**הסתברות ותהליכים סטוכסטיים (פרופ' יזהר בר-גד)**

**09/05/2020**

**אופטימיזציה**

*נושא האופטימיזציה איננו חלק שעוסק בעיבוד אותות, אלא כלי שימושי שמשתלב בהרבה כלים חשובים המלווים את הקורס בהרבה נושאים.*

*מקורות המידע עליהם מתבססת ההרצאה:*

*קורס בביוסטטיסטיקה של Goncalo Abecasis-*

[*http://csg.sph.umich.edu/abecasis*](http://csg.sph.umich.edu/abecasis)

*קורס באלגוריתמים מטא-היוריסטיים של Wendy Williams-*

[*http://www.dbai.tuwein.ac.at/*](http://www.dbai.tuwein.ac.at/)

*הספר של Numerical Recipes in C++*

*Practical optimization- P. Gill, W. Murray, M. Wright, 1981*

*Practical optimization: Algorithms and Engineering Applications- A. Antoniou and W-S. Lu, 2007*

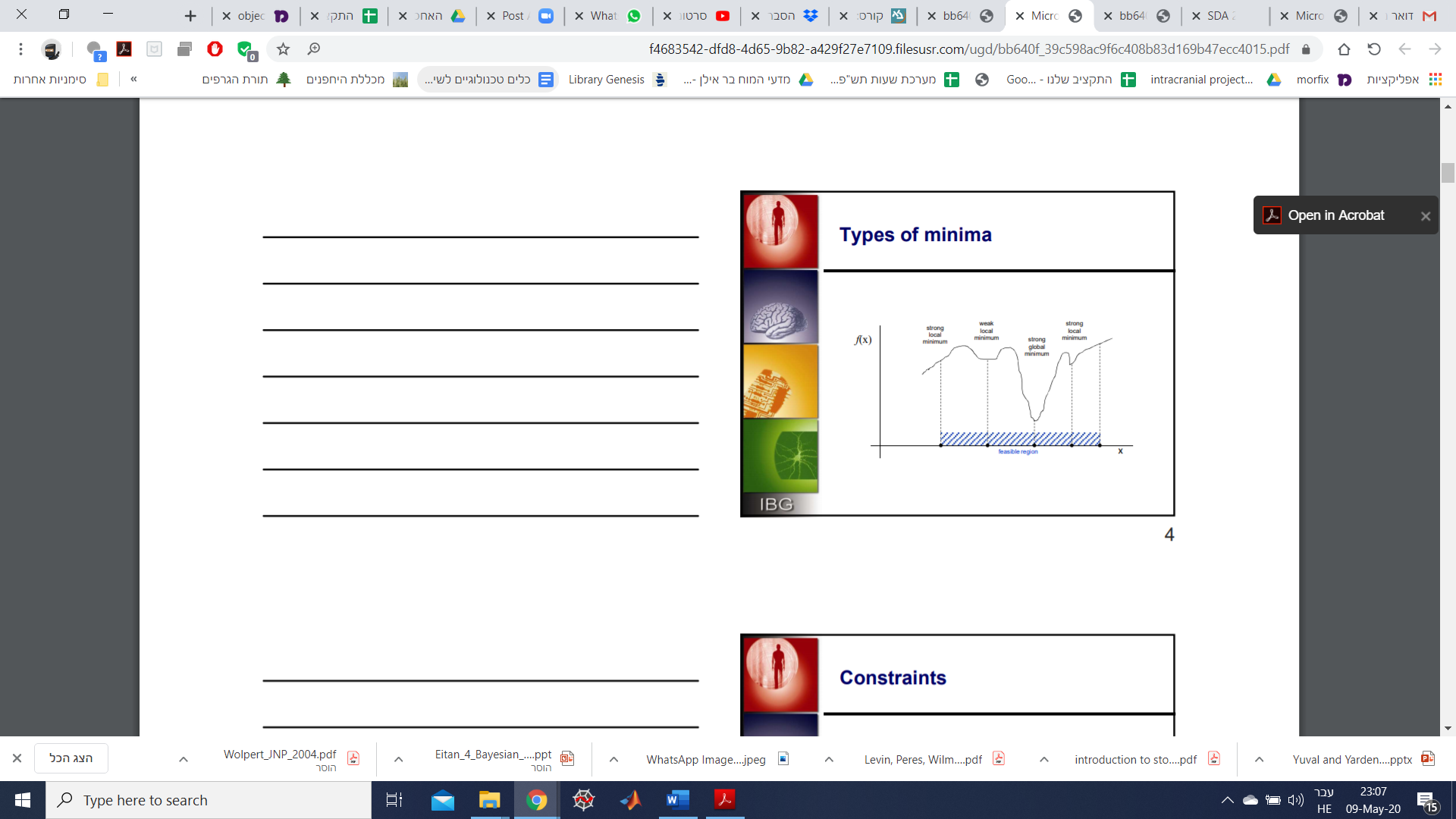
*אופטימיזציה- התהליך של מציאת התנאים שנותנים את המינימום או המקסימום של פונקציה המייצגת את המערכת שלנו. אנחנו נדון במינימיזציה בלבד, ואם יהיה צורך במקסימיזציה נעבוד עם הפונקציה .*

*מטרות האופטימיזציה:*

* *מציאת מינימום לוקאלי- הערך הנמוך ביותר במסגרת סביבה סגורה.*
* *מינימום גלובאלי- הערך הנמוך ביותר של הפונקציה.*
* *פתרונות תחת אילוצים או ללא אילוצים.*

*סוגים של נקודות מינימום:*

*לרוב מחלקים ל-weak minimum לעומת strong minimum, הראשון מתייחס לכך שיש סדרה של נקודות מינימום בעלות ערך מינימלי והשני לנקודה אחת ספציפית המקיימת זאת. נקודות הקצה תמיד יהיו נקודות אופטימום- מינימום או מקסימום (לעיתים גם גלובליות). הן יכולות להיות weak או strong, ולכן יש לנו את מגבלת ה-feasible region.*

