

Im Versuch werden die dielektrischen Eigenschaften von zwei Gasen, sowie von einem Polymer untersucht. Die Abhängigkeiten der komplexen dielektrischen Funktion $\varepsilon(\omega)$ von Druck, Temperatur und Frequenz werden vermessen um Rückschlüsse auf die Polarisations-eigenschaften der untersuchten Systeme zu schließen. Ein Hauptaugenmerk liegt auf der systematischen Auswertung der Messdaten hin zu aussagekräftigen, physikalisch relevanten, Materialeigenschaften.

Messaufgaben

Mit Hilfe eines Drehkondensators und eines RLC-Meters werden folgende Messungen an zwei Gasen (Luft, CO_2) durchgeführt:

1. Bestimmung der Kapazität des Messkondensators in Vakuum, in Luft und in CO_2 . Machen Sie jeweils 10 Messungen zur Bestimmung des Fehlers. (Schaltkapazität von 4,5 pF abziehen)
2. Bestimmen Sie aus den Ergebnissen die Permittivität ε von Luft und Kohlendioxid bei Luftdruck und Raumtemperatur. Vergleichen Sie mit den Literaturwerten und diskutieren Sie evtl. Abweichungen. Beachten Sie Ihre Handkapazität, die das Ergebnis erheblich verfälschen kann.
3. Die Druckabhängigkeiten $\varepsilon(p)$ sind für Luft und CO_2 zu messen (je 10 Druckwerte $p=1 \dots 1000$ mbar in gemeinsamem Diagramm). Begründen Sie den gefundenen funktionellen Zusammenhang gaskinetisch. Welchen Einfluß kann die Luftfeuchtigkeit auf ε_{rel} haben?

Es ist desweiteren die dielektrische Funktion von einem unausgehärtetem Epoxidharz als Funktion der Frequenz (1 Hz bis ≈ 100 kHz; min. 5 Werte pro Frequenzdekade) und der Temperatur ($-10 \dots +10^\circ C$; $\Delta T = 5K$ - stimmen Sie dies mit dem Betreuer ab) zu bestimmen. Diskutieren Sie mit dem Betreuer des Versuchs verschiedene Auswertemöglichkeiten.

Alle Rohdaten dieses Versuchsteils sind mittels der PICOSCOPE-Software als *Tabulatorgetrenntes Text-Dokument* der *derzeitigen Wellenform* zu speichern, wie in Abb. III dargestellt. Verwenden Sie dabei zwingend eine Nomenklatur für die Datei in der Form `YYYYMMDD_10kHz_-10deg.txt`, welche eindeutig Messfrequenz f , Temperatur T und Messdatum kennzeichnet. Aus diesen Rohdaten bestimmen Sie in ihrer Auswertung auch die relevanten Messgrößen.

4. Bestimmen Sie für diese Messungen zunächst die Übertragungsfunktion des Digitizers mit und ohne Elektrometerverstärker im vollen Frequenzbereich. Ersetzen

Sie dafür den mit Epoxidharz gefüllten Messkondensator durch eine frequenzunabhängige Festkapazität und interpretieren Sie ihre Beobachtungen. Stellen Sie das Ergebnis in einem Bode-Diagramm dar.

5. Das Epoxidharz befindet sich in einem Luftdrehkondensator mit nominell $C_{Luft} = 530 pF$. Messen Sie zur Bestimmung von $\varepsilon(\omega, T) = \varepsilon' + i\varepsilon''$ die Amplitude und Phasenverschiebung über den Drehkondensator, welcher Teil eines kapazitiven Spannungsteilers ist (vgl. Aufbau Abb. I und Abb. II). Zur Messung verwenden Sie den Digitizer. Beachten Sie bei der Auswertung die Leitungskapazitäten der koaxialen Anschlusskabel. Die Bestimmung insbesondere der Phasenverschiebung ist mit dem Betreuer abzustimmen.
6. Stellen Sie die Ergebnisse am Epoxidharz als Doppeldiagramm dar (ε' und $\varepsilon'' = f(\omega)$ mit T als Parameter einzelner Kurven). Vergleichen Sie ihre Ergebnisse mit Literaturdaten und diskutieren Sie die Physik der dielektrischen Eigenschaften des Epoxidharzes.
7. Gewinnen Sie aus den gemessenen Daten von $\varepsilon(\omega)(T)$ die α -Relaxationszeiten τ_α aus einem DEBYE-Modell. Stellen Sie $\tau_\alpha(T)$ dar und vergleichen Sie ihre Ergebnisse mit Literaturdaten. Unter Umständen können auch die schnelleren β -Relaxationszeiten τ_β ermittelt werden. Stellen Sie sodann $\tau_\alpha(T)$ und $\tau_\beta(T)$ gemeinsam dar und diskutieren Sie die (unterschiedlichen) Temperaturabhängigkeiten.

Hinweise: Schließen Sie die Druckflasche nach Beendigung des Versuches und belüften Sie die Drehschieberpumpe vor dem Ausschalten.

Kontrollfragen

1. Wie misst man das Vakuum im Bereich von 1 mbar ... 1 bar? Durch welche Pumpenarten kann solch ein Vakuum realisiert werden?
2. Welche Arten der Polarisation kennen Sie? Welche Relaxations- und Resonanzprozesse sind damit verbunden?
3. Wie ist vereinfacht ein Epoxydharz aufgebaut?
4. Was ist ein Elektrometerverstärker? Welche Funktion hat er im 2. Versuchsteil?
5. Wie kann man die Phasenverschiebung zweier Messsignale bestimmen?
6. Überlegen Sie sich durch welche Messgrößen Sie mit Hilfe des Aufbaus in Abb. I die komplexe Größe $\varepsilon(\omega)$ des Epoxydharzes bestimmen können.

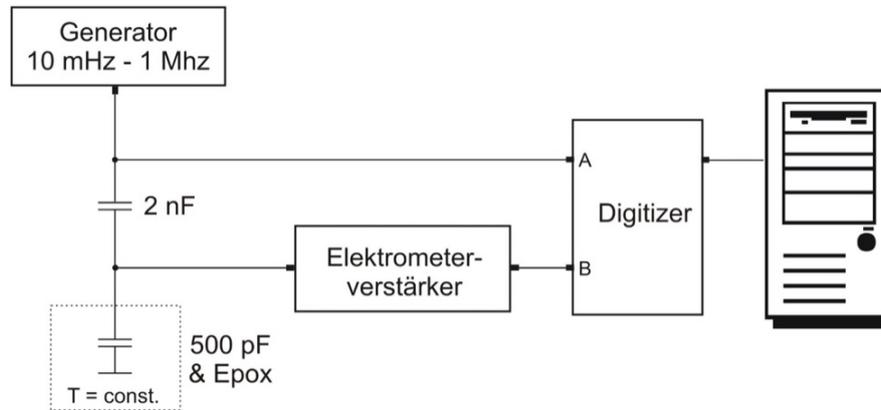


Abbildung I: Blockschaltbild für die Bestimmung der dielektrischen Funktion des Epoxydharzes.

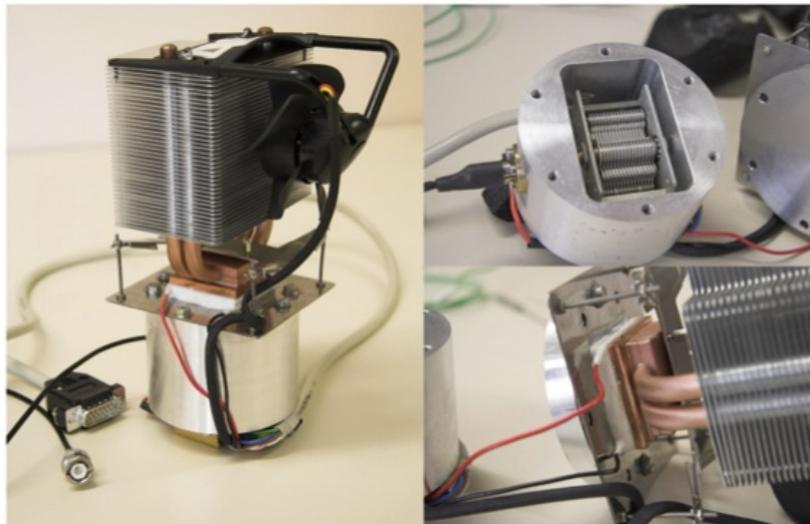


Abbildung II: Kryostat für die Bestimmung der Temperaturabhängigkeit der dielektrischen Funktion des Epoxydharzes. Ober rechts ist der (ungefüllte) Drehkondensator zu sehen. Der große Kühlkörper (links) führt die Wärme vom Peltierkühler ab (rechtes unteres Bild).

Speichern

 **Durchsuchen...**

Ausgewählter Ordnerpfad

Dateiname

Dateiart

Auswahl Wellenform

Speichern einer Datei mit ungefährender Größe: 3 MB

Abbildung III: Bildschirmaufnahme zur Speicherung der Messdaten mittels der PICSCOPE Software.

Literaturverzeichnis

- [1] P. Debye, "Polar Molecules", Dover Publications (1929)
- [2] H. Fröhlich, "Theory of Dielectrics", Oxford Clarendon Press (1949)
- [3] L. Boltzmann, "Experimentelle Bestimmung der Dielektrizitätskonstante einiger Gase", Wien. Ber. 69. S. 795—813. (1874)
- [4] R. Gross und A. Marx, "Dielektrische Eigenschaften in "Festkörperphysik", De Gruyter Oldenbourg (2014)
- [5] F. Kremer, A. Schönhals, "Broadband Dielectric Spectroscopy", Springer-Verlag (2002)
- [6] F. Kremer, A. Loidl, "The Scaling of Relaxation Processes", Springer-Verlag (2018)
- [7] S. Corezzi, M. Beiner, et al., J. of Chem. Physics Vol. 117, 2435-2448 (2002)
- [8] S.C. Santucci, S. Corezzi, et al., IEEE Transactions on Dielectrics and Electrical Insulation Vol.8, 3, (2001)
- [9] L. G. Hector and H. L. Schultz, Physics 7, 133 (1936)
- [10] C. T. Zahn, Phys. Rev. 27, 455 (1926)
- [11] A. Michels and C. Michels, Philosophical Transactions of the Royal Society of London. Vol. 231, 409-434 (1933)
- [12] K. Linkwitz, "Einige Bemerkungen zur Fehlerellipse und zum Fehlerellipsoid", Vermessung, Photogrammetrie, Kulturtechnik, 86 (1988)

Die Literatur wird Ihnen auf Nachfrage vom Betreuer auch digital zur Verfügung gestellt.

1 Elektrische Felder im Kondensator

Zwei leitende Platten stehen sich im Vakuum oder an Luft gegenüber. Dieser Kondensator wird an eine Spannungsquelle wie in Abb. 1 dargestellt angeschlossen. Ladungsgleich-

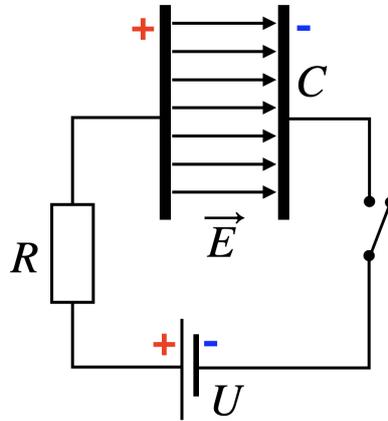


Abbildung 1: RC-Schaltkreis. Bei Anschluss an eine Spannungsquelle U fließen über den Widerstand R elektrische Ladungen auf die beiden Metallplatten des Kondensators C .

gewicht herrscht bei einer übertragenen Ladung Q . Der Ladungsstrom ist zeitabhängig. Solange ein Strom fließt, gilt nach der Maschenregel

$$U = U_R + U_C \text{ mit} \quad (1)$$

$$U = IR + \frac{Q}{C} \text{ bzw.} \quad (2)$$

$$U = R \frac{dQ}{dt} + \frac{Q}{C} \quad (3)$$

C ist die Kapazität des Kondensators. Wir differenzieren Gl. (3) nach der Zeit t

$$0 = R \frac{dI}{dt} + \frac{I}{C} \quad (4)$$

Umstellung ergibt

$$\frac{dI}{I} = -\frac{dt}{RC} \quad (5)$$

Nach Integration folgt

$$\ln \frac{I}{I_0} = -\frac{t}{RC} \text{ bzw.} \quad (6)$$

$$I = I_0 \cdot e^{-\frac{t}{RC}} \quad (7)$$

Für den Verlauf von $Q(t)$ setzen wir Gl. (7) in Gl. (2) ein

$$Q(t) = C U - C R I_0 \cdot e^{-\frac{t}{RC}}, \text{ wegen } Q(t=0) = 0 \text{ ist } C U = C R I_0. \quad (8)$$

Es folgt

$$Q(t) = C U (1 - e^{-\frac{t}{RC}}). \quad (9)$$

Durch den in Abb. 1 skizzierten und nach Gl. (9) verlaufenden Ladungsübergang wird der Raum zwischen den Platten verändert. Es entsteht ein **elektrisches Feld**, das durch Feldlinien beschrieben werden kann. Verlaufen die Feldlinien bei gleicher Liniendichte überall parallel, so ist das Feld homogen.

Als Indikator für ein elektrisches Feld dient eine Probeladung (z.B. einzelne Elektronen oder klassischer: geladene Öltröpfchen im Millikan-Experiment). Die Probeladung kann im elektrischen Feld beschleunigt werden. Es wirkt eine Feldkraft, bzw. ein Drehmoment, auf die Ladung. Die Größe der Kraft ist ein Maß für die Feldstärke. Da von elektrischen Ladungen elektrische Felder ausgehen, wir erinnern uns an $\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0}$, verändert die Probeladung den ursprünglichen Verlauf der ungestörten Feldlinien des auszumessenden Feldes: Bei metallischen Kondensatorplatten bewegen sich die Ladungen. In Metallen sind die Valenzelektronen schwach gebunden. Bei Isolatoren dagegen ist es schwer, eine homogene Ladungsverteilung und damit ein homogenes Feld zu erreichen.

