



Diagramm einlesen

gedruckte Diagramme analysieren

R. Sydow, Niederfinow (Deutschland)
(2022)

abstrakt: es wird ein Programm erarbeitet, mit welchem eine grafisch dargestellte Funktion in eine Wertetabelle überführt werden kann.



Inhaltsverzeichnis

Inhaltsverzeichnis	2
Abkürzungen	2
Einleitung	3
Programmbeschreibung	3
die Oberfläche	3
die Farbkorrektur	4
Achsenanalyse	4
Aufnahme der Punkte	5
Abschluss der Prozedur	6
Durchführung von Beispielen	7
erstes Beispiel	7
zweites Beispiel	8
drittes Beispiel	8
Anlagen	9
Anl. 1: auszuwertende Diagramme	9
Anl. 2: Programmablaufplan	10
Anl. 3: Daten	11
Literatur	14

Abkürzungen

GPA Gravity-Probe-A Experiment
PAP Programmablaufplan
s/w schwarz/weiß

Einleitung

Wozu das Ganze? Wozu muss man Diagramme einlesen? Gibt es beim heutigen Stand der Technik nicht schon immer die Funktionen als Gleichung und Grafik, sodass sich Wertetabellen berechnen lassen?

Das mag schon sein. Hat man es aber mit alten Quellen zu tun, kann man von solchen Voraussetzungen nicht mehr ausgehen. Insbesondere dann, sollte es sich um Versuchsbeschreibungen in alten wissenschaftlichen Veröffentlichungen handeln, kann man davon ausgehen, dass dort lediglich die Zusammenfassung von ermittelten Daten in einem Diagramm zu finden ist. Der Zugriff auf die Ursprungsdaten ist oftmals nicht möglich.

Das Unterfangen scheint also solange als nutzlos oder widersinnig, bis man genau das folgende Problem hat: Man möchte die Funktion, die in einem Diagramm als Graph dargestellt wurde, in eine Wertetabelle umwandeln, um dann mit den ermittelten Werten mittels Regressionsanalyse eine passende Gleichung zu finden.

Programmbeschreibung

die Oberfläche

Die Lösung von allem: man schreibt sich ein kleines Programm:

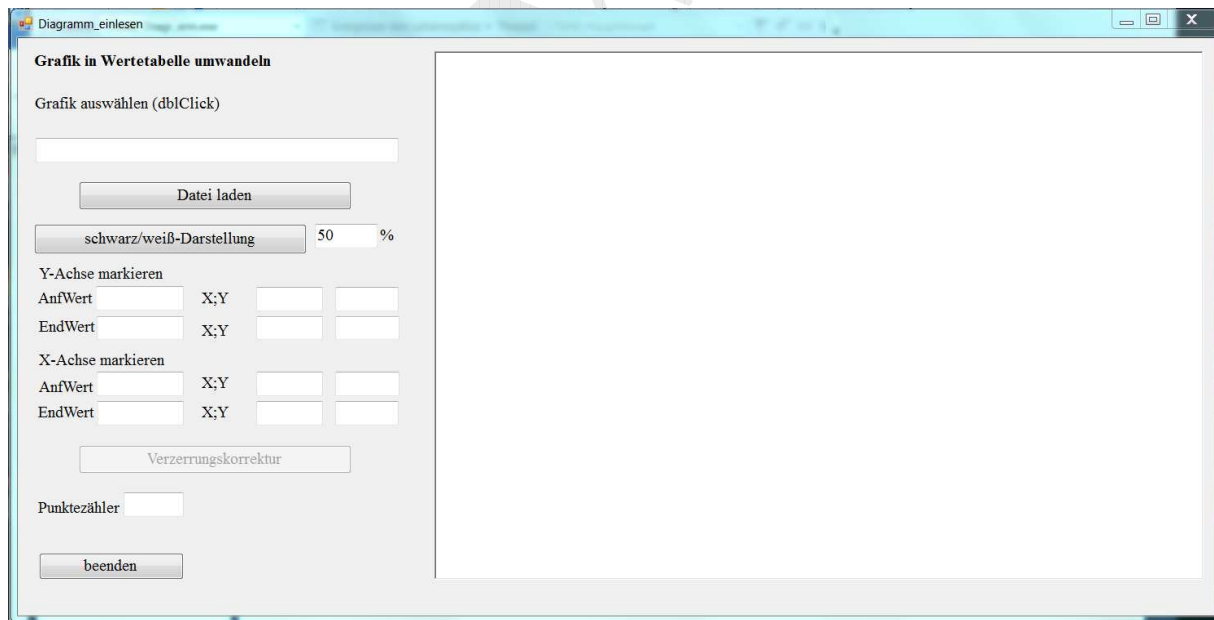


Bild 1: Oberfläche des Programms zum Einlesen von Werten aus einem gedruckten Diagramm

Das Programm basiert darauf, in einem Bild nach Farbübergängen von weiß zu schwarz und wieder zu weiß zu suchen. Damit ist es eine Voraussetzung, dass ein schwarz-weiß-Bild

Quellenangabe: Sydow, R. Diagramm einlesen, gedruckte Diagramme analysieren Niederfinow (Deutschland) 04.06.2022
<https://rolfswelt.de/software/#diagramm-analyse>

Revision: 1.1.0.2 vom 16.07.2023

copyright ©: alle Rechte vorbehalten, 2022, Rolf Sydow

vorhanden ist, in welchem nach Möglichkeit nur eine Linie ist, deren Breite durch die Detektion der beiden Übergänge ermittelt werden kann. Durch Doppelklick in der oberen Textbox wird ein File-Open-Fenster geöffnet. Damit sind Bilddateien auswählbar. Mit dem Schließen der Box wird der vollständige Pfad der gewünschten Datei in die Textbox übernommen.