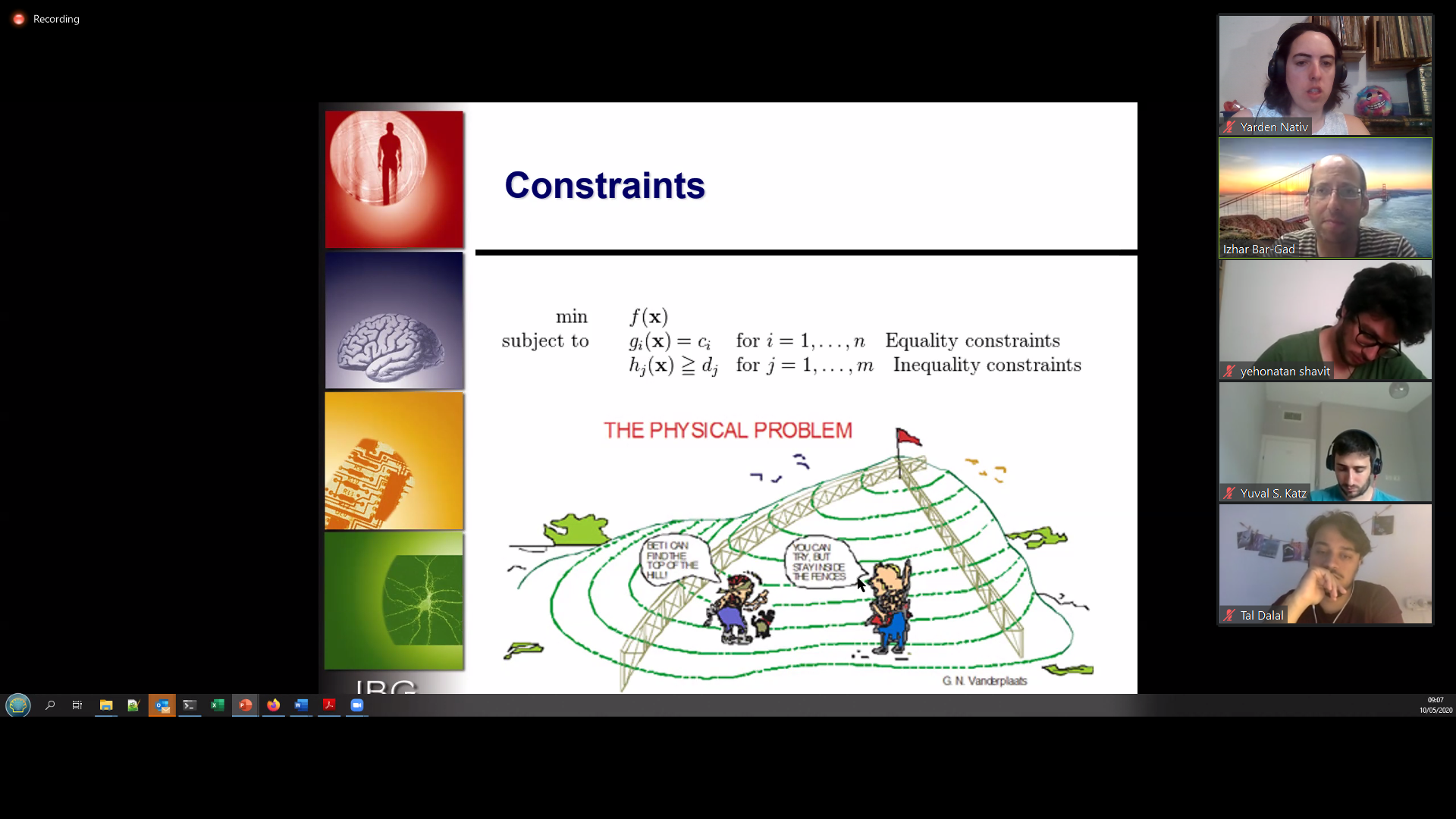


*דוגמא לאילוץ:*

*מינימום של תחת שני סוגי אילוצים-*

*אילוצי שוויון-*

*אילוצי אי-שוויון-*



***שיטות יוניוואריות:***

*בהינתן משתנה אחד בקבוצת אלטרנטיבות סופית (בדרך כלל קטנה), נשתמש בשיטות דירוג. לדוגמא- אם רוצים לצמצם את עלות הקורס וצריך לבחור בין הקורסים SDA ו-MATLAB, הבחירה יכולה להיות נתונה ע"י פונקציית העלות:*

|  |  |  |  |
| --- | --- | --- | --- |
| *קורס* | *מידת השקעה* | *רווח* | *מחיר (מידת השקעה \* 7 – רווח)* |
| *SDA* | *2* | *10* | *4* |
| *MATLAB* | *1* | *0* | *7* |

*במצב שבו נתונה קבוצה אינסופית (או קבוצה סופית אך גדולה מאוד), צריך להשתמש בשיטות אנליטיות ונומריות לאופטימיזציה. לדוגמא אם נתונה הפונקציה תחת האילוץ .*

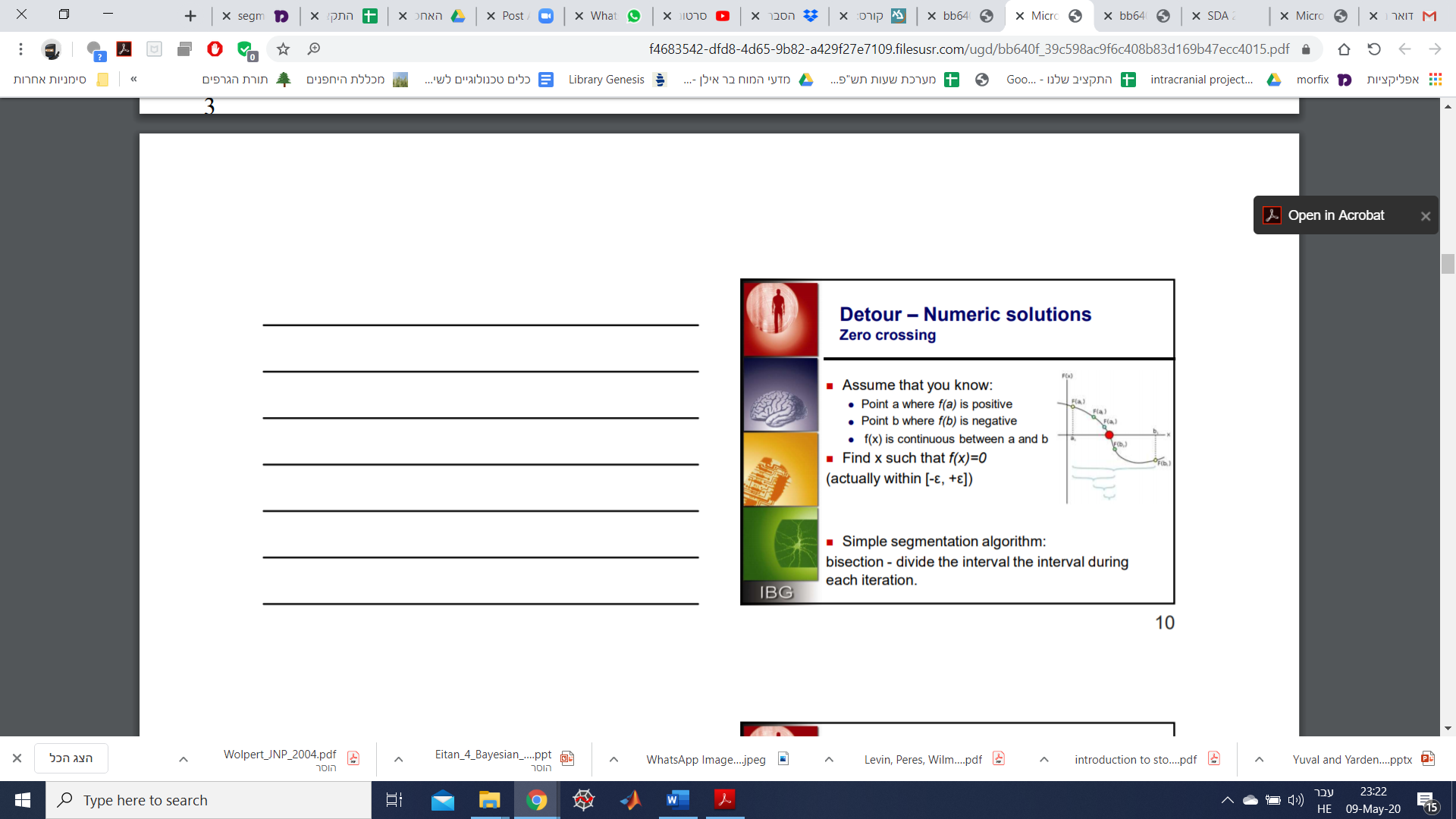
*הפתרון האנליטי לחישוב מינימום מבוצע בשני צעדים:*

* *לחשב את הנגזרת, לזהות את כל הערכים בהם הנגזרת מתאפסת, לחשב את הנגזרת השנייה עבור ערכים אלה ולזהות נקודות מינימום לוקאליות.*
* *להשוות את הערכים של כל נקודות המינימום הלוקאליות על מנת לזהות נקודת מינימום גלובאלית.*