Es folgen einige Gedankenexperimente zu möglichen Messungen. Man nehme eine möglichst kleine Probeladung auf einem entsprechend kleinen Probekörper. Im Idealfall ist dies eine Elementarladung $e = 1,6 \times 10^{-19} \text{As}$ mit der Elektronenmasse $m_0 = 9,11 \times 10^{-31} \text{kg}$. In einer Messung mit Probeladung, Torsionswaage und Spannungsmessgerät stellt man fest, dass die Feldstärke \vec{E} proportional U/l bzw. proportional Q/A ist (l : Plattenabstand; A : Plattenfläche). Die jeweiligen Proportionalitätsfaktoren werden natürlich nicht übereinstimmen. Formal lassen sich zwei Feldvektoren einführen: die **Elektrische Feldstärke**: $\vec{E} = \frac{U}{l} \vec{e}$ (\vec{e} : Einheitsvektor, orientiert von + nach -); und die **Dielektrische Verschiebung**: $\vec{D} = \frac{Q}{A} \vec{e}$. Für das Vakuum zwischen den Kondensatorplatten gilt für die Beträge der beiden Feldvektoren: $\vec{D} = \epsilon_0 \vec{E}$ mit der Permittivität des Vakuums, auch Influenzkonstanten, $\epsilon_0 = 8,854 \times 10^{-12} \text{As/Vm}$. Für den Plattenkondensator in Abb. 1 ist somit

$$\frac{U}{l} = \frac{Q}{\epsilon_0 A} \quad (10)$$

Ein Kondensator hat die Fähigkeit, Ladungen zu speichern. Diese Speicherfähigkeit heißt Kapazität C des Kondensators und ist definiert zu

$$C = \frac{Q}{U} = \epsilon_0 \frac{A}{l} \quad (11)$$

Öffnet man den Schalter in der Anordnung nach Abb. 1 zeigt ein Spannungsmessgerät die Spannung U an. Man schiebt nun einen Isolator (Dielektrikum) zwischen die Kondensatorplatten. Die Spannung am Messgerät wird geringer: $U_m = U/\varepsilon$. $\varepsilon = \varepsilon_0\varepsilon_r$ ist die Permittivität (veraltet und falsch: Dielektrizitätskonstante) des Dielektrikums, ε_r die materialspezifische relative Permittivität. In Diskussionen, wie auch z.T. hier, wird o.B.d.A. oft die Influenzkonstante vernachlässigt und $\varepsilon = \varepsilon_r$ normiert.

Die elektrische Feldstärke \vec{E}_m im Dielektrikum ändert sich gegenüber jener im Vakuum per Definition zu $\vec{E}_m = \vec{E}/\varepsilon$. Auf den Kondensatorplatten hat sich bei unserem Experiment die Ladungsdichte und damit die dielektrische Verschiebung im Zwischenraum des Kondensators nicht geändert. Es bleibt $\vec{D}_m = \vec{D}$.

Die Kapazität des Kondensators erhöht sich mit Dielektrikum um den Faktor ε_r zu

$$C_m = \frac{Q_m}{U_m} = \varepsilon_r \frac{Q}{U} = \varepsilon C \quad (12)$$

Die Relation

$$C_m = \varepsilon_r \varepsilon_0 \frac{A}{l} \quad (13)$$

nutzen wir in unserem Experiment zur Bestimmung der Materialkonstanten ε , indem wir die Kapazitäten C_m und C messen. Die Verknüpfung der Feldgrößen \vec{D}_m und \vec{E}_m ist gegeben im Vakuum durch

$$\vec{D} = \varepsilon_0 \vec{E}, \text{ in Materie durch} \quad (14)$$

$$\vec{D}_m = \varepsilon_r \varepsilon_0 \vec{E} \quad (15)$$

Ein anderes Experiment nach Abb. 1 führt bei geschlossenem Schalter zum gleichem Resultat für C_m . Nach dem Einbringen der Materie zwischen die Kondensatorplatten bleibt U natürlich erhalten. Es gilt $\vec{E}_m = \vec{E}$. Aber der Kondensator *zieht* aus der Spannungsquelle das ε -fache an Ladungen auf die Platten. Daraus ergibt sich $C_m = \varepsilon \frac{Q}{U} = \varepsilon C$. Das identische Resultat erhält man natürlich beim Experiment mit offenem Schalter. Dieser Einfluss der Materie auf die Feldgrößen begründet, warum man \vec{D} und \vec{E} zur Zustandsbeschreibung des Raumes verwendet. Im Vakuum würde offensichtlich ein Feldvektor ausreichen. Die Stärke des Verschiebungsfeldes \vec{D} wird bestimmt durch die beweglichen (freien) Ladungen auf den Kondensatorplatten. Die Stärke des elektrischen Feldes \vec{E} wird nicht durch sämtliche, für \vec{D} maßgebliche, Ladungen festgelegt. Der Anteil der Ladungen beträgt das $(1/\varepsilon)$ -fache.

2 Polarisation

Ursache für die Schwächung des elektrischen Feldes im gefüllten Kondensator ist die Polarisation des Dielektrikums. In Abb. 2a) ist die Veränderung der Ladungen und Felder im Plattenkondensator skizziert. Zur besseren Verdeutlichung der Polarisation sind zwei Luftspalte der Gesamtdicke $2d$ eingefügt. Wir fordern $2d \ll 1$ und vernachlässigen deshalb zunächst ihren Einfluss auf die Kapazität. Von den 5 Ladungen auf den Kondensatorplatten bleiben 2 übrig. Sie sind die Quellen für das elektrische Feld E_m . Die Permittivität ϵ beträgt also in unserem Beispiel 2,5. Die anderen Ladungen werden durch Polarisationsladungen in ihrer Feldwirkung eliminiert.

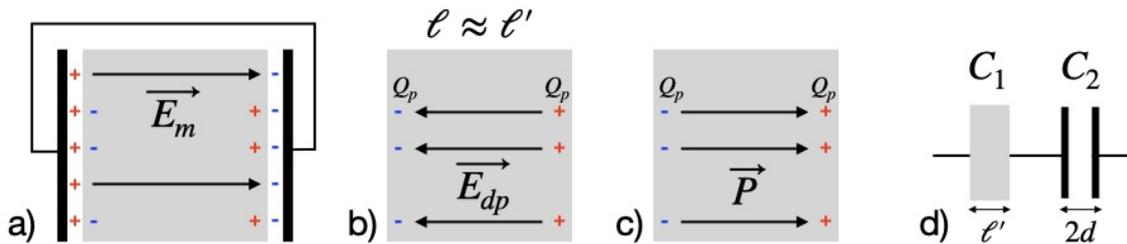


Abbildung 2: Schematische Darstellung eines Kondensators gefüllt mit einem bekannten Dielektrikum. Veranschaulichung zur Entstehung und Wirkung der a) elektrischen Feldstärke \vec{E}_m im Dielektrikum, b) dem Depolarisationsfeld \vec{E}_{dp} und c) der Polarisation \vec{P} . In d) ist die Reihenschaltung des Dielektrikums und zweier "Luftspalte" symbolisiert. Weiterende Erklärungen im Text.

Wie in Abb. 2b) dargestellt entsteht ein dem Erregerfeld entgegengesetzt gerichtetes Depolarisationsfeld \vec{E}_{dp} . Die für den Verschiebungsvektor \vec{D} maßgeblichen Plattenladungen bleiben unverändert, also bleibt es auch \vec{D} selbst. Den 3 Polarisationsladungen Q_P des Dielektrikums kann man über Q_p/A formal eine dielektrische Verschiebung zuordnen, die als Polarisation bezeichnet wird. Die Polarisation wird als Vektor \vec{P} , wie ein elektrischer Dipol, von - nach + orientiert, weil ihre Entstehungsursache, die Oberflächenladungen des Dielektrikums (siehe Abb. 2c)) Dipole bilden. Der Betrag der Polarisation P ist demnach

$$P = Q_P \frac{l}{A \cdot l} = \frac{Q_P \cdot l}{V} \quad (16)$$

Hierbei ist Gl. (16) das Verhältnis von Dipolmoment und Volumen, also eine Dipoldichte. Im Dielektrikum entsteht nach Abb. 2c) ein Polarisationsfeld. Vergleicht man die Darstellungen in Abb. 2a) und c), so sieht man, dass das elektrische Feld im Dielektrikum und der Polarisationsvektor die gleiche Richtung haben.

Aus den Darstellungen in den Abb. 2a)-c) entnimmt man, dass

$$\vec{P} = \vec{D}_m - \frac{1}{\varepsilon_r} \vec{D}_m \quad (17)$$

bzw.

$$\vec{P} = \varepsilon_0(\varepsilon_r - 1)\vec{E}_m \quad (18)$$

oder mit der elektronischen Suszeptibilität $\chi_e = 1 - \varepsilon_r$

$$\vec{P} = \varepsilon_0 \chi_e \vec{E}_m \quad (19)$$

gilt.

Das Depolarisationsfeld ist nun

$$\vec{E}_{dp} = -\frac{\vec{P}}{\varepsilon_0} . \quad (20)$$

Bisher wurden die Luftspalte zwischen den Platten und dem Dielektrikum vernachlässigt. Wir brauchen sie aber, um Ladungsübergänge zwischen Dielektrikum und Metall zu verhindern. Das Ersatzschaltbild für symmetrische Luftspalte, wie in Abb. 2d) dargestellt, ergibt als Reihenschaltung zweier Kondensatoren

$$\frac{1}{C_R} = \frac{1}{C_1} + \frac{1}{C_2} \quad (21)$$

Wegen $C_1 = \varepsilon\varepsilon_0\frac{A}{l}$ und $C_2 = \varepsilon_0\frac{A}{2d}$ ist

$$C_R = C_m \frac{1}{1 + (\varepsilon_r - 1)\frac{2d}{l}} \quad (22)$$

Der Bruch sollte möglichst nahe bei Eins liegen. Für einen realistischen Plattenkondensator mit $l = 3mm$ und einem Luftspalt von $d = 10^{-3}mm$ ist mit $\varepsilon_r = 2,5$ der Bruch = 1,001 - der Fehler unserer Annahme entspricht d/l , liegt also im Promille-Bereich.

3 Atomistische Beschreibung der Polarisation

Wir wollen nun zuerst den mikroskopischen Ursprung der im vorherigen Abschnitt elektrostatisch und phänomenologisch abgeleiteten Polarisation widmen. O.B.d.A. gehen wir dabei zuerst von einem kristallinen Festkörper aus und fokussieren uns später auf einzelnen Moleküle oder Polymere.

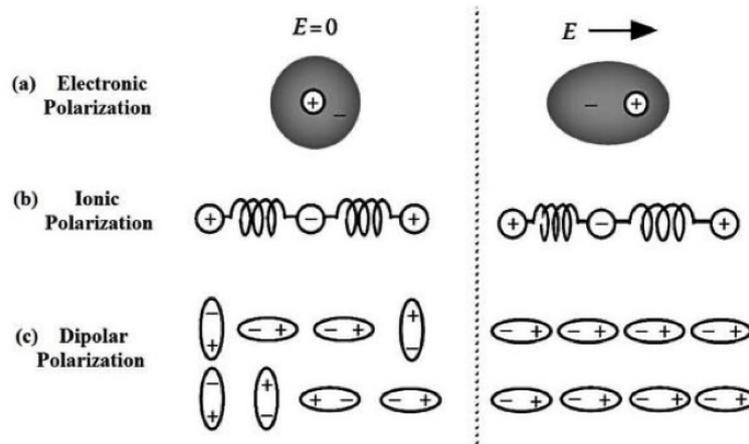


Abbildung 3: Verschiedene Beiträge zur Gesamtpolarisation in einem Festkörper bei angelegtem und nicht angelegtem Feld. a) elektronische Polarisation, b) ionische (Gitter-) Polarisation und c) Ausrichten der Dipolbeiträge im Feld als Beitrag zur Orientierungspolarisation. Bild adaptiert nach [DOI:10.5772/intechopen.97720](https://doi.org/10.5772/intechopen.97720).

Die Polarisation \vec{P} von Materialien im äußeren elektrischen Feld entsteht durch Verschiebung ihrer mikroskopischen *Bestandteile* (Atomrümpfe, Elektronen, Ionen) aus ihrer Gleichgewichtslage. Diese Gleichgewichtslage ist durch das Minimum der freien Enthalpie bestimmt. Die Verschiebung ändert sich zeitlich periodisch in einem elektrischen Wechselfeld $\vec{E} = \vec{E}_0 e^{-i\omega t}$. Wir unterscheiden aus atomistischer Sicht Elektronenpolarisation, Ionen- und Orientierungspolarisation wie in Abb. 5 dargestellt. Für die Elektronenpolarisation werden die Elektronendichten (vereinfacht und falsch "Bohrsche Elektronenbahnen") relativ zu den Atomrümpfen verschoben bzw. deformiert und führen zu einer Netto-Polarisation.

Die Ionenpolarisation, auch Gitterpolarisation, beruht auf der Veränderung der relativen Lage der Ionen gegenüber ihren Nachbarn. Ein Beispiel ist durch eine atomare Kette in Abb. 3b) gegeben. Das Gleichgewicht im äußeren Feld wird durch die Veränderungen der inneren Felder hergestellt um ein neues Energieminimum zu erreichen. Vor dem Einschalten des äußeren Feldes fallen die Schwerpunkte von positiver und negativer Ladung der Ionen zusammen, es gibt kein permanentes Dipolmoment.

Das entsteht aber z.B. bei der Orientierungspolarisation der stilisierten Dipole in Abb. 3c). Eine Vielzahl solcher Dipole orientieren sich statistisch zur resultierenden Polarisation Null. Ein äußeres Feld \vec{E} wird optimal abgeschwächt durch bevorzugte Orientierung der Dipole in Feldrichtung. Als konkurrierender Vorgang wirkt dabei die thermisch stimulierte Entropieproduktion. Bei höheren Temperaturen wirkt die thermische Unordnung der Dipole der feldinduzierten Ausrichtung entgegen. In der Materie überlagern sich i.A. mindestens die drei Arten der Polarisation.

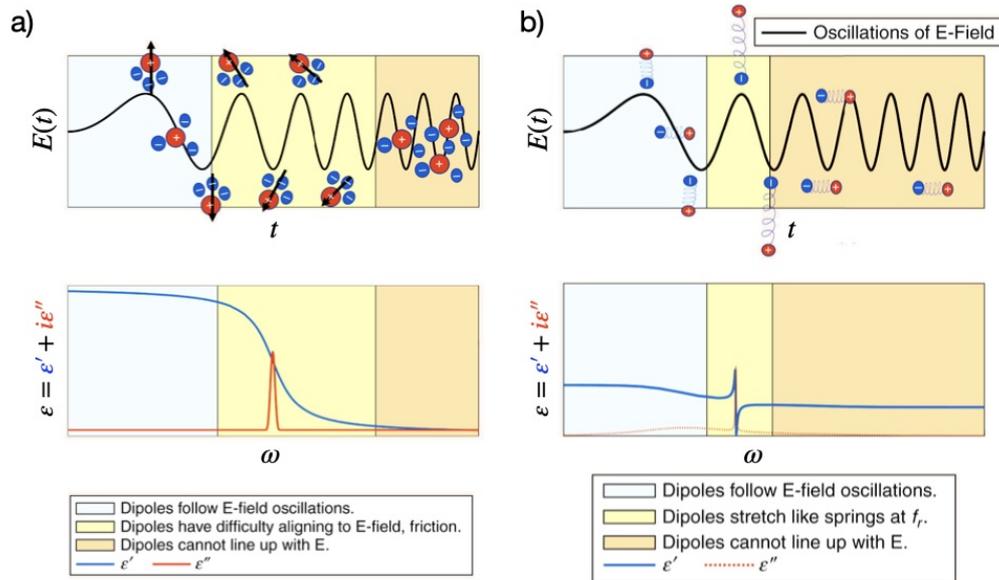


Abbildung 4: Veranschaulichung zu den Beiträgen a) der Dipol-Relaxation und b) der ionischen Resonanz. Ein elektrisches Feld $E(t)$ mit variierender Frequenz ω wirkt auf das dielektrische Material. Der Verlauf des Real- (blau) und Imaginärteils (rot) der dielektrischen Funktion ist gegeben. Bild adaptiert nach [DOI:10.1109/MMM.2022.3233510](https://doi.org/10.1109/MMM.2022.3233510).