Nun kann diese Datei durch Drücken des „Datei laden“-Buttons in den Bildbereich übernommen werden. Zu beachten ist, dass der Pfad für die Bilddatei korrekt sein muss, da es sonst zu einer Fehlermeldung kommt.

die Farbkorrektur

Da nicht grundsätzlich davon auszugehen ist, dass Bilder im schwarz-weiß-Modus vorliegen, bietet der „schwarz/weiß-Darstellung“-Button die Möglichkeit, Bilder in den entsprechenden Modus umzurechnen. Dabei werden die Farben in Grautöne umgerechnet. Über die Eingabe in % kann die Entscheidung, ob der Grauton dann zu weiß oder schwarz wird, beeinflusst werden.

Im Ergebnis sollte dann ein Bild vorliegen, in welchem der Hintergrund weiß ist und die zu analysierende Funktion im Diagramm schwarz dargestellt wird.

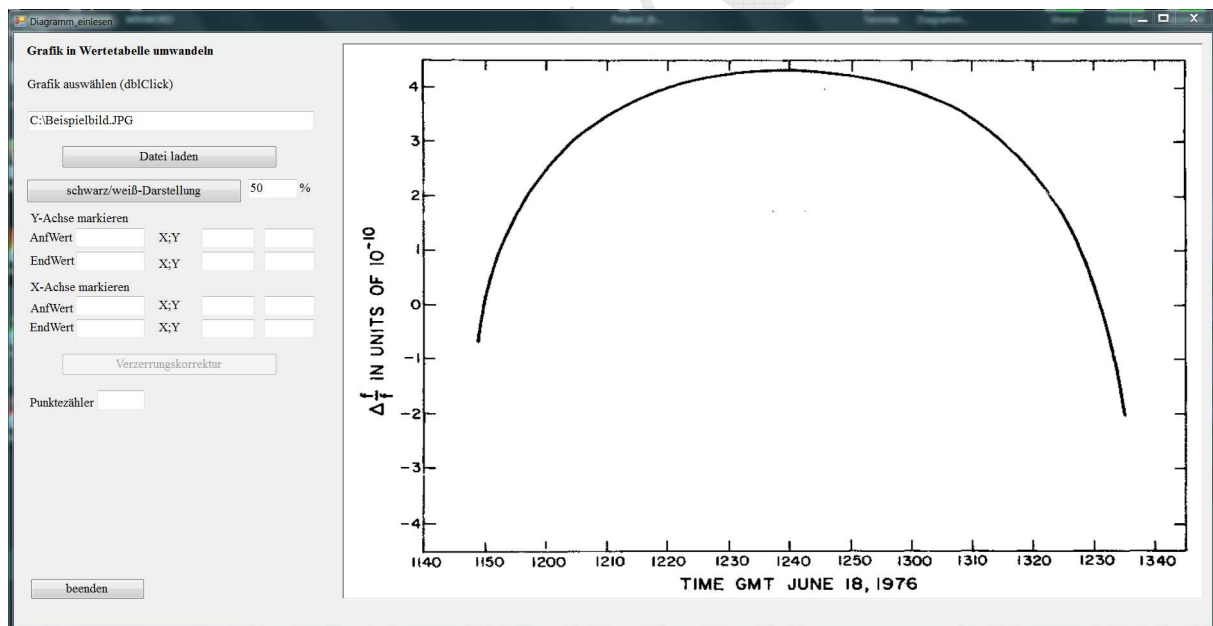


Bild 2: Ausführung des Programms bis zur Vorbereitung des Bildes (Beispiel [Ves1] S. 41)

Achsenanalyse

Im zweiten Schritt der Bearbeitung des Bildes sind die Diagrammachsen einzugeben. Beginnend mit der Ordinate ist der Anfangswert der y-Achse im dafür vorgesehenen Feld zu notieren. Um die höchste Genauigkeit zu erreichen, sollte dieser Punkt soweit es geht unten

liegen. Hier wäre der Wert -4 sinnvoll einzutragen. Dann ist mittels Mausklick im Bild der Punkt (0; -4) anzuklicken. Die Koordinaten des Bildes werden automatisch übernommen. In dieser gleichen Weise sind die weiteren Werte zur Bestimmung der Diagrammachsen aufzunehmen. Dabei ist darauf zu achten, dass jeweils der Anfangs- und Endwert der Achsen soweit wie möglich auseinander liegen.

