בהכרח נקודת המינימום.

כעת, ניגש להוכחת המשפט.

<u>הוכחת משפט 8.3:</u>

תהא $a \in M$, ונניח כי:

$$\nabla f(a) \notin \operatorname{span}\{\nabla g_1(a), \nabla g_2(a), \cdots, \nabla g_m(a)\}$$

אנו נראה כי ב-a לא מתקבלת נקודת מינימום. זה כמובן יוכיח את המשפט.

נגדיר $F: U \to \mathbb{R}^{m+1}$ על ידי:

$$F(x) = \big(f(x), g(x)\big) = \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}$$

$$J_F(x) = \begin{bmatrix} \nabla f(x) \\ \nabla g_1(x) \\ \vdots \\ \nabla g_m(x) \end{bmatrix}$$

כלומר $\operatorname{rank} J_F(a) = m + 1$ ולכן $\operatorname{rank} J_F = m + 1$ בסביבת a כי הדרגה של מטריצה הינה פונקציה רציפה (אפשר להוכיח זאת בכמה דרכים, למשל ניתן להשתמש בעובדה מאלגברה לינארית, שהדרגה של מטריצה מלבנית מתקבלת בגודל המינור הגדול ביותר). מכאן נקבל כי F העתקה פתוחה (ממשפט 8.2) בסביבת a.

כמו כן, $F(a) = (f(a), 0)$ נמצא בפנים של הקבוצה הפתוחה $F(U)$ מכאן שגם $(f(a) - \varepsilon, 0) \in F(U)$ עבור בחירה של ε קטן מספיק.

עתה, מהגדרת $F(U)$, יש $x \in U$ כך שמתקיים $F(x) = (f(a) - \varepsilon, 0) = (f(x), 0)$. מכאן נובעות שתי מסקנות:

א) $f(x) < f(a)$

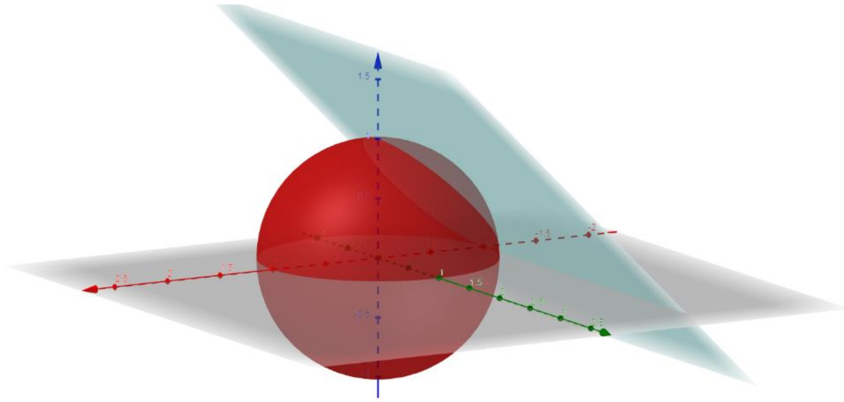
ב) $g(x) = 0$.

מהגדרת M, $x \in M$. ולכן a לא נקודת מינימום ב - M. מש"ל.

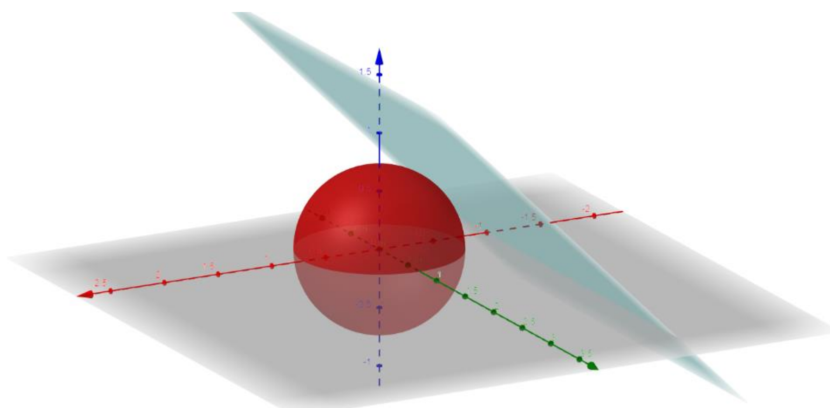
ניתן להבין אינטואיטיבית את משמעות המשפט באופן הבא –

באופן כללי, גרדיאנט של פונקציה במספר משתנים, נותן את הכיוון של וקטור הניצב למישור המשיק למשטח הרמה של הפונקציה בנקודה.

כשאנחנו מעוניינים לבצע מינימיזציה של פונקציה על איזשהו משטח במרחב, אנחנו מעוניינים בעצם למצוא נקודה, המשותפת למשטח רמה של הפונקציה וכן למשטח, שבה המישור שמשיק למשטח הוא אותו מישור (מבחינת כיוון) שמשיק למשטח הרמה של הפונקציה. מכאן שהגרדיאנט של הפונקציה ושל פונקציית היריעה צריכים להיות באותו כיוון בדיוק, למעט אולי כפל בסקלרים – הלא אלו בדיוק הם כופלי לגרנז'. איורים 5 ו-6 מדגימים את העיקרון של משפט כופלי לגרנז' במקרה של הדוגמה שחישבנו לעיל – חישוב הנקודה על מישור שהיא הכי קרובה לראשית. מומלץ לצייר דוגמה דו-ממדית: שם המשטח M הינו עקום במישור ומשטחי הרמה של הפונקציה הינם קווי גובה.



איור 5 תיאור של משטח רמה $f=1$ של הפונקציה $f(x,y,z)=x^2+y^2+z^2$ (באדום) ומשטח M במרחב (במקרה הזה המשטח M הינו מישור). הניצב למשטח הרמה (כלומר הניצב לכדור) אינו מקביל לניצב למשטח באף נקודה בחיתוך. לכן 1 אינו הערך המינימלי של f על המישור



איור 6 כאן מופיע אותו משטח כחול ומשטח הרמה $f(x,y,z)=\|(x,y,z)\|^2=1/2$. כאן הניצב למשטח הכחול מקביל לניצב למישור המשיק של משטח הרמה, ונקודת החיתוך היא המינימום של הפונקציה f על המישור

פרק 9

בהרצאות הקרובות נעסוק בתחומים $U \subseteq \mathbb{R}^n \times \mathbb{R}^m \cong \mathbb{R}^{n+m}$ פתוחים. נסמן $(x,y) \in \mathbb{R}^n \times \mathbb{R}^m$ כאשר:

$$x \in \mathbb{R}^n \quad y \in \mathbb{R}^m$$

המשפט הבא נקרא "משפט הפונקציה הסתומה", והוא אחד המשפטים הקשים בקורס, מבחינה קונספטואלית. אולי כדאי לפני שנדון במשפט, נאמר מה זו "פונקציה סתומה". המילה "פונקציה" מתארת התאמה בין שתי קבוצות. במתמטיקה, לא תמיד יש צורך לכתוב באופן מפורש מה פונקציה עושה. למשל, אם לשתי קבוצות יש את אותה העוצמה (למשל אם שתיהן בנות מנייה) אזי יש התאמה חד-חד ערכית ועל מאחת לשנייה, אבל לפעמים מספיק לנו לדעת שקיימת כזו התאמה, ואין לנו צורך בנוסחה ספציפית שמאפשרת לנו לחשב לאיזה איבר בקבוצה ב' עלינו להעתיק איבר כלשהו בקבוצה א'.

בשימושים הפשוטים של החשבון הדיפרנציאלי והאינטגרלי, לרוב אנו עובדים עם פונקציות **מפורשות**. למשל הפונקציה $f(x) = x^2 + 1$ אומרת לנו בדיוק לאן כל x צריך להיות ממופה על-ידי f. בחדו"א ישנה מסורת לזהות פונקציה כזו עם הגרף שלה ולכתוב $y = x^2 + 1$, ולומר ש- y היא פונקציה של x –, כלומר y תלויה ב-x .

כדי להבין מה זו **פונקציה סתומה**, נגדיר קודם כל **משוואה סתומה. משוואה סתומה** היא ביטוי מהצורה $F(x,y) = 0$ כאשר F היא פונקציה המוגדרת על תת קבוצה U של $\mathbb{R}^n \times \mathbb{R}^m$. נקודה $(a,b) \in U$ נקראת **פתרון** של המשוואה הסתומה אם מתקיים $F(a,b) = 0$. אנו נאמר שהמשוואה הסתומה $F(x,y) = 0$ מגדירה את y כפונקציה סתומה של x אם קיימת איזו קבוצה $V_1 \times V_2 \subseteq U$ כך שלכל $x \in V_1$ קיים $y \in V_2$ **<u>יחיד</u>** כך ש- $F(x,y) = 0$. זה מגדיר את y כפונקציה של x, כאשר ההתאמה היא

$$(*) \quad x \mapsto \text{ the unique } y \text{ for which } F(x,y) = 0$$

המושג "פונקציה סתומה" אינו לגמרי מוגדר היטב, משתי סיבות. ראשית, אפשר לומר שההוראות שמופיעות ב - $(*)$ הם די מפורשות, ואומרות לנו איך למצוא את y בהינתן x : פשוט צריך למצוא את ה-y שמקיים $F(x,y)$. אולי תתווכחו ותגידו "אבל לא אומרים לנו **<u>איך</u>** למצוא את ה y הזה!", אבל לכך אפשר לענות שגם לא אומרים לכם איך למצוא את השורש של מספר נתון, ולכולנו ברור ש- $y = \sqrt{x}$ אמורה להיחשב כפונקציה מפורשת. כאן אנו מגיעים לבעיה השנייה של המושג "פונקציה סתומה", והוא שהמושג "פונקציה מפורשת" גם לא לגמרי מוגדר היטב. למשל המשמעות של הביטוי המפורש $y = \sqrt{x}$ הוא ש- y הוא הפתרון (החיובי) למשוואה הסתומה $y^2 - x = 0$.

9.1 משפט הפונקציה הסתומה – תהא $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ פתוחה, תהא $f \in C^1(U, \mathbb{R}^m)$, וכן תהא $(a,b) \in U$ כך ש-$f(a,b) = 0$. נניח, בנוסף, כי $\det \frac{\partial(f_1, f_2, \cdots, f_m)}{\partial(y_1, y_2, \cdots, y_m)}\Big|_{(a,b)} \neq 0$. **אזי קיימת סביבה** $(a,b) \in V$ **וקיימת פונקציה** $g \in C^1$ **שמוגדרת בסביבה של** a **ומקבלת ערכים ב-** $\mathbb{R}^m$ **כך שמתקיים:**

$$\forall (x,y) \in V \quad f(x,y) = 0 \Leftrightarrow y = g(x)$$

ובפרט, $g(a) = b$ ו- $f\big(x, g(x)\big) = 0$ לכל x קרוב ל-a.

<u>משמעות המשפט:</u> תחת ההנחות אוסף הפתרונות של המשוואה $f = 0$ מהווה קבוצה "יפה": גרף של פונקציה גזירה ברציפות ב- n משתנים. דרך אחרת להסתכל על המשפט: עבור (x,y) המקיימים $f(x,y) = 0$, ניתן "לחלץ" את המשתנה y כפונקציה של x. יתר על כן, y תלוי ב x בצורה גזירה ברציפות, ובפרט באופן רציף. הביטוי $f(x,y) = 0$ מגדיר את y כפונקציה של x , אבל זה נקרא "פונקציה סתומה", לעומת הביטוי $y = g(x)$ הנותן את y כפונקציה של x, הנקרא "פונקציה מפורשת".

דוגמה:

תהא הפונקציה:

$$f: \mathbb{R}^2 \to \mathbb{R} \quad f(x,y) = x^2 + y^2 - 1$$

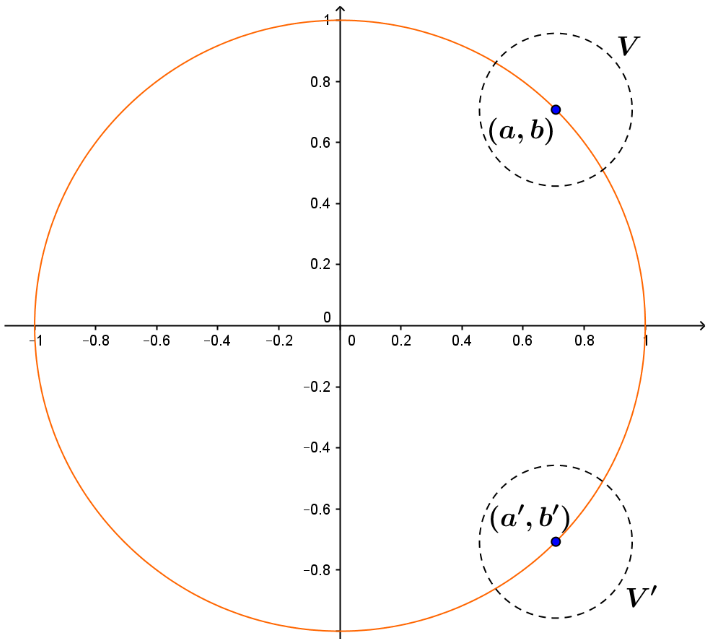
ונתבונן בנקודה $(a,b) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$. אכן מתקיים $f(a,b) = 0$, וכן

$$\frac{\partial f}{\partial y}(a,b) = 2y|_{(a,b)} = 2b = \sqrt{2} \neq 0.$$

עתה, נבחר סביבה V של הנקודה (a,b) , ונגדיר את הפונקציה $g(x) = \sqrt{1-x^2}$, נשים לב, שבסביבה קטנה של (a,b):

$$f(x,y) = 0 \Leftrightarrow y = \sqrt{1-x^2}$$

ובעבור $(a',b') = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ לדוגמה, באותו אופן מתקיים $f(a',b') = 0$ וכן $\frac{\partial f}{\partial y}(a',b') \neq 0$ ובבחירת סביבה V' של הנקודה, הפונקציה $g(x) = -\sqrt{1-x^2}$ מקיימת את הנדרש.



איור 7 – תיאור הפונקציה $0 = f(x,y) = x^2 + y^2 - 1$

אנו נתקל בבעיה רק כאשר $y = 0$, במקרה הזה לא נוכל לחלץ את y כפונקציה של x. למשל, בנקודה $(1,0)$ נשים לב כי $\frac{\partial f}{\partial y}(1,0) = 0$ ואכן אין כאן עמידה בתנאי המשפט. אך לעומת זאת, מתקיים $\frac{\partial f}{\partial x}(1,0) \neq 0$ ולכן על פי המשפט נוכל לבטא את x כפונקציה של y בסביבה של $(1,0)$ לקבלת הפתרונות. ואכן, בסביבת $(1,0)$ אוסף הפרונות מתואר על ידי הגרף $x = \sqrt{1-y^2}$.

הוכחת משפט הפונקציה הסתומה:

נגדיר $F: U \to \mathbb{R}^n \times \mathbb{R}^m$ באופן הבא:

$$F(x,y) = \big(x, f(x,y)\big) = (x_1, x_2, \cdots, x_n, f_1, \cdots, f_m)$$

ונשים לב כי מתקיים:

$$J_F = \begin{bmatrix} I_n & 0 \\ \dfrac{\partial(f_1,\cdots,f_m)}{\partial(x_1,\cdots,x_n)} & \dfrac{\partial(f_1,\cdots,f_m)}{\partial(y_1,\cdots,y_m)} \end{bmatrix}$$

והיות ודרשנו כי $\det \frac{\partial(f_1,\cdots,f_m)}{\partial(y_1,\cdots,y_m)}\Big|_{(a,b)} \neq 0$ אזי נקבל כי J_F הפיכה בנקודה (a,b). נפעיל את משפט הפונקציה ההפוכה על F סביב (a,b). נקבל סביבה V של (a,b) וסביבה W של $F(a,b)$ כך ש- $F: V \to W$ הפיכה, עם הפונקציה ההפוכה המתאימה $F^{-1}: W \to V \in C^1$. נוכל להניח ש- $W = W_1 \times W_2$ כאשר W_1 סביבה של a ו- W_2 סביבה של $0 \in \mathbb{R}^m$.

נסמן:

$$F^{-1}\Big(\underset{\in\mathbb{R}^n}{u}, \underset{\in\mathbb{R}^m}{v}\Big) = \big(\varphi_1(u,v), \varphi_2(u,v)\big) \quad \begin{matrix} \varphi_1: W \to \mathbb{R}^n \\ \varphi_2: W \to \mathbb{R}^m \end{matrix}$$

יש לנו בידיים פונקציה

$$\varphi_2: W \to \mathbb{R}^m \quad \varphi_2 \in C^1$$

נגדיר, אפוא, את הפונקציה:

$$g(x) = \varphi_2(x, 0)$$

אשר מוגדרת היטב בסביבה W_1 של a, ונראה שהפונקציה הזו מקיימת את הנדרש.

לכל $(x, y) \in V$ מתקיים:

$$f(x, y) = 0 \Leftrightarrow \big(x, f(x, y)\big) = (x, 0) \Leftrightarrow F^{-1}\big(x, f(x, y)\big) = F^{-1}(x, 0) = \big(\varphi_1(x, 0), \varphi_2(x, 0)\big)$$

אבל $F^{-1}\big(x, f(x, y)\big) = F^{-1} \circ F(x, y) = (x, y)$. לכן קיבלנו

$$f(x, y) = 0 \Leftrightarrow (x, y) = \big(\varphi_1(x, 0), \varphi_2(x, 0)\big) = \big(\varphi_1(x, 0), g(x)\big)$$

אך תמיד מתקיים $\varphi_1(x, z) = x$ לכל $(x, z) \in W_1 \times W_2 = W$ זה נובע מההגדרות. אפשר לראות זאת גם מהשוויונות:

$$\Big(\varphi_1\big(x, f(x, y)\big), \varphi_2\big(x, f(x, y)\big)\Big) = F^{-1}\big(x, f(x, y)\big) = F^{-1} \circ F(x, y) = (x, y)$$

כאשר אנו משתמשים בעובדה ש $(x, z) \in W$ הינו מהצורה $(x, z) = F(x, y) = \big(x, f(x, y)\big)$ עבור $(x, y) \in V$.

בסך הכל, קיבלנו ש - $f(x, y) = 0 \Leftrightarrow (x, y) = \big(x, g(x)\big)$ ולכן נסיק כי $f(x, y) = 0$ אם ורק אם $y = g(x)$, וזה מש"ל.

עבור המשפט הבא, אנו צריכים כעת להכיר כמה סימונים חדשים. המרחב הוקטורי $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$ הוא מעין "סכום ישר" של המרחבים $\mathbb{R}^m$ ו- $\mathbb{R}^n$: כל וקטור $v \in \mathbb{R}^{n+m}$ אפשר לרשום באופן הבא: $v = (u, w)$ כאשר $u \in \mathbb{R}^n$ ו- $w \in \mathbb{R}^m$. בהינתן העתקה לינארית $T: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$, אנו יכולים לחשוב על T כאופרטור לינארי של שני משתנים: $Tv = T(u, w)$. נסמן ב- T_1 את האופרטור הלינארי $T_1: \mathbb{R}^n \to \mathbb{R}^k$ הנתון על-ידי $T_1 u = T(u, 0)$ וב- T_2 את האופרטור $T_2: \mathbb{R}^m \to \mathbb{R}^k$ הנתון על-ידי $T_2 w = T(0, w)$. אז האופרטור T נקבע ביחידות על-ידי T_1, T_2 ולהיפך, על-ידי הקשר

$$Tv = T(u, w) = T_1 u + T_2 v$$

העניין נהיה שקוף יותר לחלק מן הסטודנטים, כאשר משתמשים בשפה של מטריצות. אם A, A_1 ו - A_2 הן המטריצות המייצגות של T, T_1 ו - T_2 בהתאמה, אזי ל- A יש צורת בלוקים

$$A = [A_1 \quad A_2]$$

וכפל המטריצה בוקטור מתבצע כך:

$$Av = [A_1 \quad A_2]\begin{bmatrix} u \\ w \end{bmatrix} = A_1 u + A_2 w$$

כעת, אם $U \subseteq \mathbb{R}^n \times \mathbb{R}^m \cong \mathbb{R}^{n+m}$ פתוחה, $(a, b) \in U$, ו- $f: U \to \mathbb{R}^k$ גזירה ב- (a, b), אז הנגזרת בנקודה (a, b) היא העתקה לינארית $Df(a, b): \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$. אנו מסמנים את הנקודות ב U – כזוגות (x, y) כאשר

$$(x, y) = (x_1, \dots, x_n, y_1, \dots, y_m)$$

ולכן אנו נרשום את הפירוק (לפי $\mathbb{R}^n\times\mathbb{R}^m\cong\mathbb{R}^{n+m}$) של ההעתקה הלינארית $T=Df(a,b)$ כ- $T_1=D_xf(a,b)$ ו- $T_2=D_yf(a,b)$.

אם מסתכלים על מטריצות היעקוביאן, העניין צריך להיות ברור:

$$J_f(a,b)=\begin{bmatrix}\frac{\partial(f_1,\dots,f_k)}{\partial(x_1,\dots,x_n)} & \frac{\partial(f_1,\dots,f_k)}{\partial(y_1,\dots,y_m)}\end{bmatrix}$$

9.2 משפט (נגזרת של פונקציה סתומה) – תהא $U\subseteq\mathbb{R}^n\times\mathbb{R}^m$ פתוחה, תהא $f\in C^1(U,\mathbb{R}^m)$, וכן תהא $(a,b)\in U$ כך ש-$f(a,b)=0$. נניח, בנוסף, כי $\det\frac{\partial(f_1,f_2,\cdots,f_m)}{\partial(y_1,y_2,\cdots,y_m)}\Big|_{(a,b)}\neq 0$. תהי $g\in C^1$ מוגדרת בסביבה של a ומקבלת ערכים ב- $\mathbb{R}^m$ כמו במסקנה של משפט הפונקציה הסתומה המקיימת $f\big(x,g(x)\big)=0$ לכל x בסביבה של a. אזי הנגזרת של g בנקודה של a נתונה על-ידי

$$g'(a)=-\big[D_yf(a,b)\big]^{-1}D_xf(a,b)$$

<u>**הערה**</u> – הנוסחה הנ"ל שימושית מאוד כיון שהיא מאפשרת "למצוא" את $g'(a)$ מבלי לדעת מהי הפונקציה $g(x)$.

<u>**הוכחה:**</u> נגדיר פונקציה $h(x)=f\big(x,g(x)\big)$, כלומר $h=f\circ\phi$ כאשר $\phi(x)=\big(x,g(x)\big)$. לפי משפט הפונקציה הסתומה, מתקיים $h(x)=0$ בסביבה של a. נגזור את h בנקודה a, ונקבל לפי כלל השרשרת:

$$0=Dh(a)=Df\big(\phi(a)\big)D\phi(a)$$

נוח כעת לעבור למטריצות מייצגות כדי לבצע את החישוב המכריע. המשוואה האחרונה הופכת למשוואה מטריצית

$$J_f\big(\phi(a)\big)J_\phi(a)=0$$

כעת נחשב $J_\phi(a)=\begin{bmatrix}I\\J_g(a)\end{bmatrix}$ ואילו $J_f\big(\phi(a)\big)=J_f(a,b)=\begin{bmatrix}\frac{\partial(f_1,\dots,f_k)}{\partial(x_1,\dots,x_n)}\Big|_{(a,b)} & \frac{\partial(f_1,\dots,f_k)}{\partial(y_1,\dots,y_m)}\Big|_{(a,b)}\end{bmatrix}$ ואם נציב את הנגזרות הללו במשוואה לעיל נקבל

$$\frac{\partial(f_1,\dots,f_k)}{\partial(x_1,\dots,x_n)}\Big|_{(a,b)}+\frac{\partial(f_1,\dots,f_k)}{\partial(y_1,\dots,y_m)}\Big|_{(a,b)}J_g(a)=0$$

לפי ההנחה, המטריצה $\frac{\partial(f_1,\dots,f_k)}{\partial(y_1,\dots,y_m)}\Big|_{(a,b)}$ הינה הפיכה, לכן נקבל

$$J_g(a)=-\left(\frac{\partial(f_1,\dots,f_k)}{\partial(y_1,\dots,y_m)}\Big|_{(a,b)}\right)^{-1}\frac{\partial(f_1,\dots,f_k)}{\partial(x_1,\dots,x_n)}\Big|_{(a,b)}$$

זהו בדיוק מה שצריך להוכיח, רק בשפה של מטריצות מייצגות.

אמנם כאן סיימנו את ההוכחה, אבל כדאי לעצור רגע ולראות איך אפשר להוכיח את הנוסחה לנגזרת בלי להשתמש במטריצות מייצגות. נחזור לזהות $Df\big(\phi(a)\big)D\phi(a)$. קל לחשב שהנגזרת של ϕ פועלת על כל $v\in\mathbb{R}^n$ לפי $D\phi(a)v=\begin{bmatrix}v\\g'(a)v\end{bmatrix}$, ולכן

$$0=Df\big(\phi(a)\big)D\phi(a)v=Df\big(\phi(a)\big)\begin{bmatrix}v\\g'(a)v\end{bmatrix}=D_xf(a,b)v+D_yf(a,b)g'(a)v$$

לכן

$$D_y f(a,b) g'(a) = -D_x f(a,b)$$

ומכאן מכילים ב- $\left[D_y f(a,b)\right]^{-1}$ ומקבלים את הדרוש. מש"ל.

<u>לתרגול/תרגיל בית:</u> נגזרת של פונקציה סתומה, נגזרות חלקיות, דוגמאות ושימושים.

פרק 10

שורשי פולינומים

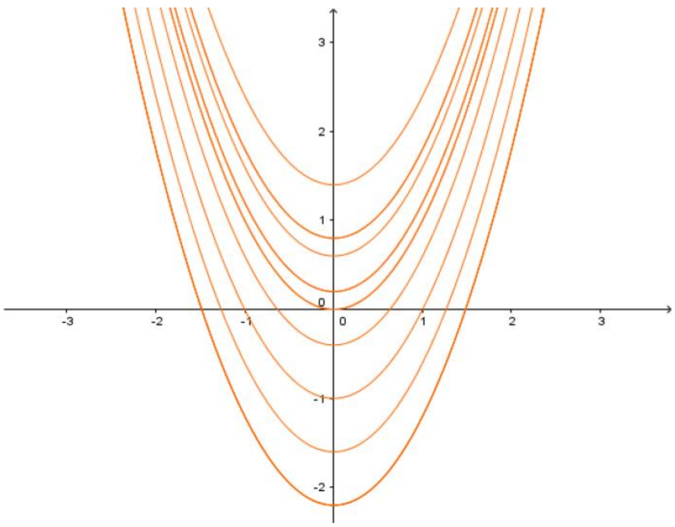
נניח ונתון לנו פולינום מהצורה $p(x) = \sum_{k=0}^{n} a_k x^k$ ונניח כי קיים אלגוריתם למציאת פתרון כלשהו.

דוגמה:

עבור הפולינום $P_\alpha = x^2 - \alpha$ אנו יודעים כי אכן קיים אלגוריתם כנ"ל. הפתרון, כתלות ב-α, נתון על ידי:

$$x_\alpha = \pm\sqrt{\alpha}$$

באיור 8 משורטט הגרף של הפולינום ערכים שונים של α . באיור 9 רואים גרף של הפתרונות כפונקציה של α .



איור 8 – מספר פולינומים מהצורה $P_\alpha = x^2 - \alpha$ עבור ערכי α שונים.

ניטיב לראות כי הפונקציות:

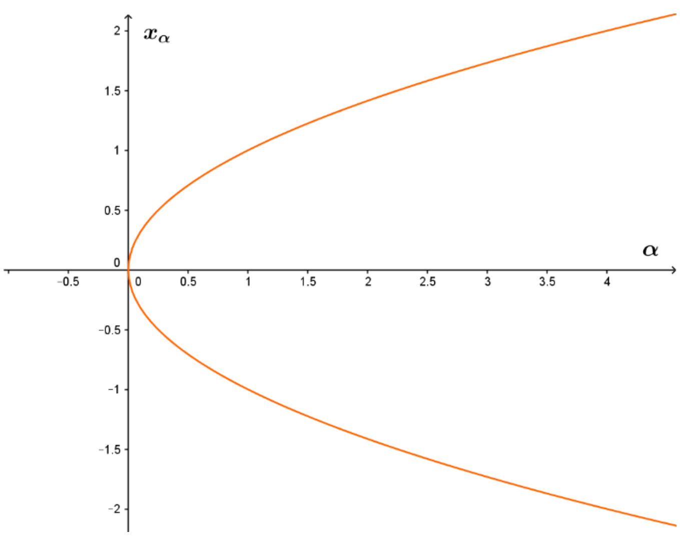
$$f_1(\alpha) = \sqrt{\alpha} \quad f_2(\alpha) = -\sqrt{\alpha}$$

גזירות ברציפות בחצי הציר החיובי. באופן כללי, שורשים של פולינום תלויים במקדמים שלהם באופן רציף, במובן מסוים. מטרתנו להבין את ה"מובן המסוים" הזה טוב יותר.

נרצה להראות – אם $\bar{x}$ שורש פשוט של $P_{\bar{a}}(x) = \sum_{k=0}^{n} \bar{a}_k x^k$ כאשר $\bar{a} \in \mathbb{R}^{n+1}$ מהצורה $\bar{a} = (\bar{a}_0, \cdots, \bar{a}_n)$, אזי קיימת סביבה של $\bar{a}$ ב-$\mathbb{R}^{k+1}$ כך שאחד השורשים של

$$P_{\bar{a}}(x) = \sum_{k=0}^{n} \bar{a}_k x^k$$

נתון כפונקציה גזירה ברציפות של $\bar{a}$.



איור 9 – פתרונות הפולינום P_α כתלות בערכו של α - $x_\alpha = \pm\sqrt{\alpha}$

נגדיר פונקציה :

$$f(a_0, \cdots, a_n, x) = P_a(x)$$
$$f\colon \underset{\cong \mathbb{R}^{n+1} \times \mathbb{R}}{\mathbb{R}^{n+2}} \to \mathbb{R}$$

המשוואה $P_a(x) = f(a, x) = 0$ מגדירה את השורשים כפונקציה סתומה של וקטור המקדמים a.

נניח כי $\bar{a} \in \mathbb{R}^{n+1}$ וקטור המקדמים של הפולינום $P_{\bar{a}}$ ונניח כי $\bar{x}$ הינו שורש של $P_{\bar{a}}$ כלומר $P_{\bar{a}}(\bar{x}) = 0$.

לכן ישנה נקודה $(\bar{a}, \bar{x})$ עבורה מתקיים:

$$f(\bar{a},\bar{x}) = P_{\bar{a}}(\bar{x}) = 0$$

וכיון שהשורש פשוט מתקיים גם:

$$\left.\frac{\partial f}{\partial x}\right|_{(\bar{a},\bar{x})} = \frac{d}{dx}P_{\bar{a}}\bigg|_{x=\bar{x}} \neq 0$$

נשים לב שהתנאי להפעלת משפט הפונקציה הסתומה הוא שיתקיים $P'_{\bar{a}}(\bar{x}) \neq 0$, וזהו תנאי שקול לכך שהשורש $\bar{x}$ הינו שורש מריבוי 1 בלבד.

10.1 מסקנה – **אם $\bar{x}$ שורש פשוט של $P_{\bar{a}}$ אזי קיימת סביבה $(\bar{a},\bar{x}) \in V$ וקיימת פונקציה $g \in C^1$ מוגדרת בסביבה של $\bar{a}$ כך שמתקיים:**

$$\boxed{\forall(a,x) \in V \quad P_a(x) = 0 \Leftrightarrow x = g(a)}$$

כלומר, קיבלנו שבסביבת שורשים מריבוי 1 של פולינומים, קיימת פונקציה גזירה ברציפות של המקדמים של הפולינום אשר מבטאת את השורש x.

למסקנה הנ"ל משמעות חשובה: אם אנו מעוניינים לפתור משוואה פולינומיאלית באמצעות איזשהו אלגוריתם נומרי, וקטור המקדמים נתון כסדרה של שברים עשרוניים (או בינאריים) הידועים עד דיוק סופי כלשהו. המסקנה מבטיחה לנו שאם אנו יודעים את המקדמים בקירוב מאוד טוב, אזי השורשים (הפשוטים) יושפעו בצורה רציפה ואפילו חלקה.

פרק 11

בחשבון דיפרנציאלי במשתנה אחד, הפונקציות אותן אנו חוקרים מוגדרות בדרך כלל על קטע. קטע פתוח הוא פשוט "חלק" נחמד, קשיר וריף של הישר: כל תת-קבוצה פתוחה וקשירה של הישר היא קטע פתוח. וקטע סגור הוא פשוט הסגור של קטע פתוח. בקורס הזה התחלנו ללמוד על פונקציות בכמה משתנים, והפונקציות שלנו הוגדרו לרוב על המרחב כולו או על תת קבוצה פתוחה של המרחב $\mathbb{R}^n$. אך במרחב הרב ממדי ישנן קבוצות מעניינות ויפות שגל עליהן אפשר לעשות חשבון דיפרנציאלי ואינטגרלי. בצורה מעט תמימה אפשר לומר שהמרחב האוקלידי הינו מרווח מספיק כדי להכיל קבוצות משמעותיות "ממימד יותר נמוך".

לא נוכל להתקדם אם נרצה לדון ב**כל** תתי הקבוצות של $\mathbb{R}^n$. בפרק זה אנו נגדיר את המושג "יריעות משוכנות גזירות ברציפות" (או בקיצור "יריעות") – אלו הן הקבוצות שעליהן אפשר לפתח תורה סבירה של גזירה ואינטגרציה. הסיבה לכך שאפשר לפתח על קבוצות אלו גזירה ואינטגרציה היא שיריעות "דומות מאוד", לפחות באופן "מקומי", לתתי קבוצות פתוחות של מרחב אוקלידי כלשהו. אנו נגדיר את המושג באופן מדויק להלן, אך לפני ההגדרה של יריעות, בואו נזכר במושג של "תת-מרחב וקטורי".

בתוך המרחב $\mathbb{R}^n$ ישנם תתי מרחבים לינאריים ממימדים קטנים יותר. תת-מרחב של $\mathbb{R}^n$ זוהי תת-קבוצה של $V \subseteq \mathbb{R}^n$ הסגורה לכפל בסקלר ולחיבור. אנו יודעים שכל תת מרחב הוא "איזומורפי" ל- $\mathbb{R}^k$ עבור $k \leq n$ כלשהו. בהינתן תת-מרחב ספציפי V ממימד k, ישנן ארבע דרכים לתאר אותו בצורה **מפורשת**:

1. אפשר להציג את V כתת-המרחב הנפרש על-ידי k וקטורים בלתי תלויים לינארית $\{v_1, \dots, v_k\}$. אפשר להשלים את הבסיס הזה של V לבסיס של כל המרחב $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ ואז בבסיס זה רואים שתת מרחב תמיד "נראה כמו" (כלומר עד כדי החלפת בסיס) $\mathbb{R}^k := \mathbb{R}^k \times \{0_{n-k}\} \subseteq \mathbb{R}^n$.
2. כגרעין של העתקה לינארית: $V = \ker T$ כאשר $rank(T) = n - k$ (באופן שקול, V הינו אוסף הפתרונות של מערכת משוואות לינאריות)
3. V נתון כגרף של פונקציה לינארית (למשל $y = ax$ או $z = ax + by$ או $\begin{cases} y = ax \\ z = bx \end{cases}$).
4. V נתון על-ידי פרמטריזציה, או, באופן שקול כתמונה של העתקה לינארית מדרגה k.

המושג של יריעה משוכנת גזירה ברציפות מהווה הכללה מושג של תת-מרחב, וכל אחת מארבע הדרכים להגדיר תת-מרחב מקבלת תרגום בשפה של יריעות.

יריעות C^1 משוכנות

11.1 הגדרה – יהיו $k, n \in \mathbb{N}$, $k \leq n$ (המקרה $k < n$ הוא המעניין יותר). תת קבוצה $M \subseteq \mathbb{R}^n$ נקראת **יריעה C^1 ממימד k** אם לכל נקודה $a \in M$ קיימות קבוצות פתוחות $U, V \subseteq \mathbb{R}^n$ כך ש- $a \in U$ ו- $V \cap \mathbb{R}^k \times \{0\} \neq \emptyset$, וכן פונקציה הפיכה ורגולרית $C^1 \ni f: U \to V$ כך שמתקיים:

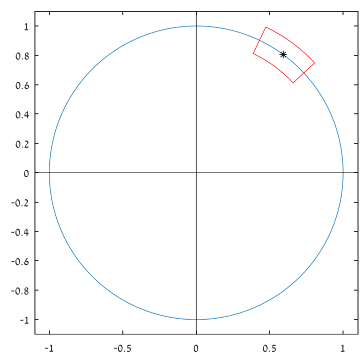
$$f(M \cap U) = V \cap (\mathbb{R}^k \times \{0_{n-k}\}) = \{x \in V \mid x_{k+1} = \dots = x_n = 0\}$$

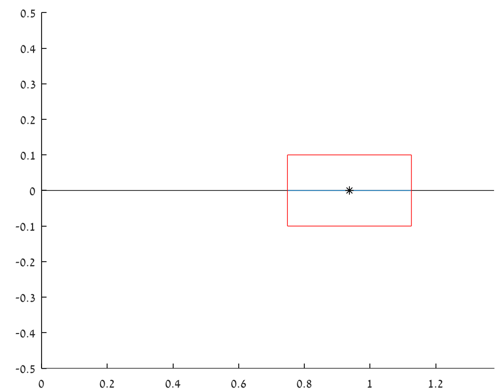
להזכירכם: פונקציה f נקראת **רגולרית** אם בכל נקודה הנגזרת שלה היא העתקה לינארית על (זאת אומרת שהדרגה של מטריצת היעקוביאן שווה למספר השורות). במקרה של ההגדרה לעיל, כיון ש-f היא העתקה בין קבוצות פתוחות באותו מימד, f היא **רגולרית** אם ורק אם הנגזרת הפיכה בכל נקודה (בהמשך נצטרך גם העתקות רגולריות בין ממדים שונים). פונקציה גזירה ברציפות הפיכה ורגולרית נקראת **דיפאומורפיזם**. ממשפט הפונקציה ההפוכה, לדיפאומורפיזם יש פונקציה הפוכה שגם היא דיפאומורפיזם.

הערה – בחלקים החישוביים של הקורס אנו נתמקד בעיקר ביריעות דו-ממדיות (ב- $\mathbb{R}^3$) וביריעות חד ממדיות (ב- $\mathbb{R}^3$ וב- $\mathbb{R}^2$). כלומר אצלנו בדרך כלל $n = 2,3$ ו- $k = 1,2$. יריעה 1-ממדית נקראת **עקום גזיר ברציפות** (או **עקום חלק**) ויריעה 2-ממדית נקראת **משטח גזיר רציפות** (או **משטח חלק**).

דוגמה:

נתבונן במעגל היחידה $M=\{(x,y): x^2+y^2=1\}$. בואו נראה יחדיו מדוע M הינו עקום גזיר ברציפות, כלומר מדוע M יריעה C^1 ממימד אחד. באיור שלהלן אפשר לראות איך סביב נקודה מסוימת על M (מסומנת בכוכבית) אפשר למצוא סביבה U (מסומנת עם גבולות אדומים) ועל U אפשר להגדיר פונקציה גזירה ברציפות $f: U\to\mathbb{R}^2$ כך ש- f לוקחת את $M\cap U$ לתת קבוצה פתוחה של ציר ה-x, כלומר ההעתקה f "מיישרת" את $M\cap U$.





בדוגמא הפשוטה הזאת, אנו יכולים לבחור U שמתוארת באופן הבא:

$$U=\left\{(x,y): x>0, \operatorname{atan}\frac{y}{x}\in(\theta_1,\theta_2), \sqrt{x^2+y^2}\in(r_1,r_2)\right\}$$

כלומר בקוורדינטות פולריות אפשר לתאר את U לפי התנאים $\theta_1<\theta<\theta_2$ ו - $r_1<r<r_2$. העתקה f ש"מיישרת" את $M\cap U$ ניתנת על-ידי $(\theta, r-1)=\left(\operatorname{atan}\frac{y}{x}, \sqrt{x^2+y^2}-1\right)=f(x,y)$. נשאיר לקורא לוודא שהעתקה זו הינה חד-חד ערכית ועל מ- U על $V=f(V)=(\theta_1,\theta_2)\times(r_1-1,r_2-1)$, ושהיא רגולרית. כיון שאפשר למצוא פונקציה כזו בסביבה של כל נקודה על המעגל, אנו רואים שמעגל היחידה אכן מקיים את ההגדרה של יריעה גזירה ברציפות.

המשפט הבא מציג מספר תנאים שקולים להגדרה 11.1, ולכן בהמשך אנו נתייחס אל כל אחד מהתנאים הללו כהגדרה. ישנן סוגים שונים של יריעות במתמטיקה: למשל יריעות טופולוגיות, יריעות C^∞, יריעות אנליטיות, וכדו'. אנו נדון רק ביריעות C^1 (שנקראות גם יריעות **גזירות ברציפות**) ולכן נקרא להן, בפשטות, **יריעות**. למעשה, קיים מושג מופשט של יריעות שאינן נתונות כלל כתת קבוצה של מרחב אוקלידי $\mathbb{R}^n$ כלשהו – אנו לא נדון כלל במושג המופשט בקורס זה.

הערה: הביטוי **יריעות משוכנות** נועד להבדיל בין מושג היריעות המופשטות (אותו לא נלמד בקורס זה) לבין היריעות בהן אנו נדון, שהן תתי קבוצות של המרחב האוקלידי (הן "משוכנות" במרחב האוקלידי). כאמור, בדרך כלל לא נוסיף את מילת התיאור "משוכנות" ונאמר רק "יריעות" כיון שזהו המושג בו נתקל בקורס זה (כל הקורס שלנו מתרחש בתוך מרחב אוקלידי $\mathbb{R}^n$ כלשהו).

11.2 משפט – התנאים הבאים שקולים עבור $M\subseteq\mathbb{R}^n$:

א. M יריעה C^1 ממםד k.

ב. לכל $M\in a$ קיימת סביבה $a\in U$ וקיימת $g\in C^1(U,\mathbb{R}^{n-k})$ רגולרית כך שמתקיים:

$$M\cap U=\{x\in U | g(x)=0\}$$

ג. לכל $a \in M$ עד כדי פרמוטציה של המשתנים, קיימת סביבה $a \in V \times W$ כך ש- $V \subseteq \mathbb{R}^k$, $W \subseteq \mathbb{R}^{n-k}$ וקיימת $h \in C^1(V, \mathbb{R}^{n-k})$ כך שמתקיים $M \cap (V \times W) = graph(h)$ כאשר:
$$graph(h) = \{(x', h(x'))|x' \in V\}$$
ד. לכל $a \in M$ קיימת סביבה $a \in U$,קיימת קבוצה פתוחה $V \subseteq \mathbb{R}^k$, וקיימת $H \in C^1(V, \mathbb{R}^n)$ חד-חד ערכית, כך שמתקיים:
1. $rank\ DH = k$
2. $M \cap U = H(V) := \{H(v): v \in V\}$
3. $H^{-1}: H(V) \to V$ הינה רציפה (ביחס לטופולוגיה המושרית על $H(V)$ מ- $\mathbb{R}^n$).

פונקציה כמו בסעיף ד' נקראת **פרמטריזציה** (או **מערכת קואורדינטות**) של $M \cap U$.

<u>דוגמאות:</u> הגרף של הפונקציה $f(x) = |x|$ אינו יריעה חלקה, כי סביב הנקודה $(0,0)$ במישור אי אפשר להציגו כגרף של אחד המשתנים כפונקציה גזירה ברציפות של המשתנה השני. דוגמאות נוספות: העקומים הנתונים על-ידי המשוואות: $g(x,y) = y^2 - x^3 - x^2$, $g(x,y) = y^3 - x^2$, ו- $g(x,y) = y^2 - x^2 = 0$, לכולם יש נק' שאינה מקיימת את התנאים של יריעה גזירה ברציפות איפה ש-g לא רגולרית. מעניין להסתכל גם בעקום המוגדר על-ידי $g(x,y) = x^3 - y^3 = 0$, זהו פשוט קו ישר! $\nabla g(0,0) = 0$, כלומר איננה רגולרית, אבל בכל זאת העקום הזה הוא יריעה.

גם הקונוס $z^2 = x^2 + y^2$ הוא קבוצת אפסים של פונקציה גזירה ברציפות $g(x,y,z) = z^2 - x^2 - y^2$

אבל פונקציה זו אינה רגולרית בראשית, ואכן ניתן להראות שבנק' הראשית לא מתקיימים התנאים שמגדירים יריעה.

<u>דוגמה:</u>

תהא $S^n = \partial\mathbb{B}_{n+1} = \{x|\|x\| = 1\} \subseteq \mathbb{R}^{n+1}$ ספרת היחידה ה n-מימדית. נראה כי קבוצה זו הינה יריעה גזירה ברציפות.