*למשל:*

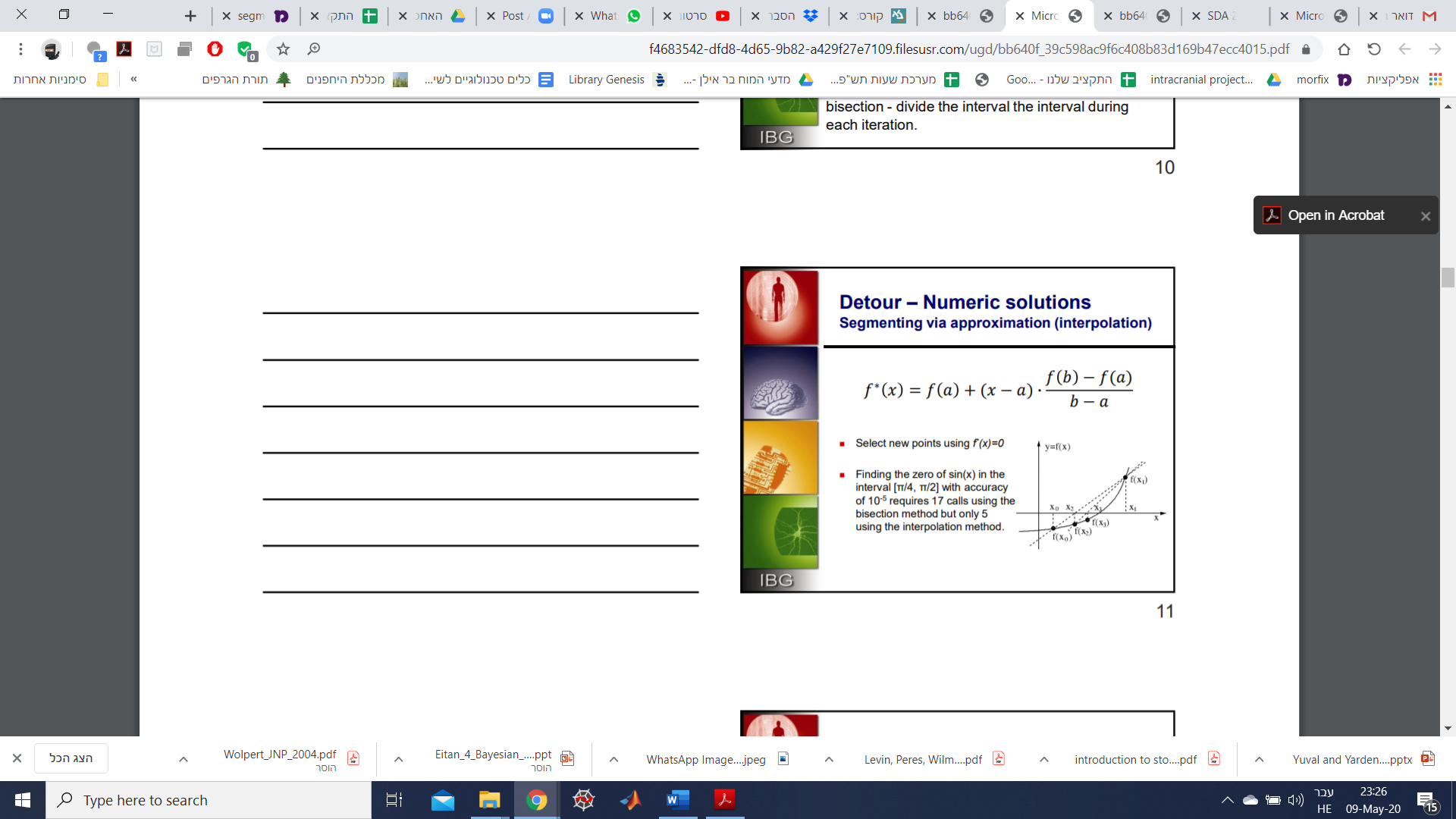
*הפתרונות הנומריים הם מעין "עיקוף" לפתרונות האנליטיים. הם משמשים אותנו במקרים בהם לא ניתן לבצע פתרון אנליטי ישיר (למשל אם איננו יודעים בדיוק מהי הפונקציה).*

*נתבונן בשיטה הראשונה- zero crossing. נניח כי ידוע לנו שקיימת נקודה בפונקציה שהיא חיובית ונקודה בפונקציה שהיא שלילית, וידוע כי רציפה בין ל-, אז צריך למצוא כך ש- (בדומה לחישוב אריה במדבר, חישוב בינארי, נחפש כמובן בטווח קטן שבין ). אלגוריתם החלוקה הפשוט הוא ביסקציה- כל פעם מחלקים את ההפרש של ההפרש בתוך כל איטרציה, בהתבסס על חיפוש בינארי. אפשר לראות להלן דוגמא לשלושה צעדים המביאים למציאת נקודת החיתוך עם ציר ה-. חשוב לשים לב שייתכנו מספר נקודות חיתוך עם הציר, והשיטה הזו מוצאת רק נקודה אחת ספציפית.*



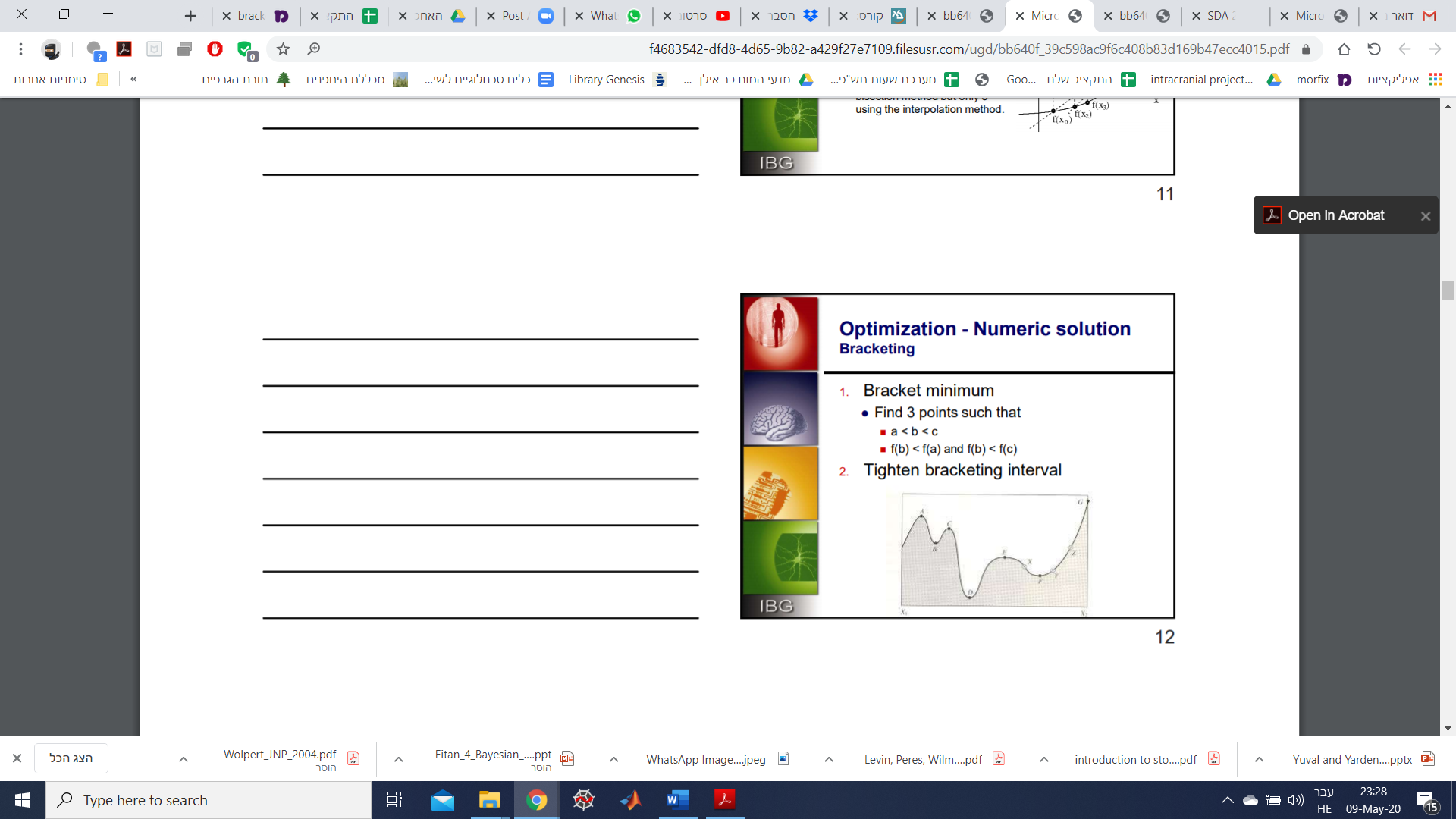
*שיטה נוספת היא חלוקה דרך הערכה (אינטרפולציה). ישנן שיטות רבות של אינטרפולציה, כאשר השיטה הפשוטה היא ע"י שימוש בקירוב מסדר ראשון של טור טיילור:*