Gehen wir nun von einem statischen äußeren elektrischen Feld hinzu einem explizit zeitabhängigen. Im Allgemeinen kann man sich sehr schnell verdeutlichen, dass schon alleine auf Grund ihrer deutlich unterschiedlichen Massen $m_e \ll m_i \ll m_{dp}$, die Polarisationsänderungen aus Abb. 3 bei deutlich unterschiedlichen Anregungsfrequenzen von $E(\omega, t)$ geschehen werden. Die mikroskopischen Polarisationsbeiträge können in elektrischen Wechselfeldern also separiert werden. Es handelt sich um **dielektrische Spektroskopie**.

In Abb. 4 ist die Anregung von Dipolen, bzw. von atomaren, ionischen Ketten dargestellt. Ein externes oszillierendes elektrisches Feld $E(\omega, t)$ bewirkt die Ausrichtung von Dipolmomenten im Material zum elektrischen Feld (Abb. 4a)). Je stärker diese Ausrichtung, desto größer das gesamte elektrische Feld und desto größer die Permittivität. Wenn die Ladungen aufgrund ihrer Massenträgheit den Schwingungen des elektrischen Feldes nicht mehr folgen können, sinkt ihr Beitrag, also der der Orientierungspolarisation, zum gesamten elektrischen Feld auf Null. Die Polarisation \vec{P} nach Gl. (18) wird kleiner, was die Abnahme von $\varepsilon(\omega)$ signalisiert. Dies zeigt sich im Spektrum als Abfall des Realteils der Permittivität und wird als dielektrische **Relaxation** bezeichnet. Wir gehen später noch detaillierter auf die Relaxationsprozesse ein. Eine Diskussion ist auch über die Phasenbeziehung möglich. Bei sehr niedrigen Frequenzen folgen die Ladungsbestandteile in Phase dem enliegenden Feld. Im Resonanzfall ergibt sich eine Phasenverschiebung

von 90° und maximale Energieaufnahme (Maximum von ε''). Bei sehr hohen Frequenzen können die Ladungen nicht mehr folgen und stehen quasi dem äußeren Feld immer entgegengerichtet mit einer Phasenverschiebung von 180° .

Leicht abgeändert kann man die Anregung von Gitterbindungen, wie in Abb. 4b) dargestellt, diskutieren. Die physikalischen Vorgänge, die den Beiträgen bei sehr hohen Frequenzen (THz-Bereich) des Spektrums zugrunde liegen, sind recht komplex. Wir versuchen jedoch, die grundlegende Idee zu beschreiben: Wenn das externe elektrische Feld mit sehr hohen Frequenzen schwingt, können sich nur noch wenige, sehr kleine Strukturen im elektrischen Feld ausrichten und so zur Polarisation beitragen. Diese kleinen Strukturen sind die Elektronen in der Elektronendichte, die ihren positiven Kern umgibt, und Atome, die in einem Molekül oder Gitterstruktur aneinander gebunden sind. Die dielektrische Spektroskopie geht im THz-Bereich nahtlos in die IR- und optische Spektroskopie über, wodurch sich der Übergang zu den charakteristischen Fingerprint-Spektren im Infrarotbereich nachvollziehen lässt, die eine eindeutige Identifikation molekularer und atomarer Strukturen ermöglichen.

Das elektrische Feld dreht sie nicht nur, sondern zieht auch die positiven und negativen Teile auseinander, woraufhin eine Rückstellkraft wirkt, die sie wieder zusammenführt; man kann die Strukturen dann vereinfacht wie Federn betrachten (siehe Abb. 3b) und Abb. 4b)). Bei ihrer Resonanzfrequenz erreicht der Abstand zwischen den positiven und negativen Teilen sein Maximum, zunächst in Richtung des externen elektrischen Felds (lokales Permittivitätsmaximum), dann, aufgrund der Verzögerung zum elektrischen Feld, in die entgegengesetzte Richtung (lokales Permittivitätsminimum). Dies ist in Abbildung Abb. 4b) dargestellt. Es handelt sich hierbei also nicht um eine dielektrische **Relaxation** sondern vielmehr um eine dielektrische **Resonanz**. Beachten Sie die unterschiedliche Form von $\varepsilon(\omega)$ für dielektrische Relaxation und dielektrische Resonanz. Wir wissen bereits, dass im Resonanzfall einer erzwungenen Schwingung eine maximale Energieumwandlung stattfindet - erinnern Sie sich an den Fall eines angeregten RLC-Schwingkreises. Wir wollen nun ein Verständnis für diese Energieumwandlungen entwickeln.

Die im äußeren Feld $\vec{E}(\omega, t)$ zu erzwungenen Schwingungen angeregten Dipole haben eine Masse, weshalb zwischen \vec{E} und \vec{D}_m nach Gl. (15) bzw. $\vec{P}(t)$ nach Gl. (18) eine Phasenverschiebung auftreten muss. Denken Sie an die Trägheit der Dipole und der damit einhergehenden zeitlichen Verzögerung zwischen Anregung und Antwort.

Wir nehmen zunächst an, dass bei niedrigen Frequenzen diese Phasenverschiebung Null ist. Die Dipole verhalten sich dann trägheitslos. Um die Bedeutung dieser Annahme zu verstehen, bestimmen wir die durch den Ladestrom Gl. (7) auf den Kondensator (Abb. 1) übertragene elektrische Energie. Bei anliegender Spannung U wird eine Ladung q quantisiert (d_q) bis zum Endwert Q gespeichert. Pro Ladungsschritt dt beträgt die

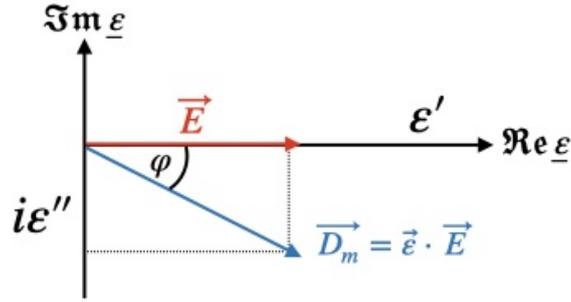


Abbildung 5: Die komplexe Permittivität ε als Maß für die Phasenverschiebung φ zwischen \vec{D}_m und \vec{E} .

Energiezunahme

$$dW = U \cdot dq - \frac{q}{C} dq \quad (23)$$

Integration führt auf

$$W = \frac{1}{2} C U^2 \quad (24)$$

Daraus folgt die Energiedichte $w = W/V$ zu

$$w = \frac{1}{2} \varepsilon_r \varepsilon_0 \vec{E}^2 = \frac{1}{2} \vec{E} \cdot \vec{D}_m \quad (25)$$

Die mittlere Energiedichte ist dann

$$\langle w \rangle = \frac{1}{4} \varepsilon_r \varepsilon_0 E_0^2 \quad (26)$$

die mittlere absorbierte Leistung verschwindet. Diese Ergebnisse folgen daraus, dass nach Gl. (15) $\frac{\vec{D}_m}{\varepsilon_0 \vec{E}} = \varepsilon_r$ eine vermeintlich reelle Größe ist. Für den Fall einer Phasenverschiebung zwischen \vec{E} und \vec{D}_m gilt aber allgemein nach Abb. 5

$$\vec{D}_m = \varepsilon_r \varepsilon_0 \vec{E} \quad (27)$$

Es wird ersichtlich, dass ε_r und damit ε im Allgemeinen komplex ist. Die mittlere absorbierte Leistungsdichte für den Fall der Phasenverschiebung ist dann

$$\left\langle \frac{dw}{dt} \right\rangle = \frac{1}{2} \varepsilon'' \omega E_0^2 \quad (28)$$

Man kann über die DRUDE-Theorie der Leitfähigkeiten zeigen, dass

$$\varepsilon'' = \frac{\sigma}{\varepsilon_0 \omega} \quad (29)$$

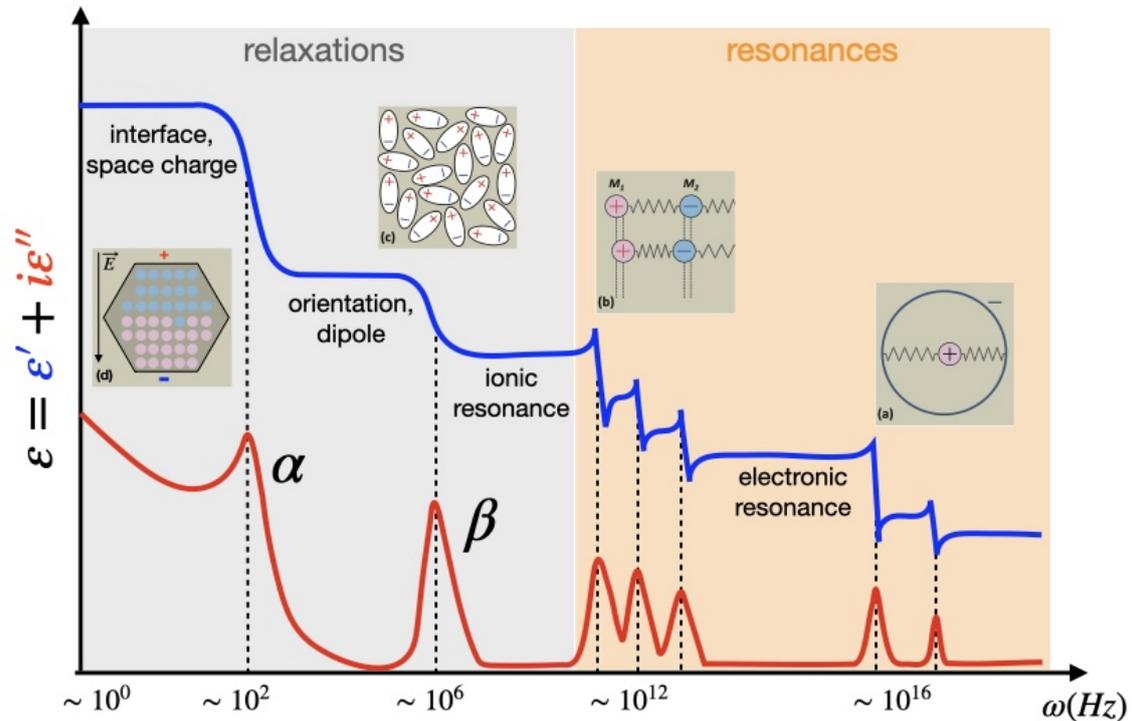


Abbildung 6: Verlauf des Real- (blau) und Imaginärteils (rot) der dielektrischen Funktion. (a-d) Verschiedene allgemeine Relaxations- und Resonanz-Beiträge im Festkörper sind symbolisiert.

ist. Dielektrische Verluste als "Reibungswärme" entstehen nach Gl. (29) also formal gleichberechtigt bei elektrischer Leitfähigkeit σ oder durch Polarisationsströme durch Verschiebungsstromverluste.

Spätestens jetzt realisieren wir, dass ε nicht als *Dielektrizitätskonstante* diskutiert werden darf, sondern im allgemeinen Fall eine komplexe und frequenzabhängige Funktion darstellt. Die **dielektrische Funktion** ist gegeben durch

$$\varepsilon(\omega) = \varepsilon' - i\varepsilon'' \quad (30)$$

Die Wahl des Vorzeichens des Imaginärteils ist hierbei arbiträr. Der Realteil ε' ist ein Maß für die Ladungsverschiebung, z.B. von Elektronen, Ionen, Dipolausrichtung und auch Ladungsverschiebungen an Grenzflächen, an denen verschiedene Materialien zusammentreffen. Der Imaginärteil ε'' beschreibt die durch Phasenverschiebung entstehenden Energietransformationen und Verluste. Eine prototypische Frequenzabhängigkeit von $\varepsilon(\omega) = \varepsilon' - i\varepsilon''$ für einen Festkörper ist in Abb. 6 detailliert dargestellt. Die Phasenverschiebung zwischen Erregerfeld und Reaktionsfeld, seine Antwortfunktion (*engl. response function*), des Körpers und deren Auswirkungen auf die Schwächung des Energieflusses in der Materie deuten eine komplizierte Wechselwirkung zwischen den elektromagnetischen Primärwellen und den Streuwellen an. Letztere werden von den zu erzwungenen

Schwingungen angeregten Ladungen emittiert. Wir erkennen im Verhalten von ε'' im Relaxations- und Resonanzfall einen maximalen Energieübergang von Leitungsstrom (Metalle) bzw. Verschiebungsstrom (Dielektrika oder Paraelektrika) zu Verlusten. Ein Verhalten welches in der dielektrischen Spektroskopie häufig ausgenutzt wird.

Es sei darauf hingewiesen, dass obige Ableitung der komplexwertigen und frequenzabhängigen Eigenschaften der Permittivität ε nur eine mögliche, z.T. phänomenologische, Herangehensweise darstellt. Alternative Ableitungen finden sich in [4], [2] und den Originalarbeiten von DEBYE [1], welche wir später noch strapazieren werden.