<u>לפי ב'</u> – נבחר $U = \mathbb{R}^{n+1} \setminus \{0\}$ ונגדיר:
$$g: U \to \mathbb{R} \quad g(x) = \sum_{i=1}^{n+1} x_i^2 - 1$$
נשים לב כי $\nabla g = 2(x_1, \cdots, x_n)$ ולכן רגולרית ב-U. וכן מתקיים, כנדרש:
$$S^n = S^n \cap U = \{x \in U|g(x) = 0\}$$
<u>לפי ג'</u> – יהא $a \in S^n$ ויהא i אינדקס כלשהו עבורו $a_i \neq 0$. על ידי החלפת משתנים נוכל להניח ש- $i = n+1$, וללא הגבלת הכלליות נניח ש- $a_{n+1} < 0$. נבחר $V = \mathbb{B}_n$ וכן $W = (-\infty, 0)$ ונגדיר $h: V \to \mathbb{R}^1$ באופן הבא:
$$h(v_1, \cdots, v_n) = -\sqrt{1 - \sum_{i=1}^{n} v_i^2}$$
ונקבל כי אכן:
$$S^n \cap (W \times V) = \{(v, h(v))\big|v \in V\}$$

לפי ד' – נטפל במקרה $n = 2$. נניח שאנו רוצים למצוא פרמטריזציה בסביבה של הנקודה $a = (1,0,0) \in S^2$. נגדיר $V = (0,\pi) \times (-\pi/2, \pi/2) \subseteq \mathbb{R}^2$, נגדיר $U = \{(x,y,z): x > 0\}$, ו-

$$H(\theta,\phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

אז אפשר לראות ש- H פרמטריזציה של $S^2 \cap U$. היא חד-חד ערכית. הדבר היחיד שיש לבדוק זה התנאי על הדרגה. אבל:

$$J_H = \begin{pmatrix} \cos\theta\cos\phi & -\sin\theta\sin\phi \\ \cos\theta\sin\phi & \sin\theta\cos\phi \\ -\sin\theta & 0 \end{pmatrix}$$

הדטרמיננטה של המטריצה 2×2 העליונה היא $\sin\theta\cos\theta$, וזה מתאפס רק ב $\theta = \pi/2$. אבל אז למטריצה יש את הצורה :

$$J_H = \begin{pmatrix} 0 & -\sin\phi \\ 0 & \cos\phi \\ -1 & 0 \end{pmatrix}$$

והמטריצה הזו בעלת דרגה 2 לכל ערך של ϕ.

דוגמאות: – כל $U \subseteq \mathbb{R}^n$ פתוחה היא יריעה מממד n . אם $V \subset \mathbb{R}^n$ תת מרחב מממד k אזי V הינו יריעה מממד k.

הוכחת משפט 11.2:

א ⇐ ב : נניח ש - f, a, U, M כמו בהגדרה 11.1 . תהי $\pi: \mathbb{R}^n \to \mathbb{R}^{n-k}$ ההטלה $\pi(x_1, \dots, x_n) = (x_{k+1}, \dots, x_n)$, ונגדיר $g: U \to \mathbb{R}^{n-k}$ על-ידי $g = \pi \circ f$. מתקיים $g'(x) = \pi f'(x)$ וכיון ש- $f'(x)$ בעלת דרגה n, ל-$g'(x)$ דרגה $n-k$, כלומר g רגולרית. כמו כן עבור $x \in U$ מתקיים

$$x \in U \cap M \Leftrightarrow f(x) \in V \cap (\mathbb{R}^k \times \{0\}) \Leftrightarrow f(x) \in V \text{ and } \pi f(x) = 0 \Leftrightarrow x \in U \text{ and } \pi f(x) = 0$$

לכן $U \cap M = \{x \in U: g(x) = 0\}$.

ב ⇐ ג : שימו לב: הגרירה הזאת נובעת מהפעלה של משפט הפונקציה הסתומה: נתונה לנו קבוצה (יריעה) המוגדרת כקבוצת אפסים של פונקציה רגולרית, ועלינו לתארה כגרף של פונקציה, כלומר עלינו "לחלץ" חלק מהמשתנים כפונקציה של האחרים. הנה הפרטים:

נניח ש- g, a, U, M כמו בתנאי ב'. כיון ש- g רגולרית $rank\ \frac{\partial(g_1,\dots,g_{n-k})}{\partial(x_1,\dots,x_n)}(x) = n-k$. נניח בה"כ, שבנק' a כל ה $n-k$ עמודות האחרונות בלתי תלויות (הסבר: אם ה $n-k$ עמודות האחרונות לא בת"ל, אז נבחר $n-k$ עמודות אחרות שהן בת"ל. היינו מחליפים את המשתנים כך שהמשתנים המתאימים לעמודות הבת"ל יהיו האחרונים, וכאן נכנס ה-"עד כדי פרמוטציה של המשתנים" שמופיע במשפט). לכן $n-k$ העמודות האחרונות הן בת"ל בסביבה של a, ונוכל להניח (על-ידי הקטנת הסביבה) שזה מתקיים בכל U.

נכתוב $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$ ו- $a = (a', a'')$. לפי משפט הפונקציה הסתומה, קיימות סביבות $a' \in V, a'' \in W$ וקיימת $h: V \to W$ גזירה ברציפות כך ש- $V \times W \subseteq U$ וכך ש-

$$(V \times W) \cap M = \{x \in V \times W : g(x) = 0\} = \{(x', h(x')) : x' \in V\} = graph(h)$$

ג ⇐ ד : נניח שתנאי ג' מתקיים. אז $V \subseteq \mathbb{R}^k$ ונגדיר סביבה $U = V \times W$ ופונקציה $H: V \to M \cap U$ על ידי

$$H(v) = (v, h(v))$$

אז $J_H = \begin{bmatrix} I_k \\ J_h \end{bmatrix}$, ולכן $rank\ DH(v) = k$ בכל נקודה v. ברור ש-H חד-חד ערכית וגזירה ברציפות. בנוסף, מתקיים שהפונקציה ההפוכה $H^{-1}: H(V) \to V$ נתונה ע"י $H^{-1}(x,y) = x$, וניכר שהיא רציפה.

<u>ד ⇐ א :</u> נניח שתנאי ד' מתקיים. $rank\ J_H(v) = k$ בכל נק', ובה"כ נניח ש- k השורות הראשונות בת"ל. נגדיר

$$F: V \times \mathbb{R}^{n-k} \to \mathbb{R}^n$$

$$F(v,u) = H(v) + (0,u)$$

חשוב לשים לב ש $F(v,0) = H(v)$, לכן F שולחת את $V \times \{0\}$ באופן חד-חד ערכי על $M \cap U$. בפרט קיימת נקודה $b \in V \times \{0\}$, כך ש- $F(b) = a$.

מתקיים $J_F = \begin{bmatrix} J_H & \begin{matrix} 0 \\ I_{n-k} \end{matrix} \end{bmatrix}$ הפיכה. לפי משפט הפונקציה ההפוכה, קיימות סביבות $U_a \subseteq U$ של a ו- U_b של b כך ש- $F: U_b \to U_a$ הפיכה ורגולרית.

מטרתנו כעת להגדיר סביבה מתאימה $U' \subseteq U_a$ של a באופן כזה שאם ונגדיר $f = F^{-1}$ על U', אז

$$f: U' \to V' \coloneqq F^{-1}(U')$$

תהיה הפונקציה הנדרשת כמו בהגדרה 11.1 (כאשר U', V' תמלאנה את התפקיד של U, V בהגדרה). בזכות משפט הפונקציה ההפוכה מובטח לנו ש- $f \in C^1$, ונותר לנו להגדיר את U', V' כך שיתקיים

$$f(M \cap U') = f(H(V) \cap U') = \{x \in V': x_{k+1} = \cdots = x_n = 0\}$$

מהרציפות של H^{-1} על $H(V)$ נובע קיימת סביבה פתוחה $W \subseteq \mathbb{R}^n$ של a כך ש –

$$H\big((V \times \{0\}) \cap U_b\big) = (H^{-1})^{-1}\big((V \times \{0\}) \cap U_b\big) = W \cap H(V)$$

ולכן

$$\{H(v): (v,0) \in U_b\} = F\big((V \times \{0\}) \cap U_b\big) = H(V) \cap W$$

נגדיר $U' = U_a \cap W$ ו- $V' = F^{-1}(U')$. כעת $f \coloneqq F^{-1}|_{U'}: U' \to V'$ הינה פונקציה גזירה ברציפות, הפיכה ורגולרית, המקיימת

$$f(M \cap U') = f(H(V) \cap U') = F^{-1}(H(V) \cap W \cap U_a) = \{x \in V': x_{k+1} = \cdots = x_n = 0\}$$

כנדרש.

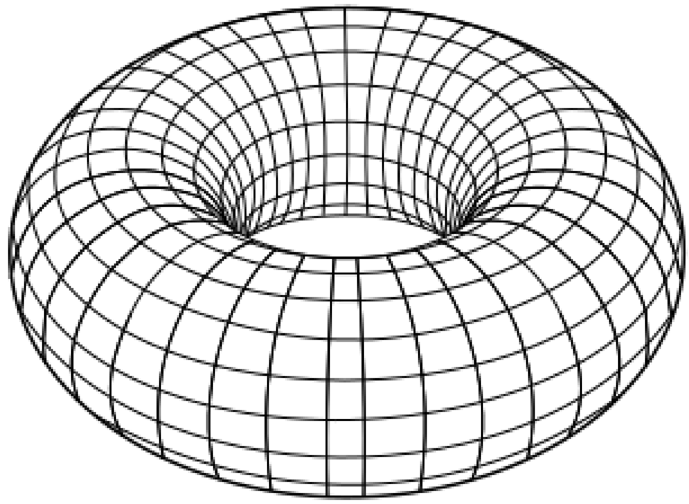
11.3 הגדרה – תהא $M \subseteq \mathbb{R}^n$ יריעה מממד k. אזי פונקציה $g: M \to \mathbb{R}^m$ נקראת **גזירה ברציפות** (ומסמנים $g \in C^1$) אם לכל $a \in M$ ולכל מערכת קואורדינטות (פרמטריזציה) ב-a מהצורה (H,V) מתקיים:

$$g \circ H \in C^1(V, \mathbb{R}^m)$$

<u>תרגיל</u> – הראו כי:

א. ההגדרה אינה תלויה בבחירת מערכת הקואורדינטות.
ב. הוכיחו, כי $g: M \to \mathbb{R}^m$ הינה C^1 ⟺ קיימת סביבה $a \in U$ ו-$G \in C^1(U, \mathbb{R}^m)$ כך ש:
$$g|_{M \cap U} = G|_{M \cap U}$$

<u>דוגמה (הטורוס):</u> הטורוס הוא הצורה הבאה:



הגוף הממשי הכי דומה לזה במציאות זוהי פנימית של גלגל, או "אבוב". יש הרואים דמיון מסוים גם למאפה מסוג "doughnut".

איך אפשר לתאר כזה דבר בצורה אנליטית?

<u>כקבוצת אפסים של פונקציה רגולרית:</u> על-ידי מעבר לקואורדינטות גליליות, אפשר לראות שהטורוס מתואר על-ידי המשוואה $z^2+(r-a)^2=b^2$ כאשר a הרדיוס הגדול ו- b הרדיוס הקטן.

<u>על-ידי פרמטריזציה:</u>

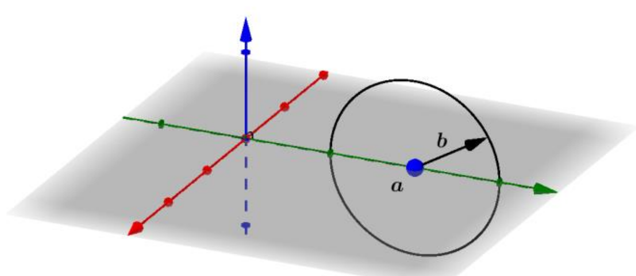
בציור לעיל

נניח כי $b<a$ כלשהם, אזי עבור $0\le u\le 2\pi$ $\quad 0\le v\le 2\pi$ נגדיר:

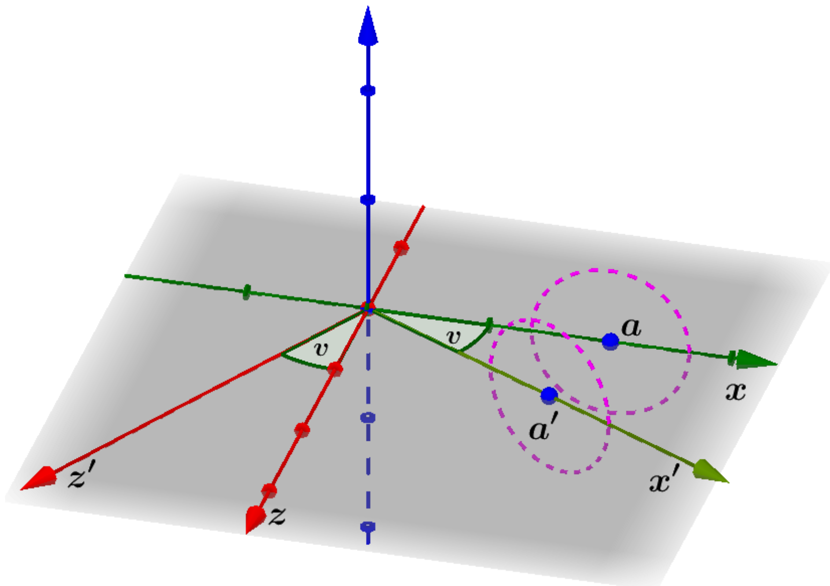
$$\varphi(u,v)=\begin{bmatrix}\cos v & -\sin v & 0\\ \sin v & \cos v & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}a+b\cos u\\ 0\\ b\sin u\end{bmatrix}=\begin{bmatrix}(a+b\cos u)\cos v\\ (a+b\cos u)\sin v\\ b\sin u\end{bmatrix}$$

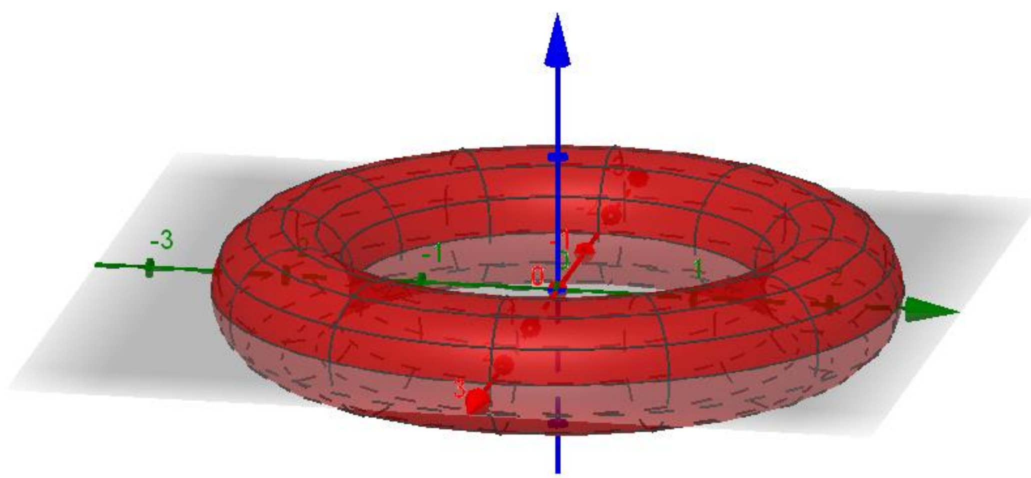
זוהי פרמטריזציה של טורוס. על מנת לקבל אותה, מבחינה אינטואיטיבית, נתבונן במישור xz, ונבחר בנקודה a על ציר ה-x. סביב נקודה זו נבנה מעגל ברדיוס b, כמתואר באיור 2. (נניח כי z הוא הציר הכחול, x הציר הירוק ו-y הוא הציר האדום).

לאחר מכן, את המעגל הנ"ל נרצה להגדיר עבור כל "סיבוב" של מערכת הצירים כשציר z הינו ציר הסיבוב. לשם כך נרצה להפעיל על נקודות מטריצת סיבוב זה סיבוב בזווית $0\le v\le 2\pi$ סביב ציר z. זו בדיוק המטריצה שמופיעה בנוסחא הנ"ל.



שלב א' בבניית הטורוס, בחירת המעגל הראשון לסיבוב (x ירוק, y כחול, z אדום)





טורוס בעל רדיוס ראשי 2 ורדיוס משני 0.5

פרק 12

העתקה $f:[a,b]\to\mathbb{R}^n$ תקרא **מסילה רציפה**. אם f גם גזירה ברציפות נאמר שזו **מסילה גזירה ברציפות**. התמונה של מסילה (כלומר הקבוצה $f([a,b])$) היא קבוצה "חד מימדית" בתוך $\mathbb{R}^n$, שתקרא **עקום** (רציף/גזיר ברציפות).

כאשר המסילה גזירה ברציפות, אפשר לחשוב על הנגזרת $f'(t)$ בכל נקודה כעל וקטור ב - $\mathbb{R}^n$, ואם $f'(t_0)\neq 0$ נאמר ש - $f'(t_0)$ הוא **הוקטור המשיק למסילה בנקודה** $\boldsymbol{f(t_0)}$. שימו לב שהפונקציה

$$t\mapsto f(t_0)+f'(t_0)(t-t_0)$$

הינה הפונקציה האפינית היחידה שמקרבת את $t\mapsto f(t)$ עם שגיאה יחסית, כלומר עם שגיאה $o(t-t_0)$ (זה נובע מהגדרת הנגזרת ומיחידות הנגזרת). מכאן אפשר לראות, שאם מתקיים התנאי $f'(t_0)\neq 0$, אז הישר $\{f(t_0)+f'(t_0)(t-t_0): t\in\mathbb{R}\}$ הוא בעצם הישר שמקרב הכי טוב את העקום המוגדר על-ידי f.

כאשר $f'(t_0)=0$ אזי לעקום שהפונקציה מגדירה אין בהכרח משיק בנקודה $f(t_0)$, גם כאשר הפונקציה הינה גזירה ברציפות (אתם תראו דוגמאות בתרגול או בתרגיל הבית).

את המקרה $n=3$ ניתן לדמיין, ואז אפשר לתת משמעות אינטואיטבית ושימושית למושג מסילה. נוכל לחשוב על המשתנה $t\in[a,b]$ כעל זמן, ועל $\mathbb{R}^3$ כעל המרחב התלת מימדי "בו אנו חיים", ואז הווקטור $f(t)$ מתאר את המיקום של חלקיק כלשהו בזמן t. העקום כולו הוא שרטוט של המסלול שעובר החלקיק, ואילו הנגזרת $f'(t)$ זהו וקטור המהירות של החלקיק בזמן t.

מטרתנו ביתרת הפרק הזה היא להגדיר את מושג המרחב המשיק עבור יריעות.

המרחב המשיק

מרחב וקטורי ניתן לתיאור כאוסף של וקטורים, וניתן לתאר אותו גם כתמונה של העתקה לינארית, וכן כגרעין של העתקה לינארית, וגם באופן מפורש כמרחב הנפרש על-ידי קבוצת וקטורים מסוימים. אנו ניתן עכשיו מספר הגדרות שונות של **המרחב המשיק** של יריעה. בהמשך נראה שהן שקולות.

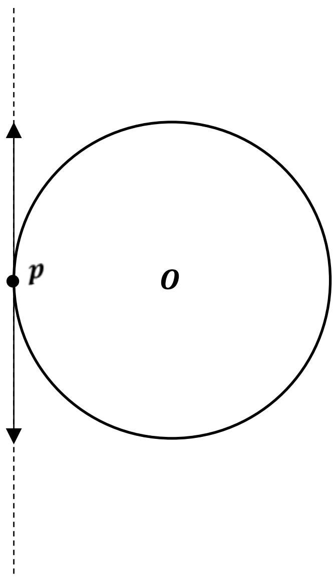
לפני שנגדיר מהו המרחב המשיק של יריעה, נאמר מה זה צריך להיות מבחינה אינטואיטבית. המרחב המשיק בנקודה כלשהי p צריך להיות אוסף כל הכיוונים שבהם אפשר לנוע כאשר אנו נעים על היריעה דרך הנקודה p. לכן, אם היריעה שלנו היא פשוט המרחב כולו (או תת קבוצה פתוחה במרחב) אנו יכולים לנוע בכל הכיוונים, והמרחב המשיק צריך להיות המרחב כולו. אנו אוהבים לחשוב על המרחב המשיק בנקודה p כאילו הוא "יושב" על הנקודה p. כדי שהעניין יהיה מדויק ומתמטי, דרושות לנו כמה הגדרות.

12.1 הגדרה – לכל $p\in\mathbb{R}^n$ נגדיר את **המרחב המשיק ל-$\mathbb{R}^n$ בנקודה p** להיות:

$$T_p(\mathbb{R}^n)=\{(p,v)|v\in\mathbb{R}^n\}$$

12.1.1 סימון - $v_p = (p,v)$ ולפעמים אפילו מזהים עם v. אם $V \subseteq \mathbb{R}^n$ תת מרחב וקטורי, אזי $V_p =$ $\{(p,v) | v \in V\} \subseteq T_p(\mathbb{R}^n)$ וניתן לזהות אותו עם המרחב המוזז:

$$V_p = p + V = \{p + v | v \in V\}$$



איור 10 – דוגמה למרחב המשיק לנקודה p על מעגל היחידה (המרחב המשיק הוא הקו המקווקו).

12.2 הגדרה – תהא $M \subset \mathbb{R}^n$ יריעה C^1 ממד k, ותהא $p \in M$ נקודה ביריעה.
המרחב המשיק ל-M ב-p מוגדר להיות:

$$\left\{ p + \gamma'(t_0) \middle| \begin{matrix} \overset{\in C_1}{\gamma} : (a,b) \to M \\ t_0 \in (a,b) \quad \gamma(t_0) = p \end{matrix} \right\}$$

12.2.1 סימון - $T_p(M) = \left(\left\{ \gamma'(t_0) \middle| \begin{matrix} \overset{\in C_1}{\gamma} : (a,b) \to M \\ t_0 \in (a,b) \quad \gamma(t_0) = p \end{matrix} \right\} \right)_p$

12.3 הגדרה - $T_p(M) = \left[Im\big(DH(q)\big) \right]_p$ כאשר $H: V \subseteq \mathbb{R}^k \to \mathbb{R}^n$ פרמטריזציה (מערכת קואורדינטות) בסביבת p ו- $q \in V$ מקיימת $H(q) = p$.

<u>דוגמה:</u>

עבור מערכת קואורדינטות $H: \left(\frac{\pi}{2}, \frac{3\pi}{2}\right) \to \mathbb{R}^2$ המוגדרת על ידי:

$$H(t) = (\cos t, \sin t)$$

(קרי, פרמטריזציה של חצי מעגל), נבחר בנקודה $p = (-1,0) = H(\pi)$ ונשים לב כי:

$$DH(\pi) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}_{t=\pi} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} : \mathbb{R} \to \mathbb{R}^2$$

ונשים לב כי:

$$Im\big(DH(0)\big) = \operatorname{span}\left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \Longrightarrow T_p(M) = p + \operatorname{span}\left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

וזה אכן, מרחב המתיישב אינטואיטיבית עם ההגדרה שלנו למרחב המשיק בנקודה הנ"ל (במקרה זה, הנ"ל דומה למרחב האפיני באיור 10).

12.4 הגדרה - תהא $p \in M$ כך שקיימת U עבורה $p \in U$. נסמן $p = (a,b)$ (כך ש-$a \in \mathbb{R}^k$ וכן $b \in \mathbb{R}^{n-k}$), עבורה:

$$U \cap M = \left\{ \big(v, h(v)\big) \middle| v \in \underset{\subseteq \mathbb{R}^k}{V} \right\} \quad h \in C^1(V, \mathbb{R}^{n-k})$$

אזי נגדיר:

$$T_p(M) = graph \begin{pmatrix} \text{הקירוב האפיני של} \\ h \text{ בנקודה } a \end{pmatrix}$$

כלומר:

$$T_p(M) = \left\{ \big(v, h(a) + Dh(a)(v - a)\big) \middle| v \in \mathbb{R}^k \right\} \subset \mathbb{R}^n$$

שימו לב שזהו מקרה מיוחד של הגדרה 12.3 כך ש-$H: V \to \mathbb{R}^n, H(v) = \big(v, h(v)\big)$.

12.5 הגדרה – אם $M \cap U = \{x | g(x) = 0\}$ עבור $g \in C^1(\mathbb{R}^n, \mathbb{R}^{n-k})$ רגולרית, אזי נגדיר:

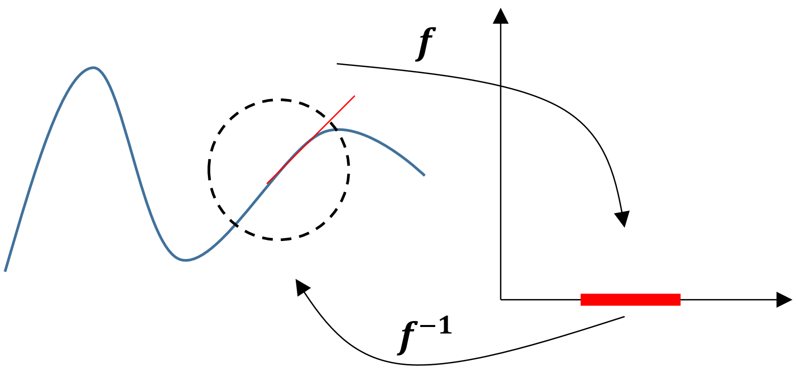
$$T_p(M) = [ker\ Dg(p)]_p = \{x \in \mathbb{R}^n | Dg(p)(x-p) = 0\}$$

12.6 הגדרה – עבור $p \in U$, כך שמוגדרת $f \in C^1(U, \mathbb{R}^n)$ גזירה ברציפות, הפיכה ורגולרית (דיפאומורפיזם) עבורה:

$$f(M \cap U) = V \times \{0\} \subseteq \mathbb{R}^k \times \mathbb{R}^{n-k}$$

נגדיר:

$$T_p(M) = \{p + Df(p)^{-1}(v, 0) | v \in \mathbb{R}^k\}$$



12.7 משפט – כל ההגדרות הנ"ל שקולות.

איור 11 תיאור סכמטי של הגדרה 13.6 כתיאור המרחב המשיק כמרחב וקטור

<u>הערה:</u> כשאנו אומרים שההגדרות שקולות, הכוונה היא שהמרחבים המוגדרים בהגדרות השונות שווים. שימו לב שמכאן נובע גם שהמרחב המשיק אינו תלוי בפרמטריזציה H, או בפונקציה g שהיריעה היא קבוצת האפסים שלה

<u>הוכחה:</u>

נוכיח רק ש- הגדרה 12.2 $\Leftrightarrow$ הגדרה 12.3 $\Leftrightarrow$ הגדרה 12.5. הגדרה 12.4 היא מקרה פרטי של 12.3, והשקילות להגדרה 12.6 מושארת כתרגיל.

<u>הגדרה 12.3 $\Leftrightarrow$ הגדרה 12.5</u> : נניח כי H, g נתונות כמתואר בהגדרות המתאימות. ותהא:

$$p \in U \cap M = \{x \in U | g(x) = 0\}$$

ונניח, כאמור, כי מתקיים בנוסף $H: V \subseteq \mathbb{R}^k \to U \cap M$ כך שקיים $q \in V$ עבור $H(q) = p$. לכן, מתקיים:

$$g \circ H \equiv 0$$

כלומר:

$$Dg(p) \circ DH(q) = 0 \Longrightarrow Im\big(DH(q)\big) \subseteq \ker\big(Dg(p)\big)$$

אבל שניהם ממימד k ולכן מתקיים שוויון.

<u>הגדרה 12.2 $\Leftrightarrow$ הגדרות 12.3 ו- 12.5</u> : תהא $p = \gamma(t_0)$ עבור $\gamma: (a, b) \to M$. אז $g \circ \gamma(t) \equiv 0$, ולכן $Dg(p)\gamma'(t_0) = 0$, ואנו ממקבלים שהמרחב המשיק לפי הגדרה 12.5 מכיל את המרחב המוגדר בהגדרה 12.2. מצד שני, אם H , V , $q \in V$ כמו קודם, ואם $(p, v) \in T_p(M)$ לפי הגדרה 12.3, כלומר $v \in Im\ DH(q)$, אז נקבל מההגדרה ש: $v = DH(q)u$ כאשר $u \in \mathbb{R}^k$. נגדיר $\gamma(t) = H(q + tu)$ בסביבת $t_0 = 0$. אז $\gamma(t) \in M$ לכל t, ומתקיים $\gamma'(t_0) = DH(q)u$, ולכן הקבוצה המוגדרת בהגדרה 12.3 מוכלת בקבוצה המוגדרת בהגדרה 12.2 . כיון שהגדרות 12.3 ו-12.5 שקולות ומגדירות את אותו מרחב, זה מוכיח את השקילות של 12.2 לשתי הגדרות אלו.

<u>תרגיל:</u> הוכיחו שהגדרה 12.6 שקולה ליתר ההגדרות.

12.8 הגדרה – יהא $M \subseteq \mathbb{R}^3$ משטח. **נורמל** ל- M בנק' $p \in M$ הוא כל וקטור שונה מאפס הניצב ל- $T_p(M)$.

נשים לב שמישור משיק למשטח המשוכן במרחב התלת-ממדי נקבע על-ידי הנורמל, ומצד שני הנורמל מוגדר ע"י מישור משיק עד-כדי קבוע כפלי (כלומר עד כדי כיוון "למעלה"/"החוצה" או "למטה"/"פנימה"). נשים לב לעוד עובדה שימושית: אם M שווה לקבוצת האפסים של פונקציה רגולרית g בסביבה של p, אז הנורמל חייב להיות כפולה של $\nabla g(p)$.
הסיבה: הן הנורמל $\hat{n}$ והן הגרדיאנט $\nabla g(p)$ שניהם מקיימים

$$T_p(M) = \{x \in \mathbb{R}^3 : \nabla g(p) \cdot (x-p) = 0\} = \{x \in \mathbb{R}^3 : \hat{n} \cdot (x-p) = 0\}$$

כיון שהניצב למרחב דו-מימדי במרחב התלת-מימדי הוא ממימד אחד, הנורמל והגרדיאנט קו-לינאריים.

יריעות עם שפה

מושג היריעות הינו יפה וחשוב, אך בפועל אנו נפגוש באובייקטים שלא עונים בדיוק על ההגדרה של יריעה. למשל: קטע סגור, או תמונה של קטע סגור על-ידי מסילה (עקום). או: חצי ספרה. המושג "יריעה עם שפה" הוא מספיק כללי כך שכמעט כל קבוצה שתעניין אותנו בפועל תהיה איחוד סופי של יריעות עם שפה.

נגדיר את **חצי המרחב ה-k מימדי** להיות:

$$\mathcal{H}_k = \{x \in \mathbb{R}^k : x_k \geq 0\}$$

12.9 הגדרה – תת קבוצה $M \subseteq \mathbb{R}^n$ נקראת **יריעה C^1 ממימד k עם שפה** אם לכל $a \in M$ קיימת סביבה $a \in U$, קיימת קבוצה פתוחה $V \subseteq \mathbb{R}^k$, וקיימת $H \in C^1(V, \mathbb{R}^n)$ חד-חד ערכית, כך שמתקיים:

1. $rank\ DH = k$
2. $M \cap U = H(V \cap \mathcal{H}_k) := \{H(v) : v \in V \cap \mathcal{H}_k\}$
3. $H^{-1} : M \cap U \to V \cap \mathcal{H}_k$ הינה רציפה (ביחס לטופולוגיה המושרית על $M \cap U$ מ- $\mathbb{R}^n$).

פונקציה H נקראת **פרמטריזציה** שלM (בסביבה של a).

שימו לב שישי שני סוגים של נקודות ביריעה עם שפה: נקודות שבאות מהפנים של $\mathcal{H}_k$ ואלו בעצם מקיימות את התנאי שנקודות ביריעות חלקות אמורות לקיים, ומנגד נקודות שבאות מהשפה של $\mathcal{H}_k$, והן נמצאות "על הקצה" של היריעה.

12.10 הגדרה – תהי $M \subseteq \mathbb{R}^n$ יריעה C^1 ממימד k עם שפה. נקודה $a \in M$ תקרא **נקודת שפה** של M אם $a = H(b)$ כאשר b נקודת שפה של $\mathcal{H}_k$, כלומר $b \in V \cap \{x \in \mathbb{R}^k : x_k = 0\}$. את אוסף כל נקודות השפה מסמנים ∂M. שאר הנקודות נראות **נקודות הפנים** של M, ומסמנים אותן $int(M)$.

<u>**היזהרו!**</u> למרות הסימונים <u>הזההים</u> (!!!) המשמעות של שפה (או נקודה פנימית) של יריעה עם שפה שונה מהמשמעות של שפה טופולוגית. לכן לפעמים מכניסים סימונים שונים כמו $bd(M)$ עבור אחת השפות כדי להבחין ביניהן. לצערנו, אין מוסכמה אחת מוסמכת, וצריך תמיד להיות ערים להקשר ולהבין למה מתכוונים.

אנו לא נוכיח את העובדות הבאות: השפה של יריעה אינה תלויה בפרמטריזציה, כלומר נקודה ביריעה היא נקודת שפה או לא בלי תלות באיזו פרמטריזציה אנו משתמשים כדי לתאר סביבה של הנקודה. כמו כן, השפה של יריעה חלקה עם שפה ממימד k היא יריעה חלקה ממימד $k-1$.

<u>דוגמאות:</u> עקום, מלבן סגור במרחב, כדור סגור, חצי ספרה.

<u>תרגיל:</u> הראו שהדוגמאות הנ"ל מקיימות את ההגדרה של יריעה עם שפה, וקבעו: מהו k?

<u>**הערה טכנית חשובה:**</u> בהגדרה 12.9 אפשר להחליף את התנאי $H \in C^1(V, \mathbb{R}^n)$ בתנאי $H \in C^1(V \cap \mathcal{H}_k, \mathbb{R}^n)$, כלומר הפרמטריזציה יכולה להיות מוגדרת וגזירה ברציפות רק על ה"חצי סביבה" $V \cap \mathcal{H}_k$. זו אכן הצורה הנוחה יותר הנפוצה בפועל. לדוגמה, כדי להציג סביבה של הנקודה $(-1,0)$ בחצי העיגול העליון $\{(x,y) : y \geq 1, x^2 + y^2 = 1\}$ כתמונה של פרמטריזציה לפי ההגדרה, אפשר לבחור $V = (-\varepsilon, \varepsilon)$, ו- $H : (-\varepsilon, \varepsilon) \to \mathbb{R}^2$ גזירה ברציפות הנתונה על-ידי

הנוסחה $H(t) = (\cos(\pi - t), \sin(t))$, ואז ההגדרה מסתכלת מה קורה ב - $V \cap \mathcal{H}_1 = (-\varepsilon, \varepsilon) \cap [0, \infty) = [0, \varepsilon)$. אבל בפועל יותר נוח פשוט להגדיר את H לפי אותה נוסחה על הקטע $[0, \varepsilon)$, ולשים לב שהפרמטריזציה H שהגדרנו היא בעצם גזירה ברציפות על הקטע הזה, במובן הזה ש - $H = (H_1, H_2)$ ו - H_1, H_2 שתיהן פונקציות גזירות ברציפות על הקטע $[0, \varepsilon)$, כלומר גזירות ברציפות במובן הרגיל בקטע הפתוח $(0, \varepsilon)$, וכך שהנגזרת החד צדדית קיימת בנקודה 0, והנגזרת המתקבלת רציפה על כל הקטע.

למעשה, בהגדרה 12.9 להחליף את V ב - $V \cap \mathcal{H}_k$, כאשר כל שעלינו לעשות הוא להבין מה הכוונה בגזירות ברציפות בתחום "חצי פתוח" מהסוג $V \cap \mathcal{H}_k$. כלומר עלינו להגדיר מה זה $C^1(V \cap \mathcal{H}_k, \mathbb{R}^n)$ (עד עכשיו הגדרנו גזירות רק בתחומים פתוחים). בעצם זה לא מסובך: נאמר ש- $H \in C^1(V \cap \mathcal{H}_k, \mathbb{R}^n)$ אם H רציפה, ואם כל הנגזרות החלקיות(מסדר ראשון) של H קיימות ורציפות בתחום זה. הנקודה העדינה היחידה היא איך להגדיר את הנגזרת החלקית $\frac{\partial H}{\partial x_k}$ ב- $V \cap \partial\mathcal{H}_k$, שהוא ה"קצה" של התחום הכולל נק' מהסוג $(x', 0)$ כאשר $x' \in \mathbb{R}^{k-1}$. ובכן, שם פשוט מגדירים את $\frac{\partial H}{\partial x_k}$ להיות הנגזרת החד צדדית: $\frac{\partial H}{\partial x_k}(x', 0) = \lim_{0<h\to 0} \frac{H(x',h) - H(x',0)}{h}$.

<u>**תרגיל:**</u> הוכיחו שאם V קבוצה פתוחה ב - $\mathbb{R}^k$ כך ש - $V \cap \mathcal{H}_k \neq \emptyset$, ואם $H \in C^1(V \cap \mathcal{H}_k, \mathbb{R}^n)$, אז אפשר להמשיך את H לפונקציה $\tilde{H} \in C^1(V, \mathbb{R}^n)$ כך ש - $\tilde{H}\big|_{V\cap\mathcal{H}_k} = H$. (רמז: חישבו קודם על המקרה $k = 1$).

בפועל, משתמשים גם בפרמטריזציות גלובליות, למשל את חצי העיגול העליון אפשר לתאר על-ידי פרמטריזציה גלובלית $H: [-1,1] \to \mathbb{R}^2$ הנתונה על-ידי $H(t) = (\cos(t), \sin(t))$. קל לראות שקיום של פרמטריזציה מהסוג הזה (המקיימת את כל התנאים) מבטיחה שחצי העיגול הוא יריעה עם שפה.

בקורס הזה אנו לא נתמקד בפרטים הטכניים שהופיעו בהערה לעיל, ומעבוד עם פרמטריזציות בלי חרדות וחששות.

<u>**לתרגול/תרגיל בית:**</u> נוסחאות מפורשות עבור מישור משיק ונורמל למשטח, בהינתן משטח נתון על-ידי פרמטריזציה או הצגה כגרף, או כאוסף פתרונות של מערכת משוואות C^1 . דוגמאות קונקרטיות. דוגמאות של יריעה עם שפה מה כן/מה לא.

פרק 13

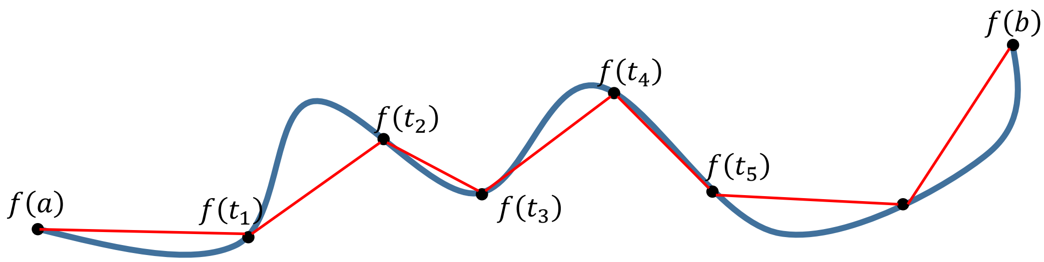
13.1 הגדרה – תהא $f:[a,b]\to\mathbb{R}^n$ מסילה רציפה, ותהא $P=\{a=t_0<\cdots<t_k=b\}$ חלוקה של $[a,b]$. נגדיר:

$$l(f,P)=\sum_{i=1}^{k}\|f(t_i)-f(t_{i-1})\|$$

ונגדיר:

$$l(f)=sup_P l(f,P)$$

אזי אם $l(f)<\infty$ נאמר כי המסילה $f:[a,b]\to\mathbb{R}^n$ היא **בעלת אורך** ו-$l(f)$ ייקרא **האורך** של המסילה.



איור 12 – דוגמה לעקום (בכחול),חלוקה של העקום (הנקודות השחורות) והסכום $l(f,P)$ (הקטעים האדומים)

13.2 משפט – אם $f\in C^1([a,b])$, אזי למסילה $f:[a,b]\to\mathbb{R}^n$ יש אורך.

<u>הוכחה:</u>

נתחיל את ההוכחה עם טיעון מהיר ולא מדויק (זהירות!).

תהא $P=\{a=t_0<t_1<\cdots<t_k=b\}$ חלוקה. אזי מתקיים:

$$\sum_{i=1}^{k}\|f(t_i)-f(t_{i-1})\|=\sum_{i=1}^{k}\|f'(t_{i-1})(t_i-t_{i-1})+o(t_i-t_{i-1})\|$$

$$=\sum_{i=1}^{k}\|f'(t_{i-1})\|(t_i-t_{i-1})+\sum_{i=1}^{k}o(t_i-t_{i-1})$$

ונשים לב, שכאשר $\max_i(t_i-t_{i-1})\to 0$ נקבל כי:

$$\sum_{i=1}^{k}\|f'(t_{i-1})\|(t_i-t_{i-1})\to\int_a^b\|f'(t)\|dt$$

ומאידך:

$$\sum_{i=1}^{k}o(t_i-t_{i-1})\overset{(\star)}{=}\sum_{i=1}^{k}(t_i-t_{i-1})\,o(1)=o(1)(b-a)\to 0$$

כאשר ($\star$) דורש הצדקה נוספת (ייתכן והצדקה נוספת תינתן בכיתה, תחת ההנחה ש-$f \in C^2$. בקיצור, אם $f \in C^2$,
אז את הביטויים לשגיאה אפשר להחליף ב- $\frac{1}{2}|f''(\xi_i)|(t_i - t_{i-1})^2$ והסכום עליהם קטן מ-
$\frac{1}{2}M \max|t_i - t_{i-1}| \sum(t_i - t_{i-1}) = \frac{b-a}{2}M \max|t_i - t_{i-1}|$).

אנו נעבור להוכחה מדויקת.

13.3 טענה – תחת כל תנאי המשפט, מתקיים $l(f,P) \le \int_a^b \|f'(t)\|dt$ לכל חלוקה P.

<u>הוכחה:</u>

נראה עבור $P = \{a < b\}$. מתקיים:

$$\|f(b) - f(a)\|^2 = \left\| \begin{pmatrix} \int_a^b f_1'(t)dt \\ \int_a^b f_2'(t)dt \\ \vdots \\ \int_a^b f_n'(t)dt \end{pmatrix} \right\|^2 = \sum_{j=1}^n \left(\int_a^b f_j'(t)dt\right)^2 = \sum_{j=1}^n \left(\int_a^b f_j'(s)ds\right)\left(\int_a^b f_j'(t)dt\right)$$

$$= \int_a^b \sum_{j=1}^n \left(\int_a^b f_j'(s)ds\right) f_j'(t)dt \overset{\substack{\text{א"ש} \\ \text{קושי–שוורץ}}}{\le} \int_a^b \underbrace{\left[\sum_{j=1}^n \left(\int_a^b f_j'(s)ds\right)^2\right]^{\frac{1}{2}}}_{\|f(b)-f(a)\|} \cdot \underbrace{\left[\sum_{j=1}^n f_j'(t)^2\right]^{\frac{1}{2}}}_{\|f'(t)\|} dt$$

ולכן, בחלוקה ב-$\|f(b) - f(a)\|$ נקבל את אי השוויון:

$$\|f(b) - f(a)\| \le \int_a^b \|f'(t)\|dt$$

ולכן נסיק כי באופן כללי עבור חלוקה כלשהי מתקיים:

$$\|f(t_i) - f(t_{i-1})\| \le \int_{t_{i-1}}^{t} \|f'(t)\|dt$$

ואם נסכום לכל i נקבל בדיוק את הטענה הכללית.

עתה, לכל $\tau \in (a,b)$ נגדיר:

$$L(\tau) = \begin{pmatrix} \text{האורך של} \\ f: [a,\tau] \to \mathbb{R}^n \end{pmatrix}$$

וכאמור $L(\tau)$ קיים וסופי לכל τ, מהטענה.

יהא $h > 0$ ונשים לב, כי:

$$\frac{\|f(\tau + h) - f(\tau)\|}{h} \le \frac{L(\tau + h) - L(\tau)}{h} \le \frac{1}{h}\int_{\tau}^{\tau+h} \|f'(t)\|dt \xrightarrow{h\to 0} \|f'(t)\|$$

(הערה: מדוע אי-השוויון נכון? ניתן להצדיקו – עשו זאת!)

אבל אגף ימין שואף גם כן ל- $\|f'(t)\|$, ולכן L גזירה מימין בכל נקודה. באופן דומה, הגבול קיים עבור $0 < h \to 0$. כלומר, קיבלנו כי $L(\tau)$ גזירה ומתקיים:

$$L'(t) = \|f'(t)\|$$

לכן, לפי נוסחת ניוטון-לייבניץ:

$$\boxed{L(b) = \underbrace{L(a)}_{=0} + \int_a^b L'(t)dt = \int_a^b \|f'(t)\|dt}$$

וזה סוף ההוכחה של משפט 13.2.

13.4 הגדרה – עקום פשוט וחלק זו תת קבוצה $C \subseteq \mathbb{R}^n$ שהינה תמונה של מסילה גזירה ברציפות:

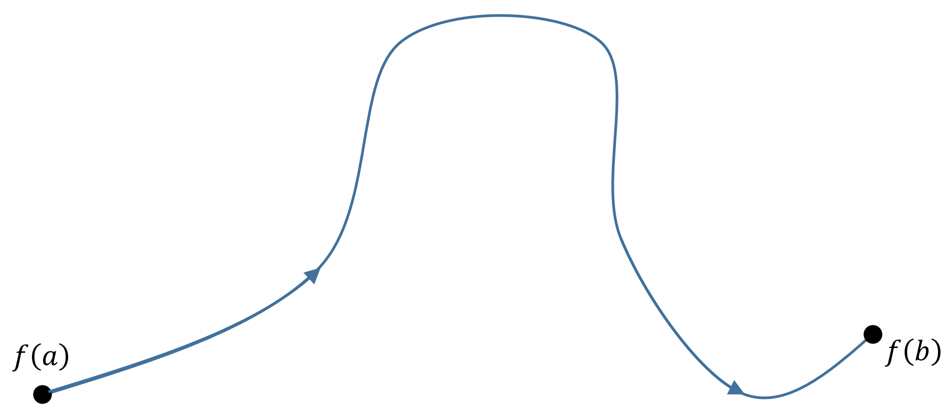
$$f: [a,b] \to \mathbb{R}^n$$

המקיימת:

א. f חד-חד ערכית למעט אולי בנקודות a,b, כלומר למעט אפשרות שבה $f(a) = f(b)$.
ב. $f'(t) \neq 0$ לכל $t \in [a,b]$.