*בשיטה זו בוחרים נקודות חדשות תוך שימוש ב-, בדיוק אותו רעיון רק עם ביסקציה רק באמצעות הפונקציה הזו. למשל אם מוצאים את הנקודות עבורן מתאפס בטווח עם דיוק של תידרש קריאה חוזרת של 17 פעמים תוך שימוש בשיטת הביסקציה, אך רק 5 פעמים בשימוש בשיטת האינטרפולציה.*



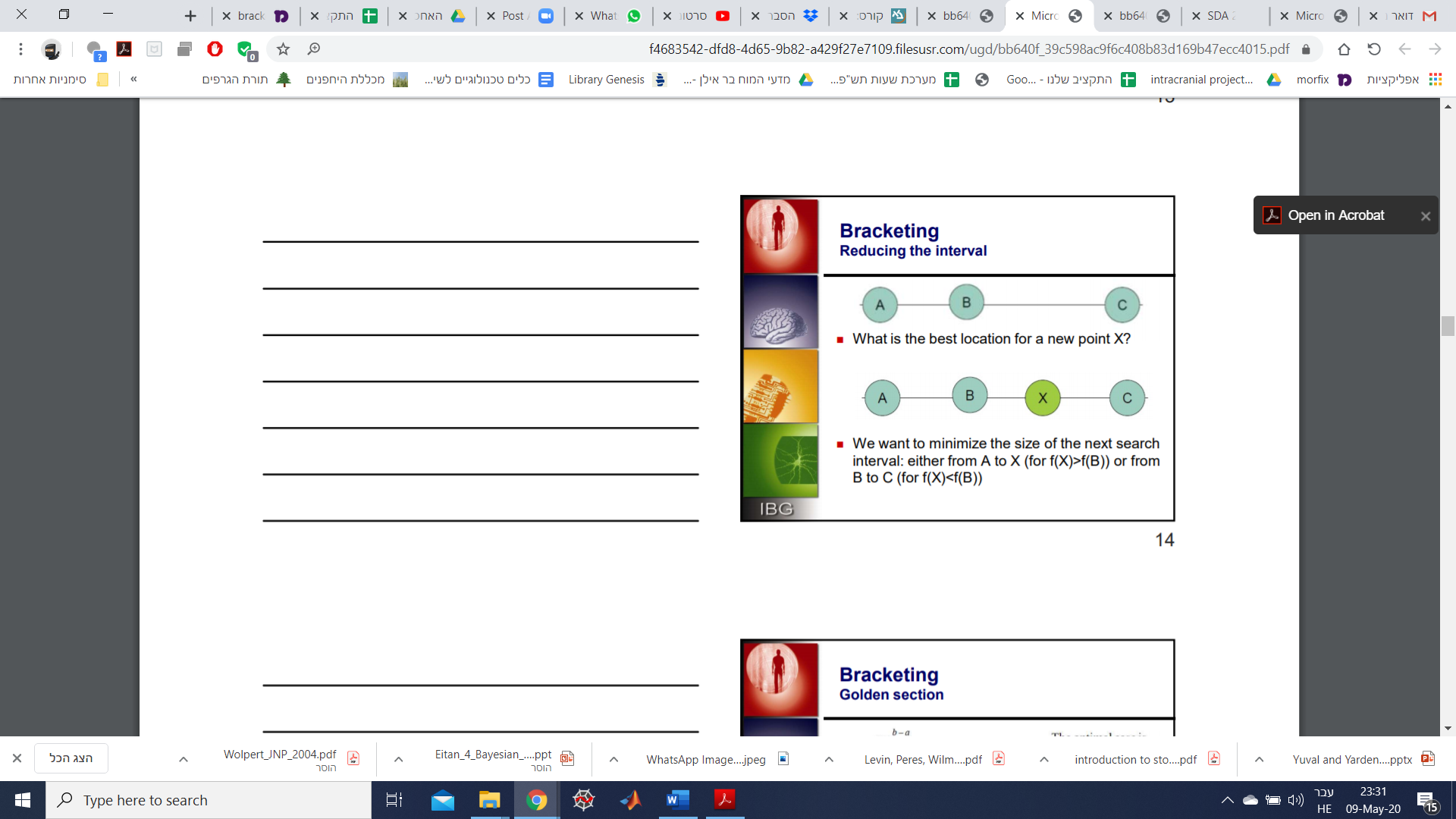
*עד עכשיו התעסקנו עם שיטות נומריות למציאת נקודות החיתוך עם הציר. כעת נעבור לשיטות שמחפשות מינימום.*

*שיטת ה-bracketing היא שיטה שבה מחפשים מינימום בתוך סוגריים. מוצאים שלוש נקודות כך ש- ומתקיים וגם ומצמצמים את הסוגריים בכל צעד.*



*שיטה דומה היא שיטת ה-initial bracketing interval. בשיטה זו צריך רק שתי נקודות כך ש-. בשיטה זו לוקחים צעדים גדולים יותר מעבר ל- עד שהפונקציה עולה.*

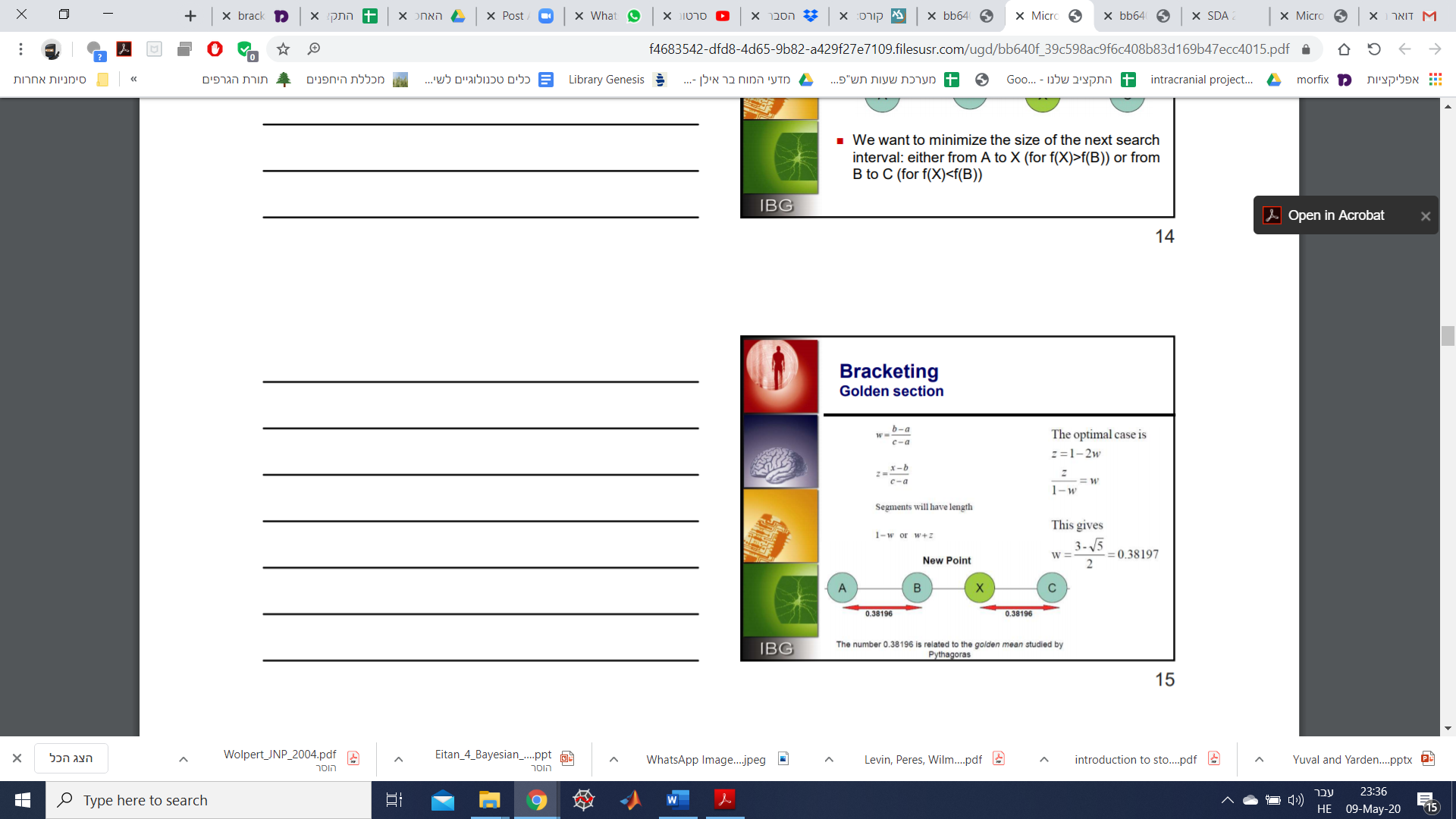
*בכל אופן בשיטות ה-bracketing צריך להפחית את האינטרוולים. למשל אם נתונות שלוש הנקודות הבאות, מהו המיקום הטוב ביותר עבור נקודה חדשה?*



