

Fra Newton til Kvantemekanik

Rasmus Nielsen

Sidst redigeret: 9. november 2021

I 1687 udgave Sir Isaac Newton sit epokegørende værk *Philosophiæ Naturalis Principia Mathematica*, hvori han beskrev sine 3 bevægelses-love, den universalle tyngde-lov, samt en udeledning af Keplers love om planeternes bevægelse. Disse opdagelser af Newton ligger uden tvivl i blandt de mest indflydelsesrige og banebrydende i hele fysikens historie. Specielt opdagelsen af Newtons 2. lov:

$$\vec{F} = m \vec{a} \quad (1)$$

gjorde det nu muligt (*i principet*) at forudsige opførslen af et hvilket som helst mekanisk system, givet et sæt af kendte begyndelsesbetingelser. Et så præcist og utvetydigt værktøj til beskrivelse af et mekanisk systems udvikling over tid var hidtil uset i fysiken. Man kan med rette sige, at opdagelsen af Newtons love (*specielt Newtons 2. lov*) er oprindelsepunktet for al moderne fysik. Selv den dag i dag, mere end 300 år efter deres opdagelse, ser Newtons love anvendelse i mange praktiske så vel som teoretiske sammenhænge.

Men er Newtons 2. lov (1) den eneste måde udviklingen af et fysisk system over tid? Svaret på dette spørgsmål viser sig at være nej! Der findes op til flere formuleringer af love, der alle forudsiger samme tidsudvikling som Newtons 2. lov, for et givent klassisk system. Man kan nu spørge sig selv hvad ideén med at indføre disse alternative fysiske love, hvis de i sidste ende giver os samme forudsigelser som Newtons 2. lov? Det kan ved første øjekast fremstå som en overflødig tilføjelse til fysiken. Der er mindst 3 oplagte grunde til at dette ikke er tilfældet.

1. For et givent system, kan en eller flere af de alternative formuleringer være mindre teknisk tunge end Newtons 2. lov. Et godt eksempel er det klassisk *dobbelt-pendul*, som er nemere beskrevet ved hjælp af så kaldt *Lagrange mekanik*. Mange andre fysiske systemer med *begrænset frihed*, så som "klods på klods systemet", er ligeledes mindre tunge at analysere i rammer af Lagrange mekaniken.
2. De alternative formuleringer kan give ny indsigt i allerede kendt fænomener. Et godt eksempel er forbindelsen mellem fysiske *bevægelses love* og *symetrierne* af et system, hvilket også kan udfoldes ved hjælp af så kaldt *Lagrange mekanik*.

3. De alternative formuleringer kan være mere hensigtsmæssige i forhold til at opdage og formulere nye fysiske love. Her er det primære eksempel *kvantemekanikken*, hvis formulering er stærkt indspireret af såkaldt *Hamilton mekanik*. Et andet eksempel er formuleringen af den gren af fysiken der sammenkobler mekanik og termodynamik; nævnlige den *statistiske fysik*, hvor mange koncepter fra Hamilton mekanik er komplet uunværlige.

Det kan desuden tilføjes, at både Lagrange og Hamilton dynamik er langt mere velegnede som fundamenter for mange grene af den moderne fysik, et par eksempler på hvilke er.

- Faststof fysik (*kan beskrive fx laserer og transistorer*)
- Partikelfysik (*førte til opdagelsen af Higgs partiklen samt resten af Standardmodellen*)
- General relativitets-teori (*kan beskrive fx sort huller og universets udvidelse*)
- Streng-teori (*en af de få kendte teori, som kan beskrive tyngdekraft kvantemekanisk*)

Disse argumenter burde være tilstrækkelige til at overbevise selv de mest konservativt stemte, om anvendeligheden af de alternative formuleringer for den klassiske fysik. Med det sagt, vil vi nu give en kort oversigt over indholdet af disse noter:

- **Kapitel 1:** Motivation for indførelsen af *variations-regning*, samt udvikling af vigtige koncepter indenfor gren af matematikken.
- **Kapitel 2:** Introduktion til *Lagrange mekanik* samt *princippet om mindste virkning*. Vi vil desuden diskutere sammenhængen mellem *symmetrier* og *bevarelses love*.
- **Kapitel 3:** Introduktion til *Hamilton mekanik* og idéen om *fase-rum*. Vi vil kort beskrive hvordan Hamilton mekanik ledte til en af de første formuleringer af *kvante mekanik*.