Die mikroskopische Beschreibung der Polarisation von Materialien geht aus vom atomaren Aufbau der Materie. In einem Festkörper z.B. werden positive Atomrümpfe (Atomkerne & stark gebundene Rumpfelektronen) durch in den "Zwischenräumen" verteilte Valenzelektronendichte in einem energetisch minimalen Ladungsgleichgewicht zusammengehalten. Durch verschiedene Anregungen nach Abb. 6 wird dieses Gleichgewicht in einem angelegten elektrischen Feld gestört. Es bilden sich induzierte Dipole mit einem elektrischen Dipolmoment \vec{m}_e als Indikator einer Verschiebungspolarisation. Dieses Dipolmoment ist proportional einer elektrischen Feldstärke analog zu \vec{P} in Gl. (18) und Gl. (16). Der Proportionalitätsfaktor ist die atomare Polarisierbarkeit α

$$\vec{m}_e = \alpha \varepsilon_0 \vec{E} \quad (31)$$

Die phänomenologische Größe "Polarisation" ergibt sich daraus über die Anzahl der Dipole (Atome oder Moleküle) pro Volumeneinheit (vgl. Gl. (16))

$$\vec{P} = n \vec{m}_e \quad (32)$$

Die Kombination der Gleichungen (31), (32), (18) und (16) liefert für die dielektrische Suszeptibilität

$$\chi_e = n\alpha \quad (33)$$

Sie ist proportional der atomaren Dichte n und der atomaren Polarisierbarkeit.

Gesucht wird nun der polarisationswirksame Anteil eines äußeren Feldes, der auf das Atom bzw. Molekül einwirkt. Bei Stoffen mit geringer Massendichte, z.B. Gase, sind die Gleichungen (31) und (32) ohne Einschränkung gültig. In Flüssigkeiten und Festkörpern setzt sich das polarisationswirksame elektrische Feld aus dem angelegten Feld und dem Polarisationsfeld der Nachbaratome zusammen. Wenn der Abstand zu den nächsten Nachbarn groß ist, wie in Gasen, erreicht das induzierte Feld nicht den Ort des interessierenden Aufatoms. Das vom äußeren Feld induzierte Polarisationsfeld (vgl. Abb. 2c)) ist von der Form des Materials, in dem dieses induzierte Feld als negatives Depolarisationsfeld Gl. (20) entsteht, abhängig. Für eine planparallele Platte als Dielektrikum, senkrecht zum Erregerfeld orientiert, ist das Polarisationsfeld \vec{P}/ε_0 (vgl. Gl. (20)). Für eine Kugel

als Dielektrikum im "Vakuumsfeld" \vec{E} ist das Polarisationsfeld $\vec{P}/3\varepsilon_0$. Man nimmt für die Umgebung des Aufatoms eine kugelförmige Umgebung an, deren Polarisationsfeld das angelegte Feld \vec{E} verstärkt. Es gilt

$$\vec{P} = n \alpha \varepsilon_0 \left(\vec{E} + \frac{\vec{P}}{3\varepsilon_0} \right) \quad (34)$$

Aus Gl. (18) folgt dann mit Gl. (34) die **Clausius-Mosotti-Gleichung**

$$\frac{n\alpha}{3} = \frac{\varepsilon - 1}{\varepsilon + 2} \quad (35)$$

Einschub: Paraelektrische Stoffe

Im Unterschied zu dielektrischen Stoffen haben paraelektrische Materialien (aus asymmetrisch aufgebauten Molekülen, wie z.B. Toluol oder Äthylalkohol) ein permanentes elektrisches Dipolmoment. Im äußeren elektrischen Feld der Feldstärke E orientieren sich die Dipole in Feldrichtung. Diese Orientierungspolarisation ist i.A. temperaturabhängig

$$\vec{P} = \frac{n(m_e)^2 \vec{E}}{3kT} \quad (36)$$

Daraus ergibt sich das bekannte CURIE-Gesetz der Temperaturabhängigkeit der dielektrischen Suszeptibilität zu

$$\chi_e = \varepsilon_r - 1 = \frac{n(m_e)^2}{3\varepsilon_0 kT} \quad (37)$$

4 Dynamik von Relaxationsprozessen

Im vorherigen Abschnitt wurden bereits einige frequenzabhängige Relaxationen, und auch Resonanzen, von unterschiedlichen mikroskopischen Polarisationsprozessen beschrieben. Im folgenden wird ein quantitatives, analytische Model hergeleitet, welches uns einen tieferen Einblick in die Physik der dielektrischen Funktion $\varepsilon(\omega)$ ermöglicht. Wir folgen dabei einem Ansatz für die Orientierungspolarisation welcher von Peter Debye 1929 [1] in Leipzig vorgeschlagen wurde.

Wir betrachten Dipole (vgl. Abb. 4a)) in einem elektrischen Wechselfeld $E(t) = E_0 \cdot \exp i\omega t$. In einem homogenen linearen isotropen dielektrischen Medium ist die Polarisation, nach gegebener Einstellzeit, parallel und proportional zum elektrischen Feld:

$$\vec{P} = \varepsilon_0(\varepsilon_r - 1)\vec{E} \quad (38)$$

Die Dipole sind optimal zum elektrischen Feld ausgerichtet und die Polarisation P_0 ist maximal. Das elektrische Feld wird nun abrupt abgeschaltet. Nach DEBYE [1] kann man nun annehmen, dass die Polarisation in einer charakteristischen Zeit τ wieder auf Null absinkt (*engl. single relaxation time approximation*). Es gilt

$$P(t) = P_0 \cdot e^{-\frac{t}{\tau}} \quad (39)$$

Da wir an einer Frequenzabhängigkeit interessiert sind, starten wir von der vereinfachten Fouriertransformierten Gl. (18)

$$P(\omega) = (\varepsilon(\omega) - 1) E(\omega) \quad (40)$$

Die Fouriertransformierte des elektrischen Feldes ist trivial¹

$$E(\omega) = \int_0^\infty E_0 e^{i\omega t} e^{-i\omega' t} dt = E_0 \quad (41)$$

Es bleibt noch die Fouriertransformierte der Polarisation $P(t)$ zu bestimmen. Es ist

$$P(\omega) = \int_0^\infty P_0 \cdot e^{-\frac{t}{\tau}} e^{-i\omega t} dt \quad (42)$$

Integration ergibt

$$P(\omega) = \frac{P_0}{\frac{1}{\tau} + i\omega} \quad (43)$$

Man setzt nun Gl. (41) und Gl. (43) in Gl. (40) ein

$$\frac{P_0}{\frac{1}{\tau} + i\omega} = (\varepsilon(\omega) - 1) E_0 \quad (44)$$

und löst zur gewünschten dielektrischen Funktion auf

$$\varepsilon(\omega) = 1 + \frac{P_0/E_0}{1 - i\omega\tau} = 1 + \frac{\varepsilon_s}{1 - i\omega\tau} \quad (45)$$

wobei im letzten Schritt die statische Permittivität $\varepsilon(\omega \rightarrow 0) = P_0/E_0 = \varepsilon_s$ eingeführt wurde. Das Resultat Gl. (45) ist beeindruckend. Offensichtlich ist $\varepsilon(\omega)$ komplexwertig. Zerlegung in Real- und Imaginärteil $\varepsilon(\omega) = \varepsilon(\omega)' - i\varepsilon(\omega)''$ ergibt

$$\begin{aligned} \varepsilon(\omega)' &= 1 + \frac{\varepsilon_s}{1 + (\omega\tau)^2} \\ \varepsilon(\omega)'' &= \frac{\varepsilon_s \omega\tau}{1 + (\omega\tau)^2} \end{aligned} \quad (46)$$

¹Man beachte, dass das eigentliche Resultat $2\pi E_0 \delta(\omega - \omega')$ erst nach geeigneter Wahl der Delta-Distribution auf E_0 zusammenfällt.

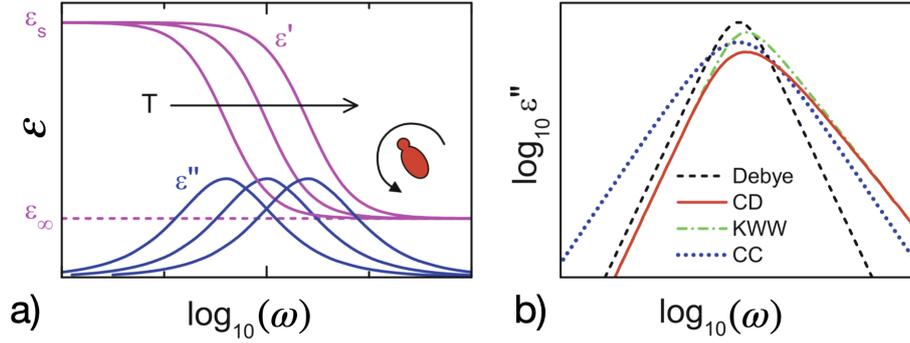


Abbildung 7: a) Typische Relaxationsspektren des Real- und Imaginärteils der Permittivität, dargestellt für drei Temperaturen. Der horizontale Pfeil zeigt steigende Temperaturen an. Die gestrichelte Linie markiert die Permittivität ϵ_∞ bei hohen Frequenzen. Ein molekularer elektrischer Dipol, der sich durch thermische Aktivierung reorientieren kann, ist angedeutet. b) Modellfunktionen des dielektrischen Verlusts in Abhängigkeit von der Frequenz in einer doppelt-logarithmischen Darstellung. Die symmetrische Debye-Funktion und abgeleitete Modellfunktionen sind dargestellt (siehe Text für Details).

Die Gleichungen 46 werden als DEBYE-Gleichungen² bezeichnet und zeigen, dass bereits in einem einfachen Modell eines relaxierenden Dipols, die dielektrische Funktion $\epsilon(\omega)$ explizit frequenzabhängig und komplexwertig sein muss.

Bisher haben wir angenommen, dass bei sehr hohen Frequenzen die Polarisation praktisch Null ist - der Dipol kann nicht folgen und $\chi(\omega \rightarrow \infty) = 0$, also $\epsilon_r(\omega \rightarrow \infty) = 1$, entsprechen dem Vakuumwert. Das ist praktisch häufig nicht zu erreichen und man führt das Limit $\epsilon(\omega \gg 1/\tau) = \epsilon_\infty$ ein. Die korrigierten Gleichungen 46 lauten nun

$$\begin{aligned}\epsilon(\omega)' &= \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + (\omega\tau)^2} \\ \epsilon(\omega)'' &= \frac{(\epsilon_s - \epsilon_\infty)\omega\tau}{1 + (\omega\tau)^2}\end{aligned}\quad (47)$$

mit dem damit verbundenen Phasenwinkel φ und dem Verlustfaktor

$$\tan(\varphi) = \frac{\epsilon''}{\epsilon'} = \frac{(\epsilon_s - \epsilon_\infty)\omega\tau}{\epsilon_s + \epsilon_\infty\omega^2\tau^2}\quad (48)$$

Die Verläufe von Real- und Imaginärteil der dielektrischen Funktion sind in Abb. 8a) durch die DEBYE-Gleichungen dargestellt. Man erkennt leicht, dass das Maximum des

²Nicht zu verwechseln mit dem DEBYE-Modell zur Beschreibung des Phononenbeitrags zur Wärmekapazität.

Imaginärteils $\varepsilon(\omega)''$ für $\omega = 1/\tau$, in der Flanke des Realteils $\varepsilon(\omega)'$ auftritt. Desweiteren erkennt man eine Temperaturabhängigkeit von $\varepsilon(\omega)$, bzw. expliziter von der Relaxationszeit $\tau(T)$. Die Temperatureigenschaften der Relaxationsprozesse werden im anschließenden Abschnitt detaillierter diskutiert.

Die analytisch hergeleiteten DEBYE-Gleichungen werden häufig durch empirische Parameter angepasst um bessere Übereinstimmung zu experimentellen Daten zu erreichen. Eine Möglichkeit ist z.B die COLE-COLE-Darstellung (siehe CC in Abb. 8b))

$$\varepsilon_{CC}(\omega) = \varepsilon_{\infty} + \frac{\varepsilon_s - \varepsilon_{\infty}}{1 + (i\omega\tau)^{1-\alpha}} \quad (49)$$

Man erkennt sofort die Erweiterung von Gl. (45) um den Exponenten $(1 - \alpha)$. Für $\alpha = 0$ reduziert sich die COLE-COLE-Darstellung zur DEBYE-Gleichung. Die COLE-COLE-Darstellung wird bevorzugt für die Beschreibung von Relaxationsprozessen in Polymeren benutzt.

Die in diesem Abschnitt angebotene Herleitung der komplexwertigen dielektrischen Funktion $\varepsilon(\omega)$ über eine Fouriertransformation war die Wahl des Verfassers. Alternative Herleitungen über Differentialgleichungsansätze finden sich in [1], Abschnitt 18 oder [2], Abschnitt 10.

4.1 Temperaturabhängigkeit von Polymer-Relaxationsprozessen

Ein wichtiges Anwendungsgebiet dielektrischer Spektroskopie ist die Untersuchung temperaturabhängiger Relaxationsprozesse in nichtkristallinen Werkstoffen, die mit Materialerweichungs- und Materialermüdungsprozessen verbunden sind. Alle amorphen Polymere (wie z.B. Polystyrol, Polycarbonat, oder auch Epoxydharze) zeigen mehrere Relaxationsprozesse, die charakteristische Eigenschaften aufweisen und auch in anderen nichtkristallinen Materialien (z.B. silikatischen Gläsern oder niedermolekularen Flüssigkeiten wie Glycerol) beobachtet werden. Polymere sind im Allgemeinen durch das Vorhandensein verschiedener dipolarer Einheiten (Hauptketten, Seitenketten etc.) in unterschiedlichen Umgebungen gekennzeichnet. Dies führt zu einer breiten Verteilung möglicher Relaxationszeiten. Die entsprechenden Verlustpeaks erstrecken sich üblicherweise über mehrere Frequenzdekaden von μHz hinzu hunderten ThZ.

Wie in Abb. 8 dargestellt finden sich dominant zwei Klassen von allgemeinen Relaxationsprozessen: ein α -Relaxationsprozess, der mit der Materialerweichung einhergeht und dessen Relaxationszeit sehr stark temperaturabhängig ist und ein (oder mehrere) Sekundärrelaxationen β (γ, δ) bei höheren Frequenzen mit deutlich schwächerer Temperaturabhängigkeit.