תרגיל: תהי $f: [a,b] \to \mathbb{R}^n$ מסילה גזירה ברציפות וחד חד ערכית (כלומר $f(t_1) \neq f(t_2)$ לכל $t_1 \neq t_2$) המקיימת גם $f'(t) \neq 0$ לכל $t \in [a,b]$. הוכיחו ש- $f(a,b)$ (כלומר העקום המוגדר כתמונה של f) הינה יריעה גזירה ברציפות.

13.5 הגדרה – עקום מכוון הינו עקום, עבורו מבחינים בין $f(a)$ כנקודת ההתחלה לבין $f(b)$ כנקודת הסוף.



13.6 הגדרה – אם $f(a) = f(b)$, אזי C נקרא **עקום סגור**.

<u>הערה</u> – קל לטפל באיחוד סופי של עקומים פשוטים וחלקים, ואנו לא נעשה זאת בשיא הפורמליות.

<u>תזכורת</u> – הגדרנו את:

$$l(f) = \sup_P l(f,P) = \int_a^b \|f'(t)\|dt$$

13.7 הגדרה – אם C עקום פשוט וחלק, נגדיר:

$$l(C) = l(f)$$

עבור f מההגדרה הקודמת.

13.8 טענה - $l(C)$ כפי שהוגדר ב-13.7 מוגדר היטב.

<u>הוכחה:</u>

דרך ההוכחה, תהיה לנסות להגדיר חלוקות של $[a,b]$ המתאימות לחלוקה של C. ההוכחה מושארת כתרגיל.

<u>דיון על פרמטריזציית "אורך קשת"</u>

תהא $f\colon [a,b] \to \mathbb{R}^n$ פרמטריזציה של עקום פשוט וחלק C. בפרק הקודם ראינו כי $L\colon [a,b] \to [0, l(C)]$ המוגדרת על ידי $L(t) = l\left(f|_{[a,t]}\right)$ גזירה:

$$L'(t) = \|f'(t)\| \Longrightarrow \begin{matrix} & L \in C^1 \\ \forall t & L'(t) \neq 0 \end{matrix}$$

וממשפט הפונקציה ההפוכה קיימת $\sigma\colon [0, l(C)] \to [a,b]$ גזירה ברציפות שהיא פונקציה הפיכה ל-L עבורה מתקיים:

$$\sigma'(s) = \frac{1}{L'\left(L^{-1}(s)\right)} = \frac{1}{\left\|f'\left(\sigma(s)\right)\right\|} (\star)$$

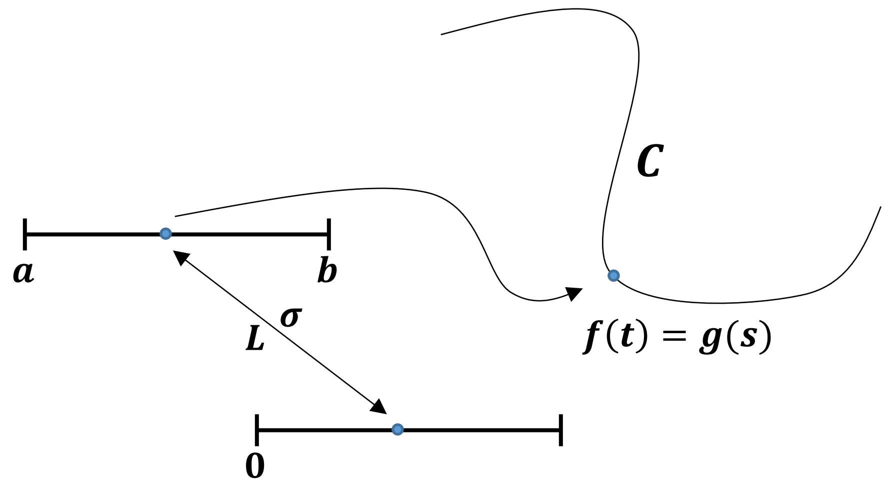
ומכאן שניתן להגדיר פרמטריזציה חדשה על ידי:

$$g\colon [0, l(C)] \to C \subseteq \mathbb{R}^n$$

כך שמתקיים:

$$g(s) = f\left(\sigma(s)\right) \quad \begin{matrix} s = L(t) \\ t = \sigma(s) \end{matrix}$$

כלומר, $g(s)$ זו הנקודה על העקום שמרחקה (כפי שהוא נמדד על העקום) מנקודת ההתחלה $f(a) = g(0)$ הוא בדיוק s, כלומר אורך המסילה $g|_{[0,s]}\colon [0,s] \to \mathbb{R}^n$ הוא בדיוק s. במילים אחרות, הפרמטר של הפרמטריזציה הוא האורך של העקום!



ולכן:

$$l(C) = \int_a^b \|f'(t)\| dt = \int_0^{l(C)} \left\|f'\big(\sigma(s)\big)\right\| \sigma'(s) ds = \int_0^{l(C)} \|g'(s)\| ds$$

$$l(C) = \int_0^{l(C)} \|g'(s)\| ds = \int_0^{l(C)} ds$$

כאשר את המעבר האחרון נצדיק על ידי:

$$\|g'(s)\| = \left\|f'\big(\sigma(s)\big)\right\| |\sigma'(s)| \overset{(\star)}{=} 1$$

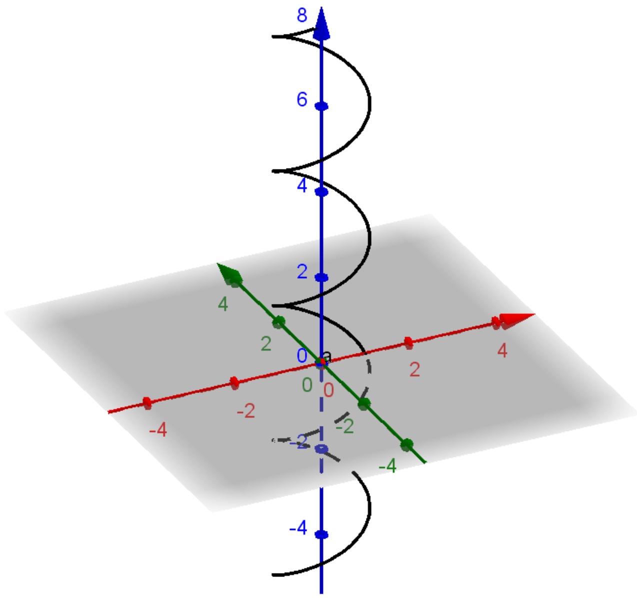
השיקול הזה גם הוא מראה את המוגדרות היטב של אורך העקום, כאשר מדובר על עקום פשוט וחלק.

13.9 הגדרה – פרמטריזציה g של עקום עבורה $\|g'(s)\| = 1$ נקראת **פרמטריזציית אורך קשת.**

לתרגול/תרגיל בית: אי תלות של אורך עקום פשוט וחלק בפרמטריזציה (כלומר במסילה המגדירה אותו).

פרק 14

<u>דוגמה:</u> הקפיץ המתואר באיור 1 נתון על ידי הפונקציה:



איור 13 - תיאור של קפיץ (Helix)

$$f:[0,L]\to\mathbb{R}^3$$

$$f(t)=(a\cos\omega t\,,a\sin\omega t\,,bt)$$

$$f'(t)=(-\omega a\sin\omega t\,,\omega a\cos\omega t\,,b)$$

נוכל לחשב את אורך הקפיץ על ידי:

$$l(C)=l(f)=\int_0^L\|f'(t)\|dt$$

$$=\int_0^L\sqrt{\omega^2a^2\sin^2\omega t+\omega^2a^2\cos^2\omega t+b^2}dt$$

$$=L\sqrt{a^2\omega^2+b^2}$$

נשים לב כי $\|f'(t)\|=\sqrt{a^2\omega^2+b^2}$ ולכן נסיק כי זו **אינה** פרמטריזציית אורך קשת כפי שהגדרנו בפרק הקודם.

נגדיר פרמטריזציה חדשה על ידי:

$$g:\left[0,L\sqrt{a^2\omega^2+b^2}\right]\to C\subset\mathbb{R}^3\quad g(s)=f\big(\sigma(s)\big)$$

כאשר $\sigma(s)=\frac{s}{\sqrt{a^2\omega^2+b^2}}$. במקרה זה, נשים לב כי:

$$\|g'(s)\|=\left\|f'\big(\sigma(s)\big)\sigma'(s)\right\|=\sqrt{a^2\omega^2+b^2}\frac{1}{\sqrt{a^2\omega^2+b^2}}=1$$

וקיבלנו כי g הינה פרמטריזציית אורך קשת מתאימה לקפיץ שמתואר בבעיה.

<u>דוגמה:</u>

תהא הפונקציה:

$$f(t)=\left(\cos t\,,\sin t\,,\frac{1}{2}t^2\right)$$

ונשים לב כי:

$$\|f'(t)\|=\|(-\sin t\,,\cos t\,,t)\|=\sqrt{1+t^2}$$

כלומר:

$$l(f)=\int_0^L\sqrt{1+t^2}dt=\frac{1}{2}\left[L\sqrt{1+L^2}+\ln\left(L+\sqrt{1+L^2}\right)\right]$$

אם נרצה למצוא עבור פרמטריזציה זו את פרמטריזציית אורך הקשת המתאימה:

$$\sigma:[0,l(L)]\to[0,L]\quad g(s)=f\big(\sigma(s)\big)\quad\|g'(s)\|=\left\|f'\big(\sigma(s)\big)\right\||\sigma'(s)|$$

כדי למצוא פרמטריזציית אורך קשת, נצטרך למצוא פונקציה הפוכה לפונקציית האורך של המסילה. כפי שניתן לראות, בניגוד לדוגמה הראשונה, במקרה הזה לא פשוט למצוא את הפונקציה ההפוכה שתיתן לנו את פרמטריזציית הקשת. אך אנו יודעים שהיא קיימת.

אינטגרלים קוויים

14.1 הגדרה – נניח ש - $C \subseteq \mathbb{R}^n$ עקום פשוט וחלק עם פרמטריזציה $f: [a, b] \to \mathbb{R}^n$ גזירה ברציפות. אם ρ פונקציה רציפה על C אז נגדיר את **האינטגרל הסקלרי** של ρ על C (או **האינטגרל הקווי מסוג ראשון**) כ:

$$\int_C \rho ds = \int_a^b \rho\big(f(t)\big)\|f'(t)\|dt$$

שימו לב שכאשר הפונקציה הסקלרית ρ שווה לקבוע אחד, נקבל פשוט את אורך העקום. משתמשים באינטגרל סקלרי על עקום כאשר יש איזשהו גודל רציף (שמקבל ערכים סקלריים) שרוצים ל"סכום" לאורך עקום. למשל, כאשר העקום מתאר אובייקט שהינו בקירוב חד מימדי (למשל חוט ברזל, או תיל), ויש בידינו פונקציה ρ המתארת את צפיפות המסה בכל נקודה על העקום, אזי האינטגרל הסקלרי נותן לנו את המסה הכוללת. כפי שעשינו כשחישבנו אורך עקום, אפשר להגיע אל הביטוי האינטגרלי של אינטגרל סקלרי על עקום הזה כגבול של "סכומי רימן" מהצורה $\sum\rho\big(f(t_i)\big)\ell(f([t_{i-1}, t_i]) = \sum\rho\big(f(t_i)\big)\int_{t_{i-1}}^{t_i}\|f'(t)\|dt$, ואפשר להוכיח שבתנאים מסוימים הסכומים הללו אכן מתכנסים לביטוי האינטגרלי. אנו לא נעשה זאת, ונשתמש בביטוי האינטגרלי בתור הגדרה.

<u>תרגיל</u>: האינטגרל הסקלרי אינו תלוי בפרמטריזציה.

<u>דוגמה</u>: נניח כי צפיפות המסה של הקפיץ מהדוגמה הראשונה נתונה על ידי הפונקציה:

$$\rho(x, y, z) = z$$

מהי, אם כן, מסת הקפיץ?

$$\int_C \rho ds = \int_0^L \rho\big(f(t)\big)\|f'(t)\|dt = \int_0^{L\sqrt{a^2\omega^2+b^2}} \rho\big(g(s)\big)ds$$

כאשר לעיתים מסמנים $\rho(s)$ במקום $\rho\big(g(s)\big)$ שכן פרמטריזציית אורך הקשת נתפסת כ"פרמטריזציה" טבעית של אורך הקטע. נשתמש בפרמטריזציה זו, ונקבל כי:

$$g(s) = \left(a\cos\left(\frac{\omega s}{\sqrt{a^2\omega^2+b^2}}\right), a\sin\left(\frac{\omega s}{\sqrt{a^2\omega^2+b^2}}\right), \frac{bs}{\sqrt{a^2\omega^2+b^2}}\right)$$

$$\int_0^{l(C)} \frac{bs}{\sqrt{a^2\omega^2+b^2}}ds = \frac{b}{\sqrt{a^2\omega^2+b^2}}\frac{1}{2}L(a^2\omega^2+b^2) = \frac{b}{2}L^2\sqrt{a^2\omega^2+b^2}$$

14.2 הגדרה – תהי $S \subseteq \mathbb{R}^n$. פונקציה $F: S \to \mathbb{R}^n$ נקראת **שדה וקטורי.**

(באופן יותר פורמלי, אפשר לחשוב על שדה וקטורי כעל פונקציה $F: S \to \bigcup_{p\in S} T_p(\mathbb{R}^n)$ אם $F(p) \in T_p(\mathbb{R}^n)$. אם $M \subseteq \mathbb{R}^n$ יריעה חלקה אז מגדירים גם שדה וקטורי מהסוג $F: M \to \bigcup_{p\in M} T_p(M)$, אך זה נושא לקורסים מתקדמים יותר).

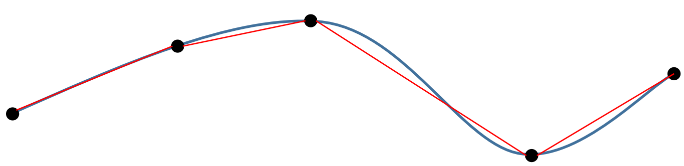
הגדרה 14.3 – תהא $f:[a,b]\to\mathbb{R}^n$ פרמטריזציה גזירה ברציפות של עקום C מכוון, פשוט, וחלק, ו-$F:C\to\mathbb{R}^n$ שדה וקטורי רציף. **האינטגרל הקווי של F לאורך C** (או **האינטגרל הקווי מסוג שני**) מוגדר להיות:

$$\int_C \vec{F}\cdot d\vec{r}=\int_C\sum_{i=1}^{n}F_i dx_i := \int_a^b F\big(f(t)\big)\cdot f'(t)dt$$

<u>הערות:</u>

א. השדה שמורגש בכל נקודה בכיוון העקום הוא בדיוק ההיטל של השדה על המשיק לעקום, ולכן ההיטל הזה מתקבל כמכפלה סקלרית של השדה עם הווקטור המשיק בכל נקודה, כמתואר באינטגרל האחרון. לכן אפשר לחשוב על האינטגרל הקווי של F לאורך C כאינטגרל שסוכם את התרומה של השדה הוקטורי בכיוון העקום לאורך העקום כולו.

ב. אפשר להגיע גם לאינטגרל זה על ידי ביצוע חלוקה כמתואר באיור להלן, וקבלת קירוב פוליגונלי לעקום. כל קו ישר מקרב את האינטגרל הקווי על הקטע המתאים.



איור 14 – דוגמה לחלוקה של עקום

ג. אם $g:[a,b]\to C$ פרמטריזציית אורך קשת, נגדיר $T(s)=g'(s)$ ונקבל כי $\|T(s)\|=1$ ואז נקבל כי:

$$\int_C\sum F_i dx_i=\int_a^b F\big(g(s)\big)\cdot g'(s)ds$$
$$=\int_C \vec{F}\cdot\vec{T}ds$$

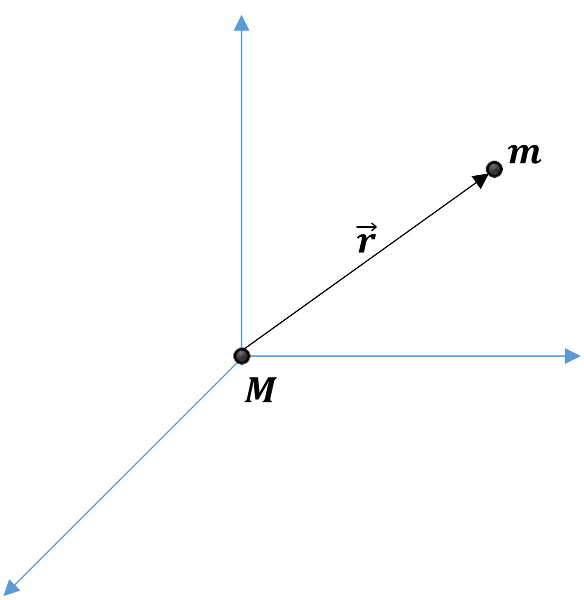
כאשר לעיתים מקובל לסמן וקטור באורך יחידה כדוגמת T על ידי $\hat{T}$ (וקטור כיוון באורך 1). היתרון של סימון זה הוא שהוא איננו משתמש בפרמטריזציה כלשהי (מסילה), והוא כתוב במונחים של העקום בלבד.

בפיזיקה, משתמשים באינטגרל קווי מסוג שני כדי לחשב את העבודה ששדה כוח עושה לאורך מסלול.

<u>דוגמה:</u>

$$F:\mathbb{R}^3\setminus\{0\}\to\mathbb{R}^3$$

יהיה הכח הגרביטציוני שגוף בעל מסה m , והנמצא במיקום המתואר על ידי $\vec{r}$, מרגיש כתוצאה מנוכחות של גוף בעל מסה M במרחב.



מתקיים:

$$\vec{F}=G\frac{Mm}{\|\vec{r}\|^3}(-\vec{r})$$

בהנחה ש-M,m קבועים נוכל לכתוב:

$$F(x,y,z)=-K\frac{(x,y,z)}{\|(x,y,z)\|^3}$$

נחשב את העבודה של הכוח הזה לאורך המסלולי הקפיצי מתחילת הפרק. לשם פשטות נניח $K=1$.

נזכר ש: $f(t)=(a\cos\omega t\,,a\sin\omega t\,,bt)$, $f'(t)=(-\omega a\sin\omega t\,,\omega a\cos\omega t\,,b)$. לכן

$\|f(t)\|=\|(a\cos\omega t\,,a\sin\omega t\,,bt)\|=\sqrt{a^2+b^2t^2}$ וכן $f(t)\cdot f'(t)=b^2t$.

$$\int_0^L F\big(f(t)\big)\cdot f'(t)dt=-\int_0^L\frac{f(t)\cdot f'(t)}{\left(a^2+b^2t^2\right)^{\frac{3}{2}}}dt=-\int_0^L\frac{b^2t}{\left(a^2+b^2t^2\right)^{\frac{3}{2}}}dt=\frac{1}{\sqrt{a^2+b^2t^2}}\Bigg|_0^L$$
$$=\frac{1}{\sqrt{a^2+b^2L^2}}-\frac{1}{a}$$

פרק 15

אינטגרציה במלבן

תזכורת קצרה על אינטגרל רימן ב – $\mathbb{R}^2$ (למדתם באינפי 2). נניח כי נתון תחום מהצורה:

$$R = [a,b] \times [c,d] \subseteq \mathbb{R}^2$$

ותהא $f: R \to \mathbb{R}$ חסומה. נגדיר חלוקה של R על ידי:

$$P_1 = \{a = x_0 \le x_1 \le x_2 \le \cdots \le x_n = b\} \quad P_2 = \{c = y_0 \le y_1 \le \cdots \le y_m = d\}$$

נזהה את P_1, P_2 עם הקטעים המוגדרים על ידי הנקודות. נגדיר $P = P_1 \times P_2$, כך שאחרי הזיהוי עם מלבנים נקבל:

$$P = P_1 \times P_2 = \left\{ \underset{\substack{1 \le i \le n \\ 1 \le j \le m}}{\forall} [x_{i-1}, x_i] \times [y_{j-1}, y_j] \right\}$$

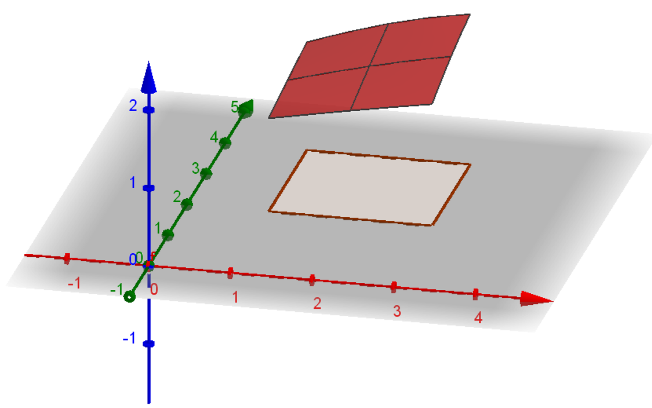
נסמן לכל $R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$:

$$M_{ij} = \sup_{R_{ij}} f$$
$$m_{ij} = \inf_{R_{ij}} f$$

ונוכל עתה להגדיר סכום עליון וסכום תחתון של f עבור החלוקה P, על ידי:

$$U(f,P) = \sum_{i,j} M_{ij} A(R_{ij})$$

$$L(f,P) = \sum_{i,j} m_{ij} A(R_{ij})$$



איור 15 פונקציה $f(x,y)$ במרחב $\mathbb{R}^3$ בתחום מלבני לדוגמה

כאשר השטח של המלבן R_{ij} מוגדר להיות

$$A(R_{ij}) = (x_i - x_{i-1}) \times (y_j - y_{j-1})$$

נאמר ש - f **אינטגרבילית** (או **אינטגרבילית רימן**, ליתר דיוק) אם קיים $I \in \mathbb{R}$ יחיד עבורו:

$$L(f,P) \le I \le U(f,P) \quad \text{for every} \quad P$$

למשל, עבור פונקציה קבוע $f(x,y) \equiv C$, נקבל כי:

$$U(f,P) = \sum_{i,j} C \cdot A(R_{ij}) = CA(R) \quad L(f,P) = \sum_{i,j} C \cdot A(R_{ij}) = C \cdot A(R)$$

ולכן מתקיים כנדרש:

$$I = C \cdot A(R)$$

אינטגרציה בתיבה

את ההגדרה של אינטגרל רימן ב-$\mathbb{R}^2$ ("אינטגרל כפול") קל להכליל ל- $\mathbb{R}^n$.

תיבה ב- $\mathbb{R}^n$ זו קבוצה $R \subset \mathbb{R}^n$ מהצורה:

$$(*)\ \ R = \prod_{i=1}^{n} [a_i, b_i] = [a_1, b_1] \times [a_2, b_2] \cdots \times [a_n, b_n] \subset \mathbb{R}^n$$

עבור מספרים $a_i \le b_i$ $i = 1,2, \dots, n$. אם $a_i = b_i$ עבור i כלשהו נאמר שהתיבה **מנוונת**. אנו מגדירים **נפח** (או **נפח -** n**-מימדי**) תיבה על ידי:

$$vol(R) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$$

חלוקה של תיבה מוגדרת באופן דומה למקרה הדו-מימדי, אך בגלל המימדים המרובים הסימונים נהיים די כבדים. נניח ש- R תיבה כמו ב - $(*)$. אז **חלוקה** P **של** R מוגדרת להיות אוסף כל התיבות מהצורה

$$\left[t^1_{i_1-1}, t^1_{i_1}\right] \times \left[t^2_{i_2-1}, t^2_{i_2}\right] \times \cdots \times \left[t^n_{i_n-1}, t^n_{i_n}\right]$$

כאשר $\left\{a_i = t^i_0 < t^i_1 < \cdots < t^i_{n_i} = b_i\right\}$ חלוקה של הקטע $[a_i, b_i]$. כמו שבמקרה החד-מימדי אפשר לחשוב על חלוקה גם כאוסף עולה של נקודות וגם כאוסף של תתי-קטעים סגורים, אפשר לחשוב על חלוקה של תיבה גם כאוסף של תתי-תיבות, וגם כאוסף הנקודות שמגדירות אותם. בדרך כלל במקרה הרב מימדי נוח יותר דווקא לחשוב על החלוקה כעל אוסף התיבות.

אם P חלוקה של המלבן R, נסמן $Q \sim P$ כאשר Q הינה תת-תיבה בחלוקה (אמנם אפשר לסמן $Q \in P$, אך כיון שבהגדרה של חלוקה לפעמים מתכוונים לנקודות ולא לתתי-התיבות, הסימון $Q \sim P$ הינו חד משמעי).

15.1 הגדרה - תהי f פונקציה חסומה על התיבה R ותהי P חלוקה של R לתיבות קטנות יותר. **סכומי דארבו העליון והתחתון** מוגדרים להיות, בהתאמה:

$$U(f, P) = \sum_{Q \sim P} M_Q \cdot vol(Q)$$
$$L(f, P) = \sum_{Q \sim P} m_Q \cdot vol(Q)$$

כאשר הסכימה מתבצעת על כל $Q \sim P$ כלומר על כל תת-תיבה ששייכת לחלוקה, וסימנו $M_Q = \sup\{f(x) : x \in Q\}$ ו- $m_Q = \inf\{f(x) : x \in Q\}$.

15.2 הגדרה – תהי f פונקציה חסומה על התיבה R. נאמר כי $\boldsymbol{f}$ **אינטגרבילית על** $\boldsymbol{R}$ (ולפעמים אומרים **אינטגרבילית רימן**) אם מתקיים

$$\sup L(f, P) = \inf U(f, P) = I$$

כלומר (באופן שקול) אם קיים I יחיד עבורו:

$$\forall P \quad L(f, P) \le I \le U(f, P)$$

את הסופרמום והאינפימום הנ"ל מכנים לעתים **האינטגרל העליון והאינטגרל התחתון** (בהתאמה) של f על התיבה R, ואנו רואים שפונקציה חסומה היא אינטגרבילית על תיבה אם ורק אם האינטגרל העליון שווה לאינטגרל התחתון. במקרה זה מסמנים

$$I = \int_R f dV$$

סימונים נוספים:

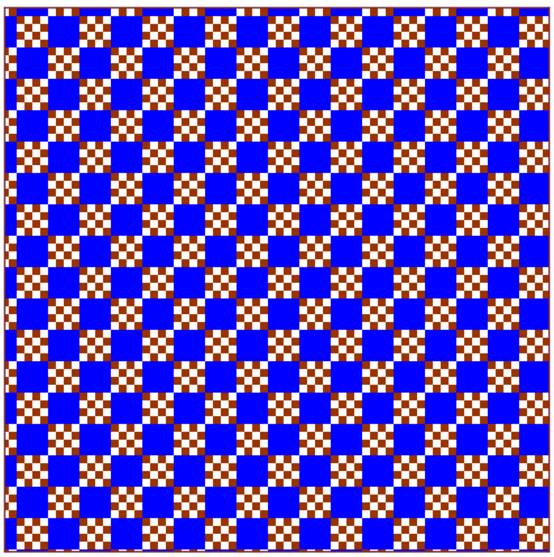
$$I = \int_R f = \int_R f(x)dx$$

בפרט, עבור $C \equiv f$, קל לראות, כי f אינטגרבילית ומתקיים:

$$\int_R f dV = C \cdot vol(R)$$

העובדה שהאינפימום והסופרמום בהגדרה מתקבלים נובעת מלמה 15.4 ומסקנה 15.5 להלן.

15.3 הגדרה – יהיו P, P' שתי חלוקות של R. נאמר כי P' עידון של P אם לכל תיבה $Q' \in P'$, קיים $Q \in P$ כך ש-$Q' \subset Q$.



איור 16 – דוגמאות לשתי חלוקות, כך שאחת מהווה עידון של השנייה. בדוגמה זו כל מלבן שני בחלוקה מסוימת מחולק ל-16 מלבנים נוספים. חישבו: מהם Q ו- Q' המתאימים להגדרת העידון?

15.4 למה – יהיו P, P' חלוקות של R ונניח כי P' עידון של P. נניח כי f חסומה על R. אזי מתקיים:

$$L(f,P) \leq L(f,P') \leq U(f,P') \leq U(f,P)$$

הוכחה:

לכל $Q \in P$, קיימים $\{Q_1, Q_2, \cdots, Q_s\} \subset P'$ שהאיחוד שלהם הוא Q. אזי מתקיים, כמובן:

$$(\star) \sum_{i=1}^{s} M_i \cdot vol(Q_i) \leq M \cdot vol(Q)$$

וכן מתקיים:

$$(\star) \geq \sum_{i=1}^{s} m_i \cdot vol(Q_i) \geq m \cdot vol(Q)$$

נסכום על כל התיבות ב-P, ונקבל את הדרוש. מ.ש.ל.

15.5 מסקנה – לכל P', P'' חלוקות של R, מתקיים:

$$L(f,P') \leq U(f,P'')$$

את המסקנה ניתן להוכיח ביתר קלות על ידי שימוש בחלוקה שמהווה הן עידון משותף הן של P' והן של P'', ואז הפעלת הלמה.

15.6 טענה (אפיון לאינטגרביליות) – תהא $f: R \to \mathbb{R}$ חסומה. אזי f אינטגרבילית על R $\Longleftrightarrow$ לכל $\varepsilon > 0$ קיימת חלוקה P כך שמתקיים $U(f,P) - L(f,P) < \varepsilon$.

הוכחה:

($\Rightarrow$) נניח כי f איננה אינטגרבילית. אזי קיימים $I_1 \neq I_2$ כך שלכל חלוקה P מתקיים:

$$L(f,P) \leq I_i \leq U(f,P)$$

נניח בלי הגבלת הכלליות כי $I_2 > I_1$ ונסמן $\boxed{\varepsilon = I_2 - I_1} > 0$.

לכל P מתקיים:

$$L(f,P) \leq I_1 < I_2 \leq U(f,P) \Rightarrow U(f,P) - L(f,P) \geq I_2 - I_1 = \varepsilon$$

הוכחנו שאם f איננה אינטגרבילית, אז התנאי לא מתקיים עבור $\varepsilon > 0$ מסוים, לכן אם התנאי מתקיים עבור כל $\varepsilon > 0$, f חייבת להיות אינטגרבילית.

($\Leftarrow$) נניח כי f אינטגרבילית, ויהא $\varepsilon > 0$. אזי קיימת חלוקה P' כך שמתקיים:

$$U(f,P') - I < \frac{\varepsilon}{2}$$

(אחרת נקבל סתירה ליחידות I). כמו כן, קיימת חלוקה P'' כך שמתקיים $I - L(f,P'') < \frac{\varepsilon}{2}$ מאותו שיקול. עתה, נבחר P שתהא עידון של P', P'', ויתקיים:

$$L(f,P'') \leq L(f,P) \leq U(f,P) \leq U(f,P')$$

ומתקיים, לכן:

$$U(f,P) - L(f,P) \leq U(f,P') - L(f,P'') < \varepsilon$$

כנדרש. מ.ש.ל.

15.7 טענה – **נניח כי** f, g **אינטגרביליות על** R. **אזי** $f + g$ **אינטגרבילית ומתקיים:**

$$\int_R (f+g)dV = \int_R f dV + \int_R g dV$$

<u>הוכחה:</u>

תהא P חלוקה של R. נשים לב כי מתקיים

$$\sup_Q (f+g) \leq \sup_Q f + \sup_Q g$$

לכל $Q \sim P$. לכן מתקיים:

$$(\star) \quad U(f+g,P) \leq U(f,P) + U(g,P)$$

מאותו שיקול, נוכל להסיק כי הנ"ל מתקיים גם עבור החסמים מלרע וסה"כ:

$$L(f,P) + L(g,P) \leq L(f+g,P) \leq (\star)$$

ומכאן נקבל כי:

$$U(f+g,P) - L(f+g,P) \leq [U(f,P) - L(f,P)] + [U(g,P) - L(g,P)]$$

יהא, עתה, $\varepsilon > 0$. אזי קיימות P', P'' עבורן מתקיים:

$$U(f,P') - L(f,P') < \frac{\varepsilon}{2} \quad U(g,P'') - L(g,P'') < \frac{\varepsilon}{2}$$

ולכן עבור P שתהא עידון של P', P'' ונקבל את הדרוש, קרי:

$$(**)\ \ U(f+g,P)-L(f+g,P)<\varepsilon$$

כנדרש.

בשלב זה, לאחר שהראינו את האינטגרביליות של $f+g$ (קרי, את קיום האינטגרל), נרצה להראות כי אכן ערכו הוא כמתואר בטענה. נסמן $I=\int_R (f+g)dV$, $I_1=\int_R f dV$, $I_2=\int_R g dV$.

אנו יודעים כי עבור החלוקה P שבחרנו לעיל (המקיימת את $(**)$) מתקיים:

$$L(f,P)\le I_1\le U(f,P) \quad L(g,P)\le I_2\le U(g,P)$$

נרצה להראות כי $I=I_1+I_2$. נניח בשלילה, אם כן, כי $I<I_1+I_2$.

נסמן $\varepsilon=(I_1+I_2)-I>0$. ממה שראינו עד כה, מתקיים:

$$L(f,P)+L(g,P)\le L(f+g,P)\le I<I_1+I_2\le U(f,P)+U(g,P)$$

ולכן, יתקיים:

$$[U(f,P)-L(f,P)]+[U(g,P)-L(g,P)]\ge\varepsilon$$

אך זו סתירה לאיך שבחרנו את החלוקה P.

באופן דומה, מראים שלא יתכן שיתקיים $I>I_1+I_2$. מ.ש.ל.

לסיום, נביא עוד שלוש טענות ללא הוכחה. ייתכן וראיתם הוכחות לטענות אלו במקרה הדו-ממדי באינפי 2, ואם לא יהיה קל לכם להוכיח אותן כתרגיל. את הטענות 15.7 – 15.10 צריך לזכור בתור העובדה שאינטגאל רימן הוא לינארי, אדיטיבי ומונוטוני.

15.8 טענה – תהא f אינטגרבילית על R ותהא $\alpha\in\mathbb{R}$. אזי $\alpha\cdot f$ אינטגרבילית בתחום ומתקיים:

$$\int_R(\alpha f)dV=\alpha\cdot\int_R f dV$$

15.9 טענה – אם $R=R'\cup R''$ איחוד של שני מלבנים שחיתוכם הוא לכל היותר בשפה של המלבנים. אזי f אינטגרבילית על $R \Leftrightarrow f$ אינטגרבילית על R', R'' ומתקיים:

$$\int_R f dV=\int_{R'} f dV+\int_{R''} f dV$$

15.10 טענה – f,g אינטגרביליות על R ונניח כי $f\le g$. אזי מתקיים:

$$\int_R f dV\le\int_R g dV$$

<u>תרגיל חשוב:</u> נסחו והוכיחו את "**<u>תכונת הערך הממוצע</u>**" ואת "**<u>אי שוויון המשולש</u>**" שאינטגרל רימן הרב מימדי מקיים.

לפני שנמשיך בפיתוח תורת האינטגרל, "נזכיר" בקצרה איך מחשבים אינטגרלים בפועל ומה עושים איתם.

שימושים של האינטגרל

א. בהינתן תחום $\Omega \subset \mathbb{R}^3$, האינטגרל $\iiint_\Omega 1dv$ הינו נפח האובייקט התלת ממדי. באופן כללי, נפח n–ממדי ניתן על ידי $\int_\Omega 1dv$ עבור כל $\Omega \subset \mathbb{R}^n$.

ב. עבור $\Omega \subset \mathbb{R}^3$ ופונקציה $\rho(x,y,z)$ שהינה פונקציית צפיפות (למשל במקרה של פיזיקה, פונקציה המתארת את המסה פר יחידת נפח, כלומר ביחידות של $\frac{kg}{m^3}$), מתקיים שמסת גוף שמתואר על ידי התחום Ω היא:

$$m = \iiint_\Omega \rho(x,y,z)dv$$

ג. שימוש פיזיקלי נוסף הינו מומנט האינרציה / חישובי מרכז מסה.

ד. ועוד ועוד...

להזכירכם, באינפי 2 חישבתם אינטגרלים בפועל לא לפי ההגדרה אלא לפי משפט ניוטון-לייבניץ (כאשר האינטגרל הוא חד ממדי) ועל ידי שימוש במשפט פוביני (במקרה הרב ממדי).

משפט פוביני (מקרה דו ממדי) – נניח כי f אינטגרבילית על מלבן $[a,b]\times[c,d] \subset \mathbb{R}^2$ ונניח כי לכל $x \in [a,b]$, הפונקציה:

$$F(x) = \int_c^d f(x,y)dy$$

מוגדרת ואינטגרבילית על $[a,b]$. אזי מתקיים:

$$\iint_R f(x,y)dxdy = \int_a^b F(x)dx = \int_a^b \left(\int_c^d f(x,y)dy\right)dx$$

אנו נראה שקיים משפט דומה במימדים גבוהים יותר, והוא מאפשר לחשב אינטגרלים באופן הבא:

משפט פוביני (מקרה רב ממדי) – אם $R = [a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_n,b_n]$ ונניח כי כל אחת מהפונקציות הבאות $F_1, F_2, \dots$ אינטגרביליות:

$$\int_{a_n}^{b_n} f(x)dx_n = F_1(x_1,x_2,\cdots,x_{n-1})$$

$$\int_{a_{n-1}}^{b_{n-1}} dx_{n-1}\int_{a_n}^{b_n} f(x)dx_n = F_2(x_1,x_2,\cdots,x_{n-2})$$

$$\vdots$$

$$\int_{a_1}^{b_1} dx_1 \cdots \int_{a_{n-1}}^{b_{n-1}} dx_{n-1}\int_{a_n}^{b_n} f(x)dx_n = I$$

אזי מתקיים: $I = \int_R fdV$.

כדי לחשב נפח של תחום Ω אנו נחשב את האינטגרל $\int_\Omega f dv$ עבור הפונקציה $f = 1$. עבור $\Omega \subset \mathbb{R}^3$ תחום פשוט, חישוב האינטגרל יהיה, בפועל, חישוב של n אינטגרלים נשנים לדוגמה:

$$I = \iiint_\Omega f(x,y,z)dxdydz \quad \text{in the domain} \quad \begin{cases} 0 \le x \le b \\ y_1(x) \le y \le y_2(x) \\ z_1(x,y) \le z \le z_2(x,y) \end{cases}$$

ונקבל במקרה זה:

$$I = \iiint_\Omega f(x,y,z)dxdydz = \int_a^b dx \int_{y_1(x)}^{y_2(x)} dy \int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)dz$$

כך שכפי שניתן לראות באיור לעיל, אנחנו בפועל סוכמים "משטחים" המרכיבים את הנפח (האיור הינו הפשטה של המקרה התלת ממדי לצורך המחשה בלבד. באיור המשטחים לדוגמה צבועים באדום, והסכימה היא עבור כל המשטחים המוכלים בתחום הכחול החצי שקוף).

<u>דוגמה:</u>

נחשב את $I = \iiint_\Omega (xyz)dxdydz$ עבור Ω התחום החסום על ידי $x + y + z = 1$ והצירים $x = 0, y = 0, z = 0$.

התחום Ω הוא בעצם הפירמידה המשולשת החסומה על ידי 4 מישורים.

על מנת לעשות זאת, נקבע את z ולכל z נקבל משולש מהצורה:

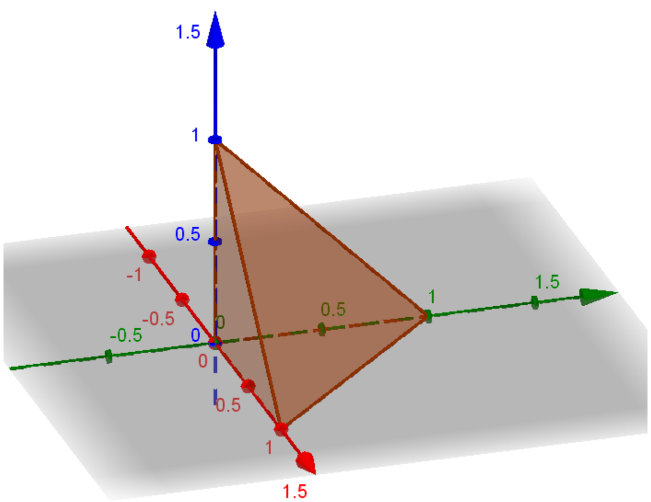
$$x + y = 1 - z$$

D_z תחום פשוט ביחס לציר y. הוא התחום:

$$D_z = \begin{Bmatrix} 0 \le x \le 1 - z \\ 0 \le y \le 1 - x - z \end{Bmatrix}$$

וסה"כ נסיק כי התחום שלנו, Ω, ניתן להגדרה על ידי:

$$\Omega = \begin{Bmatrix} 0 \le z \le 1 \\ 0 \le x \le 1 - z \\ 0 \le y \le 1 - x - z \end{Bmatrix}$$



ולכן מתקיים:

$$I = \int_0^1 dz \int_0^{1-z} dx \int_0^{1-x-z} xyzdy = \cdots = \frac{1}{720}$$

פרק 16

16.1 טענה – אם $f: R \to \mathbb{R}$ רציפה על תיבה R, אזי f אינטגרבילית.

<u>הוכחה:</u>

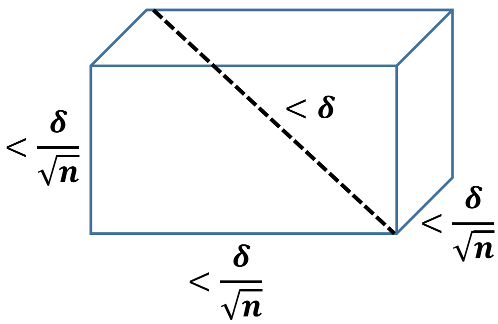
יהא $\varepsilon > 0$, ונמצא חלוקה P כך שמתקיים:

$$U(f,P) - L(f,P) \leq \varepsilon$$

לשם כך, נזכיר כי לכל P מתקיים:

$$U(f,P) - L(f,P) = \sum_{Q} \left(M_Q - m_Q\right) \cdot vol(Q)$$

עתה, היות ו-f רציפה במלבן שהינו קבוצה קומפקטית, נקבל כי f רציפה במידה שווה. כלומר, קיים $\delta > 0$ כך שלכל $\|x-y\| < \delta$ מתקיים $|f(x)-f(y)| < \frac{\varepsilon}{vol(R)}$. נבחר, אם כן, חלוקה P כלשהי כך שכל תיבה תהיה בעלת צלעות שאורכן קטן ממש מ - $\frac{\delta}{\sqrt{n}}$, וכך נקבל כי בנורמה 2 מתקיים:



$$\|x-y\| < \delta$$

לכל $x,y \in Q$ הנמצאים באותה תיבה $Q \sim P$ בחלוקה. עבור חלוקה P מסוג זה, יתקיים:

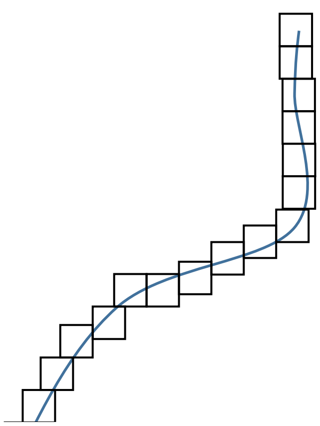
$$U(f,P) - L(f,P) = \sum_{Q} \underbrace{\left(M_Q - m_Q\right)}_{\leq \frac{\varepsilon}{vol(R)}} vol(Q) \leq \varepsilon$$

כנדרש. מ.ש.ל.

<u>נפח אפס</u>

16.2 הגדרה – נאמר כי $A \subset \mathbb{R}^n$ הינה **קבוצה בעלת נפח 0** אם לכל $\varepsilon > 0$ קיימים מספר סופי של תיבות $R_1, R_2, \dots, R_k$ כך שמתקיים:

$$\sum_{i=1}^{k} vol(R_i) < \varepsilon \quad , \quad A \subset \bigcup_{i=1}^{k} int(R_i)$$



איור 17 – עקום רציף לדוגמה שהוא קבוצה בעלת נפח אפס. רואים שניתן להקטין את המלבנים כרצוננו ועדיין להיות מסוגלים לכסות את כל העקום.

<u>תרגיל:</u> הוכיח שכל נקודה היא קבוצה בעלת נפח אפס. הוכיחו שלאיחוד סופי של קבוצות בעלות נפח אפס יש נפח אפס.

16.3 טענה – תהא $f: R \to \mathbb{R}$ חסומה, ונניח כי:

$$X = \left\{ x \in R \middle| \begin{matrix} f \text{ אינה} \\ x \text{ רציפה ב} - \end{matrix} \right\}$$

היא קבוצה בעלת נפח אפס. אזי f אינטגרבילית ב-R.