Vi er nu klar til begynde vores rejse fra Newton til kvante mekanik. Vi tager udgangspunkt i et interessant princip, udtænkt af den franske matematiker og advokat Pierre de Fermat.

Kapitel 1

Fermat's princip og variationsregning

I 1662, 25 år før udgivelsen af *Philosophiæ Naturalis Principia Mathematica*, arbejdede matematikeren Pierre de Fermat på at forstå afbøjning af lys, passerende fra et medium til et andet. Det havde på dette tidspunkt været kendt i mere end 600 år, at normal-vinklen for en indkommende lysstråle og normal-vinklen for den tilhørende udgående lysstråle er relateret på følgende måde (*se fig. 1.1 for illustration*).

$$\frac{\sin(\theta_1)}{\sin(\theta_2)} = \frac{n_2}{n_1} \quad (1.1)$$

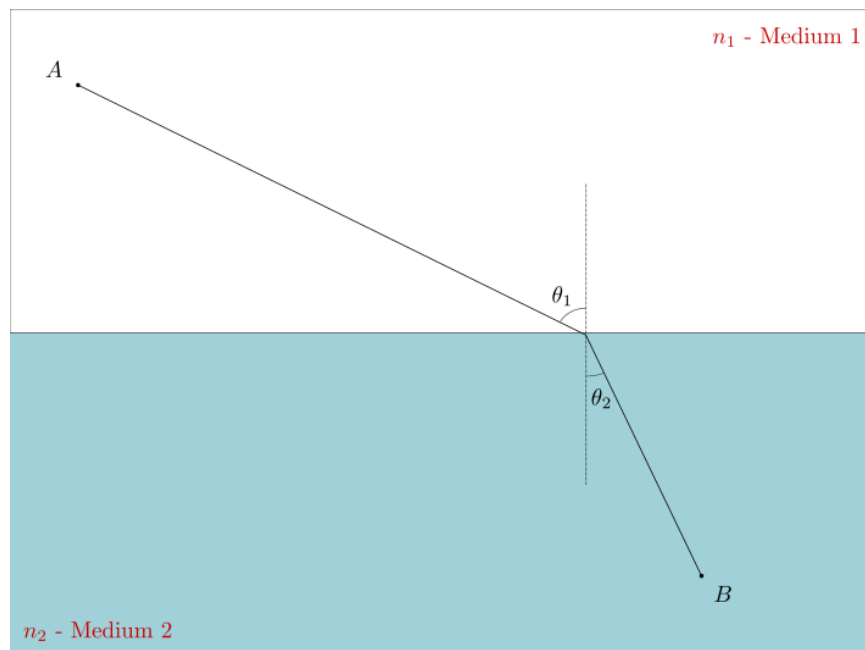
Størrelserne n_1 , n_2 er de såkaldte *brydnings-indekser* for medium 1 og 2, og disse kan bestemmes på baggrund af de elektro-magnetiske egenskaber for de givne medier. Brydningsindekserne relaterer lyshastighederne i medium 1 og 2 til lyshastigheden i vakuum.

$$v_1 = \frac{c}{n_1} \quad , \quad v_2 = \frac{c}{n_2} \quad (1.2)$$

Hvor $c \approx 10^8$ m/s, er lyshastigheden i vakuum. Sammenhængen (1.1) er typisk kendt under navnet *Snell's lov*. Fermat's bidrag til forståelsen af brydning idet lys passerer fra et medium til et andet, var formuleringen af følgende princip:

Princip (Fermat). *En lysstråle der udsendes fra punktet A til tiden t_A og dedekteres ved punktet B til tiden t_B , vil følge den vej fra A og B som minimerer størrelsen $T \equiv t_B - t_A$. Med andre ord: lyset følger den vej fra A til B som tager kortest tid.*

Lad os i første omgang påpeje, at det ovenstående Fermat's princip er konsistent med vores erfaring om at lys følger rette linjer gennem rummet, hvis det ikke passerer fra et medium



Figur 1.1: Eksempel på afbøjning af lys, passerende fra et medium til et andet.

til et andet. Dette er nemt at se, siden der for enhver vej gælder at lyshastigheden v i det givne medium er konstant. Dermed er tiden T det vil tage at gennemløbe en hver vej i mediet, givet ved følgende.

$$T = \frac{L}{v} \quad (1.3)$$

Hvor L er længden af vejen. Vi ser dermed at T er minimal når L er minimal, og siden den korteste vej mellem to punkter er en ret linje har vi overensstemmelse. Det er ligeledes muligt at udlede Snell's lov (1.1) ved brug af Fermat's princip om minimering af rejsetiden. Dette efterlades, med lidt vejledning, som en opgave til læseren 1.2.