Isotherm beginnend bei niedrigen Frequenzen findet sich in der Regel die α -Relaxation,

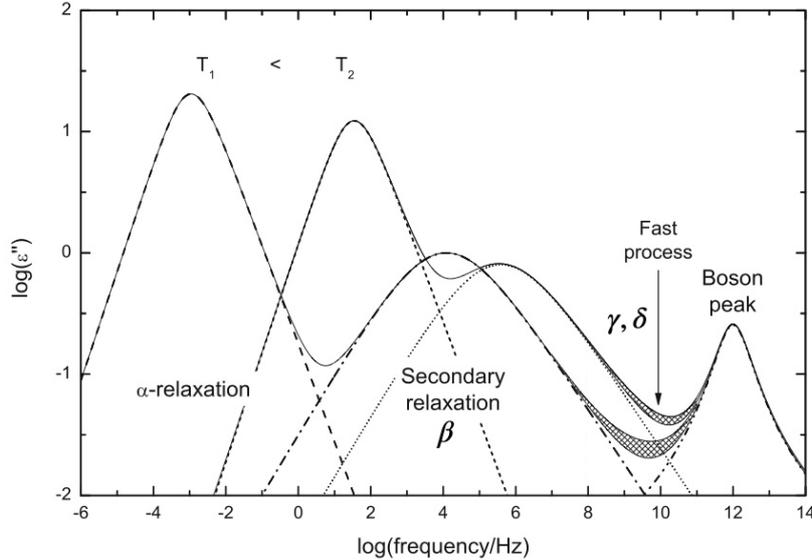


Abbildung 8: Darstellung der dynamischen, dielektrischen Verluste ε'' im Spektralbereich für ein prototypisches Polymer. (i) die langsame α -Relaxation, (ii) (schnellere) sekundäre β -Relaxationen, (iii) sehr schnelle Absorptionsprozesse die als γ, δ klassifiziert werden können. Die Temperaturverschiebung ist für zwei verschiedene Temperaturen $T_1 < T_2$ dargestellt. Gestrichelt angedeutet sind Anpassungen nach der DEBYE-Gleichung.

die mit dem Glasübergang T_g im amorphen Bereich des Materials einhergeht, also mit dem Einsetzen der segmentalen Bewegungen der Polymerketten. Der Verlustpeak der α -Relaxation weist zumeist eine leicht asymmetrische Form auf. Seine Temperaturabhängigkeit folgt dem empirischen VOGEL-FULCHER-TAMMANN-Gesetz (VFT-Gesetz), das aus der Theorie des Glasübergangs abgeleitet werden kann

$$\tau_\alpha(T) = \tau_0 \exp\left(\frac{T_A}{T - T_V}\right) \quad (50)$$

Dabei sind τ_0 und die Aktivierungstemperatur T_A Anpassungs-Parameter, während T_V die VOGEL-Temperatur darstellt, die mit der dynamischen Glasübergangstemperatur $T_V \approx 0.6 \dots 0.7 T_g$ korreliert ist. Die asymmetrische Form und das VFT-Verhalten sind typische Merkmale eines kooperativen Prozesses. Im Allgemeinen weist die α -Relaxation die höchste Intensität auf und wird daher häufig als **primäre oder Hauptrelaxation** bezeichnet.

Bei höheren Frequenzen tritt die β -Relaxation auf, die mit lokalen intramolekularen Bewegungen verbunden ist. Sie wird mitunter von der γ -Relaxation noch kleinerer molekularer Einheiten gefolgt. In Polymeren sind dies oft lokale Umordnungen von einzelnen Seitengruppen oder kurzen Stücken der Hauptkette des Polymers, die nur wenige Bindungen erfassen. Diese β, γ -Relaxationen werden als **sekundäre Relaxationen** bezeichnet. Schließlich kann am hochfrequenten Ende des Spektrums auch eine

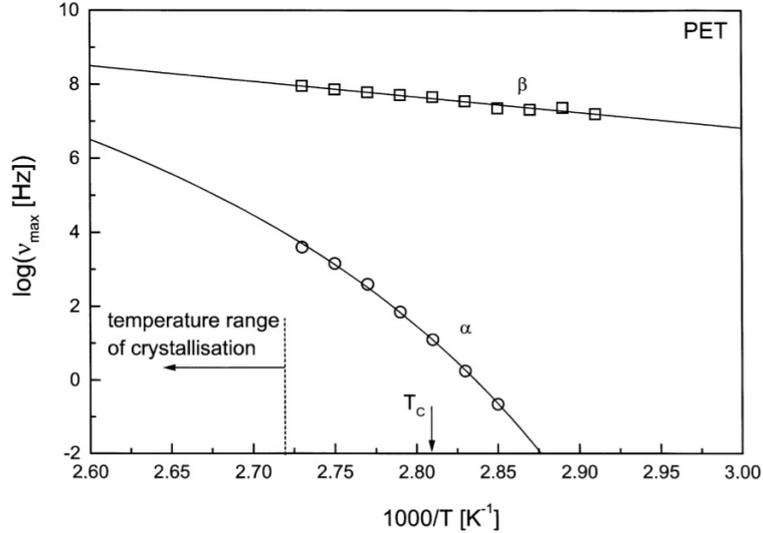


Abbildung 9: Aktivierungsdiagramm $\tau(T)$ für ein prototypisches Polymer (amorphes PET). Die durchgezogenen Linien entsprechen Fits gemäß der VFT-Gleichung bzw. der Arrhenius-Gleichung für α und β -Relaxationsprozesse.

δ -Relaxation auftreten, die auf isolierte Moleküle von Verunreinigungen zurückzuführen ist. Die Verlustpeaks dieser sekundären Prozesse sind symmetrisch und zeigen eine Arrhenius-Temperaturabhängigkeit

$$\tau_{\beta}(T) = \tau_0 \exp\left(\frac{E_A}{k_B T}\right) \quad (51)$$

mit einer charakteristischen Aktivierungsenergie E_A als typisches Merkmal eines nicht-kooperativen Prozesses.

Die deutlich unterschiedliche Temperaturabhängigkeit der Primär- und Sekundärprozesse, wie in Abb. 9 für das Polymer Polyethylenterephthalat (PET) dargestellt, machen eine separate Analyse der Relaxationsprozesse einfach zugänglich.

Die dielektrische Spektroskopie kann lediglich die Relaxationsparameter dieser Prozesse bestimmen. Ein Zusammenhang mit Bewegungen bestimmter molekularer Spezies muss durch detailliertes Wissen über die chemische Synthese und zusätzliche nicht-dielektrische Untersuchungen hergestellt werden. Beispielsweise müssen bei teilkristallinen Polymeren separate α - und sogenannte α -c-Relaxationen unterschieden werden, wobei letztere auf Polymerketten in amorphen Bereichen zurückzuführen sind, die durch Anheftung an Kristallbereiche teilweise immobilisiert sind.

Am niederfrequenten Ende des Spektrums tritt häufig eine Ladungsträgerrelaxation auf, insbesondere bei Vorhandensein ionischer Verunreinigungen. Es sei darauf hingewiesen, dass die Bewegung geladener Teilchen zwischen Grenzflächen (Elektroden oder Phasen-

grenzen) besonders effizient ist, da sie aufgrund größerer Ladungsverschiebungen im Vergleich zu rotierenden Dipolen ein großes effektives Dipolmoment erzeugt und aufgrund starker innerer Reibung hohe Verluste verursacht. Diese Ladungsträgerrelaxation kann die dipolaren Prozesse überlagern, ist jedoch nicht charakteristisch für ein bestimmtes Polymer. Daher sollte sie durch sorgfältige Reinigung der Substanz und präzise Präparation der elektrischen Kontakte vermieden werden. Die Ladungsträgerrelaxationen folgen einer Arrhenius-Temperaturabhängigkeit, da sie eng mit der exponentiellen Temperaturabhängigkeit der Leitfähigkeit verknüpft sind.

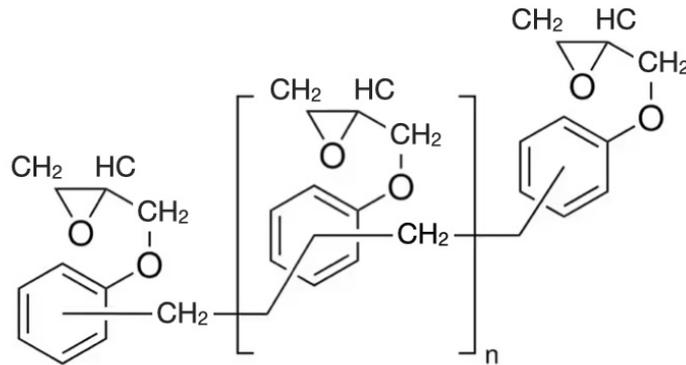


Abbildung 10: Strukturformel des Epoxydharzes Poly[(phenylglycidylether)-co-formaldehyd]. Die Epoxidgruppen $CH_2 - O - CH$ sind einem Polymer als dreigliedrige Ringe angehängen.

5 Eigenschaften des Epoxydharzes PPGE

Das im Versuch verwendet Polymer Poly[(phenylglycidylether)-co-formaldehyd] (kurz: PPGE) ist ein Harz, das durch die Copolymerisation von Phenylglycidylether und Formaldehyd entsteht. Copolymerisation bedeutet hierbei, dass die Phenylglycidylether-Einheiten über Formaldehyd-Verknüpfungen zu dem Polymer PPGE vernetzt werden. Es hat eine molare Masse von durchschnittlich $M_n = 345 \text{ g/mol}$ und ist damit ein eher leichtes, bzw. kurzes Polymer. Im Allgemeinen könnte man es als ein Oligomer bezeichnen, d.h. ein Prepolymer - also ein kurzkettiges Zwischenprodukt, das noch weiter polymerisiert werden kann. Die Massen liegen hier oft im Bereich von einigen Hundert bis wenigen Tausend g/mol. PPGE ist ein fragiles Polymer, d.h. seine Viskosität (und damit primäre Relaxationsprozesse) zeigt eine starke Temperaturabhängigkeit nahe der Glasübergangstemperatur T_g . Jene Glasübergangstemperatur T_g beträgt $258 \text{ K} \pm 1$ [7].

PPGE gehört zur Gruppe der epoxidhaltigen Polymere. Ein spezifisches Merkmal sind hierbei die typischen Epoxidgruppen $CH_2 - O - CH$ (siehe Abb. 10). Eine Epoxidgruppe

(auch Oxiranring genannt) ist eine chemische Funktionalität, die aus einem dreigliedrigen Ring mit einem Sauerstoffatom besteht. Die allgemeine Struktur $C - O - C$ gibt hierbei nur die zyklische dreigliedrige Grundstruktur wieder. Ein Standard-Epoxid ist Oxiran (Ethylenoxid, $C_2 - O - H_4$). Von besonderem Interesse ist die Asymmetrie der Epoxidgruppen $CH_2 - O - CH$ und sein endliches Dipolmoment welches in elektrischen Feldern wechselwirken wird.

Addendum: Bemerkungen zur Bestimmung der Phase zweier Messsignale

Ob in der Optik & Laserphysik (Interferometrie, messen kleinster Längenänderungen), Elektrischen Messtechnik & Signalverarbeitung (Phasenwinkel zwischen Spannung und Strom bestimmt die Wirkleistung), oder im vorliegenden Experiment wo die Phasenverschiebung zwischen \vec{E} und \vec{D} direkten Aufschluss über die komplexe dielektrische Funktion gibt, die genaue Bestimmung des Phasenunterschiedes zwischen zwei Messsignalen ist in der Physik essentiell. Abfolgend werden einige Möglichkeiten, i.A. in ihrer Präzision ansteigend, eingeführt.

Für alle nachfolgend dargestellten Methoden sei erwähnt, dass sich die Genauigkeit der zu bestimmenden Phasenverschiebung erheblich verbessern lässt, wenn man evtl. verrauschte Signale mit bekannter Trägerfrequenz f vorher mit einem schmalbandigen Bandpass $f \pm \Delta f$ dekoriert.

Bestimmung über Laufzeitunterschied

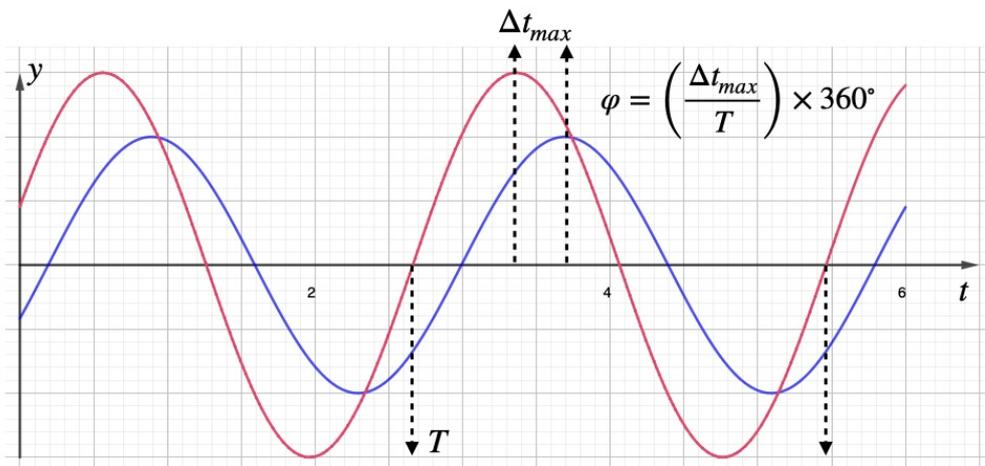


Abbildung AI: Zwei Signale gleicher Frequenz und unterschiedlicher Amplitude mit unbekannter Phasenverschiebung φ . Über die Messung des Laufzeitunterschiedes an markanten Stellen im Messsignal, hier an den Maxima, lässt sich φ direkt berechnen.

Die trivialste Methode der Bestimmung der Phasenverschiebung φ zwischen zwei periodischen Signalen ist gegeben durch eine Vermessung des Laufzeitunterschiedes zwischen ihnen. Der Phasenunterschied φ zwischen zwei periodischen Signalen **gleicher Frequenz**

kann aus dem Laufzeitunterschied Δt berechnet werden, wobei T die Periodendauer der Signale ist

$$\varphi = \frac{\Delta t}{T} \cdot 360^\circ \quad (\text{A.52})$$

Häufige Fehlerquellen sind Messungenauigkeiten, Signalverzerrungen oder eine falsche Bestimmung der Referenzpunkte für die Laufzeitmessung. Als signifikante werden Extrema der Signale, oder besser Nulldurchgänge genutzt. Ein Beispiel ist in Abb. [AI](#) dargestellt.

Bestimmung über Vermessen der Lissajous-Abbildungen

Eleganter und genauer kann die Phase φ über eine LISSAJOUS-Figur bestimmt werden, indem man zwei harmonische Signale unterschiedlicher oder gleicher Frequenz auf ein Oszilloskop gibt und zu gemeinsamen Zeitpunkten $x(t) \rightarrow y(t)$ darstellt. Die allgemeine Form der Signale in Lissajous-Form ist hierbei $x(t) = A \sin(\omega_i t)$ und $y(t) = B \sin(\omega_j t - \varphi)$.