הוכחה:

f חסומה, ולכן נסיק כי קיים $M \in \mathbb{R}$ כך שלכל $x \in R$ מתקיים $|f(x)| \leq M$. היות שכן, נשים לב כי:

$$U(f,P) - L(f,P) = \sum_{Q} (M_Q - m_Q) vol(Q)$$

מוגדר היטב וסופי. לכל $\varepsilon > 0$ קיימים $R_1, \ldots, R_s$ כך ש"הפנימים" מכסים את X, כלומר:

$$X \subset \bigcup_{i=1}^{s} int(R_i)$$

ומתקיים:

$$\sum_{i=1}^{s} vol(R_i) < \frac{\varepsilon}{4M}$$

ניקח חלוקה P של R כך שכל R_i הוא איחוד של תיבות בחלוקה, ובלי הגבלת הכלליות (על-ידי חלוקת התיבות) נוכל להניח ש $\{R_i\}_i \subseteq P$. מתקיים:

$$\sum_{i=1}^{s} (M_i - m_i) vol(R_i) < \frac{\varepsilon}{2}$$

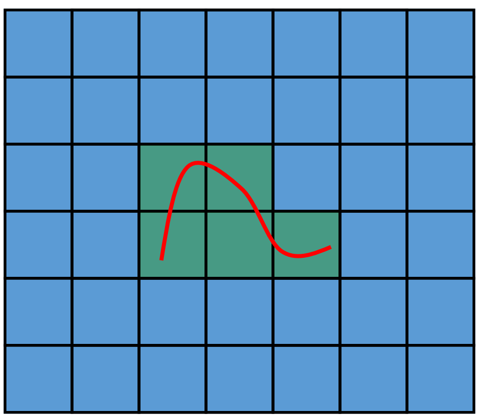
עתה נתבונן ב-$Y = R \setminus \{int(R_i)\}_i$. אנו יודעים כי f רציפה על Y ולכן רציפה במ"ש, כי Y סגורה וחסומה. על-ידי עידון נוסף נוכל להניח ש-P מקיימת:

$$U(f, P \setminus \{R_i\}) - L(f, P \setminus \{R_i\}) := \sum_{Q \sim P \setminus \{R_i\}} (M_Q - m_Q) vol(Q) < \frac{\varepsilon}{2}$$

כאשר את אי השוויון מוכיחים בדיוק כמו בטענה 16.1. נקבל שעבור החלוקה P, מתקיים, כנדרש:

$$U(f,P) - L(f,P) < \varepsilon$$

מ.ש.ל.



איור 18 – הרעיון בהוכחה הוא שאם נניח כי העקום באדום הוא קבוצת נקודות האי רציפות, אז התרומה של התיבות שמכילות את העקום תהיה קטנה ממכפלת נפח כל התיבות בחסם העליון של f שנתון כי f חסומה. וכך נוכל להזניח בסופו של דבר, כרצוננו, את התרומה של הקבוצה ה"בעייתית". על שאר התחום אפשר להראות שההפרש בין "סכומי הדרבו" קטן כמו בהוכחה של אינטגרביליות פונקציה רציפה.

מידה אפס ואינטגרציה

16.4 הגדרה – נאמר כי $A \subset \mathbb{R}^n$ הינה **קבוצה בעלת מידה 0** אם לכל $\varepsilon > 0$ קיים אוסף בן מנייה של תיבות $\{R_i\}_{i \in I}$ ($|I| \leq \aleph_0$) כך שמתקיים:

$$\sum_{i \in I} vol(R_i) < \varepsilon \quad , \quad A \subset \bigcup_{i \in I} int(R_i)$$

שימו לב: קבוצה מנפח אפס היא גם קבוצה ממידה אפס. ההיפך לא נכון!

<u>תרגיל:</u> 1) הוכיחו שהקבוצה $\mathbb{Q}^n \subset \mathbb{R}^n$ של הנקודות שכל הקואורדינטות שלהן רציונליות היא ממידה אפס, אבל איננה מפח אפס. 2) הוכיחו שאם $\{A_i\}$ הינה סדרה בת מנייה של קבוצות ממידה אפס, אזי $\cup_i A_i$ גם כן ממידה אפס.

16.5 טענה – קבוצה קומפקטית $K \subset \mathbb{R}^n$ היא מנפח אפס אם ורק אם K היא ממידה אפס.

<u>הוכחה:</u> מיידי מההגדרות שאם K מנפח אפס אזי היא ממידה אפס. נניח ש - K ממידה אפס. יהי $\varepsilon > 0$. אז יש משפחה בת-מנייה $\{R_i\}_{i\in I}$ של תיבות כך שמתקיים:

$$\sum_{i\in I} vol(R_i) < \varepsilon \quad , \quad K \subset \bigcup_{i\in I} int(R_i)$$

האוסף - $\{int(R_i)\}_{i\in I}$ מהווה כיסוי פתוח של - K. מקומפקטיות של K, אפשר למצוא תת-כיסוי סופי. לכן ישנם מלבנים $R_1, R_2, \dots, R_k$ כך שמתקיים:

$$\sum_{i=1}^{k} vol(R_i) < \varepsilon \quad , \quad K \subset \bigcup_{i=1}^{k} int(R_i)$$

כנדרש מ.ש.ל.

<u>תרגיל:</u> תת קבוצה של קבוצה ממידה/נפח אפס היא גם כן ממידה/נפח אפס.

16.6 משפט (משפט לבג) – יהי $R \subset \mathbb{R}^n$ תיבה ותהי $f: R \to \mathbb{R}$ פונקציה חסומה. אזי f אינטגרבילית על R אם ורק אם אוסף נקודות אי הרציפות של f הוא ממידה אפס.

<u>הוכחה:</u>

לפני שנתחיל בהוכחה ממש אנו זקוקים להגדרות. נגדיר את **התנודה** של f על קבוצה $B \subseteq \mathbb{R}^n$ להיות

$$\omega(f,B) = \sup\{f(x) : x \in B \cap R\} - \inf\{f(x) : x \in B \cap R\}$$

ונגדיר את **התנודה של f** בנקודה $x \in R$ להיות

$$\omega(f,x) = \inf\{\omega(f,U_r(x)) : r > 0\} = \lim_{r \searrow 0} \omega(f,U_r(x))$$

כאשר סימנו $U_r(x) = (x_1 - r, x_1 + r) \times \cdots \times (x_d - r, x_d + r)$, כלומר זהו הכדור הפתוח סביב x לפי נורמת אינסוף ברדיוס r .

<u>למה:</u> fרציפה בנקודה $x \in R$ אם ורק אם $\omega(f,x) = 0$.

<u>הוכחה:</u> מושארת כתרגיל (זה באמת צריך להיות די ברור לכם בשלב הזה בתואר, ההוכחה היא מתוך ההגדרות).

לכל $\varepsilon > 0$ נסמן

$$W_\varepsilon = \{x \in R : \omega(f,x) \geq \varepsilon\}$$

$$W = \{x \in R : \omega(f,x) \neq 0\}$$

נשים לב ש –

$$(*) \quad W = \cup_{\varepsilon > 0} W_\varepsilon = \cup_{m=1}^{\infty} W_{\frac{1}{m}}$$

למה: לכל $\varepsilon > 0$ הקבוצה W_ε היא קומפקטית.

הוכחה: כיון ש- W_ε היא תת-קבוצה של הקבוצה החסומה R (כל מלבן הוא חסום), מספיק להוכיח ש-W_ε סגורה. אנו נוכיח ש - $\mathbb{R}^n \setminus W_\varepsilon$ היא קבוצה פתוחה. נניח ש - $x \in \mathbb{R}^n \setminus W_\varepsilon$. יש שני מקרים.

מקרה ראשון: $x \in \mathbb{R}^n \setminus R$ כלומר $x \notin R$. במקרה הזה, כיון ש - R סגורה, יש ל- x סביבה פתוחה הזרה ל- R, ולכן זרה גם ל- W_ε.

מקרה שני: $x \in R$. במקרה הזה, כיון שהתחלנו מההנחה $x \notin W_\varepsilon$, מתקיים $\omega(f,x) < \varepsilon$. לכן קיים $r > 0$ כך ש –

$$\omega\big(f, U_r(x)\big) = \sup\{f(y) : y \in U_r(x) \cap R\} - \inf\{f(y) :\ x \in U_r(x) \cap R\} < \varepsilon$$

אנו נראה ש- $U_r(x) \subseteq R \setminus W_\varepsilon$, ובכך נסיים את ההוכחה ש - $\mathbb{R}^n \setminus W_\varepsilon$ פתוחה, כלומר ש - W_ε סגורה. אכן, בהינתן $z \in U_r(x)$, קיים $0 < r' < r$ כך ש - $U_{r'}(z) \subseteq U_r(x)$, לכן $\omega\big(f, U_{r'}(z)\big) < \omega\big(f, U_r(x)\big) < \varepsilon$, ולכן

$$\omega(f,z) \le \omega\big(f, U_{r'}(z)\big) < \varepsilon$$

כלומר $z \notin W_\varepsilon$. מ.ש.ל.

<u>אוסף נק' אי רציפות ממידה אפס $\Leftarrow$ אינטגרביליות</u>

נניח ש - W ממידה אפס. עלינו להראות שההפרש $U(f,P) - L(f,P)$ יכול להיות קטן כרצוננו. נתון ש - f חסומה, אז נסמן $M = \sup\limits_{x\in R} |f(x)|$. נקבע $\varepsilon > 0$. מהמשוואה $(*)$ נובע ש - W_ε ממידה אפס. מהלמה, זו קבוצה קומפקטית, לכן מטענה 16.5 נובע שישנם מלבנים $R_1, R_2, \dots, R_k$ כך שמתקיים:

$$\sum_{i=1}^{k} vol(R_i) < \varepsilon \quad , \qquad W_\varepsilon \subset \bigcup_{i=1}^{k} int(R_i)$$

מצד שני, כל נקודה $x \notin W_\varepsilon$ מוכלת בפנים של תיבה סגורה Q_x עבורה מתקיים $\sup\limits_{y\in Q_x} f(y) - \inf\limits_{y\in Q_x} f(y) < \varepsilon$. האוסף

$$\{int(R_1), \dots, int(R_k)\} \cup \{int(Q_x) : x \notin W_\varepsilon\}$$

מהווה כיסוי פתוח של הקבוצה הקומפקטית R, ונוכל להוציא ממנו תת כיסוי סופי המורכב מתיבות פתוחות, אותו נסמן

$$\{U_1, \dots, U_m\}$$

הקודקודים של התיבות $U_1, \dots, U_m$ מגדירים חלוקה P של R (כדאי לצייר את המקרה הדו-מימדי). אפשר לחלק התיבות בחלוקה Q לשתי קבוצות:

1. קבוצה ראשונה S_1 המורכבת מתיבות המוכלות באחת ה - R_i.
2. קבוצה שניה S_2 המורכבת מתיבות הזרות לפנים של כל אחת מה - R_i. כל תיבה כזו Q מקיימת $Q \cap W_\varepsilon = \emptyset$.

עבור החלוקה הזו P מתקיים

$$U(f,P) - L(f,P) = \sum_{Q\sim P} \big(M_Q - m_Q\big) vol(Q) = \sum_{Q\in S_1} \big(M_Q - m_Q\big) vol(Q) + \sum_{Q\in S_2} \big(M_Q - m_Q\big) vol(Q)$$

$$\le \sum_{Q\in S_1} 2M \cdot vol(Q) + \sum_{Q\in S_2} \varepsilon \cdot vol(Q) \le 2M\varepsilon + \varepsilon \cdot vol(R)$$

כיון שהביטוי הנ"ל קטן כרצוננו, זה מוכיח שהפונקציה f אינטגרבילית.

<u>אינטגרביליות $\Leftarrow$ אוסף נק' אי רציפות ממידה אפס</u>

נניח ש - f אינטגרבילית. לפי $(*)$ יחד עם תרגיל (איחוד בן-מנייה של קבוצות ממידה אפס הוא ממידה אפס) מספיק להראות שהמידה ש $W_{\frac{1}{m}}$ ממידה אפס לכל m. נקבע את m.

יהי $\varepsilon > 0$ ותהי P חלוקה המקיימת

$$U(f,P) - L(f,P) < \frac{\varepsilon}{m}$$

נחלק את $W_{\frac{1}{m}}$ לשני חלקים, החלק שחותך את הפנימים של התיבות והיתר:

1. $A = \{x \in W_{1/m} : x \in int(Q),\ \text{for some } Q \sim P\}$,
2. $B = W_{1/m} \setminus A$.

מספיק להראות עבור כל אחת מהקבוצות A, B שאפשר לכסות אותה עם הפנימים של מספר סופי של תיבות מנפח כולל קטן מ $-\ \varepsilon$.

אם $R_1, \dots, R_k$ התיבות בכיסוי P המקיימות שהפנים שלהן חותך את $W_{\frac{1}{n}}$, אזי הפנימים שלהם $int(R_1), \dots, int(R_k)$ מכסים את A, ומתקיים

$$\frac{1}{m}\sum_i vol(R_i) \le \sum_i \omega(f,R_i)\, vol(R_i) \le \sum_{Q\sim P} \omega(f,Q) vol(Q) = U(f,P) - L(f,P) < \frac{\varepsilon}{m}$$

לכן

$$\sum_{i=1}^k vol(R_i) < \varepsilon \quad , \quad A \subset \bigcup_{i=1}^k int(R_i)$$

מצד שני, B מוכלת באיחוד של השפות של התיבות בחלוקה: $B \subseteq \bigcup_{Q\sim P} \partial Q$, וזהו איחוד סופי. אבל השפה של כל תיבה היא מנפח אפס, לכן גם את B ניתן לכסות במספר סופי של תיבות פתוחות מנפח כולל קטן מ $-\ \varepsilon$.

בסך הכל, הראנו שלכל $\varepsilon > 0$ אפשר לכסות את $W_{1/m} = A \cup B$ על ידי מספר סופי של תיבות פתוחות עם נפח כולל קטן מ- 2ε, ולכן $W_{1/m}$ מנפח אפס. מ.ש.ל.

<u>אינטגרציה בתחום כללי</u>

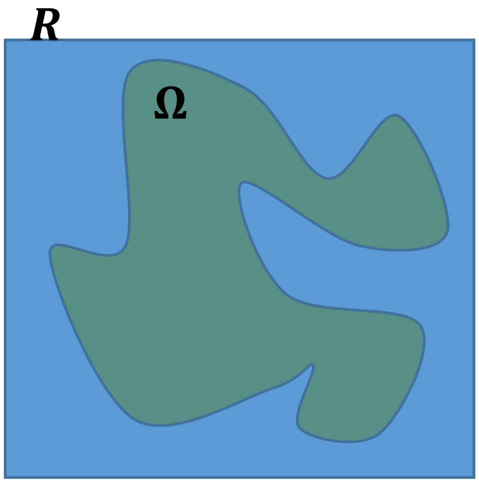
16.7 הגדרה - $\Omega \subset \mathbb{R}^n$ נקרא **תחום בעל נפח** אם הוא חסום ואם השפה של Ω היא בעלת נפח אפס.

<u>הערה</u> – שימו לב, כיון שהשפה של תחום חסום היא קומפקטית, אין זה משנה אם משתמשים במושג "נפח אפס" או "מידה אפס" כדי להגדיר תחום בעל נפח.

<u>הערה</u> – (לא נוכיח) אם Ω סגור של קבוצה פתוחה וחסומה, ואם $\partial\Omega$ ניתנת להצגה כתמונה של פונקציה חלקה על $X \subset \mathbb{R}^m$ קומפקטית $(m < n)$, שנסמנה ϕ המקיימת $\phi \in C^1$, $\phi: X \to \mathbb{R}^n$ ו-$\phi(X) = \partial\Omega$, אזי Ω הינה תחום בעל נפח. המשמעות היא שכל קבוצה "סבירה" שאפשר איכשהו לצייר היא בעלת נפח.

הגדרה 16.8 - יהא $\Omega\subset\mathbb{R}^n$ תחום בעל נפח. ותהא $R\supset\Omega$ תיבה. תהא $f:\Omega\to\mathbb{R}$ חסומה, ונגדיר:

$$\tilde{f}:R\to\mathbb{R}\quad \tilde{f}(x)=\begin{cases} f(x) & x\in\Omega \\ 0 & \text{אחרת}\end{cases}$$



נאמר כי f **אינטגרבילית על** Ω אם $\tilde{f}$ אינטגרבילית על R, ומתקיים:

$$\int_\Omega f dV=\int_R \tilde{f} dV$$

(נשים לב ש- $\tilde{f}=f\,\chi_\Omega$, כאשר χ_Ω היא הפונקציה המציינת של Ω.)

לא קשה (אבל אולי מתיש) להראות שההגדרה של אינטגרביליות והאינטגרל לא תלויה בתיבה R, ואנו נדלג על הפרטים.

16.9 טענה – אם Ω בעלת נפח ו- $f:\Omega\to\mathbb{R}$ רציפה, אזי f אינטגרבילית על Ω.

<u>הוכחה:</u>

נקודות האי רציפות של $\tilde{f}$ מוכלות ב-$\partial\Omega$ (כי בתחום של לתיבה שמחוץ ל-$\overline{\Omega}$ הפונקציה $\tilde{f}$ רציפה כי היא קבועה שם, וכמובן גם בפנים של Ω $\tilde{f}$ היא רציפה כי f רציפה שם). לכן קבוצת נק' אי הרציפות של f זו קבוצה בעלת נפח אפס ומטענה 16.3 (או במשפט לבג, אם רוצים) ומההגדרה, הפונקציה f אינטגרבילית כנדרש.

16.10 הגדרה – **נפח** של תחום בעל נפח $\Omega\subset\mathbb{R}^n$ מוגדר להיות:

$$vol(\Omega)=\int_\Omega 1dV$$

נשים לב כי הפונקציה 1 ודאי רציפה ועל כן ביטוי זה מוגדר היטב לכל תחום Ω בעל נפח.

16.11 טענה – אינטגרל רימן על תחום בעל נפח נהנה מכל התכונות הבאות: אם f,g אינטגרביליות על תחומים בעלי נפח Ω,Ω_1,Ω_2, אזי

1. **לינאריות:** לכל $\alpha\in\mathbb{R}$ הפונקציה $\alpha f+g$ גם כן אינטגרבילית ומתקיים

$$\int_\Omega \alpha f+g=\alpha\int_\Omega f+\int_\Omega g$$

2. **אדיטיביות:** אם החיתוך של Ω_1,Ω_2 בעל נפח אפס, אזי

$$\int_{\Omega_1\cup\Omega_2} f=\int_{\Omega_1} f+\int_{\Omega_2} f$$

3. **מונוטוניות:** אם $f\le g$, אזי מתקיים:

$$\int_\Omega f dV\le\int_\Omega g dV$$

4. אי שוויון המשולש:

$$\left|\int_\Omega f\right| \le \int_\Omega |f|$$

<u>הוכחה:</u>

נראה מונוטוניות, ואת היתר נשאיר כתרגיל קל. תהא $R \supset \Omega$ תיבה. אזי מתקיים $\tilde{f} \le \tilde{g}$ בתיבה הנתונה. נגדיר, אם כן $\tilde{h} = \tilde{g} - \tilde{f} \ge 0$ על R ונקבל כי $L(\tilde{h}, P) \ge 0$ לכל חלוקה P של התיבה. כאמור $\tilde{h}$ אינטגרבילית ומתקיים כנדרש:

$$0 \le \int_R \tilde{h} dV = \int_R \tilde{g} dV - \int_R \tilde{f} dV = \int_\Omega g dV - \int_\Omega f dV$$

<u>לתרגול/תרגיל בית:</u> גרף של פונקציה רציפה (או אינטגרבילית רימן) מהווה תחום בעל מידה אפס, נפח אפס אם מוגדר על תחום קומפקטי. כל חיתוך של תת-מרחב ממש עם קבוצה חסומה הוא תחום בעל נפח אפס. ליריעה יש מידה אפס, לתמונה של יריעה על-ידי העתקה חלקה יש נפח אפס. אם לא לומדים על מידה אפס בהרצאה, אפשר לדון על ההבדל בין נפח אפס למידה אפס בתרגיל.

פרק 17

אינטגרל רימן – תזכורת

הגדרנו נפח של תיבה מהצורה $R=[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_n,b_n]$ על ידי:

$$Vol(R)=\prod_{j=1}^{n}\left(b_j-a_j\right)$$

בהינתן חלוקה של תיבה R לתתי תיבות

$$P=\{Q\}\quad R=\bigcup Q$$

אם $f\colon R\to\mathbb{R}$ חסומה, נוכל להגדיר סכום (דארבו) עליון, מהצורה:

$$U(f,P)=\sum_{Q\sim P}M_Q Vol(Q)\quad,\quad M_Q=\sup_{x\in Q}f(x)$$

וכן סכום (דארבו) תחתון:

$$L(f,P)=\sum_{Q\sim P}m_Q Vol(Q)\quad,\quad m_Q=\inf_{x\in Q}f(x)$$

f נקראית <u>אינטגרבילית על R</u> אם מתקיים:

$$\overbrace{\underline{\int_R} f}^{\text{אינטגרל תחתון}}\coloneqq\sup_{P\ \text{חלוקה}}L(f,P)=\inf_{P\ \text{חלוקה}}U(f,P)\coloneqq\overbrace{\overline{\int_R} f}^{\text{אינטגרל עליון}}$$

לערך המשותף קוראים **אינטגרל (רימן) של f על R** ומסמנים $\int_R f=\int_R f(x)dx=\int_R fdV$.

קבוצה חסומה $\Omega\subseteq\mathbb{R}^n$ נקראת **תחום בעל נפח** אם היא הסגור של קבוצה פתוחה ו-$\partial\Omega$ היא קבוצה בעלת נפח אפס.

אם Ω תחום בעל נפח ו- $f\colon\Omega\to\mathbb{R}$ נגדיר נאמר ש- f אינטגרבילית על Ω אם $f\chi_\Omega$ אינטגרבילית על R עבור איזושהי תיבה $R\subseteq\Omega$, ובמקרה זה מסמנים:

$$\int_\Omega fdV\coloneqq\int_R f\chi_\Omega dV$$

כאשר χ_Ω ("חי של אומגה") היא הפונקציה המציינת של Ω :

$$\chi_\Omega(x)=\begin{cases}1 & x\in\Omega\\ 0 & x\notin\Omega\end{cases}$$

<u>הערה:</u> לא קשה להראות שההגדרה אינה תלויה במלבן R.

אם $\Omega \subseteq \mathbb{R}^n$ תחום בעל נפח, אנו מגדירים את הנפח של Ω להיות:

$$Vol(\Omega) = \int_{\Omega} 1dV = \int_{R} \chi_{\Omega} dV$$

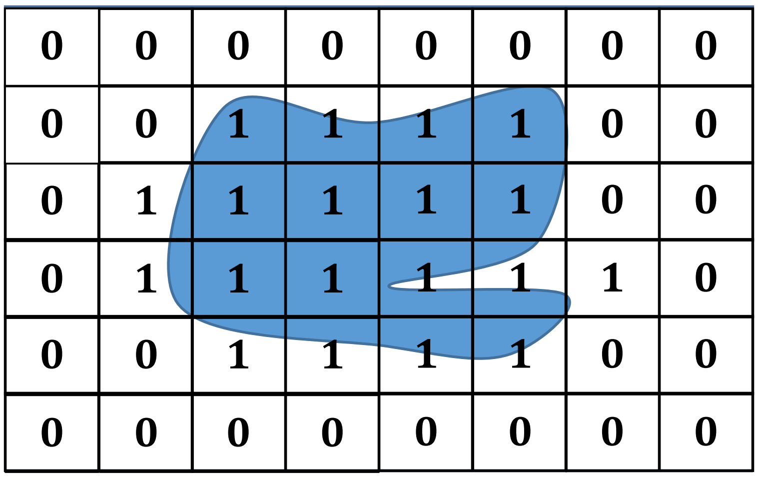
כלומר, לכל חלוקה מגדירים את סכום רימן אשר סוכם את שטח כל התיבות המכילות לפחות נקודה אחת מ-Ω. כלומר:

$$U(\chi_{\Omega}, P) = \sum_{Q\sim P_{\subseteq}} Vol(Q)$$

כאשר:

$$P_{\subseteq} = \{R \in P | R \cap \Omega \neq \emptyset\}$$

<u>הערה</u> - $Vol(\Omega) = \inf_{P} \sum_{Q\sim P_{\subseteq}} Vol(Q)$



ובאופן דומה ניתן להגדיר קירוב "מבפנים" אך זה מעט מסובך יותר (צריך להשתמש בכך שהשפה היא בעלת נפח 0, ולכן הפנים של התחום הוא בעל אותו נפח כמו הסגור, ו"יש מקום" לשים תיבות בתוך הפנים שמקרבות את הנפח מבפנים).

<u>הערה (משמעות האינטגרל כנפח מתחת לגרף)</u> – אם $f \geq 0$ אינטגרבילית בתיבה R, אזי מתקיים:

$$Vol\left(\left\{(x,y) \in R \times \mathbb{R} \middle| \begin{matrix} 0 \leq y \leq f(x) \\ x \in R \end{matrix}\right\}\right) = \int_{R} f(x)dx$$

<u>הוכחה:</u>

נניח כי $f: R \to [0, M]$ עבור:

$$\Omega \coloneqq \left\{(x,y) \in R \times \mathbb{R} \middle| \begin{matrix} 0 \leq y \leq f(x) \\ x \in R \end{matrix}\right\}$$

מתקיים לפי משפט פוביני (שנוכיח מיד):

$$Vol(\Omega) = \int_{R\times[0,M]} \chi_{\Omega} dV = \int_{R} \left(\int_{[0,M]} \chi_{\Omega}(x,y)dy\right) dx = \int_{R} \left(\int_{0}^{M} \chi_{[0,f(x)]}(y)dy\right) dx = \int_{R} f(x)dx$$

<u>משפט פוביני</u>

17.1 משפט פוביני – תהיינה $A \subset \mathbb{R}^m$, $B \subset \mathbb{R}^n$ תיבות. תהא $f: A \times B \to \mathbb{R}$ אינטגרבילית. לכל $x \in A$, נגדיר:

$$g_x: B \to \mathbb{R} \quad g_x(y) = f(x,y)$$

אם g_x אינטגרבילית על B לכל x, אזי הפונקציה $G(x) = \int_B g_x(y)dy$ הינה אינטגרבילית (על A) ומתקיים:

$$\int_{A} \left(\int_{B} g_x(y)\, dy\right) dx = \int_{A\times B} f(x,y)dxdy$$

<u>הערה</u> - הסימון dx, dy במקרה זה מתאימים ל-dV_m ו-dV_n המתאימים, וכן $dxdy$ שקול לסימון dV_{m+n}.

<u>הוכחה:</u>

נגדיר:

$$G(x) = \int_B g_x dV_n$$

לפי ההנחה, G מוגדרת היטב. תהא P_A חלוקה של A, וכן P_B חלוקה של B ונסמן:

$$P = P_A \times P_B$$

כחלוקה מתאימה של $A \times B$ המכילה את כל המכפלות הקרטזיות מהצורה $R_A \times R_B$ כאשר $P_A \sim R_A$ ו $P_B \sim R_B$.
נוכיח עתה טענת עזר:

טענה 17.2 – $L(f,P) \le L(G,P_A) \le U(G,P_A) \le U(f,P)$.

<u>הוכחה:</u>

יהא $x \in A$ השייך לתיבה $R_A \sim P_A$. נחשב:

$$G(x) \le U(g_x, P_B) = \sum_{R_B \in P_B} M_{R_B}(g_x) Vol(R_B) = (\star)$$

כאשר:

$$M_{R_B}(g_x) \coloneqq \sup\left\{ \overbrace{g_x(y)}^{f(x,y)} \middle| y \in R_B \right\}$$

מהגדרת M_{R_B} ודאי שמתקיים:

$$M_{R_B}(g_x) \le \sup\{f(\xi,y) | (\xi,y) \in R_A \times R_B\} = M_{R_A \times R_B}(f)$$

ומכאן שמתקיים:

$$(\star) \le \sum_{R_B \in P_B} M_{R_A \times R_B}(f) Vol(R_B)$$

לכן, כיון שזה נכון לכל $x \in R_A$,

$$M_{R_A}(G) \le \sum_{R_B \in P_B} M_{R_A \times R_B}(f) Vol(R_B)$$

ולכן מתקיים:

$$U(G,P_A) = \sum_{R_A \in P_A} M_{R_A}(G) Vol(R_A) \le \sum_{R_A \in P_A} \sum_{R_B \in P_B} M_{R_A \times R_B}(f) Vol(R_B) Vol(R_A) = U(f,P)$$

באופן דומה מקבלים את אי השוויון השני.

עתה, משהוכחנו טענה זו, נוכל להוכיח את המשפט.

בהינתן $\varepsilon > 0$, תהא $P = P_A \times P_B$ חלוקה כך שמתקיים

$$\int_{A\times B} f-\varepsilon < L(f,P) \le U(f,P) < \int_{A\times B} f+\varepsilon$$

מטענת העזר נקבל כי

$$\int f-\varepsilon < L(f,P) \le L(G,P_A) \le \int_A G \le U(G,P_A) \le U(f,P) < \int f+\varepsilon$$

כיון שאי שוויון זה נכון לכל $\varepsilon > 0$, נובע ש-G אינטגרבילית על A וש- $\int_A G = \int_{A\times B} f$, כנדרש. מ.ש.ל.

דוגמאות

כעת נסביר איך משתמשים במשפט פוביני כדי לחשב אינטגרלים בפועל.

דוגמה כללית:

נניח שנתון תחום בעל שטח $\Omega \subset \mathbb{R}^n$ שהינו **פשוט במשתנה האחרון** y, במובן שקיימת קבוצה בעלת נפח $\Omega' \subset \mathbb{R}^{n-1}$ וקיימות פונקציות $g,h : \Omega' \to \mathbb{R}$ כך ש:

$$\Omega = \{(x',y) \in \mathbb{R}^{n-1}\times\mathbb{R} = \mathbb{R}^n : x' \in \Omega', g(x') \le y \le h(x')\}$$

אז בהינתן $f : \Omega \to \mathbb{R}$ אינטגרבילית רימן, אפשר לחשב את האינטגרל של f באופן הבא:

$$(*)\qquad \int_\Omega f = \int_{\Omega'}\left(\int_{g(x')}^{h(x')} f(x',y)dy\right)dx'$$

ואז מבצעים קודם כל את האינטגרציה לפי y ואחר-כך נשארים עם אינטגרל $n-1$ מימדי, כלומר עישנו רדוקציה למקרה עם פחות ממדים. לפעמים צריך לבחור משתנה אחר ולא את המשתנה האחרון, ולפעמים צריך קודם לחלק את התחום לתתי תחומים. בדרך כלל לא טורחים להצדיק את השוויון ב - $(*)$, אבל אנו נעשה זאת עכשיו אחת ולתמיד: נראה איך משפט פוביני מצדיק את השוויון $(*)$.

לפי ההגדרה, אנו צריכים לחשב את $\int_{A\times B}\tilde{f}$, כאשר $A\times B$ היא תיבה המכילה את Ω , הנתונה כמכפלה קרטזית של תיבה A עם קטע (תיבה חד ממדית) B. כאן $\tilde{f}$ זו ההרחבה של f למלבן שמתאפסת מחוץ ל Ω –. נוח לכתוב $\tilde{f} = \chi_\Omega f$, כאשר χ_Ω היא הפונקציה המציינת של Ω. אז, לפי ההגדרה ופי משפט פוביני:

$$\int_\Omega f = \int_{A\times B}\tilde{f} = \int_{A\times B}\chi_\Omega f = \int_A\left(\int_B \chi_\Omega(x',y)f(x',y)dy\right)dx'$$

כעת, לכל $(x',y) \in A\times B$, מתקיים $\chi_\Omega(x',y) = \chi_{\Omega'}(x')\chi_{[g(x'),h(x')]}(y)$. האינטגרנד של האינטגרל הפנימי מתאפס כאשר $y \notin [g(x'),h(x')]$ כך שלכל $x' \in \Omega'$, אפשר לרשום את האינטגרל הפנימי באופן הבא:

$$\int_B \chi_\Omega(x',y)f(x',y)dy = \int_B \chi_{\Omega'}(x')\chi_{[g(x'),h(x')]}(y)f(x',y)dy = \int_{g(x')}^{h(x')}\chi_{\Omega'}(x')\,f(x',y)dy$$

לכן

$$\int_A\left(\int_B \chi_\Omega(x',x_n)f(x',y)dy\right)dx' = \int_A\left(\int_{g(x')}^{h(x')}\chi_{\Omega'}(x')f(x',y)dy\right)dx'$$
$$= \int_A \chi_{\Omega'}(x')\left(\int_{g(x')}^{h(x')} f(x',y)dy\right)dx' = \int_{\Omega'}\left(\int_{g(x')}^{h(x')} f(x',y)dy\right)dx'$$

וקיבלנו את הדרוש, כלומר הצדקנו את $(*)$.

באופן כללי יותר, אם אפשר להציג את התחום באופן הבא:

$$\Omega = \{(x,y) \in \mathbb{R}^m \times \mathbb{R}^n : x \in \Omega', y \in \Omega_x\}$$

אזי אפשר לקבל ממשפט פוביני את הביטוי

$$\int_\Omega f = \int_{\Omega'} \left(\int_{\Omega_x} f(x,y) dy \right) dx$$

על ידי ממש אותם השיקולים. נניח ש - $A \times B \subseteq \mathbb{R}^m \times \mathbb{R}^n$ היא תיבה המכילה את Ω . אם נשתמש בעובדה ש - $\chi_\Omega(x,y) = \chi_{\Omega'}(x)\chi_{\Omega_x}(y)$, נקבל

$$\boxed{\int_\Omega f = \int_{A\times B} \tilde{f} = \int_{A\times B} \chi_\Omega f = \int_A \left(\int_B \chi_{\Omega'}(x)\chi_{\Omega_x}(y) f(x,y) dy \right) dx = \int_{\Omega'} \left(\int_{\Omega_x} f(x,y) dy \right) dx}$$

<u>דוגמה:</u>

נסמן $D = D_n = \{x \in \mathbb{R}^n : \ \|x\|_1 \leq 1\}$, כלומר כדור היחידה בנורמת ℓ^1. נחשב את הנפח של D_n.

המרחב מתחלק ל - 2^n "רביעים", ומספיק לחשב את הנפח של החיתוך $E = E_n \coloneqq D_n \cap P_n$ כאשר $P = P_n =$ $\{x : x_1, \ldots, x_n \geq 0\}$ ואז להכפיל ב - 2^n. (הנכונות של הטריק הזה נובעת ממשפט החלפת המשתנים שנוכיח בהמשך, אך כדאי לכם לשכנע את עצמכם בנכונותו כבר עכשיו). את E_n אפשר לתאר כך:

$$\begin{aligned} E_n &= \{x : x_1 \in [0,1], x_2 \in [0,1-x_1], \ldots, x_n \in [0,1-x_1-\cdots-x_{n-1}]\} \\ &= \{(x', x_n) : x' \in E_{n-1}, x_n \in [0, 1-\|x'\|_1]\} = \{(x',x_n) : x_n \in [0,1], \|x'\|_1 \leq 1 - x_n\} \\ &\subset [0,1]^n = [0,1]^{n-1} \times [0,1] \end{aligned}$$

ולכן

$$\begin{aligned} Vol(E_n) &= \int_{[0,1]^n} \chi_{E_n} = \int_{[0,1]^{n-1}} \left(\int_{[0,1]} \chi_{E_n}(x', x_n)\, dx_n \right) dx' = \int_{[0,1]^{n-1}} \left(\int_0^{1-\|x'\|_1} \chi_{E_{n-1}}(x') dx_n \right) dx' \\ &= \int_{[0,1]^{n-1}} \chi_{E_{n-1}}(x')\,(1 - \|x'\|_1) dx' = \end{aligned}$$

זה לא נראה כל כך כיף! אז אולי נחשב את זה אחרת:

$$\begin{aligned} Vol(E_n) &= \int_{[0,1]^n} \chi_{E_n} = \int_0^1 \left(\int_{(1-x_n)E_{n-1}} dx' \right) dx_n = \int_0^1 Vol\big((1-x_n)E_{n-1}\big) dx_n \\ &= \int_0^1 (1-x_n)^{n-1} Vol(E_{n-1}) dx_n = \frac{Vol(E_{n-1})}{n} \end{aligned}$$

גם את הזהות $Vol\big((1-x_n)E_{n-1}\big) = (1-x_n)^{n-1} Vol(E_{n-1})$ אפשר להסביר באמצעות משפט החלפת המשתנים, אבל אפשר גם להבין לבד (איך מתנהג הנפח של תיבה כאשר מכפילים את התיבה בקבוע?).

לכן מאינדוקציה, תוך שימוש ב - $Vol(E_1) = 1$, נקבל $Vol(E_n) = \frac{1}{n!}$, ולכן $Vol(D_n) = \frac{2^n}{n!}$.

פרק 18

אינטגרל מוכלל

עד עתה דנו באינטגרלים של פונקציות חסומות בתחומים בעל נפח, שהם לפי ההגדרה היו תחומים חסומים. בשימושים רבים עולה הצורך לחשב אינטגרלים של פונקציות לא חסומות, או אינטגרלים של פונקציות על תחומים לא חסומים. בפרק זה נלמד איך לעשות זאת, ונחשב כמה דוגמאות.

18.1 הגדרה - תהי $U \subseteq \mathbb{R}^n$ קבוצה פתוחה ותהא $f: U \to \mathbb{R}_+ = [0,\infty)$ פונקציה שהינה אינטגרבילית רימן על כל תת-קבוצה $K \subset U$ קומפקטית ובעלת נפח. נגדיר

$$\int_U f = \sup\left\{\int_K f : K \subset U, K \text{ has volume is compact}\right\}$$

אם $\int_U f < \infty$ אז נגיד ש - f **אינטגרבילית על U.**

18.2 טענה – תהי $U \subseteq \mathbb{R}^n$ קבוצה פתוחה ותהא $f: U \to \mathbb{R}_+ = [0,\infty)$ פונקציה שהינה אינטגרבילית רימן על כל תת-קבוצה $K \subset U$ קומפקטית ובעלת נפח. תהי $\{K_m\}_{m=1}^{\infty}$ סדרת תתי-קבוצות קומפקטיות ובעלות נפח של U המקיימות:

1. $U = \bigcup_{m=1}^{\infty} K_m$.
2. לכל m, $K_m \subset int(K_{m+1})$.

אזי f אינטגרבילית על U אם ורק אם $\lim\limits_{m\to\infty}\int_{K_m} f < \infty$ ומיקרה זה מתקיים $\int_U f = \lim\limits_{m\to\infty}\int_{K_n} f$.

שימו לב שאנו כותבים $\int_U f = \infty$ כאשר f כמו לעיל איננה אינטגרבילית.

הוכחה:

ברור ש - $\int_{K_m} f$ סדרה עולה והגבול קיים במובן הרחב. יהי $r < \int_U f$, ותהי $K \subset U$ קומפקטית ובעלת נפח המקיימת $\int_K f > r$. לפי תנאי 1. ו- 2. מתקיים $U = \bigcup_m int(K_m)$, ולכן (מתכונת הכיסוי הסופי של קבוצות קומפקטיות) קיים m כך ש - $K \subset int(K_m)$, ולכן $\int_{K_m} f \ge \int_K f > r$. לכן $\lim \int_{K_m} f \ge \int_U f$. מ.ש.ל.

תרגיל: הוכיחו שלכל U פתוחה קיימת סדרה $\{K_m\}_{m=1}^{\infty}$ של תתי קבוצות קומפקטיות המקיימת את התנאים בתרגיל.

הערה: במקרים רבים נוח למצוא משפחה של קבוצות $\{K_t\}$ המאונדקסות על-ידי פרמטר רציף, והאינטגרל מחושב על-ידי גבול $\lim\limits_{t\to 0}\int_{K_t} f$ או $\lim\limits_{t\to\infty}\int_{K_t} f$. המעבר בין גבול סדרתי לגבול רציף צריך להיות שקוף לבוגרי אינפי 1, ובכל מקרה הוכחת הגרסה של טענה 18.2 עבור משפחה המאונדקסת על-ידי פרמטר רציף, מהווה תרגיל קל.

עד כאן הגדרנו אינטגרל של פונקציה אי שלילית. בהינתן פונקציה f נוכל להציגה כהפרש של שתי פונקציה אי-שליליות

$$f = f_+ - f_-$$

כאשר $f_+(x) = \frac{|f(x)|+f(x)}{2} = \max(f(x),0)$ ו - $f_-(x) = \frac{|f(x)|-f(x)}{2} = -\min(f(x),0)$.

18.3 הגדרה – תהי $U \subseteq \mathbb{R}^n$ קבוצה פתוחה ותהא $f: U \to \mathbb{R}$ פונקציה שהינה אינטגרבילית רימן על כל תת-קבוצה $K \subset U$ קומפקטית ובעלת נפח. נאמר ש – f הינה **אינטגרבילית על U** אם f_+, f_- שתיהן אינטגרביליות, ובמקרה זה נגדיר

$$\int_U f = \int_U f_+ - \int_U f_-$$

18.4 טענה – תהי $f: U \to \mathbb{R}$ פונקציה שהינה אינטגרבילית רימן על כל תת-קבוצה $K \subset U$ קומפקטית ובעלת נפח. אזי f אינטגרבילית על U אם ורק אם $|f|$ אינטגרבילית. במקרה הזה מתקיים השוויון $\int_U f = \lim_{m\to\infty} \int_{K_m} f$ לכל סדרה $\{K_m\}_{m=1}^{\infty}$ המקיימת את תנאי טענה 18.2.

<u>הוכחה:</u> מתקיים $|f| = f_+ + f_-$ וכל הפונקציות הללו אי-שליליות, לכן

$$\sup_{\mathrm{K}} \int_K f = \sup_{\mathrm{K}} \int_K f_+ + \int_K f_-$$

לכן אינטגרביליות של החלקים החיובי והשלילי גוררים אינטגרביליות של הערך המוחלט. כמו כן מתקיים $|f| \geq f_+, f_-$ ולכן משיקולים של "קריטריון ההשוואה" אינטגרביליות של הערך המוחלט גורר אינטגרביליות של החלקים החיובי והשלילי. השוויון $\int_U f = \lim_{m\to\infty} \int_{K_m} f$ נובע כעת מאריתמטיקה של גבולות. מ.ש.ל.

שימו לב: באינטגרל רב מימדי אין לנו מושג של "התכנסות בתנאי". האינטגרל מוגדר עבור פונקציות אי-שליליות (ואז הוא יכול לצאת אינסוף), ועבור פונקציות שהן "אינטגרביליות בהחלט". השתמשנו בהוכחת הטענה הקודמת במעין "קריטריון השוואה". ננסח אותו להלן למעו הסדר הטוב.

18.5 טענה (קריטריון ההשוואה) – תהיינה $f, g: U \to \mathbb{R}$ פונקציות אינטגרבילית רימן על כל תת-קבוצה $K \subset U$ קומפקטית ובעלת נפח. אם f אינטגרבילית ומתקיים $|g| \leq |f|$ אזי גם g אינטגרבילית ומתקיים:

$$\left|\int g\right| \leq \int |g| \leq \int |f|$$

את ההוכחה נשאיר כתרגיל.

<u>תרגיל:</u> הוכיחו שאם U קבוצה פתוחה וחסומה שהיא בעלת נפח. אזי אפשר לחשב את האינטגרל $\int_U f$ של פונקציה על U בשתי דרכים: לפי ההגדרה הקודמת של אינטגרל רימן על תחום בעל נפח, ולפי ההגדרה של אינטגרל מוכלל. הוכיחו ששתי ההדרכים מובילות לאותו ערך, כלומר שלכל סדרת קומפקטיות שממצות את U, מתקיים:

$$\lim \int_{K_m} f = \int_U f$$

<u>דוגמה:</u> נקבע n, ונבדוק עבור אילו ערכי $p > 0$ הפונקציה $f(x) = \frac{1}{\|x\|^p}$ אינטגרבילית על הקובייה $U = (0,1)^n$.

למען הסדר הטוב, וכיון שזו הפעם הראשונה שלנו, נהיה הכי פורמליים שאפשר: נגדיר $K_m = \left[\frac{1}{m}, 1 - \frac{1}{m}\right]^n$ וזוהי סדרה שממצה את U כפי שנדרש. נחשב

$$\int_{K_m} f = \int_{1/m}^{1-1/m} \cdots \int_{1/m}^{1-1/m} \frac{1}{\|x\|^p} dx_1 \cdots dx_n$$

מאי שוויון הממוצעים $\frac{a_1+\cdots+a_n}{n} \geq (a_1 \cdots a_n)^{\frac{1}{n}}$ נובע ש –

$$\|x\| = (x_1^2 + \cdots + x_n^2)^{1/2} \geq n^{\frac{1}{2}} (x_1^2 \cdots x_n^2)^{\frac{1}{2n}} = n^{\frac{1}{2}} x_1^{\frac{1}{n}} \cdots x_n^{\frac{1}{n}}$$

ולכן

$$\int_{\frac{1}{m}}^{1-\frac{1}{m}}\cdots\int_{\frac{1}{m}}^{1-\frac{1}{m}}\frac{1}{\|x\|^p}dx_1\cdots dx_n \le \int_{\frac{1}{m}}^{1-\frac{1}{m}}\cdots\int_{\frac{1}{m}}^{1-\frac{1}{m}}\frac{1}{n^{\frac{p}{2}}x_1^{\frac{p}{2}}\cdots x_n^{\frac{p}{n}}}dx_1\cdots dx_n = n^{-p/2}\prod_{i=1}^{n}\int_{1/m}^{1-1/m}\frac{dx_i}{x_i^{p/n}}$$

אנו יודעים שהאינטגרלים החד ממדיים המוכללים מתכנסים אם ורק אם $\frac{p}{n}<1$ ולכן נקבל שתנאי מספיק עבור אינטגרביליות של f הוא ש - $p<n$.