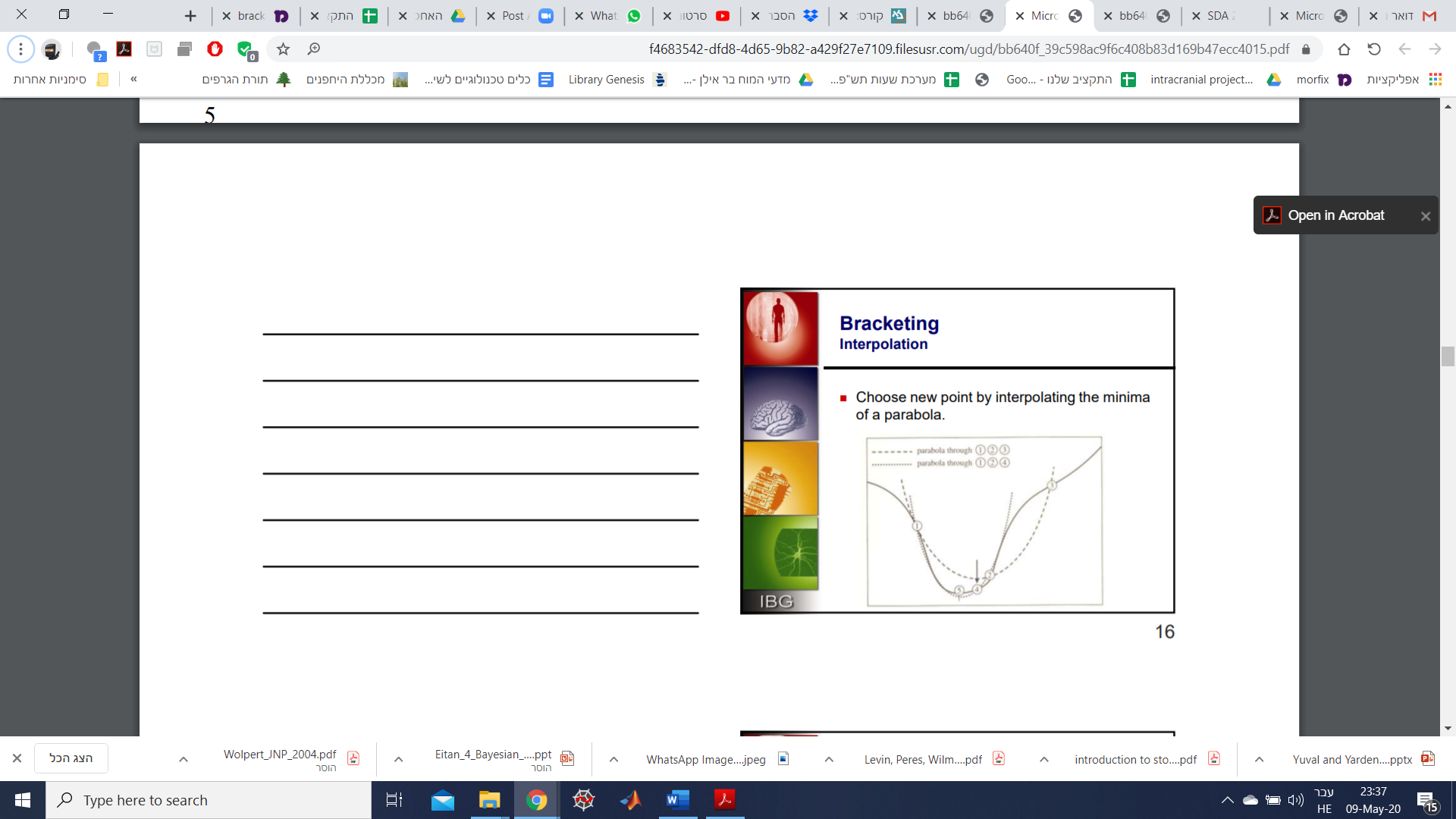
*אנחנו רוצים לצמצם את הגודל של אינטרוול החיפוש הבא- או שנצמצם את הסוגריים לטווח בין ל- אם התקבל מצב שבו או שנצמצם את הסוגריים לטווח בין ל- עבור .*

*הבחירה הטובה ביותר ל- היא חיתוך הזהב בו משתמשים ב- כאשר:*

*וחיתוך זה מקיים שם , ובמצב כזה למקטעים יהיו האורכים או מה שנותן כי . מספר זה קשור לממוצע הזהב שנחקר ע"י פיתגורס.*



*דרך אחרת לצמצום הסוגריים היא לעשות אינטרפולציה ב-bracketing ע"י בחירת נקודה חדשה בעזרת אינטרפולציה של המינימום של פרבולה. בהינתן שיש שלוש נקודות, מעבירים פרבולה:*