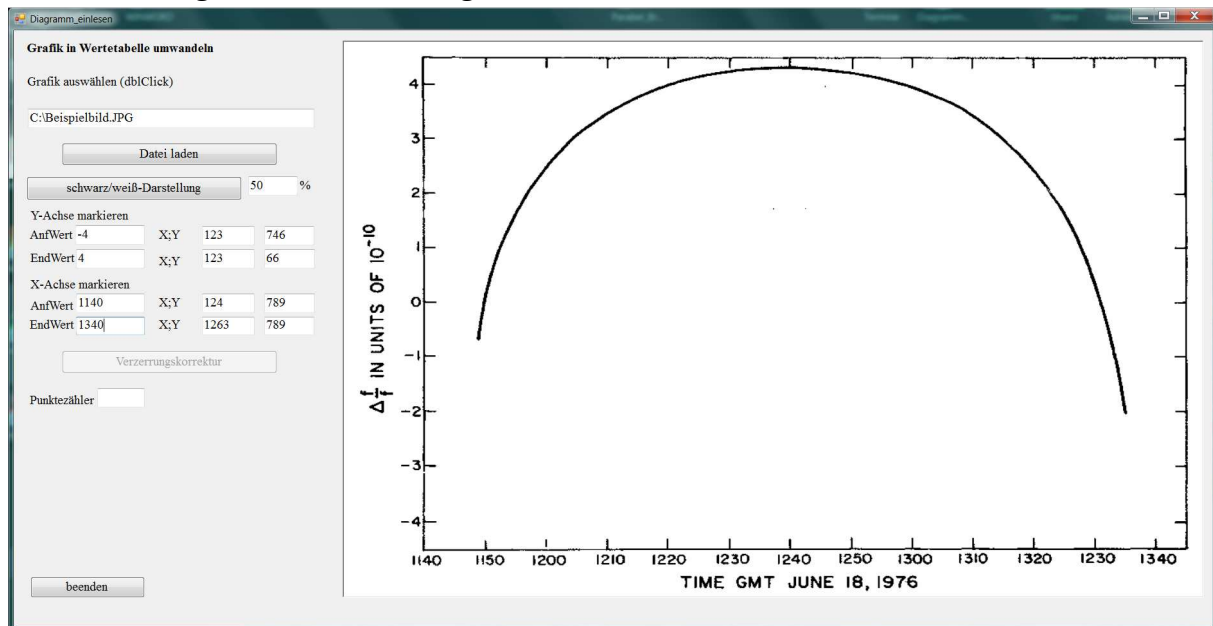


Bild 3: Aufnahme der Diagrammachsen mittels Programm

Die Aufnahme der Achsen dient dazu, eventuelle Verdrehungen des Diagrammes im Bild kompensieren zu können.

Sollte das Diagramm verzerrt erscheinen, kann ggf. die Option „Verzerrungskorrektur“ benutzt werden. Diese ist bislang nicht nötig gewesen und programmtechnisch noch nicht hinterlegt. In späteren Versionen kann diese Korrektur noch nachprogrammiert werden.

Aufnahme der Punkte

Um nun zu den Punkten der im Diagramm dargestellten Kurve zu kommen, ist es lediglich nötig, unterhalb der Kurve an der Stelle anzuklicken, wo ein Wertepaar gewünscht wird. Sinnvoll ist es die Wertepaare in aufsteigender Reihenfolge der Abszisse abzufragen.

Im Feld „Punkteähler“ wird die Anzahl der bereits aufgenommenen Wertepaare registriert. Das Wertepaar wird gefunden, indem das Programm von der Stelle des Anklickens senkrecht nach oben geht und alle Pixel analysiert, um die entsprechenden Farbänderungen zu suchen. Wird ein Wechsel von weiß nach schwarz festgestellt, sucht das Programm weiter, bis ein Wechsel von schwarz zu weiß festgestellt wird. Aus dem Mittelwert dieser beiden Positionen wird dann das Wertepaar ermittelt.

Dass ein Wertepaar gefunden wurde, zeigt sich auch in einem roten Kreis an der Stelle der Kurve, wo das registrierte Wertepaar gefunden wurde.

Automatisch werden diese Wertepaare über die eingegebenen Diagrammachsen korrigiert. Bei einer verdrehten Diagrammposition entsteht dann der blaue Kreis an der korrigierten Stelle.

Sollte die Verdrehung des Diagramms vernachlässigbar sein, werden die roten Kreise durch die korrigierten, blauen Kreise überschrieben, sodass die roten Kreise dann nicht mehr zusehen sind.