כעת אם $p \ge n$, נחסום $\|x\| \le n\|x\|_\infty$. עבור $\epsilon > 0$ קטן מספיק

$$\begin{aligned}\int_{\frac{1}{m}}^{1-\frac{1}{m}}\cdots\int_{\frac{1}{m}}^{1-\frac{1}{m}}\frac{1}{\|x\|^p}dx_1\cdots dx_n &\ge \frac{1}{n}\int_{\frac{1}{m}}^{1-\frac{1}{m}}\cdots\int_{\frac{1}{m}}^{1-\frac{1}{m}}\frac{1}{\|x\|_\infty^p}dx_1\cdots dx_n\\ &\ge \frac{1}{n}\int_{\frac{1}{m}}^{\frac{1}{m}+\epsilon}\cdots\int_{\frac{1}{m}}^{\frac{1}{m}+\epsilon}\frac{1}{\|x\|_\infty^p}dx_1\cdots dx_n \ge \frac{1}{n}\int_{\frac{1}{m}}^{\frac{1}{m}+\epsilon}\cdots\int_{\frac{1}{m}}^{\frac{1}{m}+\epsilon}\frac{1}{\left(\frac{1}{m}+\epsilon\right)^p}dx_1\cdots dx_n\\ &= \frac{\epsilon^n}{n\left(\frac{1}{m}+\epsilon\right)^p}\end{aligned}$$

הביטוי האחרון שואף ל- $n^{-1}\epsilon^{n-p}$ וזה מספיק כדי להראות שהאינטגרל לא מתכנס עבור $p>n$, כי זה אומר שהאינטגרלים גדולים יותר מ - $n^{-1}\epsilon^{n-p}$ לכל $\epsilon > 0$. עבור $p = n$ דרוש שיקול מעט עדין יותר. נרשה לעצמנו לעבוד לא בדיוק לפי ההגדרה, כפי שעושים בפועל:

$$\int_U \frac{1}{\|x\|^p}dx_1\cdots dx_n \ge \sum_{k\ge 1}\int_{\left[2^{-k-1},2^{-k}\right]^n}\frac{1}{\|x\|^p}dx_1\cdots dx_n \ge \sum_k 2^{-(k+1)n}2^{kp} = \frac{1}{2}\sum_k 1 = \infty$$

בשביל השוויון הלפני אחרון השתמשנו בכך ש - $n=p$, אך אם $p>n$ היינו יכולים לקבל אי-שוויון.

כדאי לשבת ולהבין מדוע השיקול הנ"ל מוצדק.

פרק 19

משפט החלפת המשתנים

הגענו למשפט המרכזי בחלק זה של הקורס העוסק באינטגרל רימן.

19.1 משפט החלפת המשתנים – יהא $\Omega \subseteq \mathbb{R}^n$ תחום בעל נפח, ונניח כי $\Omega \subseteq U$ פתוחה, ושנתונה $g \in C^1(U,\mathbb{R}^n)$. נניח כי g חד-חד ערכית ובעלת נגזרת הפיכה בכל נקודה. אם מתקיים:

א. $g(\Omega)$ תחום בעל נפח.
ב. $f: g(\Omega) \to \mathbb{R}$ אינטגרבילית.
ג. $f \circ g: \Omega \to \mathbb{R}$ אינטגרבילית.

אזי:

$$\int_{g(\Omega)} f(y)dy = \int_{\Omega} [f \circ g(x)]\left|det\big(Dg(x)\big)\right|dx$$

הערות:

א. הנוסחה בעצם מראה ש- g כזו לוקחת כל קבוצה מנפח אפס לקבוצה מנפח אפס. אפשר גם להראות זאת ישירות (ואכן, מצטרך להסביר עניין זה במהלך ההוכחה).
ב. מספיק להניח ש- g מקיימת את התנאים למעט בקבוצה בעלת נפח אפס, על מנת שהנוסחה עדיין תהיינה נכונה. בהרבה שימושים נשתמש בנוסחה כאשר התנאים על g מתקיימים למעט על קבוצה מנפח אפס.
ג. אפשר להראות ש-$g(\Omega)$ בעל נפח תחת ההנחה ש-Ω בעל נפח וש - g מקיימת את התנאים שבמשפט (גם זה נובע מההערה הראשונה). עם זאת, כמעט בכל השימושים (אם לא בכולם) העובדה ש - $g(\Omega)$ תחום בעל נפח לא דורשת הוכחה או הצדקה מיוחדים. לכן הוספנו הנחה זו לניסוח המשפט כדי להקל על עצמנו את מלאכת ההוכחה.
ד. בדומה להערה ב', ניתן להוכיח ש- $f \circ g: \Omega \to \mathbb{R}$ אינטגרבילית אם f, g מקיימות את התנאים במשפט, ואין צורך להוסיף את זה כתנאי נפרד. בשימושים של המשפט כמעט תמיד ברור שפונקציה זו אינטגרבילית, ולכן אנו מרשים לעצמנו להניח את זה כדי לפשט את ההוכחה.
ה. אם Ω היא עצמה קבוצה פתוחה בעלת נפח אז המשפט תקף גם כאשר g מוגדרת על Ω, כלומר פשוט לוקחים $U = \Omega$ ותנאי המשפט מתקיימים.
ו. המשפט נכון גם עבור אינטגרל מוכלל ו-Ω הינה קבוצה פתוחה. אנו לא הוספנו את זה לניסוח של המשפט כדי לא לסרבל אותו, אבל ההוכחה שלנו עבור המקרה בו Ω קבוצה פתוחה עובדת גם עבור אינטגרל מוכלל.

נוכיח את המשפט בהמשך. קודם, דרושות לנו מספר הכנות.

19.2 הגדרה – אם $v_1, \cdots, v_n \in \mathbb{R}^n$ אזי **המקבילון הנוצר על ידי וקטורים** אלו זו הקבוצה:

$$P(v_1, \cdots, v_n) = \{\sum a_i v_i : 0 \leq a_i \leq 1\}$$

אם $a \in \mathbb{R}^n$ אזי גם $a + P(v_1, \cdots, v_n)$ נקרא **מקבילון**. כאשר $v_1, \cdots, v_n$ תלויים ליניארית, המקבילון נקרא **מקבילון מנוון**.

תרגיל – השטח של מקבילית ב-$\mathbb{R}^2$ נתון על ידי בסיס × גובה. כלומר:

$$Vol\big(a + P(v_1, v_2)\big) = |\det(v_1|v_2)| = \|v_1\| \cdot \|v_2\| |\sin\angle(v_1, v_2)|$$

המשמעות של הסימון $\det(v_1|v_2)$ הוא הדטרמיננטה של המטריצה שעמודותיה הן v_1, v_2 .

(הערה: שימו לב שלא הוכחנו ששטח במובן של גיאומטריה מישורית מבית-הספר מתלכד עם המובן של שטח כפי שהגדרנו אותו. לכן יש כאן קצת יותר מה להסביר ממה שנראה לעין במבט ראשון).

<u>תרגיל</u> – אם $\Omega_2 \subseteq \mathbb{R}^n$, $\Omega_1 \subseteq \mathbb{R}^m$ תחומים בעלי נפח, אזי מתקיים:

$$Vol(\Omega_1 \times \Omega_2) = Vol(\Omega_1)Vol(\Omega_2)$$

<u>תרגיל</u> – 1) אם Q מקבילון לא מנוון אזי Q תחום בעל נפח ונפחו חיובי, ואם Q מנוון אזי ל-Q נפח אפס.

2) אם $T: \mathbb{R}^n \to \mathbb{R}^n$ ליניארית אזי אם Q מקבילון, גם $T(Q)$ הוא מקבילון, ולכן הוא בעל נפח.

19.3 משפט – אם $\Omega \subset \mathbb{R}^n$ תחום בעל נפח ו-$T: \mathbb{R}^n \to \mathbb{R}^n$ ליניארית, אזי $T(\Omega)$ בעל נפח, ומתקיים:

$$Vol\big(T(\Omega)\big) = |det(T)| \cdot Vol(\Omega)$$

<u>הערה</u> – בהינתן מקבילון $P(v_1, \cdots, v_n)$, ניתן לבטאו $P(v_1, \cdots, v_n) = T\big(P(e_1, \cdots, e_n)\big)$ עבור העתקה לינארית T המקיימת $v_i = Te_i$. נקבל:

$$[T] = [v_1| \cdots |v_n]$$

ולכן הנוסחה ממשפט 19.3 נותנת את הנפח של מקבילון :

$$Vol\big(P(v_1, \cdots, v_n)\big) = |\det T| \cdot Vol\big(P(e_1, \cdots, e_n)\big) = |\det T|$$

למי ששאל: זוהי אחת מהמשמעויות החשובות של דטרמיננטה.

<u>הוכחה:</u> נוכיח עתה את המשפט. אם T לא הפיכה אז השוויון הנדרש $Vol\big(T(\Omega)\big) = 0 = |det(T)| \cdot Vol(\Omega)$ מתקיים.

נניח כי T הפיכה. במקרה זה, T ניתנת להצגה כמכפלה של מטריצות אלמנטריות מהצורה:

$$T = E_1 \cdots E_m$$

כאשר כל E_k לכל $1 \le k \le m$ היא מטריצה מהצורה:

$$\underset{\text{הוספת שורה מוכפלת בסקלר}}{\begin{pmatrix} 1 & 0 & \cdots & C & 0 \\ \vdots & \cdots & 0 & \cdots & \vdots \\ \vdots & & 1 & & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix}} \quad \underset{\text{החלפת שורות}}{\begin{pmatrix} 1 & & \cdots & & 0 \\ & 0 & 0 & \ddots & \\ \vdots & \vdots & 0 & \vdots & 1 \\ & \ddots & \cdots & \ddots & \\ 0 & & 1 & & 0 \end{pmatrix}} \quad \underset{\text{מכפלת שורה בסקלר}}{\begin{pmatrix} 1 & & \cdots & & 0 \\ & \ddots & \cdots & \ddots & \\ \vdots & \vdots & C & \vdots & \vdots \\ & \ddots & \cdots & \ddots & \\ 0 & & \cdots & & 1 \end{pmatrix}}$$

בשלב ראשון, נראה שהנוסחה מתקיימת כאשר Ω הינה תיבה ו- T הינה מטריצה אלמנטרית.

בהינתן תיבה מהצורה:

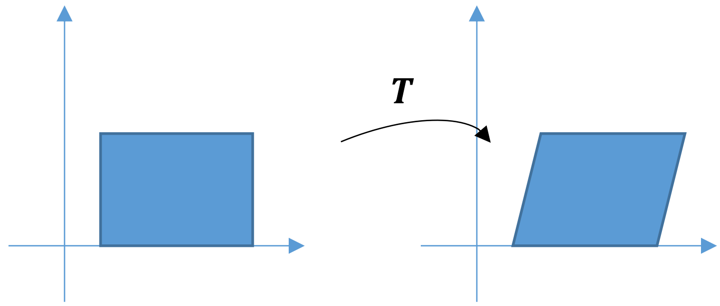
$$Q = a + [0, L_1] \times \cdots \times [0, L_n]$$

נשים לב כי אם T הינה מטריצה של החלפת שורות או הכפלת שורה בסקלר, ברור שמתקיימת המסקנה של משפט 19.3. אם T אלמנטרית מסוג הוספת כפולה של שורה אחת לאחרת, אזי מתקיים (לדוגמא):

$$T(Q) = Ta + \left(\begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix} [0, L_1] \times [0, L_2]\right) \times \cdots \times [0, L_n]$$

ובמקרה הדו ממדי, לדוגמה, נקל לראות כי:

$$Vol\left(\begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix} [0, L_1] \times [0, L_2]\right) = Vol(L_1 \times L_2)$$



איור 19 – במקרה הדו ממדי. קל לראות כי טרנספורמציה אלמנטרית של הוספת שורה מוכפלת בסקלר לא משנה את ה"בסיס" ואת ה"גובה" ולכן שטח המקבילית נשאר זהה

לכן הנוסחה עבור תיבות ומטריצה אלמנטרית נובעת מהתרגיל השני לעיל.

כעת נטפל במקרה בו Ω תחום בעל נפח כללי, והמטריצה T היא אלמנטרית. בפרק הקודם ראינו כי:

$$Vol(\Omega) = \inf_P \sum_{Q \in P_\subseteq} Vol(Q) = \inf_P U(\chi_\Omega, P) = \int_R \chi_\Omega$$

כאשר

$$P_\subseteq := \{Q \in P | Q \cap \Omega \neq \emptyset\}$$

והאינפימום רץ על כל החלוקות P של תיבה כלשהי R המכילה את Ω . כעת נסמן לכל קבוצה חסומה $\Omega \subset \mathbb{R}^n$

$$Vol^*(\Omega) = \inf_P \sum_{Q \in P_\subseteq} Vol(Q)$$

והאינפימום רץ על כל החלוקות P של תיבה כלשהי R המכילה את Ω. אפשר לקרוא לביטוי הנ"ל ה"נפח החיצוני" של Ω. לפי ההגדרה $Vol(\Omega) = Vol^*(\Omega)$ כאשר Ω בעלת נפח. אנו זקוקים למושג הזה כי נרצה להעריך את הנפח של גוף לפני שנדע להוכיח שהוא בעל נפח. שתי תכונות חשובות ברורות לעין: 1) אם $\Omega \subseteq D$, אזי $\mathrm{Vol}^*(\Omega) \leq \mathrm{Vol}^*(D)$ 2) אם $\mathrm{Vol}^*(\Omega) = 0$ אם ורק אם Ω בעלת נפח אפס.

בהינתן חלוקה P כלשהי של R , נקבל "חלוקה" של $T(R)$ למקבילונים לא חופפים מהצורה $\{T(R_i) | R_i \in P\}$, ו"חלוקה" זו מכסה את $T(\Omega)$ במובן ש- $T(\Omega) \subseteq \bigcup_{Q \in P_\subseteq} T(Q)$ ("חלוקה" כתובה עם מרכאות, כי אין זו חלוקה במובן בו השתמשנו בתורת האינטגרציה).

מההגדרה של נפח נובע ש- $Vol^*\big(T(\Omega)\big) \leq \sum_{Q \in P_\subseteq} Vol\big(T(Q)\big)$ ולכן מתקיים:

$$Vol^*\big(T(\Omega)\big) \leq \sum_{Q \in P_\subseteq} Vol\big(T(Q)\big) \overset{(\star)}{=} \sum_{Q \in P_\subseteq} |\det T| Vol(Q)$$

כאשר השוויון $(\star)$ מהנוסחה עבור תיבות (ומטריצות אלמנטריות), שהוכחנו לעיל. כאשר נבצע inf באגף ימין על כל החלוקות P נקבל כי:

$$Vol^*\big(T(\Omega)\big) \leq |\det T| Vol(\Omega)$$

לכל T אלמנטרית ולכל תחום בעל נפח Ω. אם נשכח לרגע ש - Ω בעלת נפח, אז נשים לב שבעצם הוכחנו ש –

$$Vol^*\big(T(\Omega)\big) \leq |\det T| Vol^*(\Omega)$$

לכל T אלמנטרית ולכל תחום חסום Ω. מכאן, באינדוקציה, נובע אי השוויון $Vol^*\big(T(\Omega)\big) \leq |\det T| Vol(\Omega)$ לכל T הפיכה ולכל תחום בעל נפח Ω.

כעת נשתמש בעובדה ש T – העתקה לינארית הפיכה ולכן היא **הומיאומורפיזם** – העתקה רציפה והפיכה עם הפיך רציף. לכן T משמרת את כל התכונות הטופולוגיות, ובפרט שולחת שפה ושפה. לכן אם Ω בעלת נפח,

$$Vol^*\big(\partial(T\Omega)\big) = Vol^*\big(T(\partial\Omega)\big) \leq |\det T| Vol(\partial\Omega) = 0$$

ולכן Ω בעלת נפח, כפי שהיה צריך להוכיח. כעת מותר לכתוב את מה שהוכחנו בצורה הבאה:

$$Vol\big(T(\Omega)\big) \leq |\det T| Vol(\Omega)$$

כיון שהוכחנו את האי-שוויון הזה **לכל** T ו- Ω נפעיל אותו עבור T^{-1} ו- $T(\Omega)$ ונקבל

$$Vol(\Omega) = Vol\big(T^{-1}(\mathrm{T}(\Omega))\big) \leq |\det T|^{-1} Vol\big(T(\Omega)\big)$$

וזה נותן לנו את האי-שוויון השני. בכך סיימנו את הוכחת משפט 19.3. מ.ש.ל.

משפט 19.3 הוא בעצם משפט החלפת המשתנים (משפט 19.1) עבור המקרה המיוחד בו מבצעים אינטגרציה על פונקציה f קבועה, ו"החלפת המשתנים" g הינה טרנספורמציה לינארית. האסטרטגיה שלנו להוכחת משפט 19.1 היא לחלק את התחום להרבה חלקים קטנים, כאשר הודות לגזירות של "החלפת המשתנים" g נוכל לקרב אותה היטב על-ידי הנגזרת שלה (שהיא טרנספורמציה לינארית). לפני שניגש להוכחה, אלינו להבין איך נפחים מתנהגים תחת הפעולה של העתקה גזירה.

נגדיר עתה סימונים – $\|x\|_\infty = \max_i |x_i|$ וכן לצורך ההוכחה נסמן:

$$\|T\| = \|T\|_{\infty,op} = \sup_{x\neq 0} \frac{\|Tx\|_\infty}{\|x\|_\infty} \quad T: \mathbb{R}^n \to \mathbb{R}^n$$

זוהי נורמת האופרטורים על M_n המושרית על-ידי הנורמה $\|\cdot\|_\infty$.

19.4 למה – תהא $g \in C^1$ מוגדרת והפיכה עם הפיך גזיר (גלומר דיפאומורפיזם) בסביבה של תיבה $Q \subseteq \mathbb{R}^n$ עם מרכז ב-a. יהא $\varepsilon > 0$. נניח כי:

$$\forall x \in Q. \quad \|Dg(a)^{-1} Dg(x) - I_n\| < \varepsilon$$

אזי $g(Q)$ קבוצה בעלת נפח ומתקיים:

$$Vol\big(g(Q)\big) \leq (1+\varepsilon)^n |det\, Dg(a)| Vol(Q)$$

<u>הערה:</u> המשמעות של הלמה היא, שאם העתקה $g \in C^1$ פועלת על תיבה Q מספיק קטנה, הנפח של $g(Q)$ יהיה לא הרבה יותר גדול ממה שהיינו מקבלים אם היינו מחליפים את g בהעתקה הלינארית $Dg(a)$. כלומר, הלמה נותנת לנו **הערכה כמותית** של הביטוי הלא מדויק "g היא בקירוב העתקה ליניארית".

<u>הוכחה:</u>

בדומה להוכחת משפט 19.3, העובדה ש – $g(Q)$ בעלת נפח נובעת מקיום האי שוויון

$$Vol^*\big(g(Q)\big) \leq (1+\varepsilon)^n |det\, Dg(a)| Vol(Q)$$

כי אז נקבל שהשפה של $g(Q)$ הינה בעלת נפח אפס. אנו נקצר את ההוכחה בכך שניקח כעובדה את זה ש - $g(Q)$ בעלת נפח, ונתמקד בהוכחת אי השוויון.

נניח כי Q קוביה עם צד באורך $2r$ ונגדיר $\phi:[-r,r]^n\to\mathbb{R}^n$ על ידי:

$$\phi(x)=Dg(a)^{-1}\big(g(a+x)-g(a)\big)$$

אזי מתקיים:

$$D\phi(x)=Dg(a)^{-1}Dg(x+a)$$

כמו כן $\phi(0)=0$. אם נחבר זאת עם:

$$\|D\phi\|_{\infty,op}\le\|I_n\|+\|Dg(a)^{-1}Dg(a+x)-I_n\|<1+\varepsilon$$

נקבל (ראו התרגיל אחרי טענה 7.2)

$$\phi([-r,r]^n)\subseteq[-(1+\varepsilon)r,(1+\varepsilon)r]^n$$

נשים לב כי לכל $x\in[-r,r]^n$ מתקיים $x+a\in Q$ וכן:

$$g(x+a)=Dg(a)\phi(x)+g(a)$$

ובאופן כללי נוכל לכתוב:

$$g(Q)=Dg(a)\cdot\phi([-r,r]^n)+g(a)\subseteq Dg(a)([-(1+\varepsilon)r,(1+\varepsilon)r]^n)+g(a)$$

ומכאן שמתקיים:

$$Vol\big(g(Q)\big)\le|\det Dg(a)|(1+\varepsilon)^nVol(Q)$$

מ.ש.ל.

<u>תרגיל</u> – הוכחנו את אי השוויון עבור קובייה (תיבה שכל הצדדים שלה באורך שווה). הוכיחו את אי השוויון עבור תיבה כללית.

עתה, נוכיח את משפט החלפת המשתנים (משפט 19.1):

<u>הוכחת משפט 19.1:</u> אנו נוכיח את המשפט לשני מקרים חשובים: עבור Ω קומפקטית ובעלת נפח, ובסוף עבור Ω פתוחה.

נניח תחילה ש - Ω קומפקטית ובעלת נפח. נרצה להראות כי:

$$\int_{g(\Omega)}f(y)dy=\int_\Omega f\circ g(x)\,|\det Dg(x)|dx$$

לכל פונקציה אינטגרבילית f . כיון שכל פונקציה אינטגרבילית f הינה הפרש של שתי פונקציות אי-שליליות ואינטגרביליות

$$f=\frac{1}{2}(|f|+f)-\frac{1}{2}(|f|-f)=f_+-f_-$$

נוכל להסתפק בהוכחת המשפט עבור $f\ge 0$. לכן נניח מעתה ש- f הינה אי שלילית.

תהא P חלוקה (של איזושהי תיבה גדולה שמכילה את Ω) כך שכל תיבה שתורמת לסכום $U(f\circ g|\det Dg|,P)$ נמצאת בתוך U. נסמן $P_\subseteq=\{R_i\}=\{N\sim P : N\cap\Omega\neq\emptyset\}$ וכן נסמן:

$$Q = \bigcup_{R_i \cap \Omega \neq \emptyset} R_i$$

יהא $\varepsilon > 0$. כמו כן, יהא M כך שמתקיים $|f| \leq M$ וכן $\|(Dg)^{-1}\| \leq M$ ב-Q.

נניח כי P עדינה מספיק כך שמתקיים:

$$\begin{matrix} \forall x, y \in R_i \\ \forall R_i \end{matrix} \qquad \begin{matrix} \|Dg(x) - Dg(y)\|_\infty < \dfrac{\varepsilon}{M} \\ \left|\det\big(Dg(x)\big) - \det\big(Dg(y)\big)\right| < \dfrac{\varepsilon}{M} \end{matrix}$$

(אם P לא עדינה מספיק, נוכל לקחת עידון שמקיים זאת – זה רק יקטין את הקבוצה Q). כעת אם a_i המרכז של התיבה R_i, אזי מתקיים:

$$\|Dg(a_i)^{-1}Dg(x) - I\|_\infty \leq \|Dg(a_i)^{-1}\|\|Dg(x) - Dg(a_i)\| < \varepsilon$$

עתה נשים לב, כי:

$$\int_{g(\Omega)} f(y)dy \leq \sum \int_{g(R_i)} f(y)dy \leq \sum M_i Vol\big(g(R_i)\big) \leq (\star)$$

כאשר סימנו:

$$M_i = \sup_{y \in g(R_i)} f(y) = \sup_{x \in R_i} f \circ g(x)$$

לפי הלמה מתקיים:

$$(\star) \leq \sum M_i (1+\varepsilon)^n |\det Dg(a_i)| Vol(R_i) \leq \sum \widetilde{M_i}(1+\varepsilon)^n Vol(R_i) + \varepsilon(1+\varepsilon)^n \sum Vol(R_i)$$

וזאת עבור הסימון:

$$\widetilde{M_i} = \sup_{s \in R_i} f \circ g(x) |\det Dg(x)|$$

ונעזרנו בכך שמתקיים:

$$M_i |\det Dg(a_i)| \leq \widetilde{M_i} + \varepsilon.$$

בסה"כ קיבלנו שלכל $\epsilon > 0$ קיימת חלוקה P כך ש-

$$\int_{g(\Omega)} f(y)dy \leq U(f \circ g |\det Dg|, P)(1+\varepsilon)^n + \varepsilon(1+\varepsilon)^n \cdot C$$

מכאן, על-ידי לקיחת אינפימום על כל החלוקות העדינות יותר מ- P, נובע שלכל $\epsilon > 0$

$$\int_{g(\Omega)} f(y)dy \leq (1+\varepsilon)^n \cdot \int_\Omega f \circ g(x) |\det Dg(x)| dx + \varepsilon(1+\varepsilon)^n \cdot C$$

ומכאן , ע"י התבוננות ב- $\epsilon \to 0$, נקבל

$$\int_{g(\Omega)} f(y)dy \leq \int_\Omega f \circ g(x) |\det Dg(x)| dx$$

לכל $f \geq 0$ ולכל g ו- Ω קומפקטית בעלת נפח. כמו בהוכחת משפט 19.5, מהעובדה שאי-שוויון מתקיים לכל $f \geq 0$ ולכל g ו- Ω אנו נקבל שמתקיים שוויון. אכן, אם נציב באי השוויון הזה במקום f, g, Ω את הנתונים

$$\tilde{f} = f \circ g \,|\det Dg|, \qquad \tilde{\Omega} = g(\Omega) \text{ ו- } \tilde{g} = g^{-1}$$

נקבל את השוויון בכיוון ההפוך:

$$\int_{\Omega} f \circ g(y)\,|\det Dg(y)|dy = \int_{\tilde{g}(\tilde{\Omega})} \tilde{f}(y)dy \leq \int_{\tilde{\Omega}} \tilde{f} \circ \tilde{g}(x)\,|\det D\tilde{g}(x)|dx = \int_{g(\Omega)} f(x)dx$$

ולכן קיבלנו

$$\int_{g(\Omega)} f(y)dy = \int_{\Omega} f \circ g(x)\,|\det Dg(x)|dx$$

לכל $f \geq 0$. כעת עבור f כללית נרשום $f = \frac{1}{2}(|f| + f) - \frac{1}{2}(|f| - f)$ ונקבל את התוצאה עבור כל f אינטגרבילית.

עד כאן ההוכחה עבור Ω קומפקטית ובעלת נפח.

אם Ω קבוצה פתוחה (לאו דווקא חסומה), אז נוכל למצוא סדרה $\{\mathrm{K}_m\}$ כמו בטענה 18.2 (קיום מובטח מהתרגיל שאחרי הטענה), ואז נקבל לכל m

$$\int_{g(K_m)} f(y)dy = \int_{K_m} f \circ g(x)\,|\det Dg(x)|dx$$

לפי החלק הקודם. כעת ניקח את הגבול כאשר $m \to \infty$ ונקבל בגבול

$$\int_{g(\Omega)} f(y)dy = \int_{\Omega} f \circ g(x)\,|\det Dg(x)|dx$$

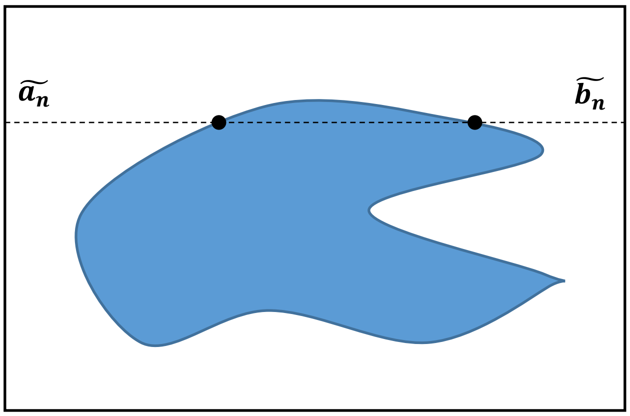
מ.ש.ל.

<u>לתרגול/תרגיל בית:</u> השלמות ותוספות בעניין נפח מקבילית ומקבילונים. הוכחת תרגילים שנדרשו לנו. דוגמאות!! נפח כדור אן-ממדי.

נספח לפרק 19 – דוגמאות

חישוב אינטגרלים באמצעות אינטגרל נשנה

$$\int_R f dV = \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \cdots \int_{a_n}^{b_n} f dx_n$$



אם f אינטגרבילית ב"תחום פשוט", כלומר בתחום $\Omega \subset \mathbb{R}^n$ מהצורה:

$$\Omega = \begin{cases} a_1 \le x_1 \le b_1 \\ a_2(x_1) \le x_2 \le b_2(x_1) \\ a_3(x_1, x_2) \le x_3 \le b_3(x_1, x_2) \\ \vdots \\ a_n(x_1, \cdots, x_{n-1}) \le x_n \le b_n(x_1, \cdots, x_{n-1}) \end{cases}$$

ומקיימת את תנאי משפט פוביני במלבן $\Omega \subset R$ אזי מתקיים:

$$\int_\Omega f dv = \int_R \tilde{f} dv = \int_{\widetilde{a_1}}^{\widetilde{b_1}} dx_1 \int_{\widetilde{a_2}}^{\widetilde{b_2}} dx_2 \cdots \int_{\widetilde{a_n}}^{\widetilde{b_n}} \tilde{f} dx_n$$
$$= \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \cdots \int_{a_n}^{b_n} f dx_n$$

<u>דוגמאות:</u>

נגדיר:

$$f(x,y) = \begin{cases} 1 & x = 0, y \in \mathbb{Q} \\ 0 & \text{אחרת} \end{cases}$$

עבור המלבן $R = [0,1] \times [0,1]$. נשים לב כי f אינטגרבילית על R (שכן הפונקציה מתאפסת בכל מקום למעט על ציר ה-y בקבוצה בעלת שטח אפס), ומתקיים $\int_R f dV = 0$. אך נשים לב כי הנ"ל לא מתקבל ממשפט פוביני בניסוח הפשוט שאנו נתנו (משפט 17.1) שכן:

$$F(x) = \int_0^1 f(x,y) dy$$

לא מוגדר עבור $x = 0$.

מאידך, נתבונן בפונקציה:

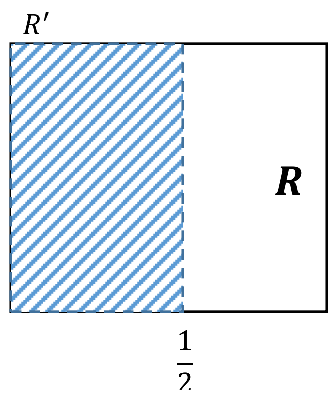
$$f(x,y) = \begin{cases} 1 & y \in Q \\ 2x & y \notin Q \end{cases}$$

עבור המלבן $R = [0,1] \times [0,1]$. לכל y הפונקציה $f(\cdot, y)$ היא אינטגרבילית. נשים לב כי:

$$\forall y \quad F(y) = \int_0^1 f(x,y) dx = 1 \Longrightarrow \int_0^1 F(y) dy = 1$$

$$\Longrightarrow \int_0^1 dy \int_0^1 f(x,y) dx = 1$$

אך נראה עתה, כי על אף הקיום של האינטגרל הנשנה, f איננה אינטגרבילית במלבן.



אם f אינטגרבילית במלבן R אזי היא בהכרח אינטגרבילית גם במלבן $R' = [0,1/2\,] \times [0,1]$. לכל חלוקה P של R' מתקיים:

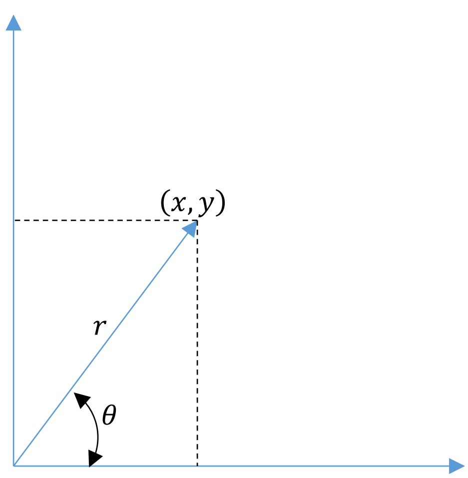
$$U(f,P) = \frac{1}{2}$$

מצד שני,

$$\int_0^{1/2} f(x,y)dx = \begin{cases} 1/2 & y \in Q \\ 1/4 & y \notin Q \end{cases}$$

וזו לא פונקציה אינטגרבילית, בניגוד למסקנה ממשפט פוביני (מומלץ לחזור למשפט 17.1 ולהבין לאיזה חלק מהמשפט אנו מתכוונים כאן.

החלפת משתנים: <u>תזכורת</u> – בקואורדינטות פולריות, הגדרנו את הטרנספורמציה:



$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

כאשר $r = \sqrt{x^2+y^2}$ זהו המרחק מן הראשית. ביטוי זה מוגדר היטב בכל $\mathbb{R}^2$ ובפרט $r \geq 0$ לכל (x,y).

θ זוהי הזווית ביחס לציר x ולכן נעה על מקטע שאורכו 2π. ניתן להגדיר למשל מטעמי נוחות $0 \leq \theta < 2\pi$. זווית זו מוגדרת לכל $\mathbb{R}^2 \setminus \{(0,0)\}$.

כאשר אנו מבצעים החלפת משתנים, אנו מעתיקים את $\mathbb{R}^2$ לפס:

$$\begin{cases} r \geq 0 \\ 0 \leq \theta \leq 2\pi \end{cases}$$

וזוהי העתקה C^1 וחד-חד ערכית, למעט קבוצה ממידה אפס. בנוסף, נשים לב כי מתקיים:

$$J = \left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| = \begin{vmatrix} x_r & y_r \\ x_\theta & y_\theta \end{vmatrix} = r \neq 0$$

למעט קבוצה ממידה אפס. משפט החלפת המשתנים (משפט 19.1):

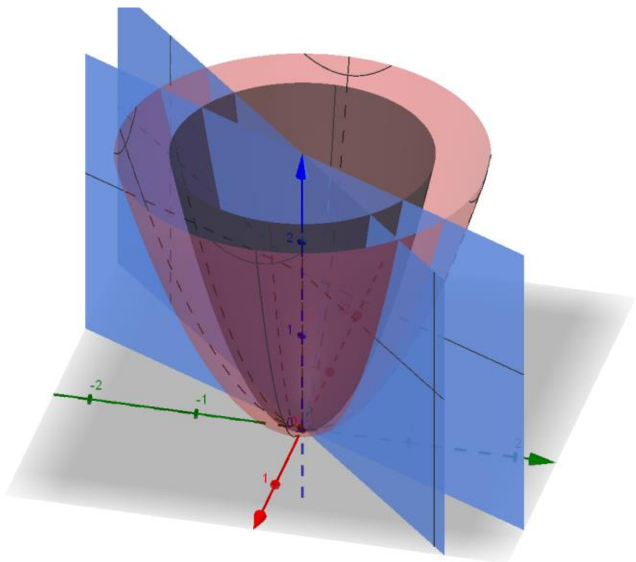
$$\iint_D f(x,y)dxdy = \iint_{\tilde{D}} f(r\cos\theta\,, r\sin\theta)|r|drd\theta$$

כפי שראיתם באינפי II ואולי גם בקורסים בפיזיקה.

<u>החלפת משתנים ב-$\mathbb{R}^3$:</u>

יהא $\Omega \subset \mathbb{R}^3$ הגוף ברביע הראשון החסום על ידי:

$$y = x \quad y = 2x \quad z = x^2+y^2 \quad z = 2(x^2+y^2)$$



וכן:

$$z=h,\quad h>0$$

ונרצה לחשב את הנפח של Ω עבור $x,y,z\geq 0$:

$$V=\iiint_{\Omega}1dxdydz$$

נעבוד בקואורדינטות גליליות, r,θ,z.
הטרנספורמציה במקרה זה תהיה:

$$\begin{cases}x=r\cos\theta\\ y=r\sin\theta\\ z=z\end{cases}$$

נחשב את היעקוביאן של העתקה זו ונקבל כי:

$$J=\frac{\partial(x,y,z)}{\partial(r,\theta,z)}=\begin{vmatrix}x_r & y_r & z_r\\ x_\theta & y_\theta & z_\theta\\ x_z & y_z & z_z\end{vmatrix}$$

$$J=\begin{vmatrix}\cos\theta & \sin\theta & 0\\ -r\sin\theta & r\cos\theta & 0\\ 0 & 0 & 1\end{vmatrix}=r\neq 0$$

למעט המקרים של הנקודות שנמצאות בדיוק על ציר z אך במקרה של חסימות על ידי $z<h$ נקבל כי זוהי קבוצה בעלת נפח אפס ולכן לא מהווה בעיה.

נתבונן בתמונת Ω על ידי הצגת החלפת המשתנים באמצעות אילוצים:

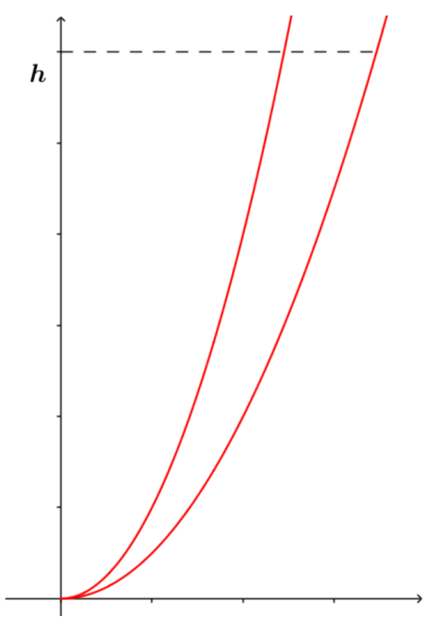
$$y=x\Longrightarrow r\sin\theta=r\cos\theta\Longrightarrow\tan\theta=1\Longrightarrow\theta=\frac{\pi}{4}$$

$$y=2x\Longrightarrow r\sin\theta=2r\cos\theta\Longrightarrow\tan\theta=2\Longrightarrow\theta=\arctan 2$$

$$z=x^2+y^2\Longrightarrow z=r^2$$

$$z=2(x^2+y^2)\Longrightarrow z=2r^2$$

$$z=h$$



כאשר למעט דרישות אלו נזכור כי תנאי לטרנספורמציה הינו $r\geq 0$ וכן $0\leq\theta\leq 2\pi$.

סה"כ נקבל כי האילוצים שנקבל הם:

$$\frac{\pi}{4}\leq\theta\leq\arctan 2\qquad\sqrt{\frac{z}{2}}\leq r\leq\sqrt{z}\quad 0\leq z\leq h$$

ולכן האינטגרל החדש שיתקבל יהיה:

$$V=\iiint_{\tilde{\Omega}}rdrd\theta dz=\int_{\frac{\pi}{4}}^{\arctan 2}d\theta\int_0^h dz\int_{\sqrt{\frac{z}{2}}}^{\sqrt{z}}rdr$$

$$= \left(\arctan 2 - \frac{\pi}{4}\right)\int_0^h \left[\frac{z}{2} - \frac{z}{4}\right] dz = \frac{h^2}{8}\left(\arctan 2 - \frac{\pi}{4}\right)$$

<u>הערה</u> – למעשה, שימוש בקואורדינטות גלילות שימושי במיוחד כאשר עוסקים באובייקטים המתקבלים מנפח גוף סיבוב (או חלק ממנו).

<u>דוגמה נוספת:</u>

נרצה לחשב את נפחו של כדור בעל רדיוס R. לשם כך נתבונן בכדור מהצורה:

$$\Omega = \{(x,y,z) | (x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2 \le R^2\}$$

נפח הכדור, אינו תלוי בנקודת המרכז. אכן, נגדיר:

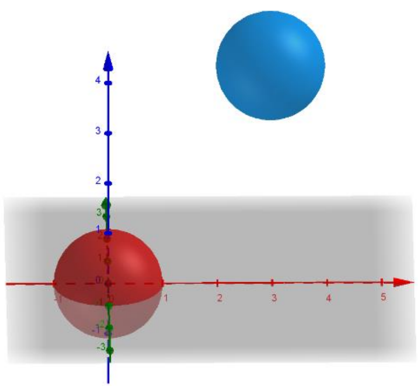
$$\begin{cases} \tilde{x} = x - x_0 \\ \tilde{y} = y - y_0 \\ \tilde{z} = z - z_0 \end{cases}$$

ותמונת הכדור $B(x_0, R)$ על ידי העתקה זו תהיה פשוט הכדור $B(0,R)$ במרחב $\tilde{x}, \tilde{y}, \tilde{z}$.

נחשב את היעקוביאן של העתקה זו:

$$J^{-1} = \frac{\partial(\tilde{x},\tilde{y},\tilde{z})}{\partial(x,y,z)} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1$$

(זהו ההופכי של היעקוביאן שכן היעקוביאן מחושב על ידי גזירת המשתנים הישנים לפי משתני "ההעתקה" החדשים).
לכן נקבל כי:



איור 20 –הכדור הישן בכחול והחדש באדום, לדוגמה

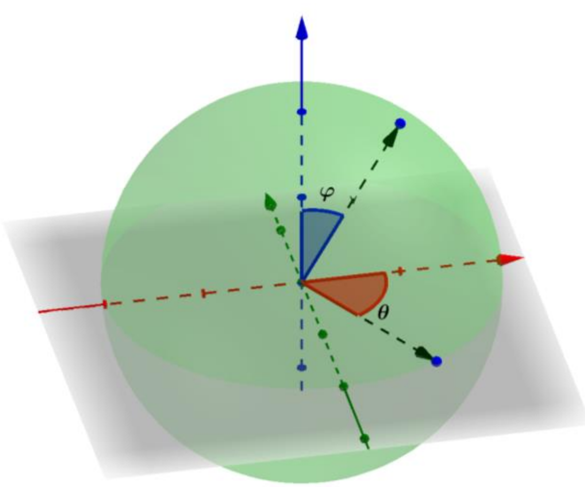
$$J = (J^{-1})^{-1} = 1 \neq 0$$

נקבל כי:

$$V_{B(x_0,R)} = \iiint_{B(x_0,R)} 1 dxdydz = \iiint_{B(0,R)} 1 J d\tilde{x}d\tilde{y}d\tilde{z}$$
$$= \iiint_{B(0,R)} 1 d\tilde{x}d\tilde{y}d\tilde{z} \ (\star)$$

ניעזר בקואורדינטות ספריות (כדוריות) המתקבלות על ידי הטרנספורמציה הבאה:

$$\begin{cases} \tilde{x} = r\cos\theta\sin\varphi & r \ge 0 \\ \tilde{y} = r\sin\theta\sin\varphi & 0 \le \theta \le 2\pi \\ \tilde{z} = r\cos\varphi & 0 \le \varphi \le \pi \end{cases}$$



היעקוביאן של העתקה זו הינו:

$$J=\frac{\partial(\tilde{x},\tilde{y},\tilde{z})}{\partial(r,\theta,\varphi)}=\begin{vmatrix}\cos\theta\sin\varphi & \sin\theta\sin\varphi & \cos\varphi\\ -r\sin\theta\sin\varphi & r\cos\theta\sin\varphi & 0\\ r\cos\theta\cos\varphi & r\sin\theta\cos\varphi & -r\sin\varphi\end{vmatrix}=$$

$$=\cos\varphi\,[-r^2\sin\varphi\cos\varphi]-r\sin\varphi\,[rsin^2\varphi]=-r^2\sin\varphi$$

$$|J|=r^2\sin\varphi$$

תמונת $B(0,R)$ מתקבלת תחת האילוץ:

$$r^2\le R^2$$

ולכן:

$$\tilde{\Omega}=\begin{cases}0\le\theta\le 2\pi\\ 0\le\varphi\le\pi\\ 0\le r\le R\end{cases}$$

כלומר, במקרה זה קיבלנו כי $\tilde{\Omega}$ מלבן ולכן:

$$(\star)=\iiint_{\tilde{\Omega}} r^2\sin\varphi\,drd\theta d\varphi=\int_0^R r^2dr\cdot\int_0^\pi\sin\varphi\,d\varphi\int_0^{2\pi}d\theta=\frac{4}{3}\pi R^3$$

<u>הערה</u> – קואורדינטות כדוריות שימושיות כאשר Ω הינו פלח תחום על סימטריה סיבובית סביב הראשית.

פרק 20

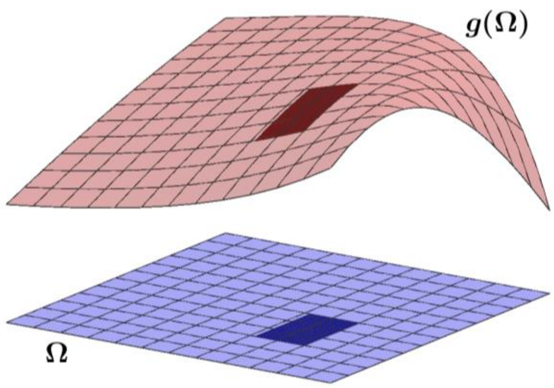
תזכורת – משפט החלפת המשתנים:

ראינו כי תחת התנאים הנדרשים במשפט זה מתקיים:

$$\int_{g(\Omega)} f dV = \int_{g(\Omega)} f(y) dV = \int_{\Omega} f \circ g |J| dV = \int_{\Omega} f \circ g |\det Dg(x)| dV(x)$$

בפרט:

$$Vol\big(g(\Omega)\big) = \int_{g(\Omega)} 1 dV(y) = \int_{\Omega} |\det Dy(x)| dV(x)$$



איור 21 – דוגמה לפרמטריזציה של יריעה כלשהי ו"פיסות" שטח על היריעה ועל קבוצת המקור שלה ביחס לפרמטריזציה (הערת המרצה: האיור הזה שייך לנושא הבא, שטח פנים של משטחים)

כפי שניתן ללמוד מן האיור **הסכמתי** שלהלן, היות ואנו יודעים כי g בסביבה לוקלית של כל נקודה מתנהגת בדומה להעתקה ליניארית, אנו יודעים שככל שנבחר תיבה קטנה יותר במקור, תמונת g שתתקבל עבור תיבה זו תהיה קרובה יותר ויותר למקבילית אשר $|\det Dg(x)|$ הוא היחס בין שטח המקבילית לבין שטח התיבה שנבחרה. (הערה: בציור הזה אין להבין את המשטח העליון כגרף של פונקציה, אלא כגרסה מעוקמת של המישור).