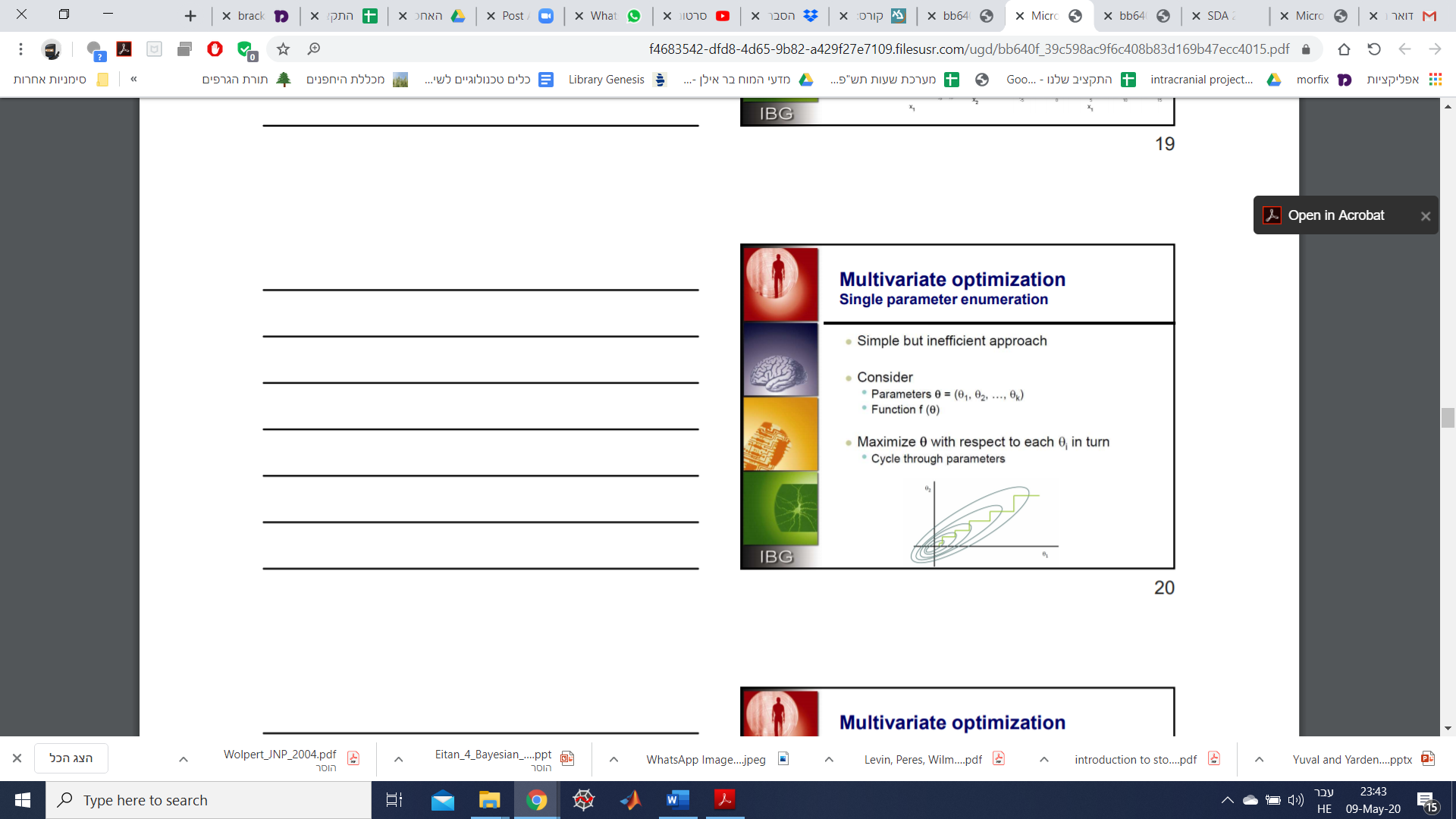


*ואפשר גם על ידי אינטרפולציה פולינומיאלית. מבצעים סדרת צעדים שחוזרת על עצמה- שמים סוגריים מסביב למינימום, לאחר מכן מתאימים פולינום מסדר רביעי או שלישי שעושה אינטרפולציה ל- בחלק מהנקודות באינטרוול. בשלב הזה מוצאים את המינימום של הפולינום, נפטרים מהנקודות הגרועות ביותר וחוזרים על כל הצעדים.*

***שיטות מולטיוואריטיות:***

*אופטימיזציה של פונקציה בעלת פרמטרים קורית באופן שונה. נתבונן למשל במצב שבו , כך שפונקציית המחיר תוכל להיראות בצורה ויזואלית בשלושה מימדים.*

*השיטה הראשונה בה נתבונן היא שיטה של אנומרציה של פרמטר יחיד. זוהי שיטה פשוטה אך לא יעילה. עבור הפונקציה בהינתן עושים מינימיזציה של ביחס לכל רכיב ומקבלים את התהליך הבא (ירידה במדרגות דרך האליפסות:*



*שיטה נוספת היא מקסום המתבסס על גזירה. בהינתן סט הפרמטרים הנתון כוקטור והפונקציה וקטור הציון נתון ע"י:*

*מוצאים את המקסימום לאורך .*

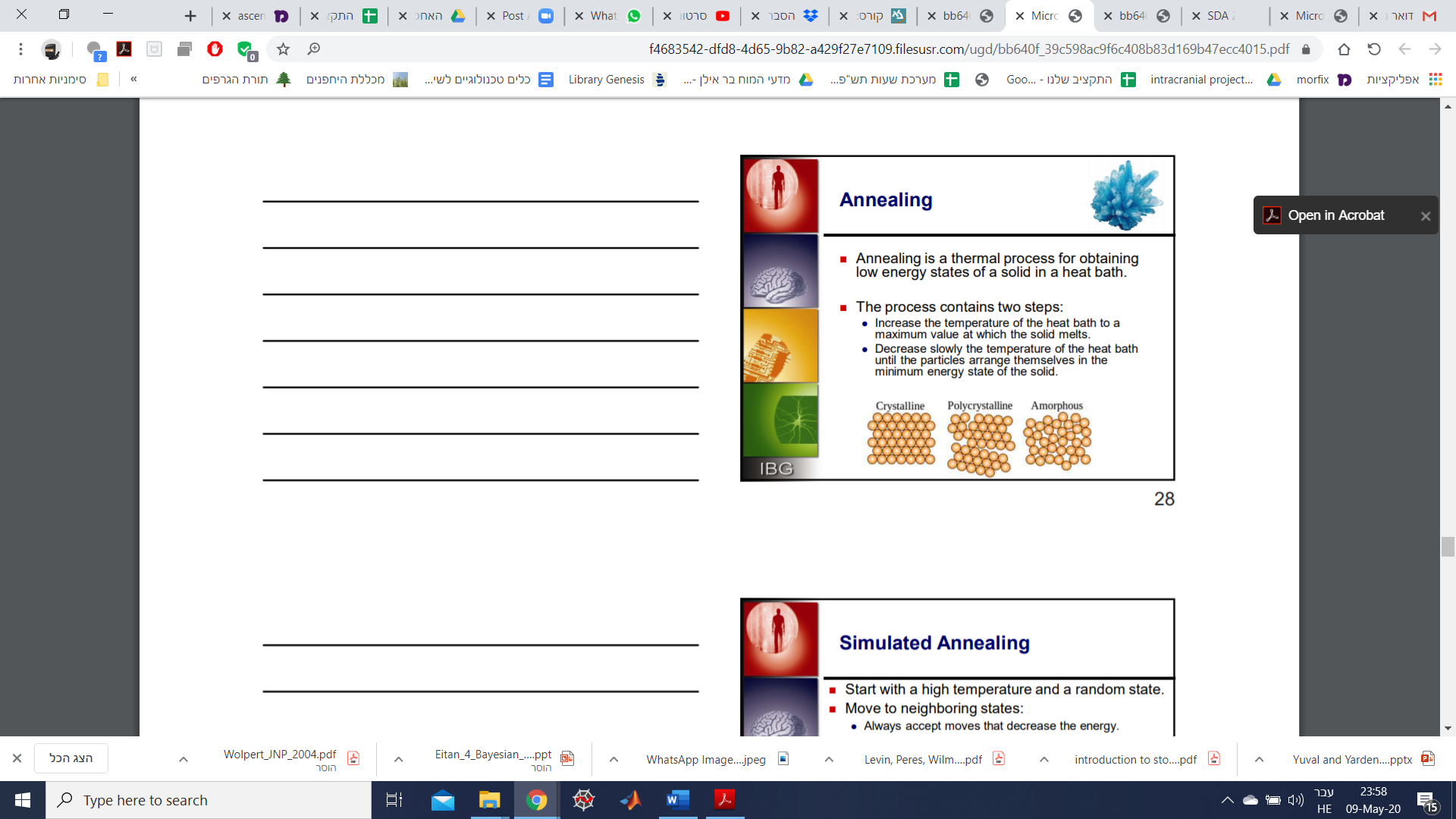
*מרעיון זה התפתחה שיטת "המדרון התלול" המתבססת על כך שניתן לעשות מינימיזציה של פונקציה בעלת מימדים באמצעות סדרה של קווים בגודל המבצעים את המינימיזציה: . בשיטת המדרון התלול בוחרים את כך שיהיה מקביל לגראדיאנט .*

***Simulated Annealing***

*כל הפתרונות שהראינו עד עכשיו לא אפשרו למצוא פתרון גלובלי, אלא רק פתרונות מינימליים. לחיפוש מינימום גלובלי ניתן להשתמש באלגוריתמים לחיפוש לוקאלי. אלגוריתמים היוריסטיים הם דרך לגשת לפתרון בעיות ע"י שימוש בכלים פרקטיים שאינם בהכרח אופטימליים או מושלמים, אבל מספיק יעילים עבור מטרות מיידיות. השיטה הזו לא תמיד תעבוד, ב-100% מהמקרים, אבל השיטה עובדת בהרבה מקרים להרבה דברים. אלגוריתמים לחיפוש לוקאלי הם גישה היוריסטית לאופטימיזציה שמתחילה מפתרון מועמד ואז באופן איטרטיבי פונה לפתרונות שכנים. דוגמאות לאלגוריתמים לחיפוש לוקאלי: Hill-climbing, Random Walk, Simulated Annealing.*

*נשים לב כי בכל האלגוריתמים שהוזכרו עד עכשיו, האלגוריתמים שימשו למציאת נקודות מינימום לוקאליות. למציאת מינימום גלובלי, הגישה הפשוטה ביותר היא לבחון מספר נקודות התחלה. כמובן במקרה של בעיה במספר משתנים, הבעיה מורכבת יותר. הבעיה בעלייה מנקודות התחלה היא שהאפקט שאנחנו מעוניינים בו הוא שנעזוב נקודות מינימום לוקלי אך אנחנו עלולים לצאת מנקודת מינימום גלובאלית לאחר שהגענו אליה.*