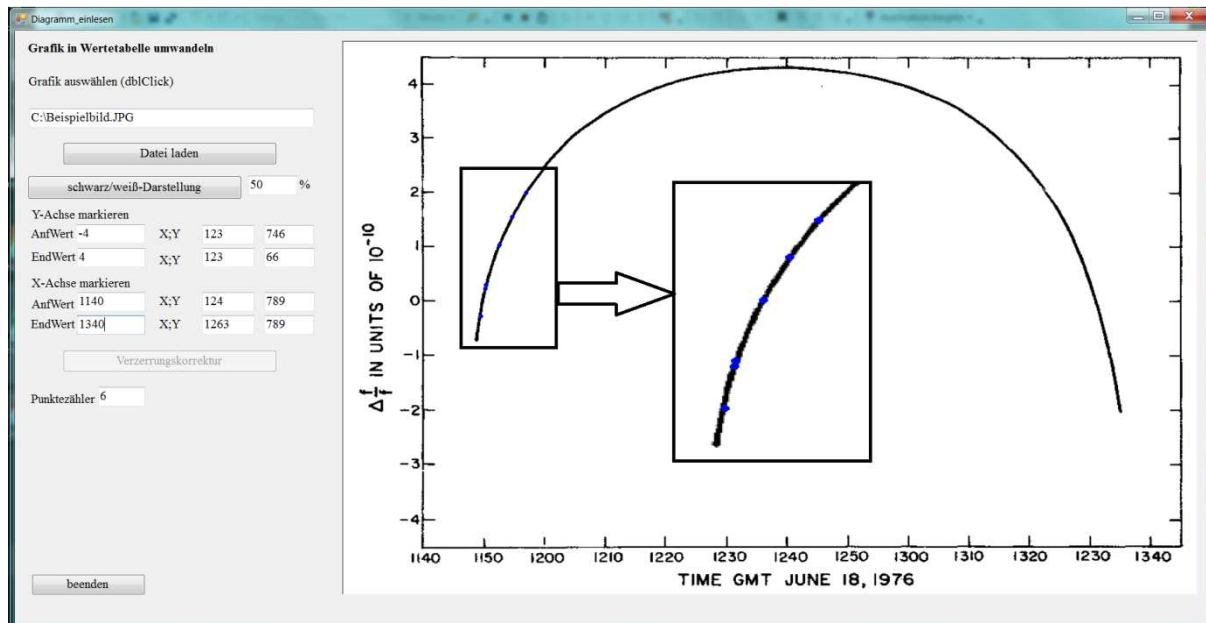


Bild 4: Aufnahme der Koordinaten der Kurve

Bei der Aufnahme der Wertepaare ist darauf zu achten, dass nur an solchen Stellen Werte der Kurve gefunden werden können, wo oberhalb der angeklickten Stelle ausschließlich die zu analysierende Kurve kommt und diese dort nicht gerade eine Unterbrechung hat.

Das Anklicken unterhalb von Gitternetzlinien oder Artefakten im Diagramm führt zu Fehlern bei der Wertfindung.

Die Markierung der gefundenen Wertepaare ist eine Kontrolle der Richtigkeit der durchgeführten Analyse.

Abschluss der Prozedur

Die gewünschten Wertepaare sind nun vom Programm registriert worden. Dabei werden lediglich die korrigierten Wertepaare gespeichert. Damit kann das Programm geschossen werden.

Die aufgenommenen Daten finden sich in einer Text-Datei mit demselben Namen des Bildes und im dessen Pfad. In dem exerzierten Beispiel (Bild 2 bis 4) würden die Daten in der Datei: „C:\Beispielbild.txt“ zu finden sein.

Zur Vorbereitung der Weiterverarbeitung in Excel sind die in der Text-Datei hinterlegten Werte als Kommazahlen abgelegt. Einer Weiterverarbeitung steht damit nichts im Wege.



Dieser Programmablauf ist in der Anlage 2 als PAP dargestellt.

Durchführung von Beispielen

Zur Kontrolle des Programms sollen an dieser Stelle einige Beispiele durchgerechnet werden.

erstes Beispiel

In Anlage 1 Bild 1 a ist die Auswertung von Versuchsdaten des GPA ([Ves1] S. 86 Bild 28) dargestellt. In der Darstellung wird eine Funktion gezeigt, die eine Störstelle bei etwa 13:25 Uhr aufweist. Damit ist die Funktion nicht mehr stetig und kann nicht sinnvoll in eine Gleichung überführt werden.

Hier soll der Versuch unternommen werden, die Funktion mithilfe des oben beschriebenen Programms doch in eine Gleichung zu überführen. Doch bevor das Programm angewendet werden kann, muss die Störstelle interpretiert und beseitigt werden.

Die Rechtfertigung, die Störstelle zwischen 13:23 Uhr bis 13:25 Uhr zu eliminieren und das verbleibende Reststück der Funktion anzuheben, ist die berechnete Annahme, dass es sich bei dem im Versuch registrierten Prozess um einen kontinuierlichen Prozess handeln muss. Es gibt keinen physikalischen Grund, dass die aufgenommene Funktion unstetig sein muss.

Insofern wurden die Änderungen mittels Bildbearbeitungsprogramm vorgenommen und im Bild A1.1b dargestellt.

Das Bild A1.1c ist das Ergebnis der Anwendung oben beschriebenen Programms. Es wurde eine möglichst große Anzahl von Punkten aufgenommen, um damit ein statistisch gutes Ergebnis der Auswertung zu erzielen.

Im Bereich um das Apogee (um 12:40 Uhr) wurden weniger Punkte aufgenommen, weil dort die Gitterlinie den Funktionsverlauf verfälschen konnte. Auch im Bereich um 13:25 Uhr wurden keine Werte aufgenommen, da dort durch das Eliminieren der Störung eine freie Stelle in der Funktion entstanden ist.

Die vom Programm registrierten Wertepaare sind in der Anlage 3 Spalte 1 und 2 zu sehen. Zu beachten ist, dass bei der Eingabe der Daten für das Programm die x-Achse nicht in der Uhrzeit angegeben wurde, da sonst die Minutenumrechnung in Dezimalschreibweise hätte durchgeführt werden müssen. Hier wurde der Einfachheit halber die x-Achse direkt in Minuten eingegeben. Das bedeutet, dass sie um 11:40 Uhr mit 0 Minuten beginnt und um 13:40 Uhr mit 120 Minuten endet.