כלומר, עבור תיבה במקור ששטחה נתון על ידי ביטוי מהצורה dx_1dx_2, ההעתקה g מבטיחה כי שטח המקבילית שתתקבל בתמונה ישאף ל-$|\det Dg(x)|dx_1dx_2$.

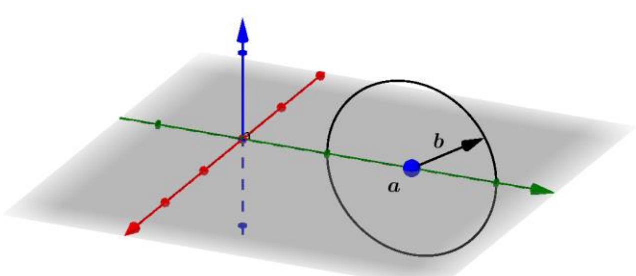
על מנת לקבל את כל השטח של $g(\Omega)$, עלינו לסכום את שטחי כל המקביליות.

דוגמה:

נניח כי $a < b$ כלשהם, אזי עבור $0 \le u \le 2\pi \quad 0 \le v \le 2\pi$ נגדיר:

$$\varphi(u,v) = \begin{bmatrix} \cos v & 0 & -\sin v \\ 0 & 1 & 0 \\ \sin v & 0 & \cos v \end{bmatrix} \begin{bmatrix} a + b\cos u \\ b \sin u \\ 0 \end{bmatrix} = \begin{bmatrix} (a + b\cos u)\cos v \\ b\sin u \\ (a + b\cos u)\sin v \end{bmatrix}$$

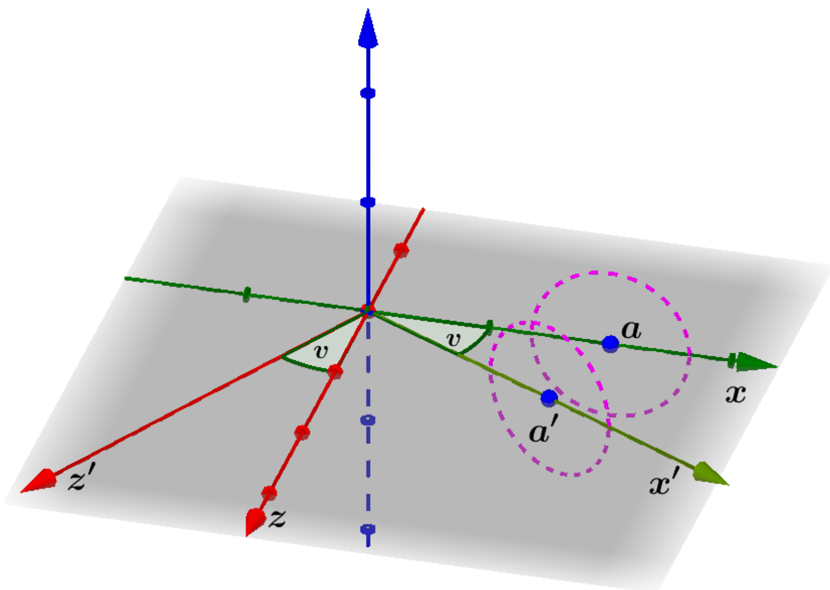
זוהי פרמטריזציה של טורוס. על מנת לקבל אותה, מבחינה אינטואיטיבית, נתבונן במישור xy, ונבחר בנקודה a על ציר ה-x. סביב נקודה זו נבנה מעגל ברדיוס b, כמתואר באיור 2. (נניח כי y הוא הציר הכחול, x הציר הירוק ו-z הוא הציר האדום).



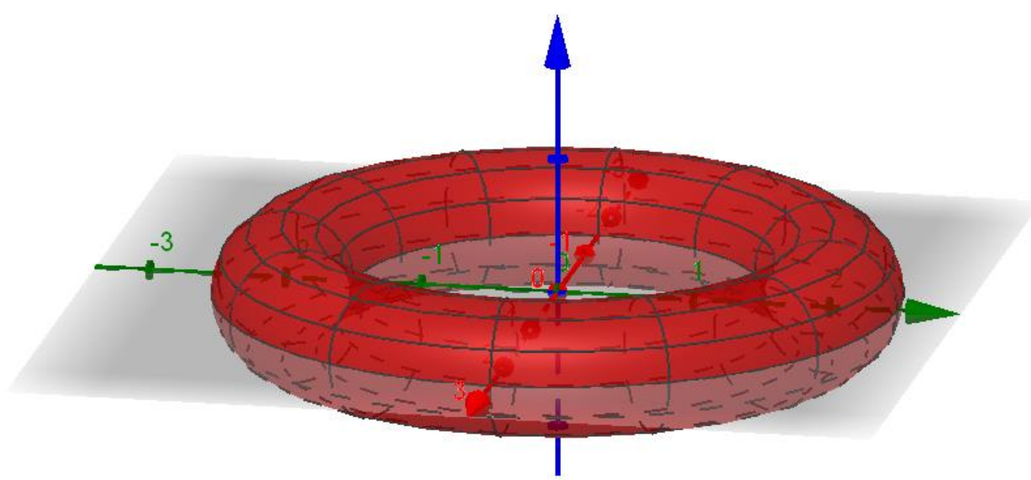
איור 22 – שלב א' בבניית הטורוס, בחירת המעגל הראשון לסיבוב (x ירוק, y כחול, z אדום)

לאחר מכן, את המעגל הנ"ל נרצה להגדיר עבור כל "סיבוב" של מערכת הצירים כשציר y הינו ציר הסיבוב. לשם כך נרצה להרכיב על נקודות מעגל זה סיבוב בזווית $0 \le v \le 2\pi$.

לשם כך נרצה להבין כיצד נראה אופרטור הסיבוב סביב ציר y במרחב תלת ממדי.



איור 24 – הדגמה של סיבוב המעגל על ידי סיבוב כל המערכת בזווית v



איור 23 – טורוס בעל רדיוס ראשי 2 ורדיוס משני 0.5

אך עתה נרצה להגדיר פרמטריזציה לפנים של הטורוס (הטורוס המלא). ולכן נוכל להגדיר:

$$g(t,u,v)=\begin{bmatrix}(a+t\cos u)\cos v\\ t\sin u\\ (a+t\cos u)\sin v\end{bmatrix}\quad \Omega=[0,b]\times[0,2\pi]\times[0,2\pi]$$

והנפח שנרצה לחשב הוא $Vol(\widetilde{\Omega})$ כאשר $\widetilde{\Omega} = g(\Omega)$. g, באופן כללי, איננה חד-חד ערכית על Ω והנגזרת שלה אף אינה רגולרית בכל מקום, אך הנ"ל מתקיים בקבוצה בעלת שטח 0. כפי שהסברנו בצמוד למשפט 19.1, אין הנ"ל משפיע על חישוב הנפח.

g אכן חד-חד ערכית על התחום $(0,b] \times [0,2\pi) \times [0,2\pi)$ ומתקיים:

$$Dg = \begin{bmatrix} \cos u \sin v & -t \sin u \cos v & -(a + t\cos u)\sin v \\ \sin u & t\cos u & 0 \\ \cos u \sin v & -t \sin u \sin v & (a + t\cos u)\cos v \end{bmatrix}$$

$detDg \neq 0$ למעט במקרה שבו $t = 0$ או $a = -t\cos u$ אך כל הנקודות בהן הנ"ל מתקיים הינן קבוצת נקודת בעלת שטח אפס.

$$|\det Dg| = \cdots = t(a + t\cos u)$$

ולכן:

$$Vol\big(g(\Omega)\big) = \int_{g(\Omega)} 1 dV_{(x,y,z)} = \int_{\Omega} 1|\det Dg(t,u,v)| dV_{(t,u,v)}$$

$$= \int_{[0,b]\times[0,2\pi]\times[0,2\pi]} t(a + t\cos u) dV \overset{\text{פוביני}}{=} \int_0^{2\pi}\int_0^{2\pi}\int_0^{b} t(a + t\cos u) dt du dv$$

$$= \int_0^{2\pi}\int_0^{2\pi} \left(\frac{b^2 a}{2} + \frac{b^3}{3}\cos u\right) du dv = 2\pi^2 b^2 a$$

יפה וטוב, חישבנו את נפח הטורוס המלא. כעת נעבור לשאלה אחרת: מה עלינו לעשות אם נרצה לחשב את שטח הפנים של טורוס? לשם כך, עלינו קודם להגדיר מהו שטח פנים. ניתן להלן טיפול מעט כללי יותר ממה שדרוש עבור המקרה הדו-ממדי.

נפח k-ממדי:

יהיו $v_1, \cdots, v_k \in \mathbb{R}^n$ ו-a , כאשר $k < n$. נרצה להגדיר נפח k-ממדי של מקבילון הנתון על ידי $a + P(v_1, \cdots, v_k)$.

נניח תחילה, כי $v_i = \begin{bmatrix} v_i^1 \\ \vdots \\ v_i^n \end{bmatrix}$ כאשר $v_i^{k+1} = \cdots = v_i^n = 0$, כלומר ניתן לכתוב:

$$v_i = \begin{bmatrix} \widetilde{v_i} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \widetilde{v_i} \in \mathbb{R}^k \quad \forall 1 \le i \le k$$

כיון שהמקבילון יושב בתוך $\mathbb{R}^k \subset \mathbb{R}^n$, טבעי להגדיר במקרה זה:

$$Vol_k\big(a + P(v_1, \cdots v_k)\big) := Vol\big(P(\widetilde{v_1}, \cdots, \widetilde{v_k})\big) = |\det[\widetilde{v_1}| \cdots |\widetilde{v_k}]| = \big|\det \widetilde{V}\big|$$

כאשר נסמן:

$$\widetilde{V} = [\widetilde{v_1}| \cdots |\widetilde{v_k}]$$

נוכל לרשום:

$$V = [v_1| \cdots |v_k] = \begin{bmatrix} \widetilde{V} \\ 0 \end{bmatrix}$$

נקבל כי V הינה מטריצת בלוקים כאשר $V^*=[\tilde{V}^* \quad 0]$ (כאשר אנו משמנים ב - $V^*=V^T$ את המטריצה המשוחלפת).

ולכן מתקיים:

$$\left|\det \tilde{V}\right| = \sqrt{\det \tilde{V}^*\tilde{V}} = \sqrt{\det V^*V}$$

ההגדרה לעיל "עושה שכל" גם כש - $v_1,\cdots,v_k$ אינם בהכרח מוכלים בתוך תת המרחב של הוקטורים בעלי הקואורדינטות האחרונות שמתאפסות.

20.1 הגדרה - יהיו $v_1,\cdots,v_k\in\mathbb{R}^n$ וקטורים כלשהם. <u>הנפח ה-k ממדי של המקבילון $a+P(v_1,\cdots,v_k)$</u> **מוגדר** להיות:

$$Vol_k\big(a+P(v_1,\cdots,v_k)\big) = \sqrt{\det V^*V} \quad V=[v_1|\cdots|v_k]$$

<u>מקרה מיוחד</u> – כאשר $k=2, n=3$, נקבל כי:

$$v_1=\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix} \quad v_2=\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}$$

$$Vol\big(P(v_1,v_2)\big) = \sqrt{\det\left(\begin{pmatrix}x_1 & y_1 & z_1\\ x_2 & y_2 & z_2\end{pmatrix}\begin{pmatrix}x_1 & x_2\\ y_1 & y_2\\ z_1 & z_2\end{pmatrix}\right)}$$

$$= \sqrt{\det\begin{pmatrix}\langle v_1,v_1\rangle & \langle v_1,v_2\rangle\\ \langle v_2,v_1\rangle & \langle v_2,v_2\rangle\end{pmatrix}} = \cdots = \sqrt{(x_1y_2-y_1x_2)^2+(x_1z_2-x_2z_1)^2+(y_1z_2-y_2z_1)^2}$$

$$= \|v_1\times v_2\|$$

וזאת כאשר המכפלה הוקטורית הנ"ל מוגדרת להיות:

$$v_1\times v_2 := \begin{vmatrix}\hat{\imath} & \hat{\jmath} & \hat{k}\\ x_1 & y_1 & z_1\\ x_2 & y_2 & z_2\end{vmatrix} = \hat{\imath}\begin{vmatrix}y_1 & z_1\\ y_2 & z_2\end{vmatrix} - \hat{\jmath}\begin{vmatrix}x_1 & z_1\\ x_2 & z_2\end{vmatrix} + \hat{k}\begin{vmatrix}x_1 & y_1\\ x_2 & y_2\end{vmatrix}$$

נשים לב כי בניגוד למכפלה הסקלרית, במקרה זה $v_1\times v_2$ הוא וקטור שאורכו הוא $\|v_1\|\|v_2\|\cdot\sin\alpha$ כאשר α הינה הזווית בין שני הווקטורים. כיוון וקטור זה יהיה ניצב למישור שוקטורים אלה פורשים (וסימנו נקבע לפי "כלל יד ימין").

מכאן ששטח המקבילית שהם פורשים נתון על ידי:

$$\|v_1\times v_2\| = \|v_1\|\|v_2\||\sin\alpha|$$

מה עושות העתקות ליניאריות למקבילונים?

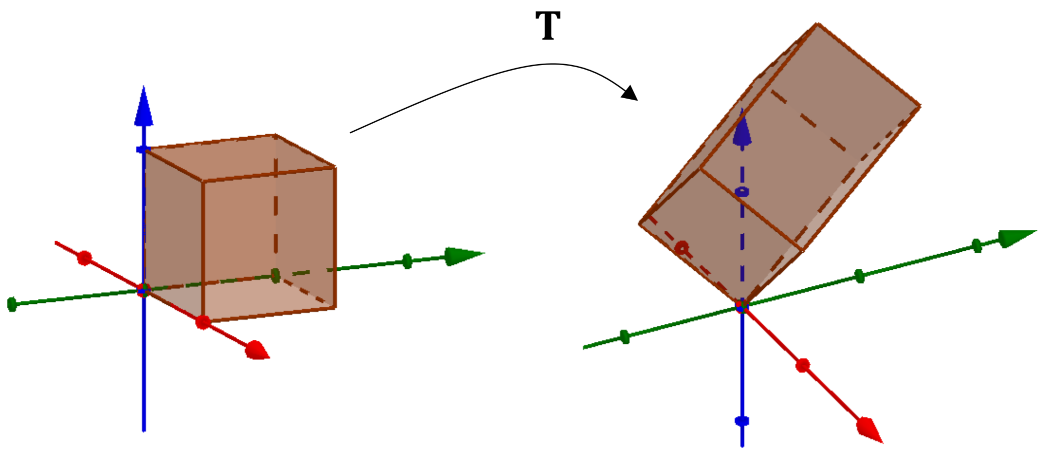
תהא $T:\mathbb{R}^k\to\mathbb{R}^n$ העתקה ליניארית כאשר $k\le n$. אם $v_1,\cdots,v_k\in\mathbb{R}^k$ אזי מתקיים:

$$T\big(P(v_1,\cdots v_k)\big) = P(Tv_1,\cdots Tv_k)$$

ואכן מתקיים:

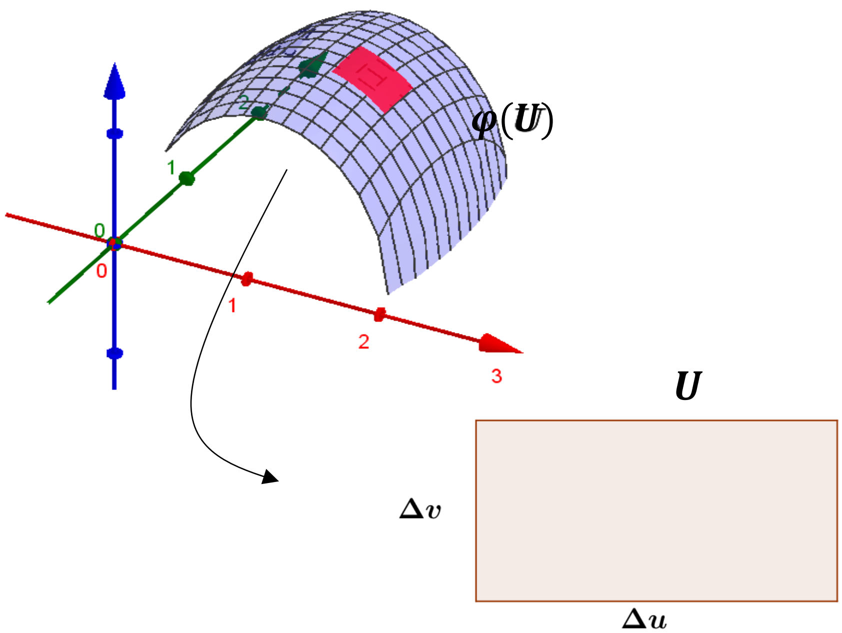
$$Vol_k\Big(T\big(P(v_1,\cdots,v_k)\big)\Big) = \sqrt{\det(TV)^*TV} = \sqrt{\det V^*\det T^*T\det V} = \sqrt{\det T^*T}\,Vol_k\big(P(v_1,\cdots,v_k)\big)$$

20.2 מסקנה – העתקה ליניארית $T:\mathbb{R}^k\to\mathbb{R}^n$ מנפחת נפח k-ממדי של מקבילון בגורם $\sqrt{\det T^*T}$.



איור25 – דוגמה להעתקה ליניארית במרחב תלת ממדי המעתיקה תיבה למקבילון

עתה, נניח $U \subset \mathbb{R}^2$ ו-$\varphi: U \to \mathbb{R}^3$ פרמטריזציה של פיסה מירִיעה $S = \varphi(U)$, בפרט מקיים ש - φ חד-חד ערכית ו-C^1 וכמו כן $\operatorname{rank} D\varphi = 2$.



איור 26 – תיאור הגדרת שטח פנים של פיסה מירִיעה

$$\text{מלבן} = (u,v) + P\left(\begin{pmatrix}\Delta u\\ 0\end{pmatrix}, \begin{pmatrix}0\\ \Delta v\end{pmatrix}\right) \to\approx \varphi(u,v) + D\varphi\left(P\left(\begin{pmatrix}\Delta u\\ 0\end{pmatrix}, \begin{pmatrix}0\\ \Delta v\end{pmatrix}\right)\right)$$

$$D\varphi\begin{pmatrix}0\\ \Delta v\end{pmatrix}=\varphi_v\Delta v, D\varphi\begin{pmatrix}\Delta u\\ 0\end{pmatrix}=\varphi_u\Delta u \Longrightarrow A(S)\approx\|\varphi_u\times\varphi_v\|\Delta v\Delta u$$

20.3 הגדרה – שטח פנים של S הנתונה כפי שתיארנו זה עתה, מוגדרת להיות:

$$A(S)=\int_U\sqrt{det\left(D\varphi^*(u,v)D\varphi(u,v)\right)}dudv=\int_U\|\varphi_u\times\varphi_v\|dudv$$

<u>מקרה חשוב</u> – אם $T\equiv\varphi$ ליניארית כלשהי, אזי $D\varphi=T$ ומתקיים:

$$A(S)=\int_U\sqrt{\det T^*T}\,dudv=\sqrt{\det T^*T}\,Vol_2(U)$$

<u>פרק 21</u>

21.1 הגדרה – תהא $U \subset \mathbb{R}^k$ וכן $\Omega \subset U$ תחום בעל נפח. תהא $\varphi: U \to \mathbb{R}^n$ פרמטריזציה (כמו בהגדרה השקולה של יריעה), כלומר $\varphi \in C^1$, חד-חד ערכית, עם הפיך רציף, כך שמתקיים $\operatorname{rank} D\varphi \equiv k$. נסמן $S = \varphi(\Omega)$. נגדיר את **שטח הפנים ה-k-ממדי של** S, להיות:

$$A_k(S) := \int_\Omega (\det(D\varphi^* D\varphi))^{\frac{1}{2}} dV$$

<u>הערה</u> – אפשר לדרוש חד-חד ערכיות או $\operatorname{rank} D\varphi = k$ למעט קבוצה בעלת נפח אפס בה זה לא יתקיים.

<u>הערה 2</u> – כאשר $k = 2$, $n = 3$, נקבל את ההגדרה של שטח פנים של (פיסה של) משטח דו ממדי במרחב התלת ממדי.

<u>דוגמה</u> – מקרה פרטי – עבור $k = 1$, למשל $\varphi: [a,b] \to \mathbb{R}^n$ כך ש-$\varphi'(t) \neq 0$ לכל $t \in [a,b]$, נקבל כי $A_1(\varphi[a,b])$ הוא "שטח פנים 1-ממדי של $\varphi[a,b]$", כלומר, מדובר על אורכו של עקום. ואכן:

$$\varphi(t) = \begin{bmatrix} \varphi_1(t) \\ \vdots \\ \varphi_n(t) \end{bmatrix} \qquad \varphi'(t) = \begin{bmatrix} \varphi_1'(t) \\ \vdots \\ \varphi_n'(t) \end{bmatrix} = D\varphi|_t$$

מכאן שמתקיים:

$$D\varphi^*(t) D\varphi(t) = [\varphi_1'(t), \cdots, \varphi_n'(t)] \begin{bmatrix} \varphi_1'(t) \\ \vdots \\ \varphi_n'(t) \end{bmatrix} = \sum_{i=1}^{n} (\varphi_i'(t))^2$$

כלומר יתקיים:

$$A_1(\varphi[a,b]) = \int_a^b \sqrt{\sum_{i=1}^{n} (\varphi'(t))^2}\, dt = \int_a^b \|\varphi'(t)\| dt$$

ואכן, הגדרה 21.1 תואמת את ההגדרה שלנו לאורך עקום שהוכחה קודם לכן (ראו פרק מס' 13).

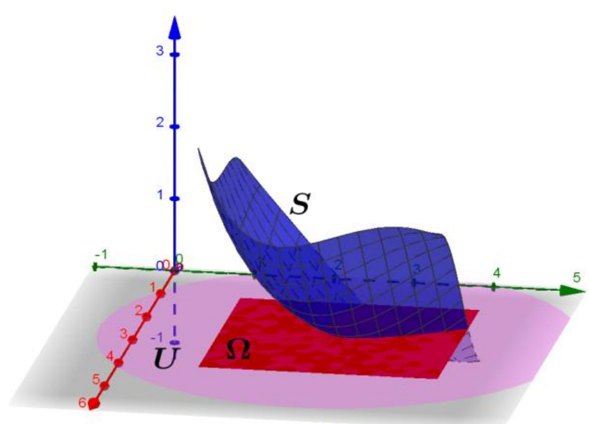
כעת, משיש לנו מושג של שטח פנים, נוכל לבצע אינטגרציה על משטחים.

21.2 הגדרה – (בסימוני ההגדרה הקודמת), נניח כי $k = 2, n = 3$. תהא $f: \overset{\varphi(\Omega)}{S} \to \mathbb{R}$ רציפה. אזי **האינטגרל הסקלרי** (או **מסוג ראשון**) של f על S מוגדר להיות:

$$\overbrace{\int_S f dA = \int_S f d\sigma = \int_S f(\sigma) d\sigma}^{\text{סינומים סינוש}}$$

$$:= \int_\Omega (f \circ \varphi) \cdot (\det(D\varphi^* D\varphi))^{\frac{1}{2}} dV$$

$$= \int_\Omega (f \circ \varphi) \|\varphi_u \times \varphi_v\| du dv$$



<u>דוגמה:</u> רוצים לחשב משקל של גלגל-ים (טורוס):

$$S = \mathbb{T}^2 = \varphi([0,2\pi] \times [0,2\pi])$$

כאשר:

$$\varphi(u,v) = \begin{bmatrix}(a+b\cos u)\cos v\\ b\sin u\\ (a+b\cos u)\sin v\end{bmatrix} = \begin{bmatrix}x\\ y\\ z\end{bmatrix}$$

נניח, כמו כן, כי צפיפות המסה של גלגל זה נתונה על ידי:

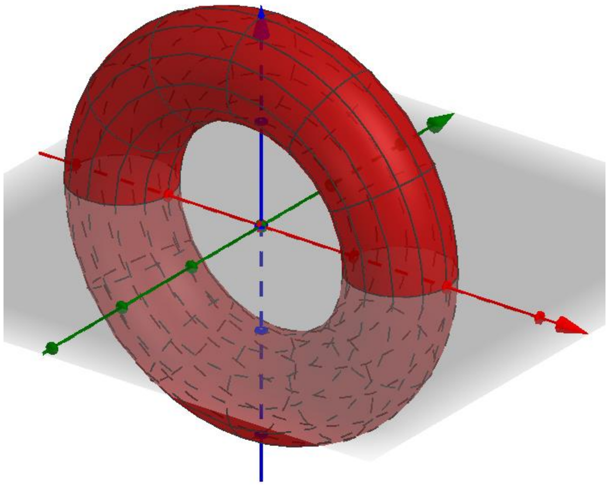
$$f(x,y,z) = \frac{1}{\sqrt{x^2+z^2}} = \rho(x,y,z)$$

(לצורך העניין, ביחידות של מסה חלקי אורך בריבוע). המסה הכוללת של גלגל הים נתונה ידי:

$$\int_S f dA = \int_{[0,2\pi]\times[0,2\pi]} f \circ \varphi(u,v)\|\varphi_u \times \varphi_v\| dudv$$

$$\int_0^{2\pi}\int_0^{2\pi} \frac{1}{a+b\cos u} b(a+b\cos u)dudv$$

$$= \cdots = 4b\pi^2$$



איור 27 – גלגל-ים, קרי $\varphi([0,2\pi]\times[0,2\pi])$

הגדרנו את האינטגרל הסקלרי על משטח (ובפרט את שטח הפנים של המשטח) באמצעות פרמטריזציה של המשטח. עלינו לשאול את עצמנו האם בבחירת פרמטריזציה אחרת נקבל ערך אחר של האינטגרל? (ובפרט, האם שטח הפנים של משטח תלוי בפרמטריזציה?)

21.3 טענה (אינטגרל משטחי אינו תלוי בפרמטריזציה) – נניח כי $U_1, U_2 \subset \mathbb{R}^2$ פתוחות ובעלות שטח, ושנתונות פרמטריזציות $\varphi_i: U_i \to \mathbb{R}^3$ (חד-חד ערכיות, C^1, $\operatorname{rank} D\varphi_i = 2$ עם הפיך רציף) עבור $i = 1,2$. נניח כי $S = \varphi_1(U_1) = \varphi_2(U_2)$.

נסמן:

$$J_i = (\det D\varphi_i^* D\varphi_i)^{\frac{1}{2}}$$

אזי מתקיים:

$$\int_{U_2} f\circ\varphi_2 \cdot J_2 dV = \int_{U_1} f\circ\varphi_1\cdot J_1 dV$$

כלומר, $\int_S f dA$ אינו תלוי בפרמטריזציה.

<u>הערה:</u> המשפט נכון גם כאשר $S = \varphi_1(\Omega_1) = \varphi_2(\Omega_2)$ כאשר $\Omega_i \subset U_i$ תחום קומפקטי בעל שטח. אותה ההוכחה תעבוד. נוח לנו לנסח את המשפט עבור קבוצות פתוחות, כי אז הניסוח שלו קצר יותר.

<u>הוכחה:</u>

בשלב ראשון, אפשר להוכיח כי קיימת $h: U_1 \to U_2$ כך ש-$h \in C^1$, חד-חד ערכית והפיכה כך שמתקיים $\varphi_2 \circ h = \varphi_1$. נדחה את הוכחת עובדה זו (ראה טענה 21.4 בהמשך), וכעת נשתמש בתוצאה כדי לסיים את הוכחת טענה 21.3.

נחשב באופן ישיר ונראה כי ערך האינטגרל לא משתנה תחת רה-פרמטריזציה:

$$\int_{\Omega_2} f\circ\varphi_2\cdot J_2 dV \overset{\text{החלפת משתנים}}{=} \int_{\Omega_1} f\circ \overbrace{\varphi_2\circ h}^{=\varphi_1}\cdot\overbrace{J_2\circ h\cdot|\det Dh|}^{(\star)}\, dv$$

נשים לב כי:

$$J_1(u,v) = \sqrt{\det\left(D\varphi_1^*(u,v)D\varphi_1(u,v)\right)}$$

ונציב בכלל השרשרת כדי לקבל:

$$D\varphi_1(u,v) = D(\varphi_2 \circ h)(u,v) = D\varphi_2\left(h(h,v)\right)Dh(u,v)$$

כלומר:

$$J_1(u,v) = \sqrt{\det\left[D\varphi_2\left(h(u,v)\right)^* D\varphi_2\left(h(u,v)\right)\right]\det[Dh^*(u,v)Dh(u,v)]} = J_2 \circ h(u,v) \cdot |\det Dh(u,v)|$$

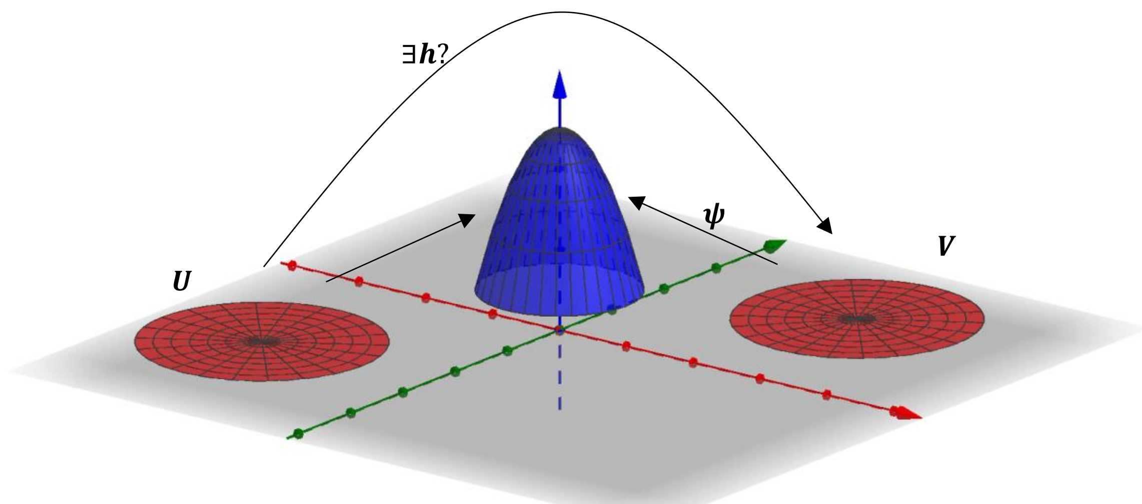
כלומר $J_1(u,v) = (\star)$ ואכן לאחר הצבה נקבל כי האינטגרלים שווים כנדרש. מ.ש.ל.

כדי להשלים את ההוכחה של טענה 21.3, מה שנותר להוכיח הוא את הטענה הבאה:

21.4 טענה – נניח ש - U,V קבוצות פתוחות במישור. תהיינה $\varphi: U \to \mathbb{R}^3, \psi: V \to \mathbb{R}^3$ אשר שתיהן C^1, חד-חד ערכיות וכן מתקיים $\operatorname{rank} D\varphi = \operatorname{rank} D\psi = 2$ בכל נקודה. כמו כן, נניח כי

$$S \coloneqq \varphi(U) = \psi(V)$$

ונניח ש- $\psi^{-1}: S \to V$ רציפה. אזי קיימת $C^1 \ni h: U \to V$ דיפאומורפיזם (כלומר פונקציה C^1 הפיכה אשר הפונקציה ההפוכה שלה היא גם כן C^1), כך שמתקיים: $\psi \circ h = \varphi$.



<u>הוכחה:</u>

$\psi: V \to S$ היא חד-יד ערכית ועל S , לכן קיימת פונקציה $\psi^{-1}: S \to V$. בשלב זה ψ^{-1} היא פשוט פונקציה והיא ההפיכה של ψ במובן של תורת הקבוצות ותו לא. נגדיר $h = \psi^{-1} \circ \varphi$ ונשים לב כי $\psi \circ h = \psi \circ \psi^{-1} \circ \varphi = \varphi$. כמו כן h הפיכה (כלומר חד-חד ערכית ועל) כהרכבה של כאלו. נרצה להראות כי $h \in C^1$ אך אנו נתקלים בבעיה: ψ^{-1} מוגדרת על התחום S ואין לנו מובן של גזירות על משטחים ($\varphi \in C^1$ מוגדרת $\varphi: U \to \mathbb{R}^2$ – לכן כאן אין בעיה).

אנחנו נראה כי בסביבת כל נקודה ב-U, h גזירה ברציפות. לשם כך, נשתמש בשיטות שלמדנו בפרק על יריעות חלקות משוכנות.

הרעיון – נמשיך את ψ^{-1} לפונקציה $\tilde{\psi}^{-1}$ שמוגדרת בסביבה של S:

$$C^1 \ni \psi: V \to \mathbb{R}^3 \quad \psi(u,v) = \begin{bmatrix} \psi_1(u,v) \\ \psi_2(u,v) \\ \psi_3(u,v) \end{bmatrix}$$

נשים לב כי מתקיים:

$$D\psi = \begin{bmatrix} \partial_u \psi_1 & \partial_v \psi_1 \\ \partial_u \psi_2 & \partial_v \psi_2 \\ \partial_u \psi_3 & \partial_v \psi_3 \end{bmatrix} \quad , \qquad \operatorname{rank} D\psi \overset{\text{נתון}}{=} 2$$

בלי הגבלת הכלליות נניח כי:

$$\det \begin{bmatrix} \partial_u \psi_1 & \partial_v \psi_1 \\ \partial_u \psi_2 & \partial_v \psi_2 \end{bmatrix} \neq 0$$

מכאן שנוכל להגדיר $\tilde{V} = \overset{\subset \mathbb{R}^2}{V} \times \mathbb{R} \subset \mathbb{R}^3$ ועבורה:

$$\tilde{\psi}: \tilde{V} \to \mathbb{R}^3 \quad C^1 \ni \tilde{\psi}(u,v,w) = \psi(u,v) + \begin{bmatrix} 0 \\ 0 \\ w \end{bmatrix}$$

ומתקיים:

$$D\tilde{\psi} = \begin{bmatrix} D\psi & \begin{matrix} 0 \\ 0 \\ 1 \end{matrix} \end{bmatrix} = \frac{\partial(\tilde{\psi}_1, \tilde{\psi}_2, \tilde{\psi}_3)}{\partial(u,v,w)} \Longrightarrow \boxed{\det D\tilde{\psi} \neq 0}$$

מתקיים:

$$\tilde{\psi}(u,v,0) = \psi(u,v)$$

והיות ו - $D\tilde{\psi}$ הפיכה, נסיק ממשפט הפונקציה ההפיכה כי $\tilde{\psi}$ הפיכה מקומית (עם הפוכה $\tilde{\psi}^{-1}$). לכל $(u_0, v_0) \in V$ יש סביבה של $(u_0, v_0, 0)$ שבה $\tilde{\psi}$ מוגדרת והפיכה ו-$\tilde{\psi}^{-1}$ מוגדרת ו-C^1 בסביבה W של הנקודה:

$$S \ni \psi(u_0, v_0) = \tilde{\psi}(u_0, v_0, 0)$$

לכן לכל נקודה $P = \varphi^{-1}\big(\psi(u_0, v_0)\big) \in \varphi^{-1}(W \cap S) \subseteq U$ מתקיים:

$$h(P) = \psi^{-1} \circ \varphi(P) = \psi^{-1} \circ \varphi\left(\varphi^{-1}\big(\psi(u_0, v_0)\big)\right) = (u_0, v_0)$$

מצד שני מתקיים

$$\tilde{\psi}^{-1} \circ \varphi(P) = \tilde{\psi}^{-1} \circ \varphi\left(\varphi^{-1}\big(\psi(u_0, v_0)\big)\right) = \tilde{\psi}^{-1}\left(\tilde{\psi}(u_0, v_0, 0)\right) = (u_0, v_0, 0)$$

ולכן, הוכחנו שבאופן מקומי מתקיים

$$h = \psi^{-1} \circ \varphi = \tilde{\psi}^{-1} \circ \varphi \in C^1$$

h היא הרכבה של גזירות ברציפות ולכן C^1 באופן מקומי (כלומר בסביבה של כל נקודה), ולכן היא C^1. מ.ש.ל.

<u>פרק 22</u>

<u>אינטגרציה וקטורית / תבניות דיפרנציאליות:</u>

<u>תזכורת</u> – יהא C עקום פשוט וחלק וכן:

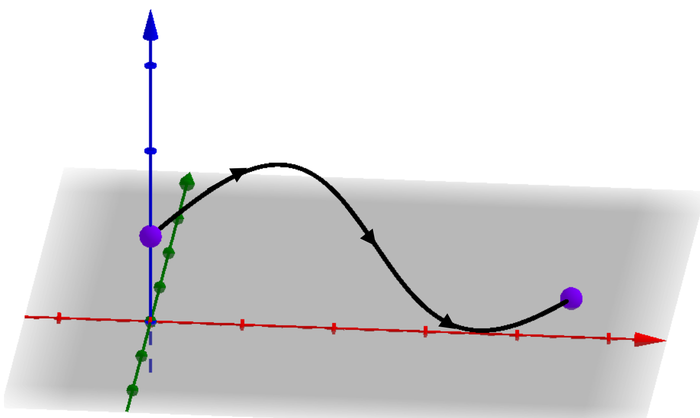
$$\gamma([a,b])=C$$

עבור:

1. $\gamma\in C^1:[a,b]\to\mathbb{R}^n$.
2. $\gamma'(t)\neq 0$ לכל t.
3. γ חד-חד ערכית למעט (אולי) בנקודות הקצה.

22.1 הגדרה – עקום C נקרא **עקום מכוון** אם "זוכרים" את כיוון התנועה. לעתים נהוג לסמן $\vec{C}$.

נגדיר:



$$-C=-\vec{C}=\left\{\begin{matrix}\tilde{\gamma}:[a,b]\to\mathbb{R}^n\\ \tilde{\gamma}(t)=\gamma(a+b-t)\end{matrix}\right\}$$

כלומר C – הוא העקום המכוון הנתון על-ידי אותה קבוצה אבל בכיוון השני.

22.2 תזכורת – בהנתן $F=(F_1,\cdots,F_n)$ שדה וקטורי, הגדרנו:

$$\int_C\sum_i F_i dx_i=\int_a^b F\big(\gamma(t)\big)\cdot\gamma'(t)dt$$

$$=\int_a^b\sum_{i=1}^n F_i\big(\gamma(t)\big)\overbrace{\gamma_i'(t)dt}^{dx_i}$$

22.3 הגדרה – לביטוי $w=\sum_i F_i dx_i$, $w(x)=\sum_i F_i(x)dx_i$ קוראים **1-תבנית דיפרנציאלית.**

<u>תרגיל</u> – הוכיחו כי מתקיים $\int_{-C}w=-\int_C w$.

כאמור יש לנו התאמה:

$$\underbrace{\sum_i F_i dx_i}_{\text{1−תבנית דיפרנציאלית}}\Leftrightarrow F=\underbrace{(F_1,\cdots,F_n)}_{\text{שדה וקטורי}}$$

22.4 הגדרה – תהא $f\in C^1$ בקבוצה פתוחה ב-$\mathbb{R}^n$. **הנגזרת החיצונית** של f זו ה-1 תבנית:

$$df=\sum_{i=1}^n\frac{\partial f}{\partial x_i}dx_i$$

כלומר התבנית שהאינטגרל עליה מוגדר באופן הבא לכל עקום מכוון עם פרמטריזציה γ:

$$\int_C df = \int_a^b \nabla f(\gamma(\mathrm{t}))\cdot \gamma'(t)dt$$

<u>הערה</u> – זו פשוט התבנית שמתאימה לשדה וקטורי הנתון ע"י הגרדיאנט של הפונקציה:

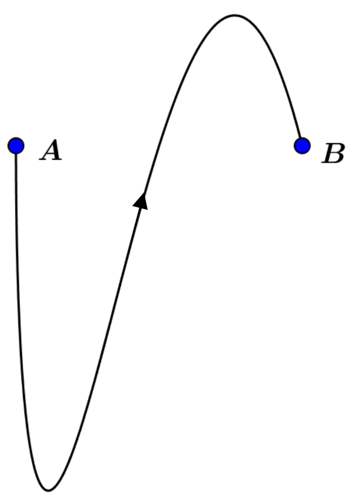
$$df = \sum_i \frac{\partial f}{\partial x_i}dx_i \leftrightarrow \mathrm{F} = \nabla f = \left(\frac{\partial f}{\partial x_1},\cdots,\frac{\partial f}{\partial x_n}\right)$$

22.5 טענה – יהא C עקום פשוט וחלק עם נקודת התחלה A ונקודת סיום B. נניח כי f גזירה ברציפות בסביבה של C. אם $w = df$ אזי מתקיים:

$$\int_C w = \int_C df = f(B) - f(A)$$

<u>בפרט</u> – עבור עקום C פשוט **וסגור**, יתקיים:

$$\int_C df = 0$$



<u>הערה:</u> שימו לב לדמיון לנוסחת ניוטון לייבניץ.

<u>הוכחה:</u>

תהא $\gamma: [a,b] \to \mathbb{R}^n$ פרמטריזציה כך ש-$\gamma(a) = A, \gamma(b) = B$. אזי:

$$\int_C w = \int_C df = \int_C \sum_i \frac{\partial f}{\partial x_i}dx_i = \int_a^b \nabla f(\gamma(t))\cdot \gamma'(t)dt$$

$$= \int_a^b \frac{d}{dt}(f\circ\gamma)(t)dt \overset{\substack{\text{המשפט}\\\text{היסודי}}}{=} f(\gamma(b)) - f(\gamma(a))$$

השתמשנו בכלל השרשרת כדי לעבור מהשורה הראשונה לשורה השנייה.
מ.ש.ל.

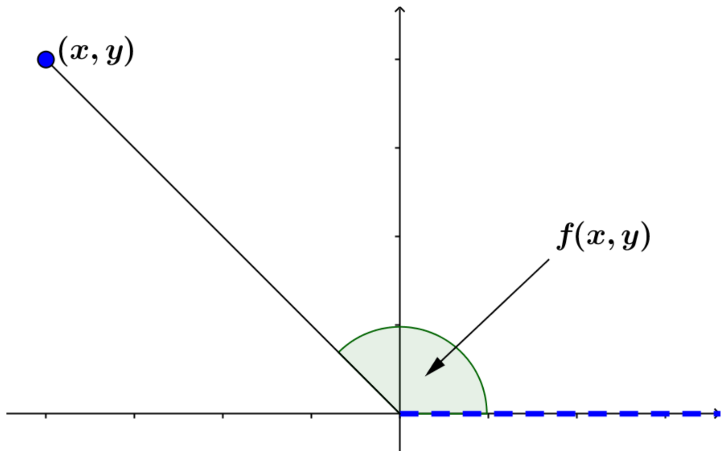
<u>דוגמה:</u>

נגדיר:

$$w = -\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy$$

מוגדר ו-C^1 ב-$\mathbb{R}^2 \setminus \{0\}$. ועתה נגדיר בנוסף:

$$f(x,y) = \arg(x+iy) \quad f: \mathbb{R}^2 \setminus \{(x,0)|x \geq 0\} \to \mathbb{R}$$



כאשר הפונקציה $f(x,y)$ היא הפונקציה שמחזירה את הזווית ש-(x,y) יוצר עם הכיוון החיובי של ציר ה-x, כמתואר בציור.

הסיבה לתחום ההגדרה ש"מסיר" את הישר $y=0$ עבור $x>0$, הוא שתפגע הרציפות של הפונקציה (ראו דוגמה בפרק מס' 7).

$$f(x,y)=\begin{cases} \arctan\left(\frac{y}{x}\right) & x>0, y>0 \\ \frac{\pi}{2}-\arctan\left(\frac{y}{x}\right) & y>0 \\ \pi+\arctan\left(\frac{y}{x}\right) & x<0 \\ \frac{3\pi}{2}-\arctan\left(\frac{y}{x}\right) & y<0 \end{cases}$$

לא קשה לבדוק שמתקיים:

$$df=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy=-\frac{y}{x^2+y^2}dx+\frac{x}{x^2+y^2}dy=w$$

ולכן מתקיים על פי הטענה, לכל עקום חלק המתחיל ב-A ומסתיים ב-B

$$\int_C w=f(B)-f(A)$$

בתנאי שהעקום כולו מוכל בתוך $\mathbb{R}^2\setminus\{(x,0)|x\geq 0\}$. בפרט, האינטגרל הקווי של התבנית w על כל עקום סגור המוכל ב- $\mathbb{R}^2\setminus\{(x,0)|x\geq 0\}$ חייב להתאפס.

<u>דוגמה:</u>

עבור $\gamma:[0,2\pi]\to\mathbb{R}^2$ המוגדרת על ידי $\gamma(t)=(\cos t,\sin t)$. נגדיר בנפרד:

$$\gamma_1=\gamma|_{[0,\pi]} \qquad \gamma_1^\varepsilon=\gamma|_{[\varepsilon,\pi]}$$
$$\gamma_2=\gamma|_{[\pi,2\pi]} \qquad \gamma_2^\varepsilon=\gamma|_{[\pi,2\pi-\varepsilon]}$$

נסמן $C=\gamma(0,2\pi)$, $C_1=\gamma(0,\pi)$, וכו'.