Falls die Signale gleiche Frequenz haben, ergibt sich eine geschlossene Kurve. Der Phasenwinkel φ kann dann über die Punkte des Schnitts mit den Achsen bestimmt werden, wie in Abb. [AII](#) dargestellt. Hierbei ist

$$\sin(\varphi) = \frac{b}{a} = \frac{y_4 - y_3}{y_2 - y_1} \quad (\text{A.53})$$

wobei a die maximale Ausdehnung in y -Richtung und b der Versatz auf der y -Achse bei $x = 0$ ist. Hierbei ist zu beachten, dass immer auch eine Eichung auf $\varphi' = 360^\circ - \varphi$ möglich ist. Die Gl. [A.53](#) ist gültig wenn das Maximum $x(t), y(t)$ im 1. Quadranten liegt (vgl. Abb. [AIVb](#)), 4. Abbildung).

Liegt das Maximum $x(t), y(t)$ im 2. Quadranten (vgl. Abb. [AIVb](#)), 2. Abbildung) gilt

$$\sin(\varphi) = \pi - \frac{y_4 - y_3}{y_2 - y_1}. \quad (\text{A.54})$$

Sind die Frequenzen unterschiedlich, erhält man eine kompliziertere Abbildung, deren Form immer noch von Frequenz- und Phasenbeziehung abhängt, aber schwerer auszuwerten ist.

Bestimmung über Kreuzkorrelation

Die **Kreuzkorrelation** zweier Signale kann direkt genutzt werden, um die Phasenverschiebung φ zwischen ihnen zu bestimmen. Anschaulich wird dabei das eine Signal

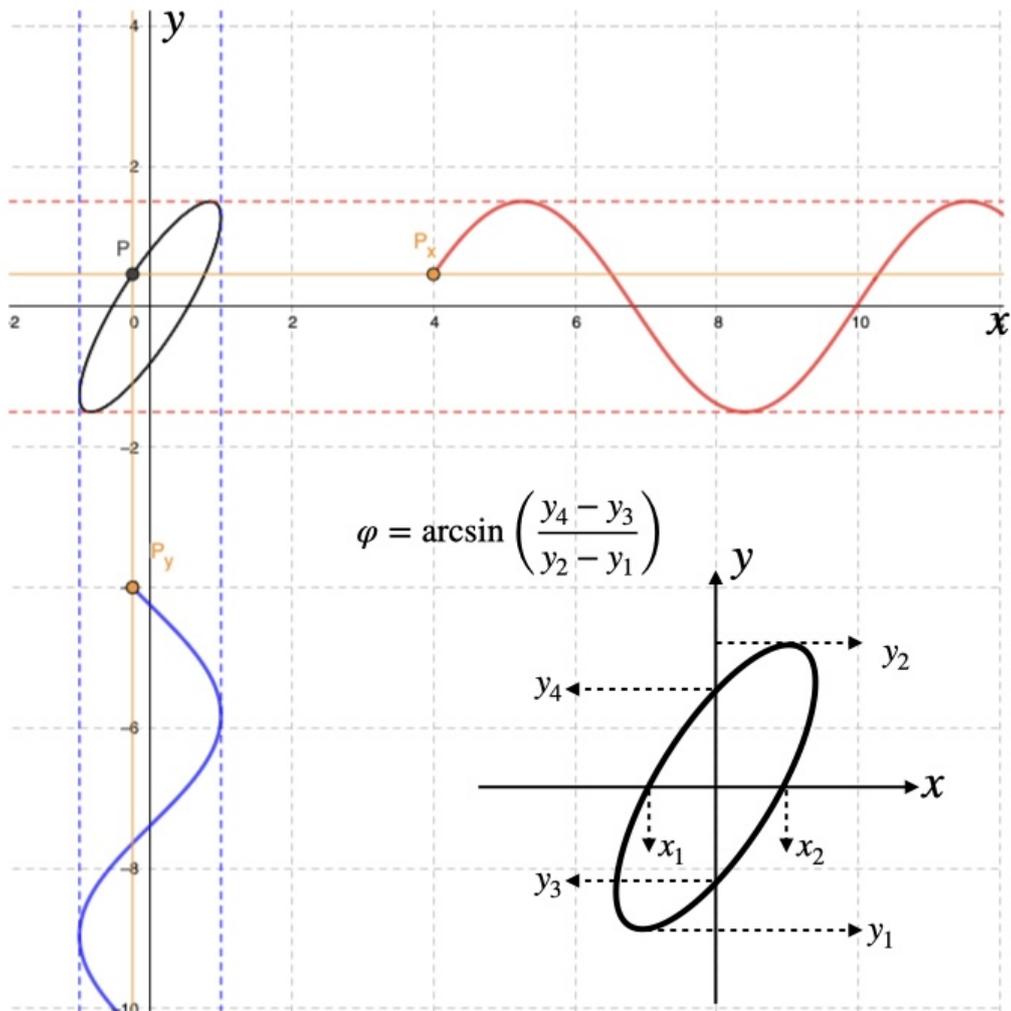


Abbildung AII: Zwei periodische Signale mit unterschiedlichen Amplituden ($A = 1.5, B = 1.0$) und Phasenunterschied von $\varphi \approx 33^\circ$ dargestellt als Signale $x(t), y(t)$ (rot,blau) und (schwarz) die parametrisierte LISSAJOUS-Figur $(x, y(t))$.

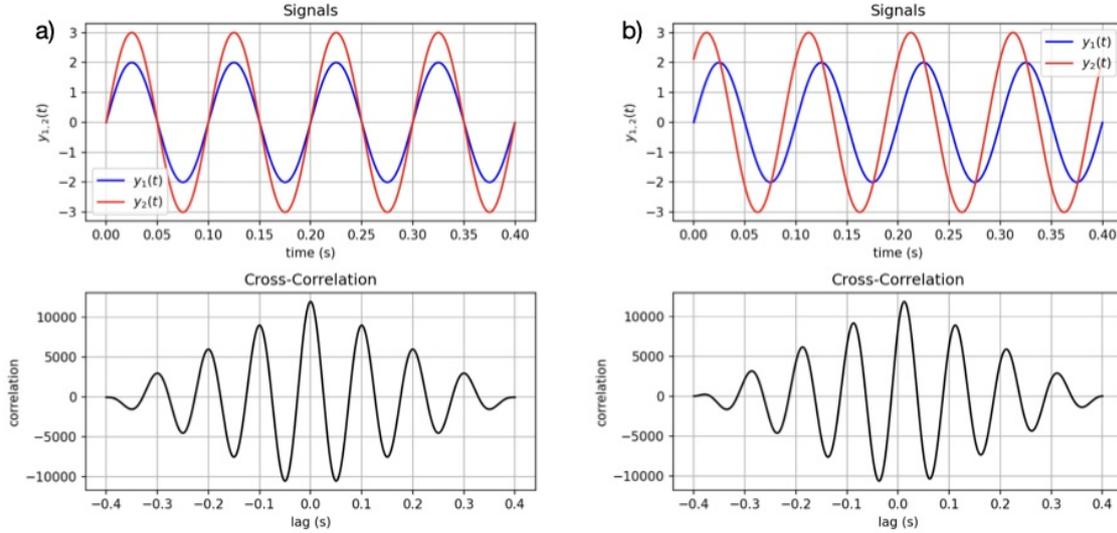


Abbildung AIII: (oben) Zwei periodische Signale mit $f = 10\text{Hz}$ und ihre (unten) Kreuzkorrelationsfunktionen R_{xy} . Gegeben ist ein Beispiel ohne a) und mit b) Phasenverschiebung zwischen den Signalen. Für den letzteren Fall findet man das Maximum von R_{xy} für $\tau_{\max} \approx 0.0122\text{s}$ ergibt sich $\varphi \approx 44^\circ$. Für den Fall a) ohne Phasenverschiebung $\varphi = 0$ sieht man deutlich das Maximum von R_{xy} bei $\tau_{\max} \approx 0$.

schrittweise gegenüber dem anderen mit τ (*time lag*) verschoben, und die Korrelation für jede Verschiebung berechnet.

Für zwei zeitdiskrete Signale $x[t_n]$ und $y[t_n]$ ist die Kreuzkorrelationsfunktion definiert als:

$$R_{xy}[\tau] = \sum_n x[t_n] \cdot y[t_n + \tau] \quad (\text{A.55})$$

Das Maximum von $R_{xy}[\tau]$ gibt den *time lag* τ_{\max} an, also die Zeitverschiebung, bei der die Signale am besten übereinstimmen. Die Phasenverschiebung ϕ kann dann mit der bekannten Abtastrate f_s bestimmt werden:

$$\varphi = \frac{\tau_{\max}}{f_s} \cdot 360^\circ \quad (\text{A.56})$$

Falls die Signale harmonische Sinusfunktionen sind, kann diese Methode direkt zur Phasenbestimmung genutzt werden. Ein Beispiel ist in Abb. AIIIa) und b) gegeben.

Determination via Covariance Matrix (Principal Component Analysis (PCA))

For the sake of simplicity, as a training for the curious student and as an excuse for the lazy author of this text, the following section will be written in English.

The covariance matrix Σ is a matrix that describes the correlation between n -dimensional data points. It contains the variances and covariances between multiple random variables. It is often used in statistics and data analysis, particularly in multivariate statistics and machine learning.

The eigenvectors of the covariance matrix indicate the principal axes of the data distribution, which is utilized in principal component analysis (*PCA*). If the data points lie along a contour line, such as in the 2D case of a circle or more generally an ellipse, the length of the principal axes will also be described by Σ . Detailed discussions can be found in [12].

We immediately recognize, if we recall the LISSAJOUS figure, that a statistical method for characterizing ellipses can provide us with another way to determine the phase shift between two measurement signals. In the following pages, a method will be explained in detail on how a general data cloud of data points \mathbf{P} in elliptical form can be described by fitting the covariance matrix and how this can then be transformed into a LISSAJOUS form to describe the phase shift of two signals representing \mathbf{P} .

Covariance Matrix

Let \mathbf{P} be a two-dimensional vector of the form

$$\mathbf{P} = [x(t) \ y(t)] \tag{A.57}$$

We assume that the data points are statistically distributed **around** the contour of the supposed ellipse and not widely distributed **only within** the area of the ellipse. Furthermore, for simplicity, let the ellipse be centered around the origin. The data points \mathbf{P} of the ellipse can then be described by

$$\mathbf{P}^T \Sigma^{-1} \mathbf{P} = c \tag{A.58}$$

where c is a constant that determines the confidence level of the confidence ellipse, i.e., its extent, and is initially of lesser interest. Σ is the covariance matrix and is given by

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} \tag{A.59}$$

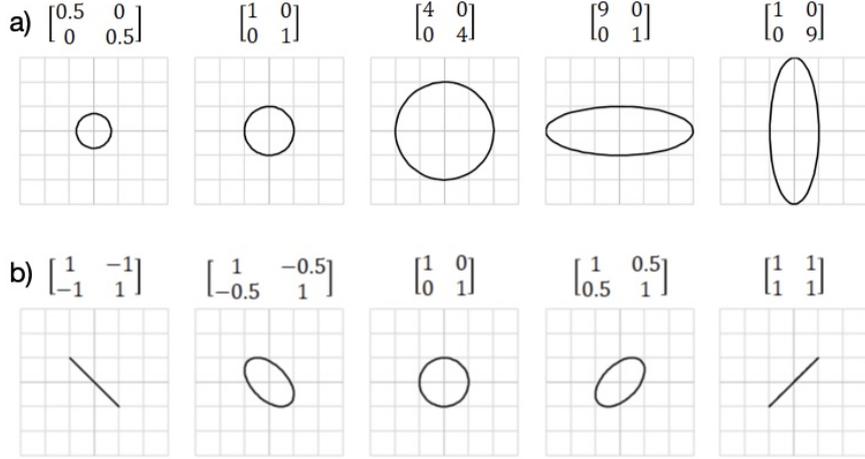


Abbildung AIV: Examples of the covariance matrix Σ . a) If $\sigma_{xy} = 0$, the data points $x(t), (y(t))$ are maximally uncorrelated. The phase shift would be e.g. 90° . σ_x^2 and σ_y^2 determine the *extent* of the data, i.e., the major and minor axes of the ellipse. b) If $\sigma_{xy} \neq 0$, the data $x(t), (y(t))$ is correlated. They are maximally positively (negatively) correlated for $\sigma_{xy} = 1, (-1)$, corresponding to a phase shift of $0^\circ, 180^\circ$. Other phase shifts have values of $\sigma_{xy} = -1 \dots 1$.

where $\sigma_x^2, (\sigma_y^2)$ is the variance of the random variables $x(t), (y(t))$, indicating how much the data scatter around the mean. σ_{xy} is the covariance between $x(t), y(t)$ and measures how strongly the two variables x, y are correlated. Illustrative examples are provided in Fig. AIV. In the most general case, with a given mean deviation μ , Eq. (A.58) resembles the Mahalanobis distance condition

$$(\mathbf{P} - \mu)^T \Sigma^{-1} (\mathbf{P} - \mu) = 1 \quad (\text{A.60})$$

which defines an ellipse representing a 1-standard deviation contour in a Gaussian distribution.

Direct relationship between Covariance Matrix & parametrized ellipses

We now want to show that the fitting of data points along the contour of an ellipse can be described directly using the covariance matrix and that the phase shift φ can be determined within the LISSAJOUS form. To do so, we proceed as follows:

1. Definition of a general, tilted ellipse in parametric form
2. Connection between the covariance matrix and the parameters of the general, tilted ellipse

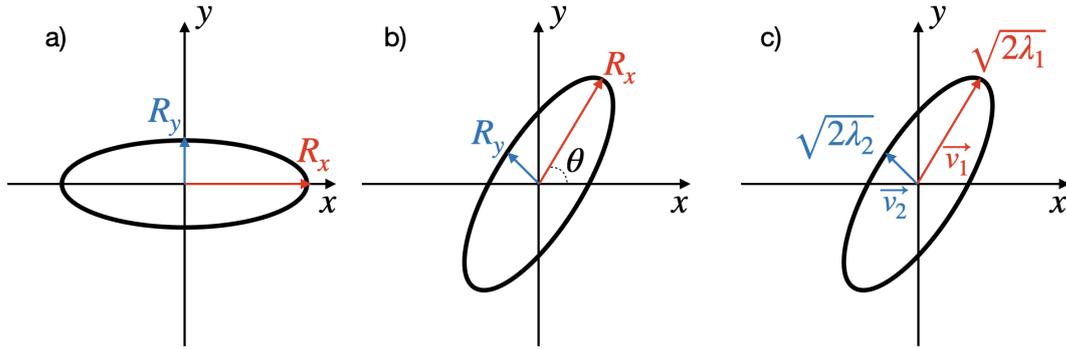


Abbildung AV: General representation of an ellipse. a) Without inclination θ , b) Generally with tilt by angle θ , c) Represented abstractly as an error ellipse described by the eigenvalues λ and eigenvectors v of the covariance matrix Σ .