*לכן שיטת ה-Annealing (ליבון) מתבססת על תהליך תרמי השומר על מצבי אנרגיה נמוכים של מוצק באמבט חום. התהליך כולל שני שלבים- בשלב הראשון מעלים את הטמפרטורה של החום באמבט לדרגה המקסימלית שבה המוצק ניתך. בשלב השני מורידים באופן איטי והדרגתי את הטמפרטורה של אמבט החום עד שהחלקיקים מסדרים את עצמם במצב האנרגטי המינימלי היציב עבור מוצק. המצב הזה מבטיח לנו אנטרופיה מינימלית (מקסימום סדר).*

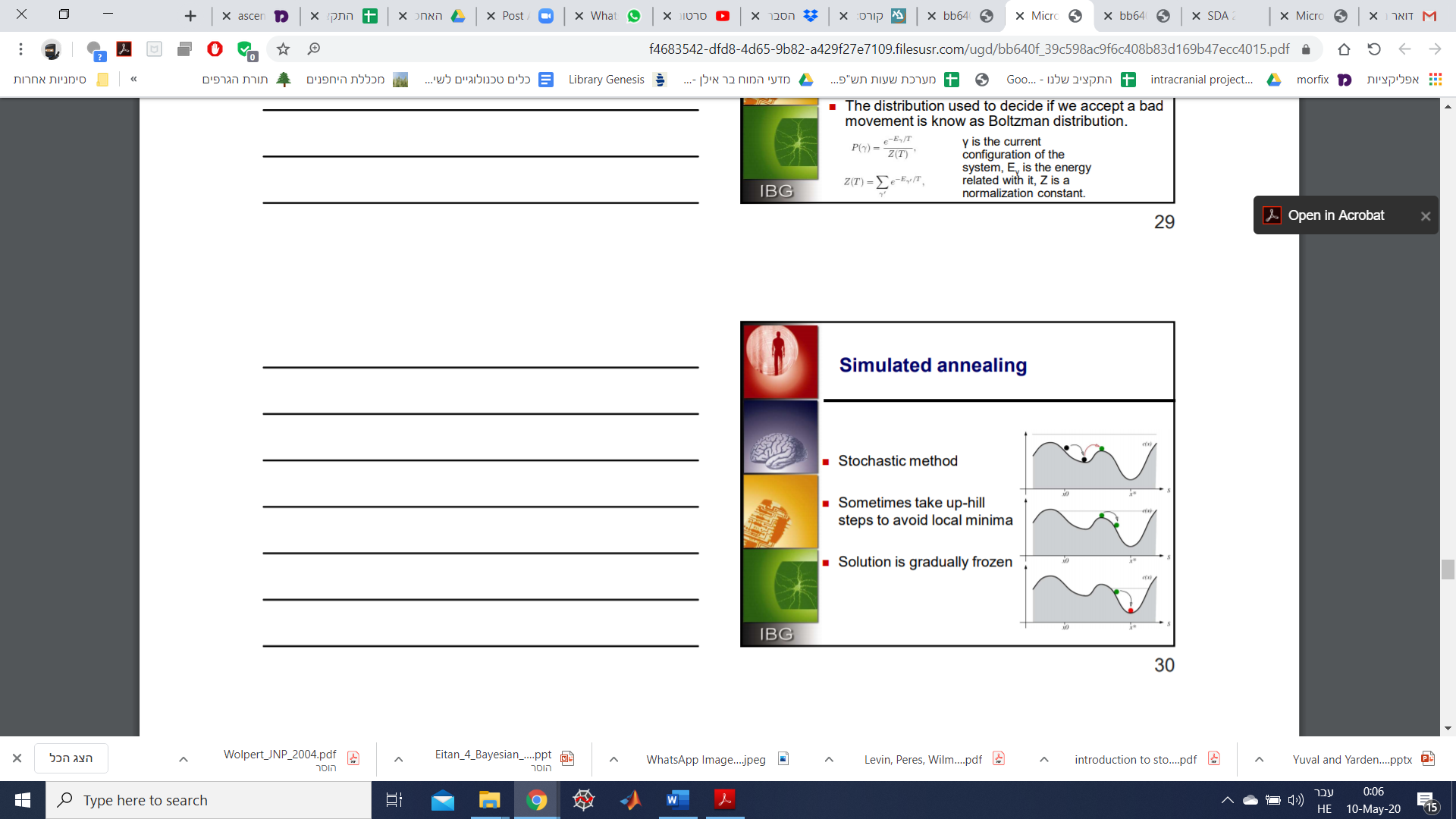


*לכן שיטת Simulated Annealing מתבססת על רעיון דומה. מתחילים בטמפרטורה גבוהה ובמצב רנדומלי. נעים למצבים שכנים- ממשיכים רק עם מצבים המקטינים את כמות האנרגיה, ומקבלים צעדים "רעים" בהתאם להתפלגות התלויה בטמפרטורה. מורידים את הטמפרטורה באופן הדרגתי, מקבלים פחות מצבים "רעים" עם כל ירידה בטמפרטורה. כאשר הטמפרטורה נמוכות מאוד מתקבל בעצם אלגוריתם hill-climbing.*

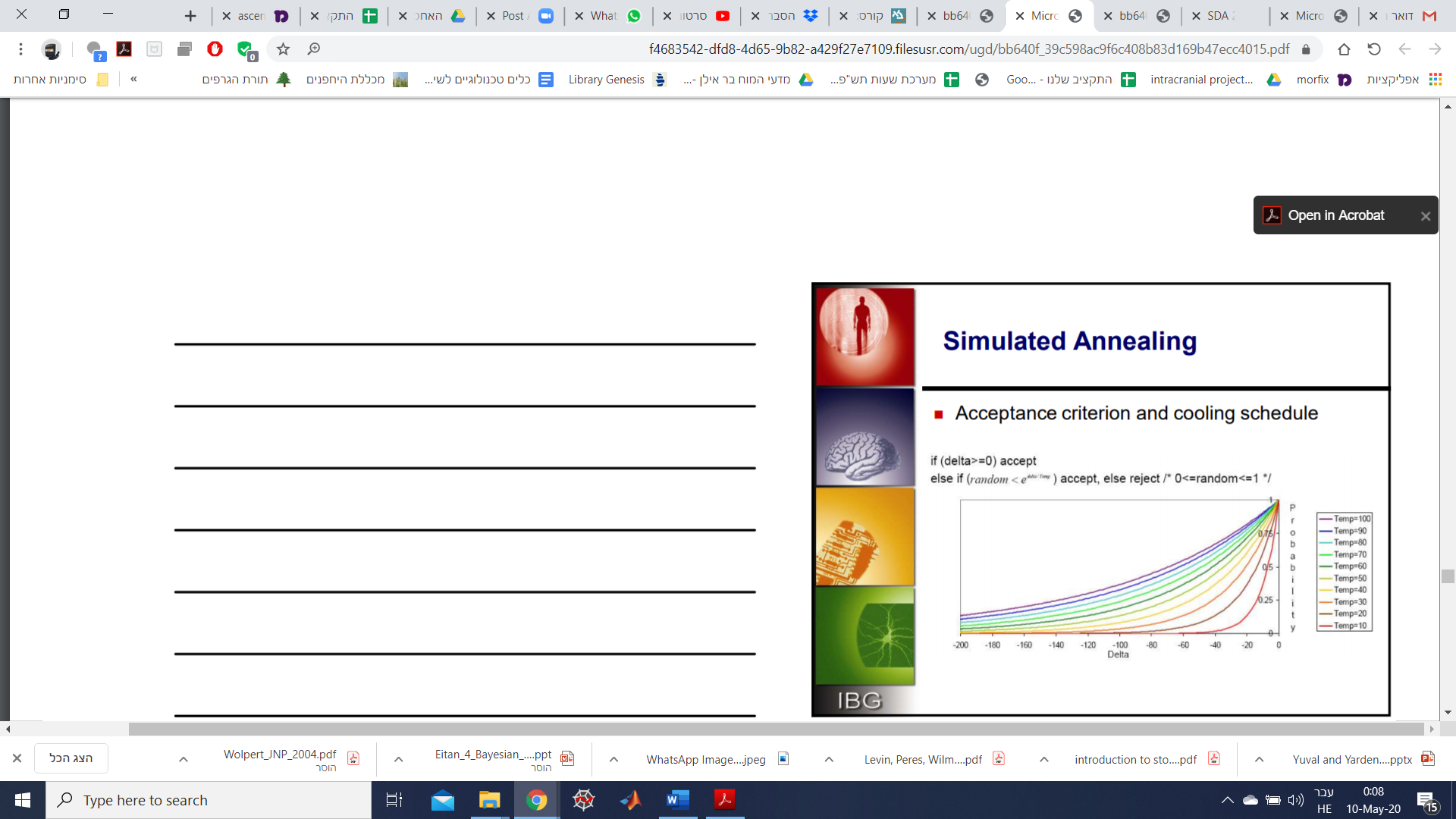
*ההתפלגות משמשת על מנת להחליט אם קיבלנו צעד רע היא התפלגות בולצמן הנתונה ע"י הנוסחא:*