נקבל כי:

$$\int_{C_1^\varepsilon}w=\int_{C_1^\varepsilon}df=f\big((-1,0)\big)-f(\cos\varepsilon,\sin\varepsilon)=\pi-\varepsilon$$

ולכן:

$$\int_{C_1}w=\pi$$

ובאותו אופן עבור העקום C_2 נקבל π כערך האינטגרל. כלומר $\int_C w=\int_{C_1}w+\int_{C_2}w=2\pi$.

כמובן, קל אפילו עוד יותר לחשב ישירות לפי ההגדרה תוך שימוש בפרמטריזציה:

$$\int_C w = \int_0^{2\pi} \left(\frac{-\sin t}{\sin^2 t + \cos^2 t}, \frac{\cos t}{\sin^2 t + \cos^2 t}\right) \cdot (-\sin t, \cos t) dt = \int_0^{2\pi} 1\, dt = 2\pi$$

קיבלנו ערך שונה מאפס עבור אינטגרל קווי של התבנית w על עקום סגור. לכאורה יש כאן סתירה לטענה 22.5, אך למעשה אין שום סתירה, כיון שלא קיימת פונקציה f המוגדרת **בסביבה של כל העקום C** המקיימת $w = df$. העקום C אינו מוכל ב - $\mathbb{R}^2 \setminus \{(x,0) | x \geq 0\}$.

משפט גרין:

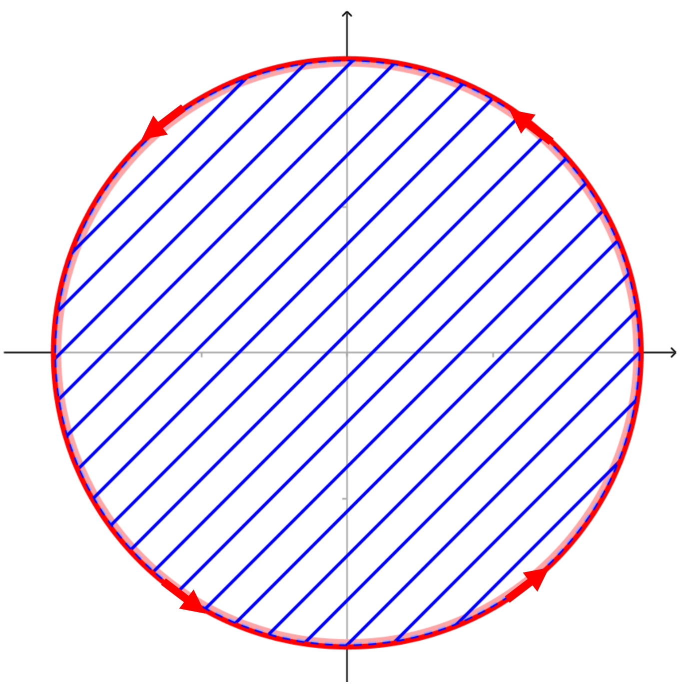
<u>רקע</u> – יהא $\Omega \subseteq \mathbb{R}^2$ תחום בעל שטח כך ש-$\partial\Omega$ הוא איחוד סופי של עקומים פשוטים, חלקים למקוטעין וסגורים.

22.6 הגדרה – הכיוון הטבעי על $\partial\Omega$ מוגדר להיות הכיוון כך שלאורך התקדמות על העקום $\partial\Omega$, Ω נמצאת "מצד שמאל" של העקום.

22.7 משפט גרין – יהא $\Omega \subseteq \mathbb{R}^2$ תחום בעל שטח כך ש-$\partial\Omega$ מורכבת ממספר סופי של עקומים פשוטים, חלקים למקוטעין וסגורים. תהיינה P, Q פונקציות C^1 בסביבת Ω. אזי:

$$\int_{\partial\Omega} Pdx + Qdy = \int_\Omega \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dV$$

כאשר על $\partial\Omega$ בוחרים את ה"כיוון הטבעי".



<u>דוגמה:</u>

נניח כי:

$$\Omega = \{(x,y) | x^2 + y^2 \leq 1\}$$

ונגדיר על $\partial\Omega$:

$$P(x,y) = 0 \quad Q(x,y) = x$$

על פי משפט גרין, יתקיים:

$$\int_{\partial\Omega} xdy = \int_{\partial\Omega} Pdx + Qdy = \int_\Omega 1dV = \text{שטח העיגול}$$

כלומר, בהתאם למשפט גרין אנו מצפים כי האינטגרל השמאלי יהיה שווה לשטח העיגול. ואכן, במקרה הזה נוכל לבדוק את השוויון באופן ישיר:

$$\int_{\partial\Omega} xdy = \int_0^{2\pi} (\cos t) \cos t\, dt$$
$$= \int_0^{2\pi} \cos^2 t\, dt = \pi$$

<u>הוכחת המשפט:</u>

משפט גרין לתחומים די כלליים נובע מרדוקציה לתחומים פשוטים יותר (על-ידי חלוקה לתתי תחומים), וכמקרה מיוחד של משפט סטוקס אותו נוכיח בהמשך. נציג את ההוכחה במקרים פרטיים.

<u>מקרה א'</u> – כחימום, נוכיח את המשפט עבור תחום מלבני $\Omega = [a,b] \times [c,d]$:

$$\int_\Omega \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dV = \int_c^d \left(\int_a^b \frac{\partial Q}{\partial x}\bigg|_{(x,y)} dx\right) dy - \int_a^b \left(\int_c^d \frac{\partial P}{\partial y}\bigg|_{(x,y)} dy\right) dx$$

$$\overset{\text{המשפט היסודי}}{=} \int_c^d \big(Q(b,y) - Q(a,y)\big)dy - \int_a^b \big(P(x,d) - P(x,c)\big)dx = \int_{\partial\Omega} Pdx + Qdy$$

<u>מקרה ב'</u> – קל להוכיח את המשפט עבור תחום שנתון באופן הבא:

$$\Omega = \{(x,y): a \le x \le b, f_1(x) \le y \le f_2(x)\} = \{(x,y): c \le y \le d, g_1(y) \le x \le g_2(y)\}$$

שימו לב שכל שניתן להציג כך קבוצות רבות, בפרט עיגול. הרעיון דומה מאוד לשלב א'. מלינאריות של האינטגרל (בעצם יש כאן שני סוגי אינטגרלים, ושניהם לינאריים בארגומנט שלהם) די להוכיח

$$\int_{\partial\Omega} Pdx = -\int_\Omega \frac{\partial P}{\partial y} dV$$

$$\int_{\partial\Omega} Qdy = \int_\Omega \frac{\partial Q}{\partial x} dV$$

נסתפק בהוכחת השוויון הראשון. לפי פוביני:

$$\int_\Omega \frac{\partial P}{\partial y} dV = \int_a^b \left(\int_{f_1(x)}^{f_2(x)} P_y(x,y)dy\right) dx = \int_a^b \big(P(x,f_2(x)) - P(x,f_1(x))\big)dx = -\int_{\partial\Omega} Pdx$$

ועלינו להצדיק את השוויון האחרון. כאן ישנם כמה מקרים, וכדאי להיעזר באיור כדי להמחיש מה קורה (ראו האיור בעמוד הקודם). אנו נטפל במקרה $f_1(a) = f_2(a)$ ו- $f_1(b) = f_2(b)$. המקרה בו אחד הגבולות (הימני או השמאלי) של Ω הינו קו ישר מטופל באופן דומה. כמו כן, נניח ש - f_1, f_2 הן גזירות ברציפות – אם הן גזירות ברציפות למקוטעין, צריך לפרק את האינטגרלים בחישוב הבא לסכום של מספר אינטגרלים.

נתבונן באיור בעמוד הקודם. אם נסמן ב - C_1 את החלק של $\partial\Omega$ שנמצא מתחת לציר ה איקס, וב - C_2 את החלק של $\partial\Omega$ שנמצא מעל (עם הכיוון הסטנדרטי), אז מתקיים $\int_{\partial\Omega} Pdx = \int_{C_1} Pdx + \int_{C_2} Pdx$. פרמטריזציה של C_1 נתונה על-ידי $a \le t \le b$, $\gamma(t) = \big(t, f_1(t)\big)$, ולכן, לפי ההגדרה של אינטגרל קווי וקטורי,

$$\int_{C_1} Pdx = \int_a^b \big(P(\gamma(t)), 0\big) \cdot \gamma'(t)dt = \int_a^b P\big(t, f_1(t)\big)\, 1dt = \int_a^b P\big(x, f_2(x)\big)dx$$

פרמטריזציה של $-C_2$ (שימו לב: זהו העקום אך בכיוון ההפוך!) נתונה על-ידי $a \le t \le b$, $\gamma(t) = \big(t, f_2(t)\big)$, ולכן

$$-\int_{C_2} Pdx = \int_{-C_2} Pdx = \int_a^b \big(P(\gamma(t)), 0\big) \cdot \gamma'(t)dt = \int_a^b P\big(x, f_2(x)\big)dx$$

כנדרש.

<u>מקרה ג'</u> – ניתן להראות כי בהינתן נכונות משפט גרין עבור Ω, אזי הוא נכון גם עבור $g(\Omega)$ כאשר g מקיימת את תנאי משפט החלפת המשתנים. בעצם אנו נקבל את שלב ג' כמקרה מיוחד של הוכחת משפט סטוקס שתינתן בפרק הבא.

מ.ש.ל, בערך, לבינתיים. (המשך יבוא...)

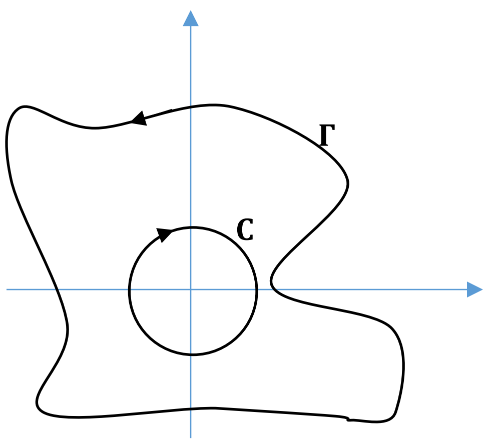
<u>מספר הליפוף של עקום:</u>

בפרק הקודם התבוננו בתבנית הדיפרנציאלית:

$$\omega = -\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy$$

ונניח עתה כי נתון לנו עקום כללי Γ, ועקום פנימי C מהצורה:

$$C = \{(x,y)|x^2+y^2=1\}$$



נסמן ב-Ω את התחום התחום על ידי Γ ו-C. כלומר מתקיים:

$$\partial\Omega = \Gamma \cup C$$

עם הכיוון הטבעי נקבל כי:

$$\partial\vec{\Omega} = \vec{\Gamma} - \vec{C}$$

ולכן נסמן:

$$Q = \frac{x}{x^2+y^2} \quad P = -\frac{y}{x^2+y^2}$$

ונשים לב כי:

$$P_y = \frac{y^2-x^2}{(x^2+y^2)^2} = Q_x$$

ולפי משפט גרין נקבל כי:

$$\int_\Gamma \omega + \int_{-C} \omega = \int_{\partial\Omega} Pdx + Qdy \overset{\text{משפט גרין}}{=} \int_\Omega (Q_x - P_y)dV = \int_\Omega 0dV = 0$$

וכיון ש- $\int_C \omega = 2\pi$ (חישבנו בפרק הקודם) נקבל ש- $\int_\Gamma \omega = 2\pi$.

22.8 הגדרה – אם Γ עקום פשוט וחלק ב-$\mathbb{R}^2 \setminus \{0\}$, מגדירים את **מספר הליפוף** (או **אינדקס**, ובאנגלית **winding number** או **index**) להיות:

$$ind(\Gamma) = \frac{1}{2\pi}\int_\Gamma \omega$$

באופן אינטואיטיבי, מספר זה סופר את מספר הפעמים שהעקום מסתובב סביב הראשית (נגד כיוון השעון).

פרק 23

אוריינטציה על משטחים:

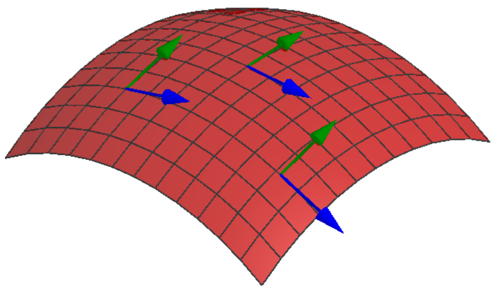
23.1 הגדרה – נאמר כי $S \subseteq \mathbb{R}^3$ הינו **משטח חלק**, אם ניתן להציג כ-S $\varphi(\Omega) = S$ כאשר:

$$\varphi: \left(\substack{\text{סביבה של}\\ \Omega} \subset \mathbb{R}^2\right) \to \mathbb{R}^3 \quad \varphi \in C^1 \quad \varphi \text{ חד חד ערכית} \quad \operatorname{rank} D\varphi \equiv 2$$

23.2 הגדרה – **אוריינטציה** ל-S הינה פונקציה רציפה:

$$S \ni p \mapsto \{v_1(p), v_2(p)\}$$

כך ש-$v_1(p), v_2(p)$ הינם זוג סדור של וקטורים משיקים ל-S בנקודה p שהינם בלתי תלוים. כלומר זהו בסיס סדור ל-$T_p(S)$.



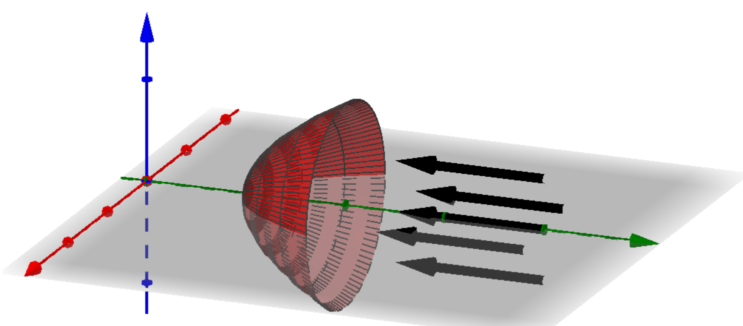
נשים לב, כי $v_1(p) \times v_2(p)$ הינו וקטור הניצב למרחב המשיק (הנורמל למשטח). לכן, בהנתן ש-$v_1(p), v_2(p)$ הם פונקציות רציפות, אזי שבהכרח $n(p) = \frac{v_1(p)\times v_2(p)}{\|v_1(p)\times v_2(p)\|}$ הוא פונקציה רציפה כהרכבה של פונקציות רציפות.

הגדרה (שקולה) – אוריינטציה ל-S זו פונקציה רציפה $S \ni p \mapsto n(p)$ כאשר $n(p)$ הינו וקטור יחידה נורמלי. משטח נקרא **אוריינטבילי** אם ניתן להגדיר עליו אוריינטציה.

לא לכל משטח הוא אוריינטבילי, למשל **לטבעת מביוס** איננה (רצוי להדביק כאן תמונה, לעת עתה אני מפנה את הסטודנטים ל – google images).

בהינתן יריעה דו-ממדית עם שפה, אוריינטציה מגדירה כיוון על השפה לפי כלל יד ימין.

שטף:



23.3 הגדרה – נניח כי $\Omega \subseteq \mathbb{R}^2$ תחום בעל שטח וכן $g \in C^1$ מוגדרת בסביבת Ω לתוך $\mathbb{R}^3$, חד-חד ערכית כך שמתקיים $\operatorname{rank} Dg \equiv 2$. נסמן:

$$S = g(\Omega)$$

אזי, אם על Ω יש אוריינטציה $\vec{\Omega}$, נגדיר על S אוריינטציה מושרית $\vec{S}$. בנקודה $g(p)$ נתאים את הזוג $\{dg(p)v_1, dg(p)v_2\}$ או הנורמל:

$$n(p) = \frac{(dg(p)v_1) \times (dg(p)v_2)}{\|dg(p)v_1 \times dg(p)v_2\|}$$

<u>הערה:</u> כאשר משטח כולו נתון על-ידי פרמטריזציה (תמונה של תחום יפה במישור) אזי יש לו אוריינטציה, שמוגדרת כמו למעלה. בפרט, משטח כזה הוא תמיד אוריינטבילי.

23.4 הגדרה – אם S מוגדר בהתאם ל-24.3, ונניח כי מוגדר $F = (F_1, F_2, F_3)$ שדה וקטורי רציף בסביבת S, אזי נגדיר את **אינטגרל השטף** של F על $\vec{S}$ להיות:

$$\int_{\vec{S}} F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2 := \int_{\vec{\Omega}} F \circ g(u,b) \cdot (g_u \times g_v) du \wedge dv$$

הסימונים $\vec{\Omega}$ ו- $\vec{S}$ מציינים את המשטח עם האוריינטציה שלו (התחום Ω מגיע עם האוריינטציה הסטנדרטית). אפשר פשוט לפרש את האינטגרל לעיל באופן הבא, כאשר על הפרמטריזציה כבר משרה אוריינטציה על המשטח:

$$= \iint_{\Omega} F \circ g(u,v) \cdot (g_u \times g_v) du dv = \iint_{\Omega} F \circ g(u,v) \cdot n\big(g(u,v)\big) dA$$

<u>הערה:</u> אין צורך לדאוג לגבי הסימנים $dx_2 \wedge dx_3$ וכולי המופיעים לעיל. אפשר להבין את הסימונים הללו פשוט לפי ההגדרה:

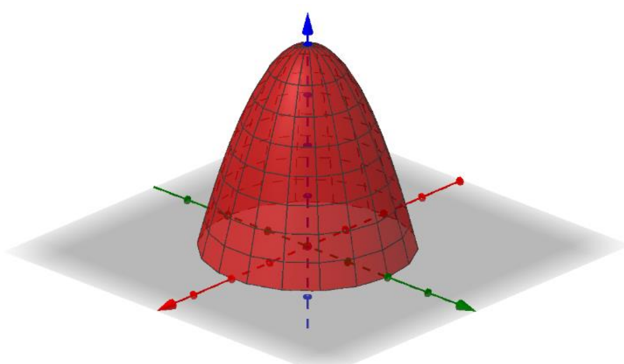
$$\int_{\vec{S}} F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2 \coloneqq \iint_{\Omega} F \circ g(u,v) \cdot (g_u \times g_v) du dv$$

אפשר גם לחשוב עליהם כעל "אלמנטי שטח מכוונים".

<u>דוגמה:</u>

נניח כי נתון המשטח:

$$S = \left\{ (x,y,z) \middle| \begin{matrix} z \geq 0 \\ z = 4 - x^2 - y^2 \end{matrix} \right\}$$



אוריינטציה – נשים לב כי לנורמל רכיב z שונה מאפס בכל נקודה במשטח. נוכל להגדיר אוריינטציה על-ידי כך שנחליט שבכל נקודה נבחר את כיוון הנורמל שיש לו רכיב z חיובי. במילים, נבחר נורמל באופן רציף כך שנכריז שאנו בוחרים את הכיוון "למעלה" (או "החוצה").

נחשב את השטף של השדה:

$$F(x,y,z) = (xz, yz, x^2 + y^2)$$

שמסומן על ידי הביטוי הבא:

$$\int_{\vec{S}} xzdy \wedge dz + yzdz \wedge dx + (x^2 + y^2) dx \wedge dy$$

על מנת לחשב את הנ"ל, נמצא פרמטריזציה של המשטח S:

$$\Omega = \{(u,v) \in \mathbb{R}^2 | u^2 + v^2 \leq 4\}$$

$$\varphi(u,v)=(u,v,4-u^2-v^2)$$

נשים לב כי:

$$D\varphi=\begin{bmatrix}1 & 0\\ 0 & 1\\ -2u & -2v\end{bmatrix}\Longrightarrow \varphi_u=\begin{bmatrix}1\\0\\-2u\end{bmatrix}\quad \varphi_v=\begin{bmatrix}0\\1\\-2v\end{bmatrix}$$

ולכן:

$$\varphi_u\times\varphi_v=\begin{bmatrix}\begin{vmatrix}0 & 1\\ -2u & 2v\end{vmatrix}\\ \begin{vmatrix}-2u & -2v\\ 1 & 0\end{vmatrix}\\ \begin{vmatrix}1 & 0\\ 0 & 1\end{vmatrix}\end{bmatrix}=\cdots=\begin{bmatrix}2u\\2v\\1\end{bmatrix}$$

אם על Ω האוריינטציה הרגילה, נקבל כי האוריינטציה המושרית על ידי φ היא האוריינטציה שהגדרנו (כי רכיב z הינו חיובי).

נוכל כעת פשוט להציב בנוסחה המופיעה בהגדרת אינטגרל השטף, ולקבל:

$$\int_{\vec{\Omega}}F\big(\varphi(u,v)\big)\cdot(\varphi_u\times\varphi_v)du\wedge dv$$

$$=\int_{\Omega}[u(4-u^2-v^2),v(4-u^2-v^2),u^2+v^2]\cdot[2u,2v,1]dudv$$

$$=\int_{\Omega}(u^2+v^2)[2(4-u^2-v^2)+1]dudv=\int_0^{2\pi}\int_0^2(9r^2-2r^4)rdrd\theta=\cdots$$

וכעת עשינו רדוקציה לאינטגרל חד ממדי, הסטודנטים מוזמנים להשלים את החישוב.

נספח – אוריינטציה ביריעות כלליות ואינטגרציה מסומנת

אוריינטציה:

טיפול שלם של נושא האינטגרציה על יריעות (ומשטחים בפרט) דורש מגדרה של מושג **האוריינטציה** ו**האינטגרציה המסומנת**. אנו נסקור נושאים אלו. **חומר זה איננו חלק מחומר החובה של הקורס**, אך עשוי לעזור בהבנה בפרק על משפט סטוקס.

כזכור, הגדרנו נפח של מקבילון להיות נתון על ידי:

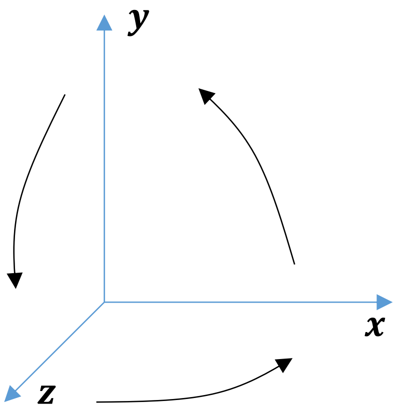
$$\mathrm{Vol}\big(\mathrm{P}(\mathrm{v}_1,\cdots,\mathrm{v}_\mathrm{n})\big)=|\det(\mathrm{v}_1|\mathrm{v}_2|\cdots|\mathrm{v}_\mathrm{n})|$$

הגדרה – יהיו $\{\mathrm{v}_1,\mathrm{v}_2,\cdots,\mathrm{v}_\mathrm{n}\}\in\mathbb{R}^\mathrm{n}$ בסיס סדור. אזי נאמר כי $\{\mathrm{v}_1,\cdots,\mathrm{v}_\mathrm{n}\}$ בעל **אוריינטציה חיובית** אם $\det(\mathrm{v}_1|\cdots|\mathrm{v}_\mathrm{n})>0$.

כלומר, אם הגודל בעל אותו סימן כמו הגודל $\det(e_1|\cdots|e_n)$ עבור הבסיס הסטנדרטי.

לדוגמה:

ב-$\mathbb{R}^2$ הבסיס הסטנדרטי נתון על ידי $\{e_1,e_2\}$. במקרה זה נקבל כי $\{e_2,e_1\}$ בעל אוריינטציה שלילית ואילו בסיס מהצורה $\{e_1+e_2,e_2\}$ הוא בעל אוריינטציה חיובית.



<u>דוגמה נוספת:</u>

במקרה של $\mathbb{R}^3$, למשל בפיזיקה, ידוע הכלל שקרוי "כלל יד ימין" שקובע את האוריינטציה של בניית מערכת צירים, גם היא בהתאם להגדרת האוריינטציה שהגדרנו זה עתה.

(אנו משמיטים ניסיון מגושם להסביר את כלל יד ימין בכתב, ומקווים כי הנושא מודגם בפרק)

עתה, נניח $\Omega \subset \mathbb{R}^n$ תחום קשיר בעל נפח.

הגדרה – אוריינטציה של Ω היא פונקציה רציפה:

$$\Omega \ni x \mapsto \{v_1(x), v_2(x), \cdots, v_n(x)\}$$

כאשר $\{v_1(x), \cdots, v_n(x)\}$ בסיס סדור של המרחב. תחום בעל אוריינטציה יסומן $\vec{\Omega}$.

טענה – $\det\left(v_1(x)\middle|\cdots\middle|v_n(x)\right)$ בעלת סימן קבוע.

<u>הוכחה:</u> זה ברור, כי מדובר בפונקציה ממשית על תחום קשיר שאינה מתאפסת.

<u>דוגמה:</u>

האוריינטציה ה"טבעית" (או "הרגילה") מתקבלת כאשר נגדיר בכל נקודה $\{v_1(x), \cdots, v_n(x)\} = \{e_1, \cdots, e_n\}$.

הגדרה – תחום עם אוריינטציה נקרא **בעל אוריינטציה חיובית** אם $\{v_1(x), \cdots, v_n(x)\}$ בעל אוריינטציה חיובית לכל x.

הגדרה – אם $\vec{\Omega}$ תחום בעל אוריינטציה אזי נגדיר:

$$\operatorname{sgn}\vec{\Omega} = \begin{cases} 1 & \text{אוריינטציה חיובית} \\ -1 & \text{אחרת} \end{cases}$$

<u>אינטגרציה מסומנת (העשרה):</u>

הגדרה – אם $\vec{\Omega}$ תחום בעל אוריינטציה נפח עם אוריינטציה ו-$f: \Omega \to \mathbb{R}$, נגדיר:

$$\int_{\vec{\Omega}} f dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n = \operatorname{sgn}(\vec{\Omega}) \int_{\Omega} f dV$$

<u>משפט החלפת המשתנים המסומן (העשרה):</u>

אם $g \in C^1$ וכן חד-חד ערכית והפיכה בסביבת Ω, אזי בהנתן אוריינטציה על Ω, g משרה אוריינטציה על $g(\Omega)$, פשוט הפעלת הנגזרת Dg על הבסיס המגדיר את האוריינטציה בכל נקודה.

אם $\vec{\Omega}$ תחום בעל נפח עם אוריינטציה, $g(\vec{\Omega})$ תחום בעל אוריינטציה מושרית, ואם g מקיימת את תנאי משפט החלפת המשתנים, אזי:

$$\int_{g(\vec{\Omega})} f(y)dy_1 \wedge \cdots \wedge dy_n = \int_{\vec{\Omega}} f\big(g(x)\big) \det\big(Dg(x)\big)\, dx_1 \wedge \cdots \wedge dx_n$$

<u>אוריינטציה מושרית על השפה (העשרה):</u>

אם $\vec{\Omega} \subseteq \mathbb{R}^2$ תחום בעל שטח ואוריינטציה, אזי נגדיר על $\partial\Omega$ אוריינטציה מושרית $\partial\vec{\Omega}$ – בוחרים משיק כך ש-$\{n(p), T(p)\}$ בעל אותה אוריינטציה כמו $\{v_1(p), v_2(p)\}$.

במקרה זה ניתן לנסח את משפט גרעין באופן הבא:

$$\int_{\partial\vec{\Omega}} \left(Pdx + Qp_y\right) = \int_{\vec{\Omega}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy$$

<u>מומלץ:</u> ציירו תחום במישור והסבירו לעצמכם למה ה"כיוון הטבעי" על שפה של תחום במישור הינה האוריינטציה המושרית, כאשר על התחום לוקחים את האוריינטציה הקבועה $\{e_1, e_2\}$.

פרק 24

משפט סטוקס:

24.1 הגדרה – תהא $F=(F_1,F_2,F_3)$ שדה וקטורי ב-$\mathbb{R}^3$. **הרוטור** של F מוגדר להיות:

$$\operatorname{rot} F=\operatorname{curl} F=\nabla\times F:=\left(\frac{\partial F_3}{\partial x_2}-\frac{\partial F_2}{\partial x_3},\frac{\partial F_1}{\partial x_3}-\frac{\partial F_3}{\partial x_1},\frac{\partial F_2}{\partial x_1}-\frac{\partial F_1}{\partial x_2}\right)$$

ניתן לקבל את הרוטור גם על ידי ביצוע המכפלה הוקטורית (באופן פורמלי):

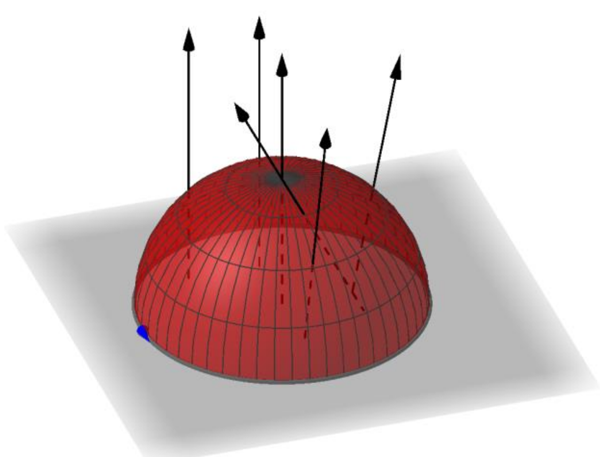
$$\nabla\times F=\begin{vmatrix}\hat{\imath} & \hat{\jmath} & \hat{k}\\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3}\\ F_1 & F_2 & F_3\end{vmatrix}$$

24.2 משפט סטוקס (Stokes) – יהא S משטח דו פרמטרי ב-$\mathbb{R}^3$ הנתון על ידי פרמטריזציה מהצורה $\varphi:\Omega\to\mathbb{R}^3$ כאשר Ω תחום בעל שטח עם שפה $\partial\Omega$ שהיא עקום פשוט סגור וחלק למקוטעין. כמו כן נניח כי φ גזירה ברציפות, חד-חד ערכית ורגולרית בסביבת Ω.

יהא F שדה וקטורי גזיר ברציפות בסביבת S. אזי:

אינטגרל השטף		האינטגרל הקווי
של $\operatorname{rot} F$	$=$	מסוג שני של
על S		F על ∂S

כלומר:



$$\int_{\partial\vec{S}}F_1dx_1+F_2dx_2+F_3dx_3=$$

$$\int_{\vec{S}}\left(\frac{\partial F_3}{\partial x_2}-\frac{\partial F_2}{\partial x_3}\right)dx_2\wedge dx_3$$

$$+\left(\frac{\partial F_1}{\partial x_3}-\frac{\partial F_3}{\partial x_1}\right)dx_3\wedge dx_1$$

$$+\left(\frac{\partial F_2}{\partial x_1}-\frac{\partial F_1}{\partial x_2}\right)dx_1\wedge dx_2$$

או במילים אחרות:

$$\int_{\partial S}F\cdot T ds=\int_S(\nabla\times F)\cdot n dA$$

כאשר על S ו- ∂S האורינטציה והכיוון המושרה על ידי g מ- Ω ו- $\partial\Omega$ (אפשר להסתפק בהבנת המקרה בו על Ω האוריינטציה הרגילה, על $\partial\Omega$ הכיוון הטבעי), ו-

dA – אלמנט שטח פנים n – נורמל יחידה ds – אלמנט אורך קשת T – וקטור משיק יחידה

הערה – כזכור, בעבר הגדרנו: הנגזרת החיצונית של הפונקציה f היא ה 1-תבנית הנתונה על ידי $df=\sum\frac{\partial f}{\partial x_i}dx_i$.
באופן דומה, נגדיר "נגזרת חיצונית" של 1-תבנית $F_1dx_1+F_2dx_2+F_3dx_3$ להיות ה "2-תבנית" הבאה:

$$d(F_1dx_1+F_2dx_2+F_3dx_3) = \left(\frac{\partial F_3}{\partial x_2}-\frac{\partial F_2}{\partial x_3}\right)dx_2\wedge dx_3+\left(\frac{\partial F_1}{\partial x_3}-\frac{\partial F_3}{\partial x_1}\right)dx_3\wedge dx_1+\left(\frac{\partial F_2}{\partial x_1}-\frac{\partial F_1}{\partial x_2}\right)dx_1\wedge dx_2$$

זו, בעצם נקראת 2-תבנית ואפשר לרשום את משפט סטוקס בצורה הבאה:

$$\int_{\partial S}\omega=\int_S d\omega$$

לכל 1-תבנית ω. שימו לב לדמיון עם נוסחת ניוטון לייבניץ, עם טענה 22.5, ועם משפט גרין. כל המשפטים הללו הם מקרים מיוחדים של משפט כללי יותר הנקרא גם כן "משפט סטוקס", הדן באינטגרציה של תבניות על יריעות "אוריינטביליות" עם שפה.

באופן כללי, בדרך זו ניתן להכליל תהליך זה על מנת לקבל מקרה כללי עבור S יריעה עם שדה ממימד $k+1$ כך ש-ω היא k-תבנית דיפרנציאלית.

הוכחה:

לצורך ההוכחה נניח ש - $\varphi\in C^2$.

אנו נניח שהשפה נתונה ע"י מסילה חלקה (אחרת מפרקים את השפה לקטעים חלקים, וצריך להסתכל על סכומים של אינטגרלים). נניח $\gamma:[a,b]\to\mathbb{R}^2$ פרמטריזציה (עם כיוון) של $\partial\vec{\Omega}$. אזי $\varphi\circ\gamma:[a,b]\to\mathbb{R}^3$ פרמטריזציה של $\overrightarrow{\partial S}$.

אזי:

$$\int_{\partial\vec{S}}\sum_i F_i dx_i=\int_a^b F\left(\varphi\big(\gamma(t)\big)\right)\cdot[\varphi\circ\gamma]'(t)dt\ (\star)$$

נסמן את הקואורדינטות ב- $\mathbb{R}^2$ כ- (u,v) , ונסמן $\gamma(t)=(u(t),v(t))$ כך שיתקיים:

$$[\varphi\circ\gamma]'(t)=D\varphi\big(\gamma(t)\big)\gamma'(t)\quad \gamma'(t)=\begin{bmatrix}u'(t)\\v'(t)\end{bmatrix}$$

$$D\varphi=[\varphi_u\quad\varphi_v]\ \text{where}\ \varphi_u=\begin{bmatrix}\frac{\partial\varphi_1}{\partial u}\\\frac{\partial\varphi_2}{\partial u}\\\frac{\partial\varphi_3}{\partial u}\end{bmatrix}\quad\varphi_v=\begin{bmatrix}\frac{\partial\varphi_1}{\partial v}\\\frac{\partial\varphi_2}{\partial v}\\\frac{\partial\varphi_3}{\partial v}\end{bmatrix}$$

$$(\star)\int_a^b F\left(\varphi\big(\gamma(t)\big)\right)\cdot\varphi_u u'(t)dt+F\left(\varphi\big(\gamma(t)\big)\right)\cdot\varphi_v v'(t)dt$$

$$=\int_{\partial\vec{\Omega}}[(F\circ\varphi)\cdot\varphi_u]\,du+[(F\circ\varphi)\cdot\varphi_v]dv$$

$$\overset{\text{גרין}}{=} \int_{\Omega}\left(\frac{\partial}{\partial u}\big((F\circ\varphi)\cdot\varphi_v\big)-\frac{\partial}{\partial v}\big((F\circ\varphi)\cdot\varphi_u\big)\right)dudv$$

עתה נשים לב כי:

$$\frac{\partial}{\partial u}\big((F\circ\varphi)\cdot\varphi_v\big)-\frac{\partial}{\partial v}\big((F\circ\varphi)\cdot\varphi_u\big)=$$

$$\frac{\partial(F\circ\varphi)}{\partial u}\cdot\varphi_v+F\circ\varphi\cdot\varphi_{vu}-\frac{\partial(F\circ\varphi)}{\partial v}\cdot\varphi_u-F\circ\varphi\cdot\varphi_{uv}$$

ועתה היות והנחנו C^2, נקבל כי $\varphi_{uv}=\varphi_{vu}$. לכן שני איברים מתבטלים וניוותר עם:

$$\int_{\Omega}\left(\frac{\partial}{\partial u}\big((F\circ\varphi)\cdot\varphi_v\big)-\frac{\partial}{\partial v}\big((F\circ\varphi)\cdot\varphi_u\big)\right)dudv=\int_{\Omega}\left[\frac{\partial(F\circ\varphi)}{\partial u}\cdot\varphi_v-\frac{\partial(F\circ\varphi)}{\partial v}\cdot\varphi_u\right]dudv$$

ועלינו להראות עתה כי האינטגרנד הוא שווה ל- $(\nabla\times F)\cdot(\varphi_u\times\varphi_v)$. נשים לב כי האינטגרנד שווה ל:

$$\big(DF(\varphi(u,v))\varphi_u\big)\cdot\varphi_v-\big(DF(\varphi(u,v))\varphi_v\big)\cdot\varphi_u$$

$$=\sum_{i=1}^{3}\left(\sum_{i=1}^{3}\frac{\partial F_i}{\partial x_j}\frac{\partial\varphi_j}{\partial u}\right)\frac{\partial\varphi_i}{\partial v}-\sum_{i=1}^{3}\left(\sum_{j=1}^{3}\frac{\partial F_i}{\partial x_j}\frac{\partial\varphi_j}{\partial v}\right)\frac{\partial\varphi_i}{\partial u}$$

ניתן כעת לאסוף את כל האיברים הכוללים את $\frac{\partial F_i}{\partial x_j}$ לכל $i,j=1,2,3$, ומקבלים שזה שווה בדיוק ל-

$(\nabla\times F)\cdot(\varphi_u\times\varphi_v)$, וזה מסיים את הוכחת משפט סטוקס. מ.ש.ל.

<u>תרגיל</u> – לאסוף איברים ולסדרם כך שניתן יהיה להגיע לביטוי המבוקש.

ניתן להוכיח את משפט סטוקס גם עבור משטחים הנתונים כתמונה של פונקציה $\varphi:\Omega\to\mathbb{R}^3$ שהיא גזירה ברציפות. בהוכחה שניתנה למעלה הנחנו ש $\varphi\in C^2$, ובאמת השתמשנו בזה בהוכחה. שימו לב שבכך הוכחנו גם את משפט גרין עבור תחום שהוא תמונה גזירה פעמיים ברציפות של תחום המקיים את משפט גרין, בפרט (כיוון שהוכחנו את משפט גרין לתחום פשוט), הוכחנו שמשפט גרין מתקיים עבור כל תחום שהוא תמונה גזירה פעמיים ברציפות של עיגול היחידה הסגור. בנספח בסוף הפרק אנו נדגים איך ניתן לטפל בתחומים ששפתם נתונה ע"י עקום גזיר ברציפות.

<u>משמעות הרוטור</u>

לרוטור של שדה וקטורי יש משמעות גיאומטרית/פיזיקלית מעניינת, שגם מסבירה את את השם "רוטור" (או curl באנגלית). נניח ש - F שדה וקטורי גזיר ברציפות ב - $\mathbb{R}^3$, תהי $p\in\mathbb{R}^3$ נקודה כלשהי במרחב, ויהי $\hat{n}$ וקטור יחידה כלשהו. אזי <u>**ההיטל של הרוטור בנקודה $\nabla\times F(p)$ בנקודה p בכיוון $\hat{n}$ שווה ל"סיבוביות ליחידת שטח של השדה F סביב דיסק אינפיניטסימלי הניצב ל- $\hat{n}$ "**</u>.

אכן (וההסבר להלן אמור גם להסביר את המשמעות של המשפט המודגש בפסקה הקודמת), אם $\overrightarrow{D_r}$ הוא דיסק (מכוון) – כלומר עיגול (מכוון) דו מימדי – שמרכזו בנקודה p , רדיוסו r, ואשר הניצב שלו בכיוון $\hat{n}$, אז לפי משפט סטוקס נקבל

$$\int_{\vec{D}}(\nabla\times F)\cdot\hat{n}dA=\int_{\overrightarrow{\partial D}}F\cdot Tds$$

כיון ש - $\nabla \times F$ רציפה, אינטגרל השטף מקיים משפט ערך ממוצע אינטגרלי, כלומר

$$\int_{\vec{D}} (\nabla \times F) \cdot \hat{n} dA = A(\vec{D}) \cdot (\nabla \times F)(q) \cdot \hat{n}$$

כאשר $q \in \vec{D}$ ו - $A(\overrightarrow{D_r}) = \pi r^2$ זה השטח של הדיסק $\overrightarrow{D_r}$. לכן נקבל

$$(\nabla \times F)(q) \cdot \hat{n} = \frac{1}{\pi r^2} \int_{\overrightarrow{\partial D}} F \cdot T ds$$

ואם נשאיף את רדיוס הדיסק לאפס הנקודה q תשאף ל- p ונקבל

$$(\nabla \times F)(p) \cdot \hat{n} = \lim_{r \to 0} \frac{1}{\pi r^2} \int_{\overrightarrow{\partial D_r}} F \cdot T ds = \lim_{r \to 0} \frac{1}{\pi r^2} \int_{\overrightarrow{\partial D_r}} F_1 dx_1 + F_2 dx_2 + F_3 dx_3$$

נספח (הערה): מה עושים עם תחומים עם פרמטריזציה שהיא גזירה ברציפות רק פעם אחת

נסמן ב- D את כדור היחידה הסגור במישור $\mathbb{R}^2 = \mathbb{C}$. המשפט הבא הוא משפט חשוב מתורת הפונקציות המרוכבות, ואנו מביאים אותו ללא הוכחה.

24.3 משפט (משפט ההעתקה של רימן) – תהי U קבוצה פתוחה ופשוטת קשר במישור. נניח גם ש - U איננה ריקה וגם איננה שווה למישור כולו. אזי קיימת פונקציה אנליטית, חח"ע ועל $\varphi: int(D) \to U$. אם ∂U הוא עקום גזיר ברציפות פשוט וסגור, אזי φ וכן כל הנגזרות הראשונות שלה ניתנות להמשכה לפונקציה רציפה חח"ע ועל $\varphi: D \to \overline{U}$.

24.4 הערה – לא חשוב כרגע להבין מהי פונקציה אנליטית (מי שלא יודע), מה שחשוב לדעת זה שפונקציה אנליטית היא גזירה אינסוף פעמים.

אנו נוכיח את הטענה הבאה:

24.5 טענה – יהי Ω תחום בעל שטח במישור, הנתון כסגור של קבוצה פתוחה ופשוטת קשר עם שפה גזירה ברציפות. אזי לכל P, Q גזירות ברציפות בסביבת Ω מתקיים:

$$(*) \qquad \int_{\partial\Omega} Pdx + Qdy = \int_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dV$$

<u>הוכחה:</u>

נציג כאן את עיקרי הרעיון. הסטודנט צריך להיות מסוגל להפוך את זה להוכחה מפורטת (ההוכחה משתמשת במשפט ההעתקה של רימן – משפט 25.3. כמובן, איננו מצפים מהסטודנט שיוכיח את משפט ההעתקה של רימן, אלא שיוכל להשלים את ההוכחה הבאה שמשתמשת במשפט רימן כקופסא שחורה).

לפי משפט 25.3, Ω בעצם נתון ע"י פרמטריזציה גזירה ברציפות $\varphi: D \to \Omega$, שמקיימת $\varphi|_{int(D)} \in C^\infty$, בפרט, הצמצום של φ לעיגול הפתוח גזיר פעמיים ברציפות. אפשר לבדוק ש- $\gamma(t) = \varphi(\cos t, \sin t)$

(עבור $t \in [0,2\pi]$), מגדיר פרמטריזציה גזירה ברציפות של $\partial\Omega$.

יהי $\epsilon > 0$. אנו נראה שהשוויון ב - $(*)$ מתקיים עד כדי $C\epsilon$ עבור איזשהו קבוע C, וכיון ש - ϵ יכול להיות קטן כרצוננו, נקבל שוויון. נזכור ש $- \varphi$ יחד עם כל הנגזרות החלקיות שלה מסדר ראשון רציפות על D, ולכן רציפות במ"ש שם. יהי $r \in (0,1)$, שהוא מאוד קרוב ל -1 (עד כמה קרוב, נקבע בהמשך).

נגדיר $\psi(z)=\varphi(rz)$ לכל $z\in D$. כיון ש- φ וגם הנגזרות החלקיות שלה מסדר ראשון רציפות על D, ניתן לבחור את r קרוב מספיק ל- 1 כך ש $\|\mathrm{D}\varphi(z)-\mathrm{D}\psi(z)\|$ וכן $\|\varphi(z)-\psi(z)\|$ יהיו קטנים כרצוננו במידה שווה לכל $z\in D$ (למשל לגבי הנגזרות, נשים לב ש- $\frac{\partial\psi_1}{\partial x}(z)=r\frac{\partial\varphi_1}{\partial x}(rz)$, ו-

$$\left|\frac{\partial\varphi_1}{\partial x}(z)-r\frac{\partial\varphi_1}{\partial x}(rz)\right|\le\left|\frac{\partial\varphi_1}{\partial x}(z)-r\frac{\partial\varphi_1}{\partial x}(z)\right|+r\left|\frac{\partial\varphi_1}{\partial x}(z)-\frac{\partial\varphi_1}{\partial x}(rz)\right|\le$$
$$\le M(1-r)+r\left|\frac{\partial\varphi_1}{\partial x}(z)-\frac{\partial\varphi_1}{\partial x}(rz)\right|$$

(כאשר $M=\max\left|\frac{\partial\varphi}{\partial x}\right|$) והביטוי האחרון קטן כרצוננו אם r מספיק קטן (הודות לרציפות במ"ש). באופן דומה מראים גם עבור שאר הנגזרות החלקיות וכן עבור φ).