3. Connection between the covariance matrix and the parameters of an ellipse in LISSAJOUS form

A general tilted ellipse

A general ellipse centered at the origin can be defined by the quadratic equation

$$\mathbf{P}^T \mathbf{Q} \mathbf{P} = 1 \quad (\text{A.61})$$

\mathbf{Q} is a symmetric positive definite 2×2 matrix that defines the shape and orientation of the ellipse. Expanding Eq. (A.61) it is

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12} & Q_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 1 \quad (\text{A.62})$$

which yields

$$Q_{11}x^2 + 2Q_{12}xy + Q_{22}y^2 = 1 \quad (\text{A.63})$$

being the matrix representation of the conic section of a general ellipse centered at the origin. The matrix \mathbf{Q} is symmetric and positive definite (i.e., it has positive eigenvalues). Looking at Eq. (A.63) students immediately recognize the resemblance of the unit circle $x^2 + y^2 = 1$.

We now want to find the parameters Q_{ij} that describe a general, tilted ellipse.

Therefore we start by the parametrization

$$\begin{aligned} x(t) &= R_x \cos(t) \cos(\theta) - R_y \sin(t) \sin(\theta), \\ y(t) &= R_x \cos(t) \sin(\theta) + R_y \sin(t) \cos(\theta). \end{aligned} \quad (\text{A.64})$$

Defining the vector of ellipse coordinates as

$$\mathbf{P}(t) = \mathbf{P} = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \quad (\text{A.65})$$

we rewrite the parametric equations as:

$$\mathbf{P} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} R_x \cos(t) \\ R_y \sin(t) \end{bmatrix}. \quad (\text{A.66})$$

Defining the orthogonal **rotation matrix** $R(\theta) = R$, which holds $R^{-1} = R^T$, and **scaling matrix** D as:

$$R(\theta) = R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \quad D = \begin{bmatrix} R_x & 0 \\ 0 & R_y \end{bmatrix}, \quad (\text{A.67})$$

using the vector of points on a unit circle

$$\mathbf{u}(t) = \mathbf{u} = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}, \quad (\text{A.68})$$

the matrix equation for the general tilted ellipse, cf. Eq. (A.64) becomes [Ban92]:

$$\mathbf{P} = R D \mathbf{u} \quad (\text{A.69})$$

which can be rewritten via $\mathbf{u} = D^{-1} R^T \mathbf{P}$ and $\mathbf{u}^T \mathbf{u} = 1$ as

$$\begin{aligned} \mathbf{P}^T (R D^{-T} D^{-1} R^T) \mathbf{P} &= 1 \\ \mathbf{P}^T (R D^{-2} R^T) \mathbf{P} &= 1 \\ \mathbf{P}^T Q_\theta \mathbf{P} &= 1 \end{aligned} \quad (\text{A.70})$$

Matrix Representation of a general tilted ellipse

We now want to show the direct connection between the previous matrix equation for the general tilted ellipse and the presentations of the same data points statistically fitted via the covariance matrix as in $\mathbf{P}^T \Sigma^{-1} \mathbf{P} = c$. Therefore we aim to derive Q_θ based on geometrical aspects of a general ellipse and connect both representations $Q_\theta \leftrightarrow \Sigma^{-1}$.

Starting again at the set of points $\mathbf{P}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$ that **surround the contour** of the ellipse centered at the origin, as deduced in Eq. A.70 the matrix representation describing the ellipse can be written as

$$\begin{aligned} Q_\theta &= R(\theta)D^{-T}D^{-1}R(\theta)^T \\ &= R(\theta)D^{-2}R(\theta)^T \end{aligned} \quad (\text{A.71})$$

Remember, that a standard axis-aligned ellipse without rotation has a matrix representation

$$\begin{aligned} Q_0 &= D^{-2} \\ &= \begin{bmatrix} \frac{1}{R_x^2} & 0 \\ 0 & \frac{1}{R_y^2} \end{bmatrix} \end{aligned} \quad (\text{A.72})$$

which with Eq. (A.61) and Eq. (A.63) resembles the well known standard form for a non-tilted ellipse

$$\frac{x^2}{R_x^2} + \frac{y^2}{R_y^2} = 1. \quad (\text{A.73})$$

We now just applied a rotation around θ in Eq. A.71 so that

$$Q_\theta = R(\theta)Q_0R(\theta)^T. \quad (\text{A.74})$$

Using $Q_0 = DD^{-1}$ from above yields

$$Q_\theta = R(\theta) \begin{bmatrix} \frac{1}{R_x^2} & 0 \\ 0 & \frac{1}{R_y^2} \end{bmatrix} R(\theta)^T. \quad (\text{A.75})$$

Expanding the multiplication it is

$$Q_\theta = \begin{bmatrix} \frac{\cos^2 \theta}{R_x^2} + \frac{\sin^2 \theta}{R_y^2} & \frac{\cos \theta \sin \theta}{R_x^2} - \frac{\cos \theta \sin \theta}{R_y^2} \\ \frac{\cos \theta \sin \theta}{R_x^2} - \frac{\cos \theta \sin \theta}{R_y^2} & \frac{\cos^2 \theta}{R_y^2} + \frac{\sin^2 \theta}{R_x^2} \end{bmatrix}. \quad (\text{A.76})$$

Comparing with Eq. (A.63) we identify

$$Q_{11} = \frac{\cos^2 \theta}{R_x^2} + \frac{\sin^2 \theta}{R_y^2} \quad (\text{A.77})$$

$$Q_{22} = \frac{\cos^2 \theta}{R_y^2} + \frac{\sin^2 \theta}{R_x^2} \quad (\text{A.78})$$

$$Q_{12} = \left(\frac{1}{R_x^2} - \frac{1}{R_y^2} \right) \cos \theta \sin \theta \quad (\text{A.79})$$

$$= \frac{1}{2} \left(\frac{1}{R_x^2} - \frac{1}{R_y^2} \right) \sin 2\theta. \quad (\text{A.80})$$

The latter expressions also can be directly derived by from the parametric form Eq. (A.64) by solving for $\cos(t)$, $\sin(t)$, eliminating t via $\cos(t)^2 + \sin(t)^2 = 1$ and some elegant algebraic transformations. Obviously the non-tilted ellipse Eq. (A.73) is recovered for $\theta = 0$. More insight might be found in John E. Davis' *Combining Error Ellipses* [Dav07].

By algebraic transformations one can show that the Eigenvalues $\lambda_2 > \lambda_1$ of Q_θ directly lead to the major semi-axis $R_x = 1/\sqrt{\lambda_1}$ and minor semi-axis $R_y = 1/\sqrt{\lambda_2}$. The angle θ as defined in Fig. AV can be obtained by the corresponding Eigenvectors, which mark the directions of the semi-axis, ν_1, ν_2 of Q_θ . Using the 2-argument arctangent it is

$$\theta = \text{atan2}(\nu_{1,y}, \nu_{1,x}). \quad (\text{A.81})$$

In terms of the standard arctangent function \arctan , whose interval is $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$, atan2 can be expressed piecewise:

$$\text{atan2}(y, x) = \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{if } x > 0, \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{if } x < 0 \text{ and } y \geq 0, \\ \arctan\left(\frac{y}{x}\right) - \pi & \text{if } x < 0 \text{ and } y < 0, \\ +\frac{\pi}{2} & \text{if } x = 0 \text{ and } y > 0, \\ -\frac{\pi}{2} & \text{if } x = 0 \text{ and } y < 0, \\ \text{undefined} & \text{if } x = 0 \text{ and } y = 0. \end{cases} \quad (\text{A.82})$$

By comparing Eqs. A.70 and A.76 with Eqs. A.58 and A.59 it is convincing that R_x, R_y, θ can be directly obtained by the Eigensystem of the statistical covariance matrix Σ with corresponding Eigenvalues $\lambda_{1,2}$, too. One [Coe08] can show that

$$\theta = \frac{1}{2} \arctan \frac{2\sigma_{xy}}{\sigma_x^2 - \sigma_y^2} \pmod{\left(\frac{\pi}{2}\right)} \quad (\text{A.83})$$

$$R_x = \sqrt{2\lambda_1} \quad \text{and} \quad R_y = \sqrt{2\lambda_2} \quad (\text{A.84})$$

With the derivation of Eqs. A.76 through A.84 we have shown, that the quantities R_x, R_y, ϑ needed to fully parametrize a general tilted ellipse can be linked to the entries of a statistically fitted covariance matrix A.59, i.e.

$$Q_{ij} \hat{=} 0.5 \cdot \Sigma_{ij}^{-1}. \quad (\text{A.85})$$

Therefore a statistical fitting of an *a priori* noisy ellipse \mathbf{P} can give us direct insight on the exact geometrical quantities of the ellipse. Thus, the covariance matrix representation encodes the **same orientation and scaling** as the explicit matrix parametrization but in terms of statistical properties.

The attentive reader might wonder about the factors 0.5 and 2, introduced in Eq. (A.85) and Eq. (A.84), respectively. When dealing with statistical fitting of error ellipses via a covariance matrix one expects the data P to be spread **within** the error ellipse given by a certain confidence interval linked to c in Eq. (A.58) [Wang15]. In our specific case we expect the data P to be spread **along the contour** of the error ellipse, so we need to introduce a scaling correction. The covariance matrix for contour points overestimates variance compared to area-distributed points by a factor of two in the eigenvalues. This difference propagates to confidence ellipses, making them about $\sqrt{2}$ times larger. Thus, one could reasonably say the covariance matrix is "off by a factor of two" depending on whether points are inside or on the boundary of the ellipse.

Matrix Representation of an ellipse in Lissajous form

Finally we aim to connect the quantities of variance and covariance as statistically defined by A.59, as well as the parameters of the general tilted ellipse R_x, R_y, θ to the special form of a LISSAJOUS ellipse with parameters A, B, φ .

In physics, or more general metrology, two periodic signals with amplitudes A, B and phase shift angle φ are often represented in the LISSAJOUS form as

$$\begin{aligned} x(t) &= A \sin(t) \\ y(t) &= B \cos(t - \varphi) \end{aligned} \quad (\text{A.86})$$

which is

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A \sin(t + \varphi/2) \\ B \cos(t - \varphi/2) \end{bmatrix} = \cos(t) \begin{bmatrix} A \sin(\varphi/2) \\ B \cos(\varphi/2) \end{bmatrix} + \sin(t) \begin{bmatrix} A \cos(\varphi/2) \\ B \sin(\varphi/2) \end{bmatrix} \quad (\text{A.87})$$

We want to highlight that the phase shift angle φ **is not identical** to the rotation angle θ of the previous general tilted ellipse. A misinterpretation often made in literature.

Reusing the unit vector $\mathbf{u} = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$ and the combined scaling and rotation matrix

$$G = \begin{bmatrix} A \sin(\varphi/2) & A \cos(\varphi/2) \\ B \cos(\varphi/2) & B \sin(\varphi/2) \end{bmatrix} \quad (\text{A.88})$$

as well as the points representing the contour of the ellipse

$$\mathbf{P} = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \quad (\text{A.89})$$

it is

$$\mathbf{P} = G\mathbf{u} \quad (\text{A.90})$$

$$\mathbf{u} = G^{-1}\mathbf{P}. \quad (\text{A.91})$$

Since again $\mathbf{u}^T\mathbf{u} = 1$, it is

$$\mathbf{P}^T G^{-T} G^{-1} \mathbf{P} = 1 \quad (\text{A.92})$$

where we immediately recognize that $G^{-T}G^{-1}$ takes the form and properties of the inverse covariance matrix Σ^{-1} as in Eq. A.58 and compares to Q_θ in Eq. A.70. We note that

$$G = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} \sin(\varphi/2) & \cos(\varphi/2) \\ \cos(\varphi/2) & \sin(\varphi/2) \end{bmatrix} \quad (\text{A.93})$$

and therefore,

$$G^{-1} = \begin{bmatrix} \sin(\varphi/2) & \cos(\varphi/2) \\ \cos(\varphi/2) & \sin(\varphi/2) \end{bmatrix}^{-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^{-1} \quad (\text{A.94})$$

$$G^{-T} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^{-1} \begin{bmatrix} \sin(\varphi/2) & \cos(\varphi/2) \\ \cos(\varphi/2) & \sin(\varphi/2) \end{bmatrix}^{-1} \quad (\text{A.95})$$

Hence, carrying out the multiplication, we obtain,

$$G^{-T}G^{-1} = \frac{1}{\cos^2(\varphi)} \begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} 1 & -\sin(\varphi) \\ -\sin(\varphi) & 1 \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix} \quad (\text{A.96})$$

which simplifies to

$$Q_\varphi = G^{-T}G^{-1} = \frac{1}{\cos^2(\varphi)} \begin{bmatrix} A^{-2} & -A^{-1}B^{-1}\sin(\varphi) \\ -A^{-1}B^{-1}\sin(\varphi) & B^{-2} \end{bmatrix} \quad (\text{A.97})$$

Now, the entries of matrix Σ_φ are known from comparison with either the fitted covariance values of Eqs. A.59 and A.85 or the parameters R_x, R_y and θ of the general tilted ellipse in Eq. A.76. Introducing the helping parameters h

$$h_1 = \frac{Q_{\varphi,11}}{Q_{\varphi,22}} = \frac{Q_{\theta,11}}{Q_{\theta,22}} = \frac{\Sigma_{11}^{-1}}{\Sigma_{22}^{-1}} = \frac{B^2}{A^2} \quad (\text{A.98})$$

and

$$h_2 = \frac{Q_{\varphi,11}}{Q_{\varphi,12}} = \frac{Q_{\theta,11}}{Q_{\theta,12}} = \frac{\Sigma_{11}^{-1}}{\Sigma_{12}^{-1}} = -\frac{B}{A\sin(\varphi)} \quad (\text{A.99})$$

Hence,

$$\sin(\varphi) = -\frac{\sqrt{h_1}}{h_2} \quad (\text{A.100})$$

which gives us the phase shift angle φ in the LISSAJOUS form. Furthermore it is,

$$Q_{\varphi,12} = -\frac{\sin(\varphi)}{\cos^2(\varphi)AB} \quad (\text{A.101})$$

and thus

$$AB = -\frac{\sin(\varphi)}{Q_{\varphi,12} \cos^2(\varphi)} \quad (\text{A.102})$$

$$= -\frac{\sin(\varphi)}{Q_{\theta,12} \cos^2(\varphi)} \quad (\text{A.103})$$

$$= -\frac{\sin(\varphi)}{0.5\Sigma_{12}^{-1} \cos^2(\varphi)} \quad (\text{A.104})$$

Therefore,

$$B = \sqrt{\frac{h_2 \tan^2(\varphi)}{Q_{\varphi,12}}} \quad (\text{A.105})$$

$$= \sqrt{\frac{h_2 \tan^2(\varphi)}{Q_{\theta,12}}} \quad (\text{A.106})$$

$$= \sqrt{\frac{h_2 \tan^2(\varphi)}{0.5\Sigma_{12}^{-1}}} \quad (\text{A.107})$$

Once we get B , we can calculate A

$$A = -\frac{\sin \varphi}{Q_{\varphi,12} \cdot \cos \varphi^2} \frac{1}{B} \quad (\text{A.108})$$

$$= -\frac{\sin \varphi}{Q_{\theta,12} \cdot \cos \varphi^2} \frac{1}{B} \quad (\text{A.109})$$

$$= -\frac{\sin \varphi}{0.5\Sigma_{12}^{-1} \cdot \cos \varphi^2} \frac{1}{B} \quad (\text{A.110})$$

We have thus derived A, B, φ directly from the general ellipse form Q_{θ} built on R_x, R_y, θ or the statistically fitted covariance data of Σ as defined in Eqs. [A.59](#) and [A.85](#).