Damit sind die gefundenen Werte in Excel zu übertragen und mittels Regressionsanalyse ist der Trend zu ermitteln:

$$y = -0,000001x^3 + 0,0004x^2 - 0,0072x - 0,6581$$

Über die Hintergründe dieses formelmäßigen Zusammenhanges zwischen der Flugzeit und den ermittelten Werten wird hier nicht weiter diskutiert. Wesentlich ist nur, dass der gefundene Zusammenhang mit einem Korrelationskoeffizienten von 99.98 % gefunden wurde.



zweites Beispiel

Im Bild A1.2a (siehe [Ves1] S. 71 Bild 24) ist eine Kurve gezeigt, die die Rotverschiebung der während des GPA aufgenommenen Signale über die Zeit zeigt.

Da es hier keine Probleme mit der Kurve gibt, kann das oben beschriebene Programm direkt angewendet werden.

Die sich ergebenden Daten sind der Anlage 3 Spalte 3 und 4 zu entnehmen.

Schön dargestellt ist im Bild A1.2a, dass das Diagramm etwas verdreht erscheint. Deshalb sind die roten Punkt (hier wegen des s/w-Bildes schwarz dargestellt) von der zu analysierenden Kurve abweichend. Durch die Korrektur der Verdrehung werden sie aber als weiße Punkte wieder auf die Kurve projiziert.

drittes Beispiel

Ein letztes Beispiel sei mit dem Bild A1.2b durchgeführt. Hier gibt es keinerlei Probleme und die ermittelten Punkte liegen direkt auf der Kurve.

Die ermittelten Daten sind in der Anlage 3 Spalte 5 und 6 abgelegt und stehen für weitere Auswertungen zur Verfügung.