נסמן $G=\psi(D)$. כמו כן נסמן $g(t)=\psi(\cos t,\sin t)$. כיון ש- φ הינה גזירה ברציפות פעמיים (למעשה אינסוף פעמים) בפנים של עיגול היחידה, נובע שהפונקציה ψ גזירה פעמיים ברציפות בסביבת D. לכן, משפט גרין מתקיים עבור G. ליתר דיוק, מההוכחה שלנו את משפט סטוקס (שעבדה תחת ההנחה של גזירות פעמיים ברציפות), מתקיימת המסקנה של משפט גרין עבור G, כלומר

$$\int_{\partial G}Pdx+Qdy=\int_G\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dV$$

לכן כדי לקבל את $(*)$ עד כדי ϵ , נותר להראות ש-

$$(**)\qquad\left|\int_{\partial G}Pdx+Qdy-\int_{\partial\Omega}Pdx+Qdy\right|<\epsilon/2$$

וש-

$$(***)\qquad\left|\int_G\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dV-\int_\Omega\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dV\right|<\epsilon/2$$

אם r קרוב מספיק ל- 1.

תהי h פונקציה רציפה כלשהי בסביבה של Ω. אזי h מוגדרת ורציפה גם בסביבה של G, כי G מוכלת ב – Ω. מתקיים

$$\int_G\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dV-\int_\Omega\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dV=\int_D\big((h\circ\psi)|\det D\psi|-(h\circ\varphi)|\det D\varphi|\big)dxdy$$

כאשר $r\nearrow 1$ הביטוי שתחת האינטגרל בצד ימין שואף במידה שווה ל-0, לכן האינטגרל שואף ל-0, ולכן קיים r_0 כך שלכל $r\in(r_0,1)$ מתקיים $(***)$. באופן מאוד דומה אפשר להראות שעבור r מספיק קרוב ל-1, מתקיים $(**)$. את הפרטים נשאיר לקוראים החרוצים.

פרק 25

משפט הדיברגנס (משפט גאוס):

משפט הדיברגנס, הידוע גם בתור משפט גאוס, הוא משפט נוסף בשרשרת המשפטים שנלמד המקשר בין האינטגרל של איזושהי נגזרת בתחום כלשהו לבין האינטגרל על השפה.

25.1 הגדרה – יהי $F \in C^1(U, \mathbb{R}^n)$ שדה וקטורי. ה**דיברגנס** של F הוא הפונקציה הסקלרית הנתונה ע"י

$$div(F) = \nabla \cdot F = \sum_{i=1}^{n} \frac{\partial F_i}{\partial x_i}$$

25.2 משפט הדיברגנס (משפט גאוס) – יהי $\Omega \subseteq \mathbb{R}^3$ תחום בעל נפח סגור המקיים ש- $\partial\Omega$ הינו משטח גזיר ברציפות למקוטעין. אזי לכל שדה וקטורי F הגזיר ברציפות בסביבה של Ω, מתקיים

$$(*) \qquad \int_{\Omega} \nabla \cdot F \, dV = \int_{\partial\Omega} F \cdot n \, dA$$

כאשר n הינו נורמל היחידה המצביע **החוצה** מהתחום.

הוכחה:

אנו נוכיח את המשפט רק עבור תחומים שהם פשוטים לפי כל המשתנים, כלומר ניתן לתאר את Ω באופן הבא:

$$\Omega = \{(x, y, z) \colon (x, y) \in G, g(x, y) \le z \le h(x, y)\}$$

כאשר G תחום בעל שטח במישור xy ו-$g, h \in C^1$. נניח גם של-Ω יש תיאורים כאלו גם בהחלפת התפקידים של המשתנים. לפני שנמשיך נעיר שכל תחום בעל נפח שהוא קמור ובעל שפה חלקה למקוטעין (כמו קובייה, כדור או אליפסה) ניתן לתאור כזה, וכן כמעט כל תחום שיצוץ בדוגמאות (וכל תחום שיצוץ בדוגמאות יהיה ניתן לפירוק לאיחוד סופי של כאלו תחומים).

כעת, כיון שהשוויון $(*)$ הוא לינארי ב- F, אפשר להוכיח בנפרד עבור כל אחד מהשדות הוקטוריים $(F_1, 0, 0)$, $(0, F_2, 0)$ ו- $(0, 0, F_3)$ לכן נניח שדה F יש את הצורה $F(x, y, z) = (0, 0, P(x, y, z))$.

נחשב

$$(**) \qquad \int_{\Omega} \nabla \cdot F dV \; = \; \int_{\Omega} \frac{\partial P}{\partial z} dV = \int_{G} \int_{g(x,y)}^{h(x,y)} \frac{\partial P}{\partial z}(x, y, z) dz \; dxdy =$$

$$= \int_{G} \left[P\big(x, y, h(x, y)\big) - P\big(x, y, g(x, y)\big)\right] dxdy$$

אנו טוענים שהביטוי האחרון שווה לשטף של השדה $F = (0, 0, P)$ דרך $\partial\Omega$. ראשית נציין, שאם יש "קירות" (חלקים של המשטח שנורמל להם ניצב לציר z), כמו למשל הצדדים המאונכים של קובייה, אז השטף דרכם הוא אפס, כי השדה ניצב לנורמל.

לכן מה שתורם לשטף הכללי זה השטף דרך המשטח התחתון $S_1 = \left\{\big(x, y, g(x, y)\big) \colon (x, y) \in G\right\}$ והשטף דרך המשטח העליון $S_2 = \left\{\big(x, y, h(x, y)\big) \colon (x, y) \in G\right\}$.

אנו נראה ש- $\int_{S_1} F \cdot ndA = -\int_{G} P\big(x, y, g(x, y)\big) dxdy$ כאשר לוקחים את הנורמל למטה.

באופן דומה מראים ש- $\int_{S_2} F\cdot ndA=\int_G P\big(x,y,h(x,y)\big)dxdy$. יחד עם (**), נקבל את (*).

נגדיר פרמטריזציה של S_1 על-ידי הנוסחה $\varphi(x,y)=\big(x,y,g(x,y)\big)$. כדי לחשב את השטף, צריך לחשב את האינטגרל הדו-ממדי:

$$\int_{S_1} F\cdot ndA=\pm\int_G F\circ\varphi\cdot\left(\varphi_x\times\varphi_y\right)dxdy$$

עד כדי סימן, כי עוד לא ברור לנו עם הפרמטריזציה הזו נותנת את הנורמל בכיוון הנכון. נחשב

$$\varphi_x\times\varphi_y=\begin{bmatrix}1\\0\\g_x\end{bmatrix}\times\begin{bmatrix}0\\1\\g_y\end{bmatrix}=\begin{bmatrix}-g_x\\-g_y\\1\end{bmatrix}$$

וזהו נורמל שיש לו רכיב כלפי מעלה, ולכן הוא מצביע פנימה. לכן

$$\int_{S_1} F\cdot n_{out}dA=-\int_G F\circ\varphi\cdot\left(\varphi_x\times\varphi_y\right)dxdy=-\int_G P\big(x,y,g(x,y)\big)dxdy$$

כפי שרצינו להראות. מ.ש.ל.

25.3 דוגמה – נשתמש במשפט הדיברגנס כדי לראות מהו הקשר בין נפח הכדור לשטח הפנים שלו. יהי B_R כדור ברדיוס R סביב הראשית. נבחר בשדה $F=\frac{1}{3}(x,y,z)$. נשים לב ש- $div(F)=\nabla\cdot F=1$. כמו כן, וקטור נורמל לשפת הכדור המצביע החוצה נתון על ידי $n=\frac{1}{R}(x,y,z)$ עבור $(x,y,z)\in\partial B_R$. לכן נקבל

$$Vol(B_R)=\int_{B_R}1dV=\int_{B_R}div(F)dV$$

לפי משפט הדיברגנס זה שווה ל-

$$\int_{\partial B_R}F\cdot ndA=\frac{R}{3}\int_{\partial B_R}dA=\frac{R}{3}A(\partial B_R)$$

כלומר קיבלנו $Vol(B_R)=\frac{R}{3}A(\partial B_R)$. זה אכן מסתדר עם הנוסחאות המוכרות

$Vol(B_R)=\frac{4\pi R^3}{3}$ ו- $A(\partial B_R)=4\pi R^2$.

<u>הערה:</u> ניתן להוכיח את משפט הדיברגנס על-ידי חלוקה של תחום לתחומים פשוטים יותר. לדוגמא: הטורוס.

<u>דוגמא:</u> נחשב את השטף של השדה $F=(x+z^2,\cos xz\,,z+x^3)$ דרך הטורוס

$$T_{a,b}=\left\{(x,y,z):\left(a^2-\sqrt{x^2+z^2}\right)^2+y^2=b^2\right\}$$

<u>הערה:</u> ניתן לנסח את משפט סטוקס בשפה של תבניות דיפרנציאליות והנגזרת החיצונית. עבור 2-תבנית דיפרנציאלית

$$\omega=F_1dx_2\wedge dx_3+F_2dx_3\wedge dx_1+F_3dx_1\wedge dx_2$$

נגדיר

$$d\omega=\left(\frac{\partial F_1}{\partial x_1}+\frac{\partial F_2}{\partial x_2}+\frac{\partial F_3}{\partial x_3}\right)dx_1\wedge dx_2\wedge dx_3=div(F)dx_1dx_2dx_3$$

כלומר, אם ω היא ה 2-תבנית המתאימה לשדה הוקטורי $F = (F_1, F_2, F_3)$ (כלומר, האינטגרל של ω על משטח שווה לאינטגרל השטף של F על המשטח), אזי $d\omega$ זה ה 3-תבנית המתאימה לדיברגנס של F. בשפה זו, משפט הדיברגנס מקבל את הצורה:

$$\int_{\Omega} d\omega = \int_{\partial\Omega} \omega$$

זוהי הצורה של משפט סטוקס הכללי, הנכונה בממדים גבוהים, ומכילה את משפט הדיברגנס, משפט סטוקס, משפט גרין, נוסחת ניוטון-לייבניץ, ועוד. השוו עם ההערה שאחרי משפט 24.2.

<u>משמעות הדיברגנס</u>

כמו הגרדיאנט והרוטור, הדיברגנס הוא אופרטור דיפרנציאלי בעל משמעות פיזיקלית/גיאומטרית פשוטה ומעניינת.

נניח ש - F שדה וקטורי גזיר ברציפות ב - $\mathbb{R}^3$ וש - p נקודה כלשהי. נסמן ב - $B_r(p)$ כדור ברדיוס r סביב p. אז לפי משפט הדיברגנס יחד עם משפט ערך הביניים האינטגרלי

$$\nabla \cdot F(q) = \frac{1}{Vol\big(B_r(p)\big)} \int_{B_r(p)} \nabla \cdot F dV = \frac{1}{Vol\big(B_r(p)\big)} \int_{\partial B_r(p)} F \cdot n dA$$

עבור נקודה $q \in B_r(p)$. לכן

$$\nabla \cdot F(p) = \lim_{r \to 0} \frac{1}{Vol(B_r(p))} \int_{\partial B_r(p)} F \cdot n dA$$

לכן הדיברגנס שווה לכמות השטף שיוצאת מיחידת נפח אינפיטסימלית.

<u>דוגמאות להרצאה, תרגול או תרגיל בית:</u>

השטף בדשה כבידה/חשמלי, חוק גאוס באלקטרוסטטיקה (או בכבידה), הסקת משוואת פואסון.

פרק 26

שדות משמרים

26.1 הגדרה – תהי $U \subseteq \mathbb{R}^n$ קבוצה פתוחה, ויהי $F: U \to \mathbb{R}^n$ שדה וקטורי רציף. F נקרא **שדה משמר** אם לכל מסילה פשוטה וחלקה למקוטעין $\gamma: [a,b] \to \mathbb{R}^n$, מתקיים שהאינטגרל $\int_{\gamma([a,b])} F \cdot dr$ תלוי רק בנק' ההתחלה $\gamma(a)$ ובנק' הסיום $\gamma(b)$. לפעמים נאמר ש - F **שדה משמר ב -** U, כדי להדגיש את העובדה שהיותו של שדה משמר או לא תלויה בקבוצה בה הוא מוגדר.

26.2 טענה – שדה F הינו משמר אם ורק אם $\int_C F \cdot dr = 0$ לכל עקום פשוט, חלק למקוטעין וסגור.

<u>הוכחה</u>: זה די ברור – לכן נשמיט הפרטים (בסמסטר חורף תש"פ זה הוכח בתרגול 12 סעיף 2).

26.3 משפט – תהי $U \subseteq \mathbb{R}^n$ קבוצה פתוחה, ויהי $F: U \to \mathbb{R}^n$ שדה וקטורי רציף. F הינו שדה משמר אם ורק אם קיימת פונקציה $f: U \to \mathbb{R}$ גזירה ברציפות כך ש - $F = \nabla f$ ב- U (פונקציה סקלרית f כזו נקראת **פונקציית פוטנציאל** עבור השדה F).

<u>הוכחה</u>: נניח ש- $F = \nabla f$. אזי לכל מסילה פשוטה וחלקה למקוטעין $\gamma: [a,b] \to \mathbb{R}^n$, מתקיים (לפי כלל השרשרת)

$$\int_{\gamma([a,b])} F \cdot dr = \int_a^b \nabla f(\gamma(t)) \cdot \gamma'(t)dt = \int_a^b \frac{d}{dt}\big(f(\gamma(t)\big)\, dt = f(\gamma(b)) - f(\gamma(a))$$

וזה תלוי רק בנק' ההתחלה והסוף של העקום (הקורא הערני ישים לב שאת הכיוון הזה כבר ראינו בטענה 22.5). שימו לב שכרגיל, אנחנו מחליפים מדי פעם את הסימונים, ועל הקורא להיות לרגיל לכך שלאותם אובייקטים יש סימונים שונים. כאן

$$\int_{\gamma([a,b])} F \cdot dr = \int_{\gamma([a,b])} F \cdot T ds = \int_{\gamma([a,b])} F_1 dx_1 + F_2 dx_2 + F_3 dx_3$$

(אני יודע שזה מעצבן, אבל ככה זה בחיים ואני רוצה שתרגישו בנוח עם כל הסימונים).

בכיוון השני, נניח ש- F הינו שדה משמר. נוכל להניח ש-U הינה קבוצה קשירה (אם לא, אז אפשר לעבוד בכל רכיב קשירות בנפרד). נבחר נקודה כלשהי $p_o \in U$, ונגדיר לכל $p \in U$

$$f(p) = \int_{p_0}^{p} F \cdot dr$$

כאשר האינטגרל מסמן אינטגרל קווי מסוג שני על עקום פשוט וגזיר ברציפות למקוטעין **כלשהו** שמחבר בין הנקודות p_0 ו- p. לפני שנראה ש- f זו היא פונקצית הפוטנציאל אותה אנו מחפשים, נסביר מדוע היא מוגדרת היטב.

ראשית, כיון ש-U פתוחה וקשירה, היא גם קשירה מסילתית וגם קשירה פוליגונלית (עובדה זו היא עובדה ידועה בטופולוגיה של המרחב $\mathbb{R}^n$. בסמסטר חורף תש"פ זה ניתן בתרגול 1, סעיף 9. ראו הערה "לתרגול/תרגיל בית" בסוף פרק מס' 1). בפרט ניתן למצוא עקום פשוט גזיר ברציפות למקוטעין המחבר בין p_0 ו- p.

בנוסף, כיון ש- F הינו שדה משמר, הערך שמתקבל מהאינטגרציה אינו תלוי במסלול הספציפי שבוחרים, מתקבל ערך יחיד ומוגדר היטב.

כעת נחשב:

$$\frac{\partial f}{\partial x_i}(p) = \lim_{h\to 0}\frac{f(p+he_i) - f(p)}{h} = \lim_{h\to 0}\frac{1}{h}\left(\int_{C_1+C_2} F \cdot dr - \int_{C_1} F \cdot dr\right) =$$

$$= \lim_{h\to 0}\frac{1}{h}\int_{C_2} F\cdot dr = \lim_{h\to 0}\frac{1}{h}\int_0^h F_i(p+te_i)dt = F_i(p)$$

כאשר C_1 הוא מסלול כלשהו בין p_0 ל- p, ו- C_2 המסילה $t \mapsto p + t\ e\_i$ בין p ו- $p + he_i$. קיבלנו $\nabla f = F$ כמבוקש, וזה סוף ההוכחה.

26.4 מסקנה – אם $U \subseteq \mathbb{R}^n$ פתוחה ו- $F \in C^1(U, \mathbb{R}^n)$ שדה משמר אזי בכל נקודה $p \in U$ ולכל i, j מתקיים:

$$(*) \quad \frac{\partial F_i}{\partial x_j}(p) = \frac{\partial F_j}{\partial x_i}(p)$$

בפרט, כאשר $n = 3$, אם F משמר אזי $\nabla \times F = 0$ בכל נקודה ב- U.

<u>הוכחה</u>: זה נובע מהמשפט הקודם שנותן $F = \nabla f$ יחד עם העובדה שעבור פונקציה גזירה פעמיים ברציפות מתקיים $\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}$. בשלושה ממדים אפשר להשתמש בזהות $\nabla \times \nabla f = 0$. הסבר נוסף נובע ממשמעות הרוטור שהוכחנו בפרק 24:

$$(\nabla \times F)(p)\cdot \hat{n} = \lim_{r\to 0}\frac{1}{\pi r^2}\int_{\overrightarrow{\partial D_r}} F\cdot T ds = \lim_{r\to 0}\frac{1}{\pi r^2}\int_{\overrightarrow{\partial D_r}} F_1 dx_1 + F_2 dx_2 + F_3 dx_3$$

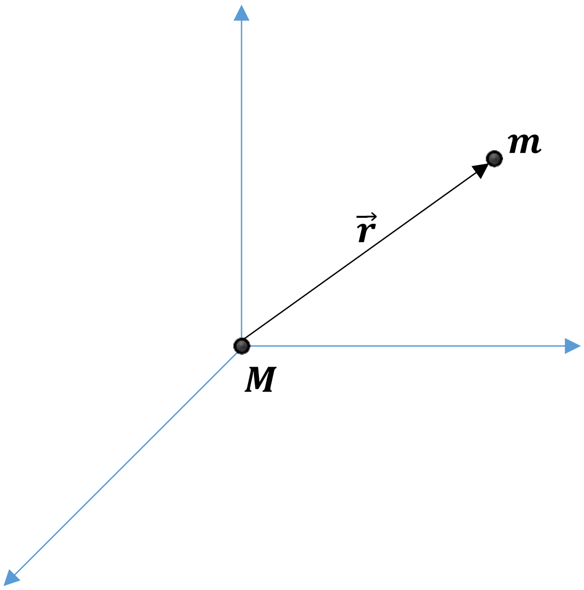
כאשר $\overrightarrow{D_r}$ דיסק מכוון שמרכזו p ורדיוסו r הניצב לוקטור יחידה $\hat{n}$ כלשהו. כיון שהשדה משמר, האינטגרל הקווי מסוג שני שלו על השפה של כל דיסק כזה היא אפס, ולכן ההטלה של $(\nabla \times F)(p)$ על כל וקטור יחידה $\hat{n}$ היא אפס. מכאן נובע ש - $(\nabla \times F)(p) = 0$. אכן: אפשר לקחת $\hat{n}$ בכיוון של $(\nabla \times F)(p)$, ואז נקבל

$$\|(\nabla \times F)(p)\| = (\nabla \times F)(p)\cdot \hat{n} = 0$$

<u>דוגמה:</u> ניזכר בדוגמה מסוף פרק 14:

יהי

$$F: \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}^3$$



שדה הכח הגרביטציוני שגוף בעל מסה m , והנמצא במיקום המתואר על ידי $\vec{r}$, מרגיש כתוצאה מנוכחות של גוף בעל מסה M במרחב.

מתקיים:

$$\vec{F} = G\frac{Mm}{\|\vec{r}\|^3}(-\vec{r})$$

בהנחה ש-M, m קבועים נוכל לכתוב:

$$F(x,y,z) = -K\frac{(x,y,z)}{\|(x,y,z)\|^3}$$

לשם פשטות נניח $K = 1$. חישבנו העבודה של הכוח הזה לאורך מסלול הקפיצי (כלומר את האינטגרל הקווי מסוג שני) וקיבלנו

$$\int_{kfitz} F \cdot d\vec{\ell} = \int_0^L F\big(f(t)\big) \cdot f'(t)dt = \cdots = \frac{1}{\sqrt{a^2 + b^2L^2}} - \frac{1}{a} = \frac{1}{\|r_1\|} - \frac{1}{\|r_0\|}$$

כאשר r_0, r_1 הן נקודות ההתחלה והסיום של המסלול. אפשר לבדוק שהינו מקבלים את אותה התוצאה אם היינו מחשבים את העבודה לאורך הקו הישר המחבר בין נקודות ההתחלה והסוף. אבל **אין צורך לבדוק**, כיון שאפשר לנחש כעת שפונקציית הפוטנציאל היא:

$$f(x,y,z) = \frac{1}{\sqrt{x^2 + y^2 + z^2}} = \frac{1}{r}$$

אכן: חישוב ישיר מראה שמתקיים

$$\frac{\partial f}{\partial x} = \frac{-x}{r^3}$$

וכולי, ולכן $\nabla f = F$, כנדרש. לפי משפט 26.3 השדה שלנו הוא משמר.

<u>הערה:</u> אותה נוסחה רק עם קבועים פיזיקליים שונים מגדירה את שדה הכח החשמלי שמושרה על-ידי מטען חשמלי נקודתי הנמצא בראשית.

הגדרה 26.5 – נניח ש- $U \subseteq \mathbb{R}^n$ פתוחה ו- $F \in C^1(U, \mathbb{R}^n)$ שדה וקטורי. נאמר ש - F הינו שדה **משמר מקומית** אם לכל נקודה $p \in U$ קיימת סביבה של p שם השדה הינו משמר.

להלן עובדות נוספות אותן נסביר בהרצאה אם יתיר הזמן:

טענה 26.6 – נניח ש- $U \subseteq \mathbb{R}^n$ פתוחה ו- $F \in C^1(U, \mathbb{R}^n)$ שדה וקטורי. אזי F הינו שדה משמר מקומית אם ורק אם מתקיימת המשוואה (*).

26.7 משפט – אם U תחום פשוט קשר אזי כל שדה משמר מקומית הוא גם משמר. בפרט, בשלושה ממדים, אם $\nabla \times F = 0$ בכל נקודה בתחום פשוט קשר U אזי F שדה משמר.

דוגמה: ניזכר ב 1-תבנית הדיפרנציאלית

$$\omega = -\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy$$

מפרק 22. תבנית זו מתאימה לשדה הוקטורי הגזיר ברציפות $F(x,y) = \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$ המוגדר בקבוצה הפתוחה $U = \mathbb{R}^2 \setminus \{0\}$. חישובים שעשינו בפרק 22 מראים ששדה זה איננו משמר, כי האינטגרל הקווי על עקום סגור סביב הראשית נותן את "מספר הליפוף" של העקום. עם זאת, זהו שדה משמר מקומית, כיון שסביב כל נקודה במישור המנוקב אפשר למצוא כדור קטן ופונקציית פוטנציאל f כך ש - $\nabla f = F$ המוגדרת היטב וחלקה בכדור קטן זה. כמו כן, אפשר לבדוק בידיים שמתקיים:

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x} = \frac{y^2 - x^2}{(x^2+y^2)^2}$$

ולכן זהו שדה משמר מקומית. דרך שלישית לראות זאת, היא לשים לב שמספר הליפוף של עקום הנמצא כולו בתוך כדור קטן סביב נקודה ב – U הוא תמיד אפס, כל עוד הכדור לא מכיל את הראשית.

נספח לפרק 26 - הוכחת משפטים 26.6 ו- 26.7

הגדרה: נניח ש - $U \subseteq \mathbb{R}^n$ קבוצה פתוחה וקשירה, וש - $\gamma_i : [0,1] \to U$ $(i = 0,1)$ שתי מסילות גזירות ברציפות המקיימות $\gamma_1(0) = \gamma_0(0)$ ו-$\gamma_1(1) = \gamma_0(1)$. נאמר שהמסילות **הומוטופיות** אחת לשנייה אם קיימת פונקציה גזירה ברציפות $H : [0,1]^2 \to U$ כך שמתקיים

1. $H(0,t) = \gamma_0(t)$ לכל $t \in [0,1]$.
2. $H(1,t) = \gamma_1(t)$ לכל $t \in [0,1]$.
3. $H(s,0) = \gamma_0(0) = \gamma_1(0)$ ו- $H(s,1) = \gamma_0(1) = \gamma_1(1)$ לכל $s \in [0,1]$.

לפעמים נדגיש ונאמר ש - **γ_0 ו- γ_1 הומוטופיות ב – U.**

דוגמא וציור!

הערה: בטופולוגיה נהוג לדבר על מסילות **רציפות** ועל הומוטופיות **רציפות.** למעשה, בין שתי מסילות גזירות ברציפות, הנמצאות בקבוצה פתוחה, קיימת הומוטופיה רציפה אם ורק אם קיימת הומוטופיה גזירה ברציפות, וזה קורה אם ורק אם קיימת הומוטופיה גזירה אינסוף פעמים ברציפות – ניתן להוכיח את זה באמצעות משפט הקירוב של ויירשטראס בשני משתנים (אותו לא מוכיחים באינפי – ראו את הקורס "מבוא לאנליזה פונקציונלית").

טענה: אם γ_0 ו- γ_1 הומוטופיות ב - U, ו- F שדה גזיר ברציפות ב - U המקיים $\frac{\partial F_j}{\partial x_i} = \frac{\partial F_i}{\partial x_j}$, אזי

$$\int_{\gamma_0} F \cdot dr = \int_{\gamma_1} F \cdot dr$$

הוכחה: בשביל ההוכחה אנו נניח ש – $H \in C^2([0,1]^2)$. לפי ההערה לעיל, זה מוצדק.

נגדיר $\gamma_s(t) = H(s,t)$. זוהי מסילה גזירה ברציפות לכל $s \in [0,1]$. כעת נגדיר

$$g(s) = \int_{\gamma_s} F \cdot dr = \int_0^1 F\big(\gamma_s(t)\big) \cdot \gamma_s'(t)dt = \int_0^1 F\big(H(s,t)\big) \cdot H_t(s,t)dt$$

מתקיים $g(0)=\int_{\gamma_0} F\cdot dr$ ו - $g(1)=\int_{\gamma_1} F\cdot dr$. אנחנו נוכיח שמתקיים $\frac{dg}{ds}\equiv 0$, ומכאן תנבע התוצאה הנדרשת $g(0)=\int_{\gamma_0} F\cdot dr=g(1)=\int_{\gamma_1} F\cdot dr$. לכן נחשב באופן פורמלי (ניתן להצדיק זאת עם משפטים מאינפי 2) על-ידי גזירה תחת סימן האינטגרל:

$$\frac{dg}{ds}=\frac{d}{ds}\int_0^1 F\big(H(s,t)\big)\cdot H_t(s,t)dt=\int_0^1\Big(DF\big(H(s,t)\big)H_s\cdot H_t+F\big(H(s,t)\big)\cdot H_{ts}\Big)dt \boxed{\equiv}$$

כעת הגיעה העת להשתמש בעובדה ש - F מקיים את התנאי $\frac{\partial F_j}{\partial x_i}=\frac{\partial F_i}{\partial x_j}$. במילים אחרות, המטריצה DF היא סימטרית! לכן

$$DF\big(H(s,t)\big)H_s\cdot H_t=H_t^T DF\big(H(s,t)\big)H_s=H_s^T DF\big(H(s,t)\big)^T H_t=H_s^T DF\big(H(s,t)\big)H_t$$

ולכן את האינטגרל לעיל ניתן לרשום כ –

$$\boxed{\equiv}\int_0^1\Big(DF\big(H(s,t)\big)H_t\cdot H_s+F\big(H(s,t)\big)\cdot H_{st}\Big)dt \boxed{\equiv}$$

כאשר השתמשנו בכך ש- $H_{st}=H_{ts}$, וכאן השתמשנו בהנחת הגזירות ברציפות. נמשיך:

$$\boxed{\equiv}\int_0^1\frac{d}{dt}\Big(F\big(H(s,t)\big)\cdot H_s\Big)dt=F\big(H(s,1)\big)\cdot H_s(s,1)-F\big(H(s,0)\big)\cdot H_s(s,0)$$

אבל $H(s,0)=\gamma_0(0)=const.$ ו - $H(s,1)=\gamma_0(1)=const.$, לכן $H_s(s,0)=H_s(s,1)=0$, ואגף ימין של המשוואה לעיל מתאפס. לכן $g'\equiv 0$ כנדרש. מ.ש.ל.

<u>**הגדרה:**</u> קבוצה פתוחה $U\subseteq\mathbb{R}^n$ נקראת **פשוטת קשר** אם כל מסילה גזירה ברציפות וסגורה הוא הומוטופית לנקודה (בהגדרה זו אנו מרשים לנקודה להיחשב כעקום סגור גזיר ברציפות).

דוגמא וציור!

<u>**הערה:**</u> בטופולוגיה עובדים עם מסילות רציפות, אך לפי משפט הקירוב של ויירשטראס, ההגדרה שלנו שקולה להגדרה עם מסילות גזירות אינסוף פעמים (כאשר מדברים על קבוצה פתוחה).

<u>**הוכחת משפט 26.6:**</u> אם F משמר מקומית, אזי התנאי $\frac{\partial F_j}{\partial x_i}=\frac{\partial F_i}{\partial x_j}$ מתקיים לפי מסקנה 26.4. בכיוון השני, אם מתקיים התנאי הזה, אז לפי הטענה, לאינטגרל הקווי של F על מסילות הומוטופיות אותו ערך. בהינתן כדור המוכל ב – U, קל להראות שכל מסילה גזירה וסגורה הינה הומוטופית לנקודה. לכן הערך של כל אינטגרל קווי סגור של F שווה ל-0, כלומר F משמר בכל כדור שמוכל ב - U. מ.ש.ל.

<u>**הוכחת משפט 26.7:**</u> צריך רק להוכיח שכל שדה משמר מקומית הוא משמר. נניח שיש לנו שדה משמר F בקבוצה פתוחה פשוטת קשר U . לפי משפט 26.6, מתקיים $\frac{\partial F_j}{\partial x_i}=\frac{\partial F_i}{\partial x_j}$. כיון שכל מסילה סגורה היא הומוטופית לנקודה, מתקיים לפי הטענה

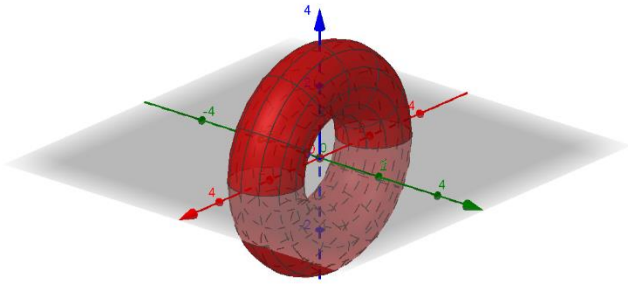
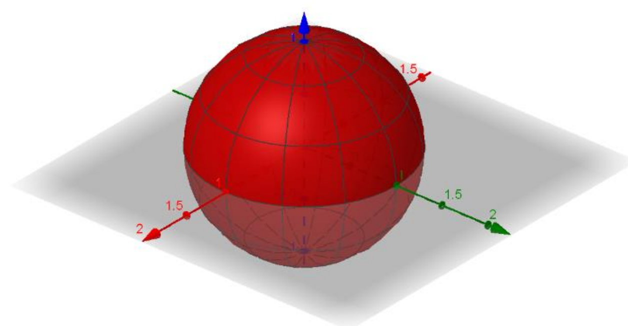
$$\int_\gamma F\cdot dr=\int_{pt.}F\cdot dr=0$$

לכל מסילה גזירה ברציפות. לכן האינטגרל הקווי של השדה מתאפס על כל מסילה (גזירה ברציפות) סגורה, ולכן השדה הינו משמר. מ.ש.ל.

<u>הרצאת העשרה</u>

נתבונן בספרה ובטורוס במרחב $\mathbb{R}^3$:

$$\mathbb{S}^2 = \partial B_3 = \{(x,y,z) | x^2+y^2+z^2=1\}$$
$$\mathbb{T}^2 = \left\{ \left((a+b\cos u)\cos v \, , b\sin u \, , (a+b\cos u)\sin v \right) \Big| \begin{matrix} 0 \le u \le 2\pi \\ 0 \le v \le 2\pi \end{matrix} \right\}$$

שאלה – האם קיים דיפאומורפיזם (כלומר העתקה גזירה ברציפות חד—חד ערכית ועל):

$$f: \mathbb{T}^2 \to \mathbb{S}^2$$

התשובה לכך, היא ש**לא** קיים דיפאומורפיזם כנ"ל (ולמעשה, על אף שזה לא יוכח בקורס, לא קיים אף הומיאומורפיזם ביניהם. כדי להראות שלא קיים הומיאומורפיזם יש להשתמש בכלים טופולוגיים, ואילו לנו יש כעת **בעיקר** כלים דיפרנציאליים).

נציג שלוש דרכים לפתרון הבעיה.

<u>**דרך ראשונה:**</u> הרעיון הכללי, הוא שבהינתן C עקום פשוט חלק וסגור ב-$\mathbb{T}^2$, עבור הנחת שלילה שקיים דיפאומורפיזם $f: \mathbb{T}^2 \to \mathbb{S}^2$, אם נגדיר $\Gamma = f(C) \subset \mathbb{S}^2$ אזי זהו עקום פשוט וחלק בספרה. זו עובדה שלכל Γ כזו מתקיים $\mathbb{S}^2 \setminus \Gamma$ לא קשיר. אבל קיים C כך ש-$\mathbb{T}^2 \setminus C$ קשיר (קל למצוא כזה), לכן $f(\mathbb{T}^2 \setminus C) = \mathbb{S}^2 \setminus f(C) = \mathbb{S}^2 \setminus \Gamma$ כתמונת פונקציה רציפה היא קבוצה קשירה תחת ההנחה שלנו, אך לפני רגע הערנו ש- $\mathbb{S}^2 \setminus \Gamma$ לא קשיר, וקיבלנו סתירה. יש בהוכחה הזו חור די רציני, כי העובדה ש - $\mathbb{S}^2 \setminus \Gamma$ לא קשיר אם Γ עקום פשוט וסגור, היא אמנם נכונה ואף אינטואיטיבית, אך היא רחוקה מלהיות טריוויאלית (זהו **משפט ז'ורדן** – באנגלית **Jordan Curve Theorem** – ולא נוכיח אותו בקורס זה).

<u>**דרך שניה להסתכל על הבעיה:**</u>

ניזכר במושג של הומוטופיה שהוגדר בפרק הקודם:

26.1 הגדרה - יהא X מרחב טופולוגי. עקום סגור ב-X זו קבוצה $C = \gamma([0,1])$ עבור X $\gamma: [0,1] \to X$ רציפה עבורה מתקיים $\gamma(0) = \gamma(1)$.

26.2 הגדרה – אם $C_i = \gamma_i([0,1])$ עבור $i = 0,1$ אזי הומוטופיה בין C_0 ל-C_1 היא פונקציה $H: [0,1] \times [0,1] \to X$ רציפה המקיימת:

$$H(s,0) = H(s,1) \quad \begin{matrix} H(0,t) = \gamma_0(t) \\ H(1,t) = \gamma_1(t) \end{matrix}$$

ב-$\mathbb{S}^2$ כל עקום סגור פשוט וחלק הינו הומוטופי לנקודה. נראה זאת, על ידי שנניח $\Gamma = \gamma([0,1])$ ונניח כי $(0,0,-1) \notin$ ונסמן γ:

$$F_s: \mathbb{S}^2 \setminus \{(0,0,-1)\} \to \mathbb{S}^2$$

$$F_s(x,y,z) = \frac{s(0,0,1) + (1-s)(x,y,z)}{\|s(0,0,1) + (1-s)(x,y,z)\|}$$

עבורה קל לראות כי מתקיים:

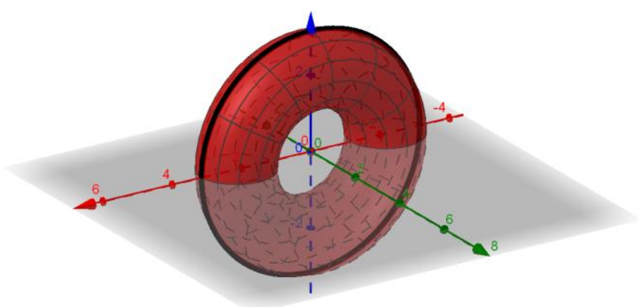
$$H(s,t) = \gamma_s(t) = F_s\big(\gamma(t)\big)$$

כלומר זו אכן הומוטופיה. עתה נרצה להראות כי בניגוד לכך, בטורוס $\mathbb{T}^2$, אין הומוטופיה שלוקחת את העקום:

$$C_1 = \{(a+b)\cos v\,, 0, (a+b)\sin v\,|v \in [0,2\pi]\}$$

לנקודה. בזאת תסתיים ההוכחה, כי אם אכן קיימת אותה f דיפאומורפיזם ונתבונן ב-$H(s,t)$ אשר לוקחת את $\Gamma =$ $f(C_1)$ לנקודה, אזי $f^{-1} \circ H$ לוקחת את γ_1 לנקודה.

נראה כי הנ"ל מוביל לסתירה – על ידי התבוננות באינטגרל:



$$\int_{C_1} \underbrace{-\frac{z}{x^2+z^2}dx + \frac{x}{x^2+z^2}dz}_{=\omega} = 2\pi$$

שאותו זכינו להכיר כבר בהרצאות הקודמות. זו ה-1-תבנית הדיפרנציאלית המוגדרת עבור:

$$F = \left(-\frac{z}{x^2+y^2}, 0, \frac{x}{x^2+z^2}\right)$$

זהו שדה משמר מקומית, כי מתקיים:

$$\nabla \times F = 0$$

אם C_1 ו-C_2 הומוטופיים אז $2\pi = \int_{C_1}\omega = \int_{C_2}\omega$ לפי הטענה בנספח של הפרק הקודם. לכן העקום C_1 איננו יכול להיות הומוטופי לנקודה.

<u>דרך שלישית ואחרונה:</u>

הרעיון המרכזי: על $\mathbb{T}^2$ קיים שדה וקטור C^1 משיק שאינו מתאפס ועל $\mathbb{S}^2$ אין.

26.3 הגדרה – שדה וקטורי על יריעה $M \subset \mathbb{R}^3$ היא פונקציה $F: M \to \mathbb{R}^3$ כך שמתקיים:

$$F(p) \in T_p(M)$$

כאן נחשוב על $T_p(M)$ כמרחב וקטורי דרך 0.

כלומר, אם ישנו שדה וקטורי $\neq 0$, אזי יש גם שדה מנורמל $\frac{F(x)}{\|F(x)\|}$. לדוגמה, $F: \mathbb{T}^2 \to \mathbb{R}^3$ המוגדרת על ידי:

$$F\big((a+b\cos u)\cos v\,, b\sin u\,, (a+b\cos u)\sin v\big) = (-(a+b\cos u)\sin v\,, 0, (a+b\cos u)\cos v)$$

הינו שדה משיק שאינו מתאפס באף נקודה.

26.4 משפט (משפט הכדור השעיר) – על $\mathbb{S}^2$ לא קיים שדה וקטורי משיק מנורמל.

<u>הערה</u> – בהקשר של הבעיה בה אנו עוסקים בשיעור, לו הייתה $f: \mathbb{T}^2 \to \mathbb{S}^2$ אזי היינו יכולים להגדיר $G: \mathbb{S}^2 \to \mathbb{R}^3$ שדה משיק שאינו מתאפס באופן הבא:

$$G\big(f(p)\big) := df(p)[F(p)]$$

בסתירה למשפט, ולכן נסיק כי אין כזו f.

ההוכחה שלנו מבוססת על המאמר:

Milnor, John. "Analytic proofs of the “hairy ball theorem” and the Brouwer fixed point theorem." *The American Mathematical Monthly* 85.7 (1978): 521-524.

מומלץ לסטודנטים לנסות לקרוא את המאמר עצמו.

נניח בשלילה כי קיים שדה וקטורי משיק מנורמל $F(x) \mapsto x \ni \mathbb{S}^2$ כלומר שדה וקטורי המקיים כמובן $\|F(x)\| = 1$.
נזכיר כי על פי הגדרה:

$$F(x) \cdot x = 0 \Leftrightarrow F(x) \in T_x(\mathbb{S}^2)$$

נסמן:

$$A = \left\{x \in \mathbb{R}^3 \middle| \frac{1}{2} \le \|x\| \le \frac{3}{2}\right\}$$

על $\mathbb{S}^2$ נגדיר את $f_t(x) = x + tF(x)$ ונרחיב את F ואת f_t ל-A על ידי:

$$\begin{matrix} F(rx) = rF(x) \\ f_t(rx) = rx + rtF(x) = rf_t(x) \end{matrix} \quad \forall \|x\| = 1$$

נקל לראות כי $f_t, F \in C^1$ בסביבת A.

26.5 למה - קיים $\varepsilon > 0$ כך שלכל $t \in (-\varepsilon, \varepsilon)$ מתקיים f_t חד-חד ערכית ורגולרית על A וכן מתקיים:

$$Vol\big(f_t(A)\big) = \begin{matrix} \text{פונקציה פולינומית} \\ t - \text{ב} \end{matrix}$$

<u>הוכחה:</u>

$F \in C^1$ בסביבת A ולכן ליפשיצית. כלומר קיים C עבורו מתקיים:

$$\forall x, y \in A \quad \|F(x) - F(y)\| \le C\|x - y\|$$

נבחר $\varepsilon_1 = C^{-1}$. עבור $t \in (-\varepsilon_1, \varepsilon_1)$

מתקיים:

$$f_t(x) = f_t(y) \Rightarrow x + tF(x) = y + tF(y) \Rightarrow \|x - y\| = |t|\|F(x) - F(y)\| \le |t|C\|x - y\| < \|x - y\|$$

ולכן מתחייב כי $x = y$ אחרת נקבל סתירה.

נסמן:

$$Df_t(x) = I + t\left[\frac{\partial F_j}{\partial x_j}(x)\right]$$

ונשים לב כי:

$$\det Df_t(x) = 1 + d_1(x)t + d_2(x)t^2 + d_3(x)t^3$$

קיים ε_2 כך שמתקיים לכל $t \in (-\varepsilon_2, \varepsilon_2)$:

$$\det Df_t(x) > 0$$

נסמן $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$ אם $t \in (-\varepsilon, \varepsilon)$ אז מתקיימים תנאי משפט החלפת המשתנים, ונקבל

$$Vol\big(f_t(A)\big) = \int_{f_t(A)} dV = \int_A \det Df_t(x)\, dV(x) = \int_A \sum_i d_i(x)t^i\, dV(x) = \sum_i a_i t_i$$

עבור הסימון $a_i = \int_A d_i(x)V(x)$.

26.6 למה – לכל $\frac{1}{2} \le r \le \frac{3}{2}$ **הפונקציה** f_t **מעתיקה את** $r\mathbb{S}^2$ **באופן חד-חד ערכי על** $\sqrt{1+t^2}r\mathbb{S}^2$.

הוכחה - $F(x) \perp x$

$$\|f_t(rx)\|^2 = \|rx + trF(x)\|^2 = r^2(1+t^2)$$

וכן ראינו כי $f_t: r\mathbb{S}^2 \to \sqrt{1+t^2}r\mathbb{S}^2$ חד-חד ערכית.

מספיק להראות כי:

$$f_t(\mathbb{S}^2) = \sqrt{1+t^2}\mathbb{S}^2$$

כלומר תחת ההנחה ש-$r = 1$ שכן לכל המקרה האחרים מדובר על מכפלה בסקלר. כבר ראינו את ההכלה $\subseteq$.

נתבונן בתמונה $f_t\left(\left\{\frac{1}{2} < \|x\| < \frac{3}{2}\right\}\right) \Leftarrow$ ממשפט ההעתקה הפתוחה מתקיים שזו קבוצה פתוחה $\mathbb{R}^3$.

בפרט מתקיים:

$$\sqrt{1+t^2}\mathbb{S}^2 \cap f_t\left(\left\{\frac{1}{2} < \|x\| < \frac{3}{2}\right\}\right) = f_t(\mathbb{S}^2)$$

וזאת משום שבמובן הטופולוגי האגף השמאלי צריך להיות קבוצה קומפקטית, סגורה, לא ריקה, ופתוחה ב-$\sqrt{1+t^2}\mathbb{S}^2$ ומכאן שמדובר בהכרח ב-$\sqrt{1+t^2}\mathbb{S}^2$ עצמה, כנדרש.

הוכחת המשפט –

תחת ההנחה שקיים שדה כזה, F, ראינו כי ניתן להרחיב אותו ל-f_t שהגדרנו זה עתה. אך נשים לב כי מתקיים:

$$f_t(A) = \sqrt{1+t^2}A$$

וזאת לפי הלמה. לכן, נקבל את היחס הבא בין הנפחים:

$$Vol\big(f_t(A)\big) = (1+t^2)^{\frac{3}{2}}Vol(A)$$

מצד שני, אנו יודעים כי $Vol\big(f_t(A)\big)$ צריך להיות פולינום מדרגה שלישית ב-t, אך זאת סתירה משום ש-$(1+t^2)^{\frac{3}{2}}$ איננו פולינום ממעלה שלישית (למשל, אם נגזור פונקציה זו ארבע פעמים היא לא תתאפס באופן זהותי).