Example data set

Assume we obtained the data as shown in Fig. [AVI](#) as a measurement of two signals.

A statistical covariance fit by the covariance matrix method gives us the approximated

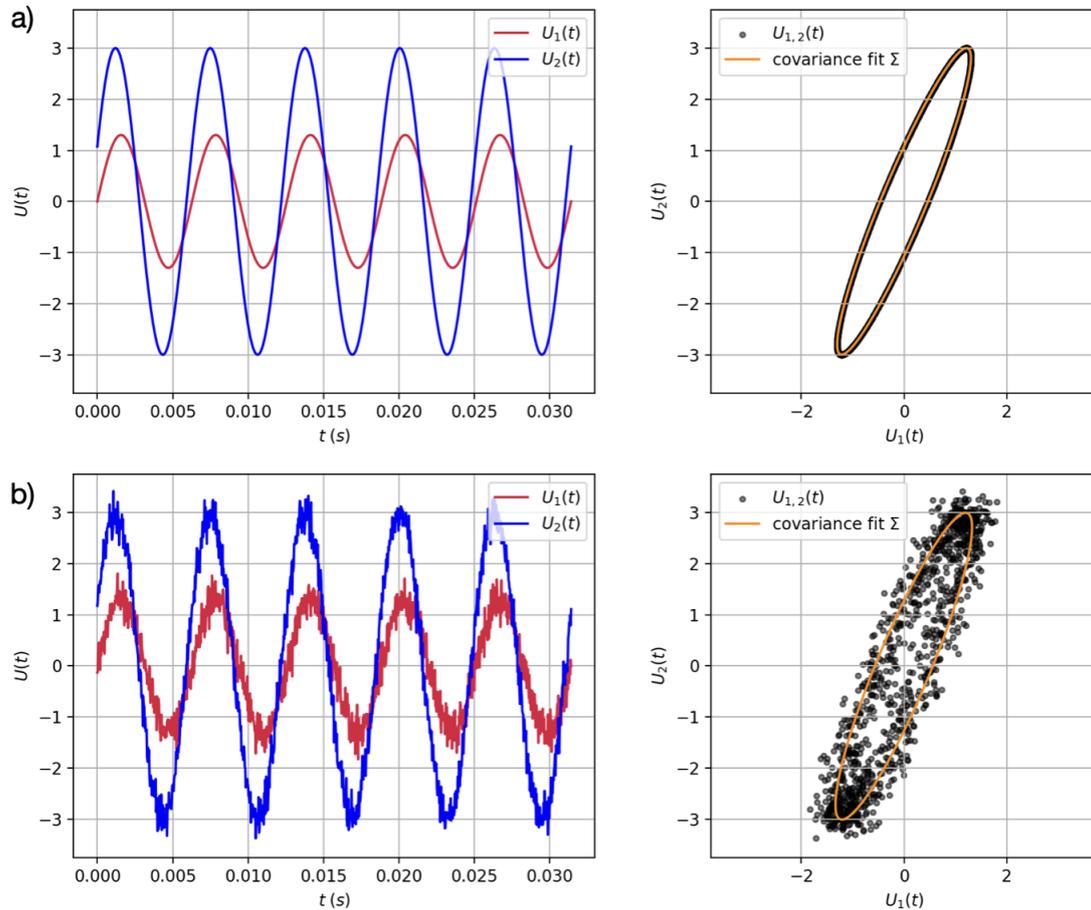


Abbildung AVI: Two simulated signals in LISSAJOUS form $U_1(t) = x(t) = A \cdot \sin(\omega t)$, $U_2(t) = y(t) = B \cdot \cos(\omega t - \varphi)$ with $A=1.3$, $B=3.0$, $\varphi = 69^\circ$. a) ideal signals and b) dressed with normal (Gaussian) distributed noise with a standard deviation of $\sigma = 200mV$. The orange lines highlight a fit by the covariance matrix. Errors in the determination of the phase are marginal even in highly noisy signals.

values for variances and covariance as

$$\begin{aligned}\sigma_x^2 &\approx 0.845 \\ \sigma_y^2 &\approx 4.502 \\ \sigma_{xy} &\approx 1.820\end{aligned}$$

By stressing Eq. A.98 and subsequent ones we directly get

$$\begin{aligned}A &= 1.3 \\ B &= 3 \\ \varphi &= 69^\circ\end{aligned}$$

and by this our desired signal amplitudes and phase shift angle in a LISSAJOUS representation.

By comparison with Eq. A.76 and more explicitly Eq. A.84 we straightforward obtain

$$\begin{aligned}R_x &= 3.241 \\ R_y &= 0.433 \\ \theta &= 67.57^\circ\end{aligned}$$

as the parameters in need for a representation as a general tilted ellipse. Note, that by accident θ and φ are numerically close. However, θ **is not** the desired physical phase shift φ .

A test dataset is made available by the author upon request.

Direct linear least squares fitting of an ellipse

The following section relies on an idea of Christian Hill as published on his [scientific blog](#).

Fitting a set of data points \mathbf{P} in the xy-plane to an ellipse is a surprisingly common problem in image recognition and analysis. In principle, the problem is one that is open to a linear least squares solution, since the general equation of any conic section can be written as:

$$F(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0, \quad (\text{A.111})$$

which is linear in its parameters a, b, c, d, e , and f . The polynomial $F(x, y)$ is called the algebraic distance of any point (x, y) from the conic (and is zero if (x, y) happens to lie on the conic).

A naive algorithm might attempt to minimize the sum of the squares of the algebraic distances,

$$d(a) = \sum_{i=1}^N F(x_i, y_i)^2, \quad (\text{A.112})$$

with respect to the parameters $\mathbf{a} = [a, b, c, d, e, f]^T$. However, there are two problems to overcome: firstly, such an algorithm will almost certainly find the trivial solution in which all parameters are zero. To avoid this, the parameters must be constrained in some way. A common choice, noting that $F(x, y)$ can be multiplied by any non-zero constant and still represent the same conic, is to insist that $f = 1$ (other algorithms set $a + c = 1$ or $\|\mathbf{a}\|^2 = 1$).

The second problem, however, is that there is no guarantee that the conic best fitted to an arbitrary set of points is an ellipse (for example, their mean squared algebraic distance to a hyperbola might be smaller). This is more likely to be the case where the points only lie near an elliptical arc rather than around an entire ellipse. The appropriate constraint on the parameters of $F(x, y)$ in order for the conic it represents to be an ellipse, is:

$$b^2 - 4ac < 0. \quad (\text{A.113})$$

Some iterative algorithms seek to minimize $F(x, y)$ in steps, ensuring this constraint is met before each step.

However, a direct least squares fitting to an ellipse (using the algebraic distance metric) was demonstrated by Fitzgibbon et al. (1999). They used the fact that the parameter vector \mathbf{a} can be scaled arbitrarily to impose the equality constraint $4ac - b^2 = 1$, thus ensuring that $F(x, y)$ is an ellipse. The least-squares fitting problem can then be expressed as minimizing $\|D\mathbf{a}\|^2$ subject to the constraint $\mathbf{a}^T C \mathbf{a} = 1$, where the design matrix

$$D = \begin{bmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & x_n y_n & y_n^2 & x_n & y_n & 1 \end{bmatrix} \quad (\text{A.114})$$

represents the minimization of F and the constraint matrix

$$C = \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{A.115})$$

expresses the constraint $\mathbf{a}^T C \mathbf{a} = 1$. The method of Lagrange multipliers (as introduced by Gander in 1980) yields the conditions:

$$S\mathbf{a} = \lambda C\mathbf{a}, \quad \mathbf{a}^T C \mathbf{a} = 1, \quad (\text{A.116})$$

where the scatter matrix is $S = D^T D$. There are up to six real solutions, $(\lambda_j, \mathbf{a}_j)$ and, it was claimed, the one with the smallest positive eigenvalue, λ_k , and its corresponding eigenvector, \mathbf{a}_k , represent the best-fit ellipse in the least squares sense.

Fitzgibbon *et al.* provided an algorithm implementing this approach, which was subsequently improved in 1998 by Halíř and Flusser for reliability and numerical stability. It is this improved algorithm that is provided in Python below.

Note that the algorithm is inherently biased towards smaller ellipses because of its use of the algebraic distance as a metric for the goodness-of-fit rather than the geometric distance. As noted by both sources for the algorithm used in this post, this issue is discussed by Kanatani in 1994 and is not easily overcome.

```
import numpy as np
import matplotlib.pyplot as plt

def fit_ellipse(x, y):
    """
    Fit the coefficients a,b,c,d,e,f, representing an ellipse described by
    the formula  $F(x,y) = ax^2 + bxy + cy^2 + dx + ey + f = 0$  to the provided
    arrays of data points  $x=[x_1, x_2, \dots, x_n]$  and  $y=[y_1, y_2, \dots, y_n]$ .

    Based on the algorithm of Halir and Flusser, "Numerically stable direct
    least squares fitting of ellipses'.

    """

    D1 = np.vstack([x**2, x*y, y**2]).T
    D2 = np.vstack([x, y, np.ones(len(x))]).T
    S1 = D1.T @ D1
    S2 = D1.T @ D2
    S3 = D2.T @ D2
    T = -np.linalg.inv(S3) @ S2.T
    M = S1 + S2 @ T
    C = np.array(((0, 0, 2), (0, -1, 0), (2, 0, 0)), dtype=float)
    M = np.linalg.inv(C) @ M
    eigval, eigvec = np.linalg.eig(M)
    con = 4 * eigvec[0]* eigvec[2] - eigvec[1]**2
    ak = eigvec[:, np.nonzero(con > 0)[0]]
    return np.concatenate((ak, T @ ak)).ravel()
```

```

def cart_to_pol(coeffs):
    """
    Convert the cartesian conic coefficients, (a, b, c, d, e, f), to the
    ellipse parameters, where  $F(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0$ .
    The returned parameters are  $x_0, y_0, ap, bp, e, phi$ , where  $(x_0, y_0)$  is the
    ellipse centre;  $(ap, bp)$  are the semi-major and semi-minor axes,
    respectively;  $e$  is the eccentricity; and  $phi$  is the rotation of the semi-
    major axis from the  $x$ -axis.

    """

    # We use the formulas from https://mathworld.wolfram.com/Ellipse.html
    # which assumes a cartesian form  $ax^2 + 2bxy + cy^2 + 2dx + 2fy + g = 0$ .
    # Therefore, rename and scale b, d and f appropriately.
    a = coeffs[0]
    b = coeffs[1] / 2
    c = coeffs[2]
    d = coeffs[3] / 2
    f = coeffs[4] / 2
    g = coeffs[5]

    den = b**2 - a*c
    if den > 0:
        raise ValueError('coeffs do not represent an ellipse: b^2 - 4ac must'
                          ' be negative!')

    # The location of the ellipse centre.
    x0, y0 = (c*d - b*f) / den, (a*f - b*d) / den

    num = 2 * (a*f**2 + c*d**2 + g*b**2 - 2*b*d*f - a*c*g)
    fac = np.sqrt((a - c)**2 + 4*b**2)
    # The semi-major and semi-minor axis lengths (these are not sorted).
    ap = np.sqrt(num / den / (fac - a - c))
    bp = np.sqrt(num / den / (-fac - a - c))

    # Sort the semi-major and semi-minor axis lengths but keep track of
    # the original relative magnitudes of width and height.
    width_gt_height = True
    if ap < bp:
        width_gt_height = False
        ap, bp = bp, ap

```

```

# The eccentricity.
r = (bp/ap)**2
if r > 1:
    r = 1/r
e = np.sqrt(1 - r)

# The angle of anticlockwise rotation of the major-axis from x-axis.
if b == 0:
    phi = 0 if a < c else np.pi/2
else:
    phi = np.arctan((2.*b) / (a - c)) / 2
    if a > c:
        phi += np.pi/2
if not width_gt_height:
    # Ensure that phi is the angle to rotate to the semi-major axis.
    phi += np.pi/2
phi = phi % np.pi

return x0, y0, ap, bp, e, phi

def get_ellipse_pts(params, npts=100, tmin=0, tmax=2*np.pi):
    """
    Return npts points on the ellipse described by the params = x0, y0, ap,
    bp, e, phi for values of the parametric variable t between tmin and tmax.

    """
    x0, y0, ap, bp, e, phi = params
    # A grid of the parametric variable, t.
    t = np.linspace(tmin, tmax, npts)
    x = x0 + ap * np.cos(t) * np.cos(phi) - bp * np.sin(t) * np.sin(phi)
    y = y0 + ap * np.cos(t) * np.sin(phi) + bp * np.sin(t) * np.cos(phi)
    return x, y

if __name__ == '__main__':
    # Test the algorithm with an example elliptical arc.
    npts = 250
    tmin, tmax = np.pi/6, 4 * np.pi/3
    x0, y0 = 4, -3.5
    ap, bp = 7, 3
    phi = np.pi / 4

```

```

# Get some points on the ellipse (no need to specify the eccentricity).
x, y = get_ellipse_pts((x0, y0, ap, bp, None, phi), npts, tmin, tmax)
noise = 0.1
x += noise * np.random.normal(size=npts)
y += noise * np.random.normal(size=npts)

coeffs = fit_ellipse(x, y)
print('Exact parameters:')
print('x0, y0, ap, bp, phi =', x0, y0, ap, bp, phi)
print('Fitted parameters:')
print('a, b, c, d, e, f =', coeffs)
x0, y0, ap, bp, e, phi = cart_to_pol(coeffs)
print('x0, y0, ap, bp, e, phi = ', x0, y0, ap, bp, e, phi)

plt.plot(x, y, 'x')      # given points
x, y = get_ellipse_pts((x0, y0, ap, bp, e, phi))
plt.plot(x, y)
plt.show()

```