Anlagen

Anl. 1: auszuwertende Diagramme

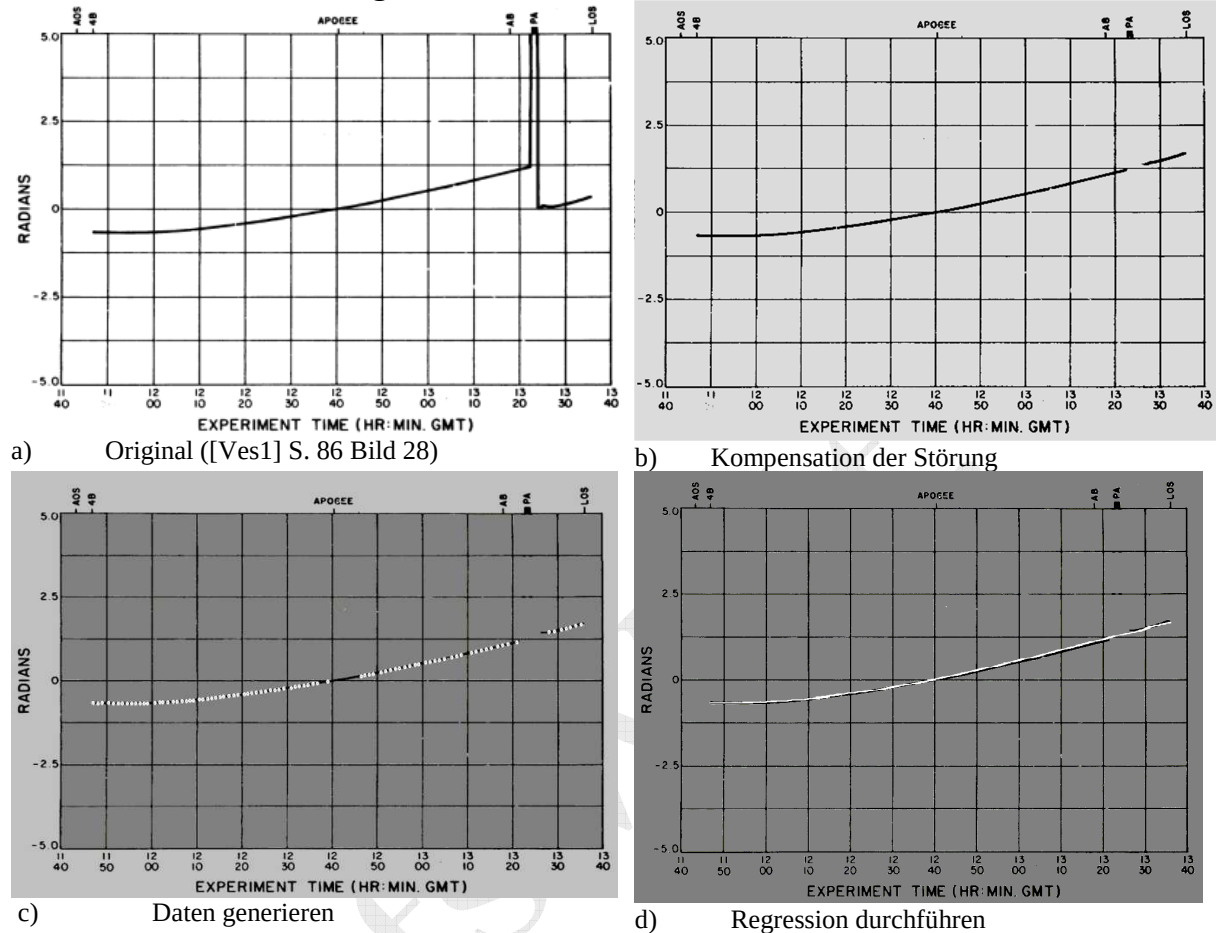


Bild A1.1: Bilderserie zur Auswertung

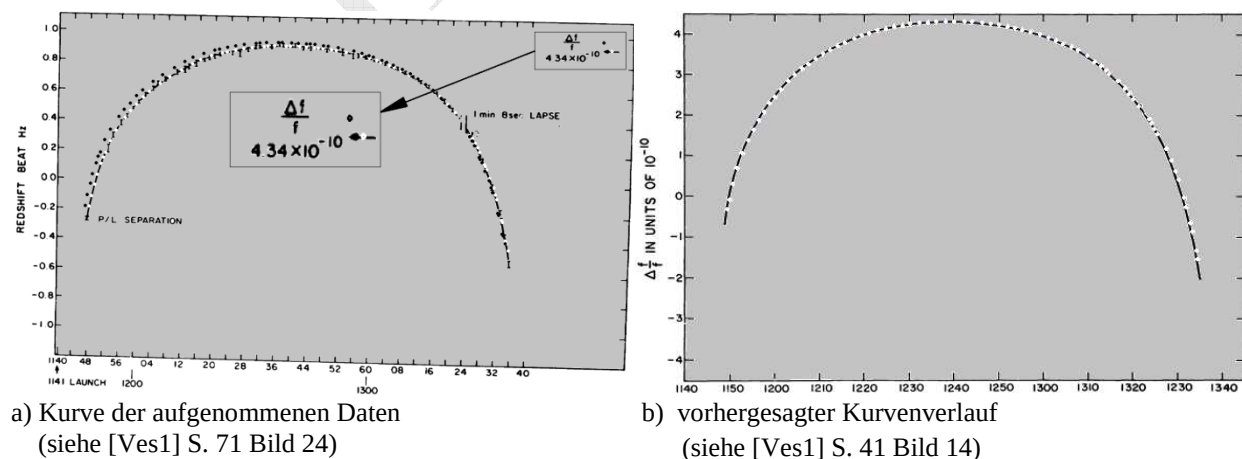


Bild A1.2: Auswertung der Daten aus dem Versuch

Quellenangabe:

Sydow, R. Diagramm einlesen, gedruckte Diagramme analysieren Niederfinow (Deutschland) 04.06.2022
<https://rolfswelt.de/software/#diagramm-analyse>

Revision:

1.1.0.2 vom 16.07.2023

copyright ©:

alle Rechte vorbehalten, 2022, Rolf Sydow

Anl. 2: Programmablaufplan

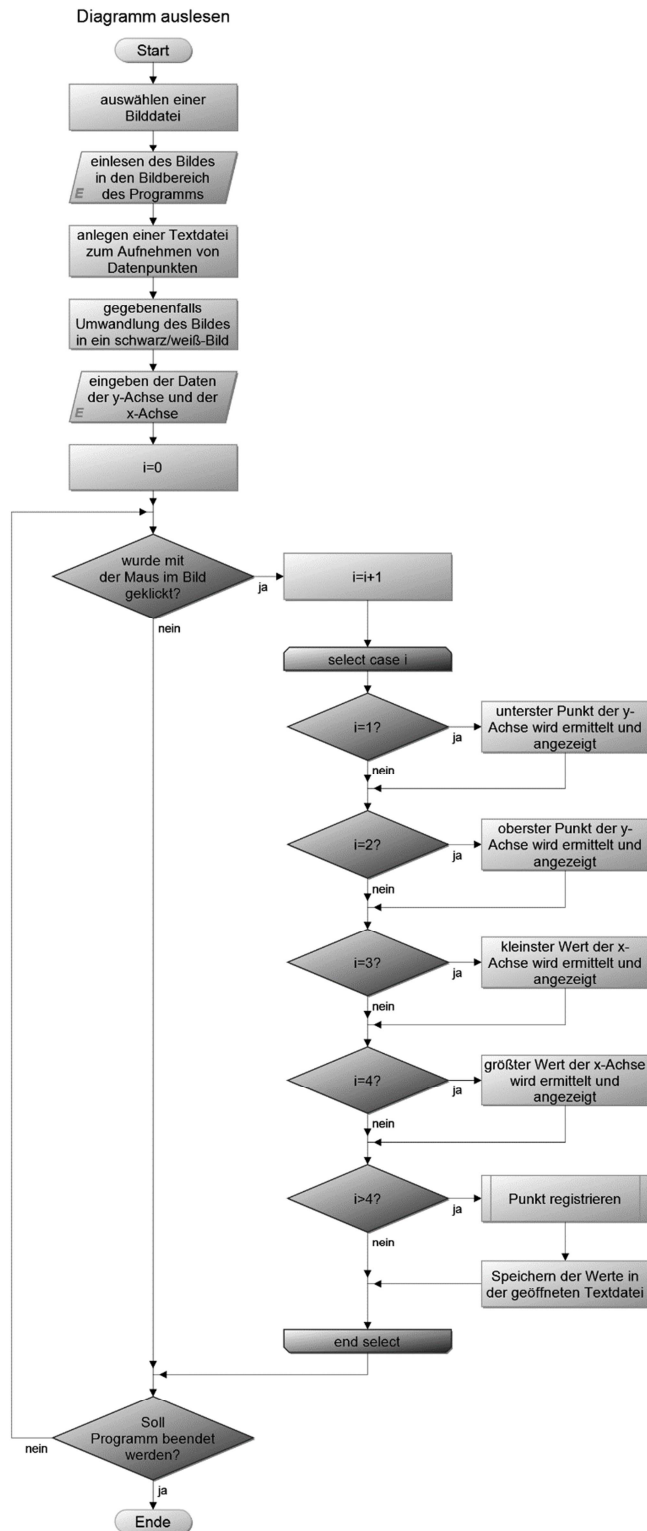


Bild A2.1: Programmablauf Hauptprogramm

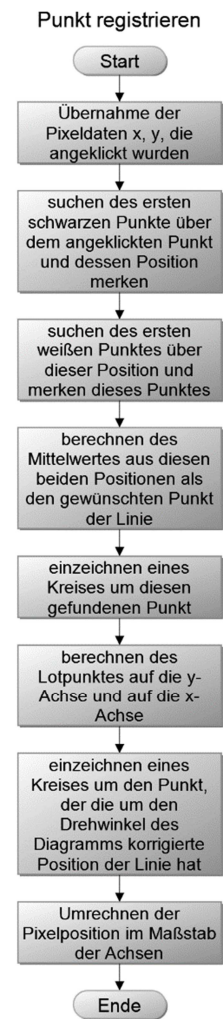


Bild A2.2: Unterprogramm

Anl. 3: Daten

	Spalte 1	Spalte 2	Spalte 3	Spalte 4	Spalte 5	Spalte 6
lfd. Nr.	Flugzeit [min]	Residuals	Flugzeit [min]	Rot- verschiebung	Flugzeit [min]	Frequenz- verschiebung
1	7,32907931	-0,67532432	8,00712063	-0,19740055	9,75813206	-0,32838448
2	8,42297174	-0,68386623	8,54513459	-0,12600173	10,3320886	-0,08483333
3	9,29808569	-0,67532432	9,32020119	-0,04880936	10,9087214	0,31548637
4	10,7201459	-0,67532432	9,86504235	0,01547509	11,9409064	0,69900945
5	11,8140383	-0,68386623	10,764097	0,08987273	13,0866209	1,05733758
6	12,9079307	-0,69240815	11,3280544	0,13423695	14,4598247	1,39047076
7	13,7830447	-0,68386623	11,9056661	0,16437242	15,6036753	1,63962078
8	14,8769371	-0,68386623	13,040408	0,24598648	16,7472869	1,87477362
9	15,8614403	-0,68386623	14,4285878	0,3163196	17,8902773	2,07353375
10	16,7365542	-0,68386623	16,6505482	0,40266149	19,2615215	2,29188995
11	17,7210574	-0,68386623	17,9393178	0,4501776	20,4038905	2,45425739
12	18,5961714	-0,68386623	19,4746982	0,49352713	21,7745135	2,63622089
13	19,3618961	-0,68386623	21,125874	0,54261918	23,3733427	2,83498102
14	21,0027347	-0,67532432	22,6667162	0,58027721	24,7434399	2,9861507
15	22,0966272	-0,65824048	25,4010141	0,63836572	26,4555119	3,14291926
16	22,7529626	-0,65824048	27,7838295	0,68349529	27,9389953	3,26049567
17	23,8468551	-0,64969857	30,7715654	0,73030285	29,4223831	3,37247321
18	25,0501367	-0,64115665	33,5208832	0,77273974	31,2479847	3,50404682
19	25,9252507	-0,64115665	35,5563123	0,80064175	32,7309902	3,59362885
20	27,0191431	-0,62407282	38,3206499	0,82742702	34,4420585	3,6916092
21	28,0036463	-0,6155309	41,0904494	0,8485208	36,2670865	3,78958955
22	28,769371	-0,59844707	42,2961942	0,85614541	38,0918755	3,8735727
23	29,5350957	-0,59844707	44,4679003	0,86844683	39,8024658	3,94355866
24	30,8477666	-0,57282132	46,52108	0,87785147	41,5129128	4,00514631
25	31,6134913	-0,5642794	48,695517	0,88730714	43,4513748	4,07233284
26	32,7073838	-0,54719557	51,7214852	0,89427422	45,1615828	4,11992329
27	33,5824977	-0,53011174	52,688812	0,89752816	47,0996147	4,16191487
28	34,4576117	-0,52156982	54,3836819	0,90108824	48,9235434	4,19550813
29	35,4421149	-0,50448599	57,4178428	0,89951808	50,6334169	4,22350251
30	36,2078396	-0,48740215	58,3879004	0,89992627	52,5712576	4,25429634
31	37,4111212	-0,47031832	60,8157753	0,89810099	54,5088593	4,27109297
32	38,5050137	-0,45323449	62,516107	0,89596957	56,3324057	4,28229072
33	39,3801276	-0,43615065	64,5829411	0,89114547	58,3838715	4,29348848
34	40,8021878	-0,41906682	65,5543642	0,89013079	60,4350506	4,2878896
35	41,5679125	-0,40198299	66,7696671	0,88779527	62,7141969	4,28509016
36	42,4430264	-0,38489915	68,9604893	0,88017645	64,4234012	4,27389241
37	43,4275296	-0,36781532	70,1771577	0,87641807	66,5882054	4,24869746
38	44,5214221	-0,35073149	72,7344824	0,86610657	68,5250903	4,22350251
39	45,5059253	-0,33364766	76,6365599	0,84497333	70,4619751	4,19830757