*כאשר היא הקונפיגורציה הנוכחית של המערכת, היא האנרגיה הקשורה אליה ו- קבוע נרמול.*

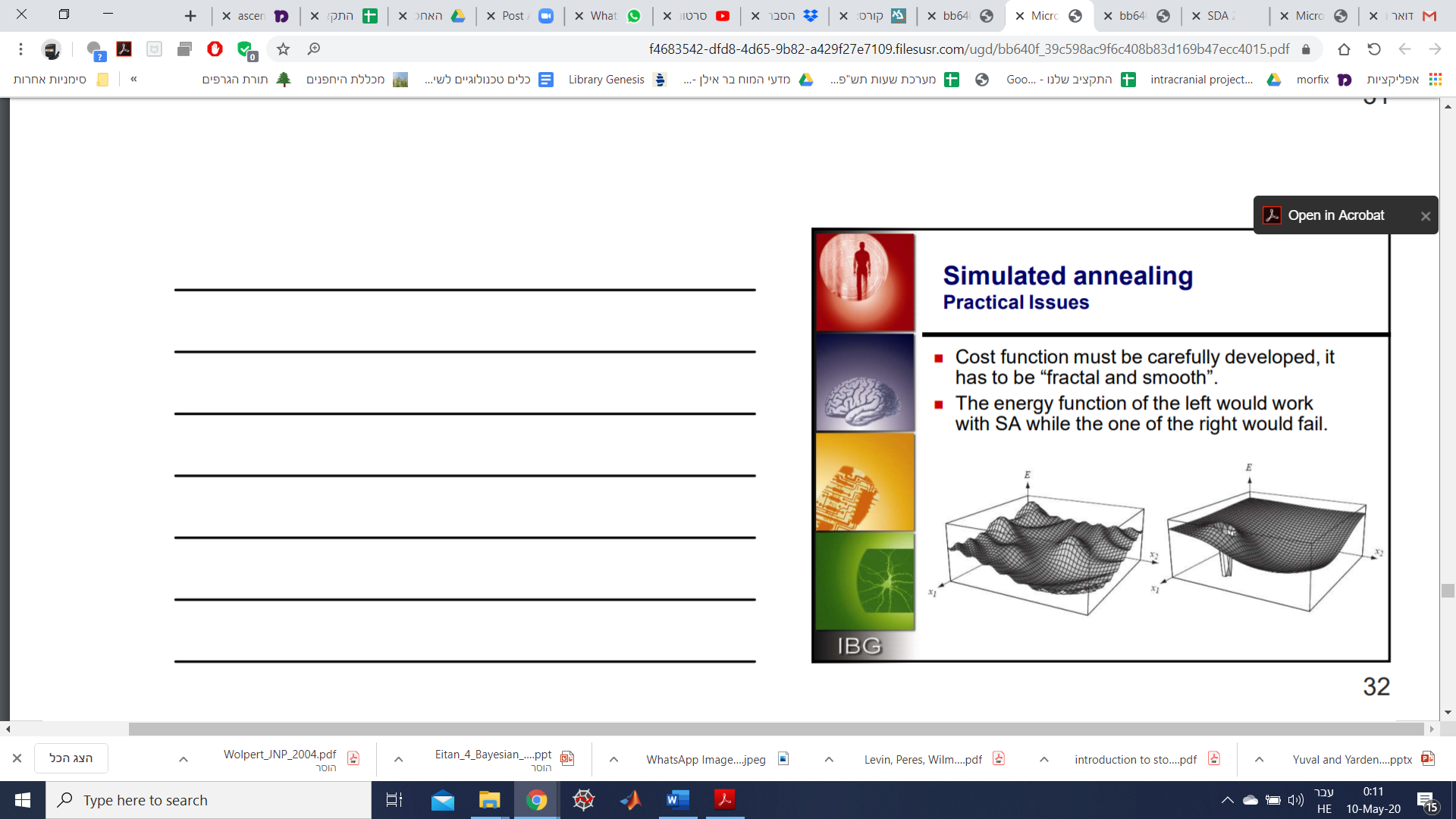
*שיטת ה-Simulated Annealing היא שיטה סטוכסטית, לעיתים לוקחת צעדים במעלה הגבעה על מנת להימנע ממינימום לוקאלי אך באופן הדרגתי הפתרון "מוקפא" באופן הדרגתי. בכל פעם נגדיר מחדש את המינימום- גם אם עלינו לעיתים כן נקבע נקודה שעלתה כמינימום אבל ככל ש"הטמפרטורה" תרד כך נקפוץ פחות מבורות מקומיים והנקודה תיעלם עם הזמן. בסדרת מחקרים בחנו איך עובד הרעיון והאם האלגוריתם באמת מבטיח מינימום גלובאלי. ברוב המקרים זה יביא למינימום הגלובאלי או לכל הפחות במינימום הדו-מימדי.*



*בשיטה זו נהוג להשתמש בקריטריון הסכמה ובלוח זמנים לקירור. אם הקריטריון קובע לקבל רק אם מתקיים בנוסף כאשר המספר הרנדומלי נבחר כך ש- .*



*לשיטת ה-simulated annealing יש בעיה פרקטית חמורה- פונקציית המחיר חייבת להיות מפותחת בזהירות, היא חייבת להיות "פרקטלית וחלקה". לכן השיטה תעבוד למשל על הפונקציה הנתונה למטה משמאל, אך לא תפעל על הפונקציה הנתונה למטה מימין. אגן הניקוז הרחב למינימום לוקאלי בפונקציה מימין יביא להגעה למינימום שהוא לא הגלובאלי, ולכן יפספס את המינימום הגלובאלי עם האגן הצר. נציין כי רוב המערכות הביולוגיות יותר מתנהגות כמו הפונקציה משמאל, ולכן מעשית זהו כלי שאפשר להשתמש בו.*



*איך קובעים את הפרמטרים ב-Simulated Annealing? הטמפרטורה ההתחלתית היא לרוב מייצגת 50% ממספר הצעדים הרעים שמוכנים לקבל. בתהליך הקירור מורידים את הטמפרטורה ב-10% בכל צעד. מספר האיטרציות עבור כל טמפרטורה נע בין 1 ל-10 פעמים עבור כל "רכיב" של המצב. הטמפרטורה הסופית לרוב צריכה להיות ללא כל "מצב רע", ושלב זה נקרא לרוב שלב ה-quenching.*

***אלגוריתמים גנטיים***

*המרכיבים של אלגוריתמים גנטיים הם רבים. הם גם משמשים ככלי לפתרון בעיות וגם כשיטה לקידוד (גן, כרומוזום), גם כתהליך אתחול (יצירה), פונקציית שערוך (סביבה), בחירה של הורים (רפרודוקציה), אופרטורים גנטיים (מוטציה, רקומבינציה) והגדרות פרמטריות (אימון). אלגוריתם גנטי כללי ייראה כך:*

*{ אתחול של האוכלוסיה;*

*שערוך של האוכלוסיה;*

*לולאת while (כל עוד תנאי הסיום לא מתקיים)*

*{בחירה של הורים לרפרודוקציה;*

*ביצוע רקומבינציה ומוטציה;*

*שערוך של האוכלוסיה;}*

*}*

*דוגמא מופשטת:*