Lad os nu påpeje en interessant konceptuel forskel på Snell's lov og Fermat's princip. Snell's lov beskriver hvad lyset gør *lokalt*; nævnligt hvordan det abøjes i passagen fra et medium til et andet. Derimod beskriver Fermats princip en *global* egenskab ved den vej lyset følger; nævnligt at denne vej minimere rejsetiden fra A til B . Fermat's princip kan altså tænkes på som en form for minimeringsproblem, så lad os forsøge at kaste det i en mere præcis matematiks form.

1.1 Matematisering af problemet

Lad os i første omgang definere en passende data-struktur til beskrivelsen af en lysstråles propagation gennem rummet.

Definition 1. En bane-kruve er en funktion $\gamma : \mathbb{R} \rightarrow \mathbb{R}^3$, der til et hvert tidspunkt $t \in \mathbb{R}$ beskriver placeringen $\gamma(t)$ af en given lysfront.

Vi vil antage at alle relevante bane-kruver er p gange kontinuert differentiabel. Vi siger at kurven γ er et element i sættet $C^p(\mathbb{R}, \mathbb{R}^3)$, bestående af p gange kontinuert differentiabel funktioner fra $\mathbb{R} \rightarrow \mathbb{R}^3$.

Vi vil indtil et senere punkt i vores analyse lade p være ubestemt. Det vil på nuværende tidspunkt være passende at indføre en ofte anvendt notation; den såkaldte "mærke notation". Vi vil med γ' , benævne den 1. afledte af funktionen γ . Vi vil ligeledes med γ'' , benævne den 2. afledte osv. Vi vil desuden med $|\gamma'(t)| := \sqrt{\gamma'(t) \cdot \gamma'(t)}$, benævne normen af hastighedsvektoren $\gamma'(t)$, defineret udfra det sædvanlige skalar-produkt i \mathbb{R}^3 . Vi er nu klar til næste definition.

Definition 2. Lad os først indføre brydning-indeks funktionen $n : \mathbb{R}^3 \rightarrow \mathbb{R}$, der som input tager et hvert punkt x i rummet (\mathbb{R}^3), og som output giver materialets brydnings-indeks $n(x) \in \mathbb{R}$ i det givne punkt $x \in \mathbb{R}^3$.

For enhver tilladt bane-kruve γ , og et hvert punkt $\gamma(t)$ på denne bane-kurve, vil hastigheden $\gamma'(t)$ i dette punkt være relateret til brydning-indekset $n(\gamma(t))$ i samme punkt:

$$n(\gamma(t)) = \frac{c}{|\gamma'(t)|} \quad \text{for alle } t \in \mathbb{R} \quad (1.4)$$

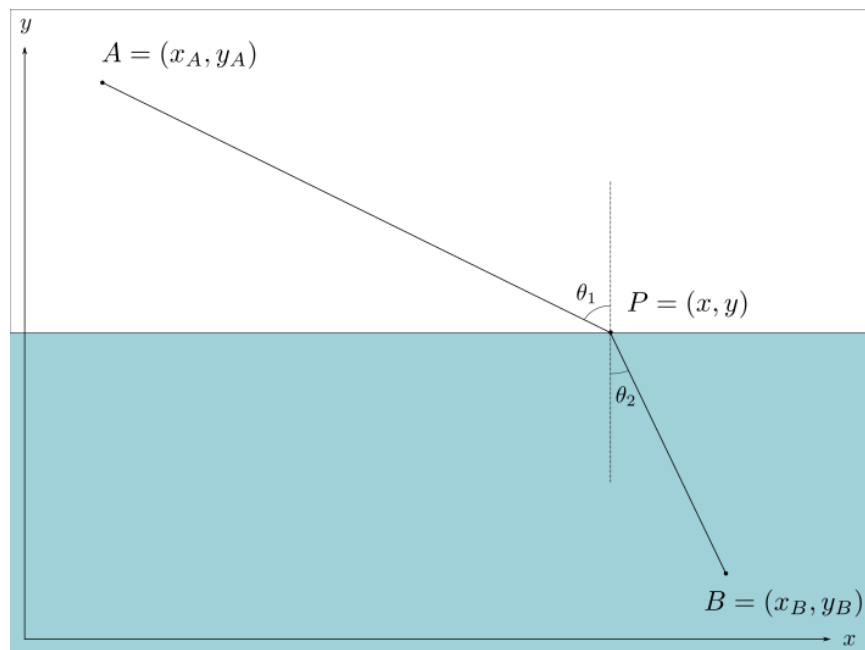
Ovenstående sammenhæng er analog til (1.2) for tilfældet med konstante brydnings-indekser.