40	46,5998177	-0,29947999	78,7143176	0,82876624	72,2847568	4,16471431
41	47,5843209	-0,29093807	80,5495609	0,81245711	73,8796191	4,13112104
42	48,4594348	-0,27385424	81,6504338	0,8029562	75,5882978	4,08912947
43	49,5533273	-0,23968657	83,8549104	0,78110864	77,5246569	4,0331407
44	51,1941659	-0,21406082	85,0811368	0,76739013	79,5748801	3,97155305
45	52,0692799	-0,18843507	86,7992192	0,74676135	81,2831286	3,90436653
46	52,9443938	-0,17135124	89,0091577	0,71922229	83,1051456	3,82598225
47	54,0382862	-0,15426741	91,1005698	0,68878646	84,6993389	3,75319685
48	54,9134002	-0,13718357	93,074821	0,65403098	86,5212125	3,66641425
49	56,0072926	-0,10301591	95,0518031	0,61642976	88,0013027	3,58523054
50	56,9917958	-0,08593208	96,9088935	0,57735464	89,9363236	3,45085749
51	59,2889699	-0,03468058	99,0235181	0,52272994	91,6434728	3,31928388
54	66,7274385	0,12761584	101,014155	0,47089998	93,1229417	3,20170746
55	67,4931632	0,15324159	105,547363	0,29772331	94,3745389	3,08693049
56	68,4776664	0,17032542	106,306311	0,26530334	95,1709186	3,00854621
57	69,3527803	0,19595117	106,941272	0,2356781	96,8773032	2,83218158
58	71,1030082	0,23866075	107,454975	0,20600184	98,4698237	2,66141583
59	72,0875114	0,27282842	108,746742	0,12398538	99,4932625	2,53264166
60	73,1814038	0,315538	109,001546	0,11128156	100,516654	2,40106805
61	74,2752963	0,34116375	109,387164	0,08866864	101,767199	2,22470343
62	75,1504102	0,3667895	110,169325	0,03205982	103,926699	1,88877081
63	76,3536919	0,39241525	111,322085	-0,03151031	104,380626	1,77679327
64	77,4475843	0,42658292	111,732281	-0,07973496	105,176671	1,67881292
65	78,650866	0,46075058	112,127458	-0,11230799	105,858183	1,54723931
66	79,4165907	0,47783442	112,127458	-0,11230799	107,56127	1,17771343
67	80,9480401	0,52908592	113,475209	-0,25266229	108,695896	0,88657183
68	81,9325433	0,54616975	113,730012	-0,26536611	109,603034	0,62062517
69	83,0264357	0,58887933	114,57908	-0,39169577	110,282968	0,39667009
70	84,0109389	0,60596317	114,704433	-0,39591337	111,529069	-0,04004231
71	85,1048314	0,63158892	114,957871	-0,40719432	112,095379	-0,24440133
72	86,74567	0,69138233	115,368068	-0,45541897	113,000366	-0,63632271
73	87,8395624	0,71700808			113,45219	-0,87147555
74	88,8240656	0,75117575			114,355886	-1,33898178
75	91,1212397	0,81951108			114,694563	-1,52654416
76	92,1057429	0,85367875	140,634418	0,89894837		
77	93,3090246	0,89638833				
78	94,4029171	0,92201408				
79	95,4968095	0,96472366				
80	96,9188696	1,00743324				
81	98,0127621	1,04160091				
82	99,2160438	1,07576858				
83	101,18505	1,12702007				

84	108,295351	1,34845733				
85	109,498633	1,36451613				
86	111,030082	1,41269254				
87	112,123975	1,45283955				
88	112,999088	1,48495715				
89	114,20237	1,52510416				
90	115,405652	1,56525117				

Tab. 3.1: ermittelte Daten der Diagramme nach dem Algorithmus (Anl. 2)

Die obenstehende Tabelle enthält Daten, die bei verschiedenen Auswertungen von Bildern ermittelt wurden. Im Folgenden sind die Daten der einzelnen Spalten beschrieben:

Spalte 1 und 2:

- Auswertung des Bildes Bild A1.1 (Quelle: [Ves1] S. 86 Bild 28)
- Beachte: Residuals sind ohne Maßeinheit

Spalte 3 und 4:

- Auswertung des Bildes Bild A1.2 a (Quelle: [Ves1] S. 71 Bild 24)
- Beachte: Rotverschiebung in Hz
- Beachte: das letzte Wertepaar ist dem Pfeil zur Umrechnung in die Frequenzverschiebung entnommen. Damit ist es möglich durch Multiplikation aller Rotverschiebungen in Hz mit dem Quotienten $4,34 \cdot 10^{-10} / 0,89894837$ auf die Äquivalenz mit der Spalte 6 zu schließen.

Spalte 5 und 6:

- Auswertung des Bildes Bild A1.2 b (Quelle: [Ves1] S. 41 Bild 14)
- Beachte: Frequenzverschiebung in Einheiten von 10^{-10}



Literatur

- [Ves1] Vessot, R. F. C.; Levine, M. W.: Gravitational Redshift Space-Probe Experiment, GP-A Project Final Report
Smithsonian Institution Astrophysical Observatory Cambridge (USA MA) (12/1979) read 24.01.2017
<https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19800011717.pdf>