Vi kan nu indføre den sidste, og mest centrale, størrelse i vores matematiserings process.

Definition 3. Vi definere rejse-tids funktionen $T : C^p(\mathbb{R}, \mathbb{R}^3) \rightarrow \mathbb{R}$, således at den som input tager en bane-kurve γ , for hvilken $\gamma(t_A) = A$ og $\gamma(t_B) = B$, og som output giver rejse-tiden fra punktet A til punktet B langs γ .

$$T[\gamma] = \int_{t_A}^{t_B} dt = \int_{t_A}^{t_B} n(\gamma(t)) \frac{|\gamma'(t)|}{c} dt = \int_{t_A}^{t_B} \frac{n(\gamma(t))}{c} \sqrt{\gamma'(t) \cdot \gamma'(t)} dt \quad (1.5)$$

Hvor vi har anvendt (1.4) til at indsætte et passende 1 i det ovenstående integral.



Figur 1.2: Propagation af lys gennem to medier, med overgangspunkt $P = (x, y)$.

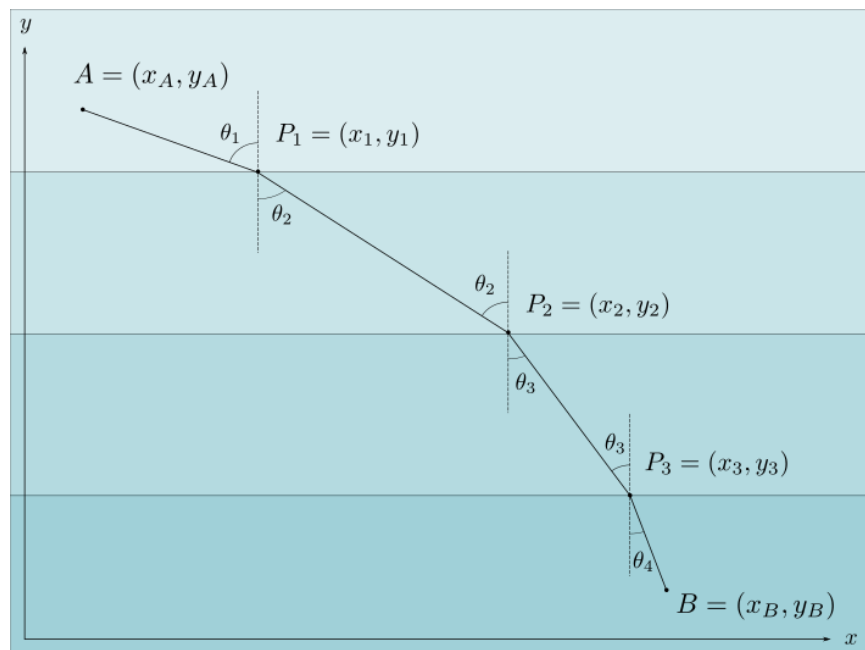
Vi kan nu give en mere matematisk præcis reformulering af Fermat's princip, ved hjælp af vores nyeligt indførte rejse-tids funktion T .

Princip (Fermat, Ref.). *En lysstråle der udsendes fra punktet A til tiden t_A og dedekteres ved punktet B til tiden t_B , vil følge kurven $\gamma \in C^p(\mathbb{R}, \mathbb{R}^3)$, således at $T[\gamma]$ er minimal.*

Problemmet om at finde den vej lyset vil rejse fra et punkt A til et punkt B , er altså ækvivalent med at minimere funktionen T over mængden af alle kurver i $C^p(\mathbb{R}, \mathbb{R}^3)$. Hvordan man rent matematisk bærer sig ad med at udføre en sådan minimering er et interessant spørgsmål, som vi snart skal se har stor relevans for alternative formuleringer af den klassiske mekanik.

1.2 Opgave: udledning af Snell's lov

Anvend Fermat's princip til at udlede sammenhængen (1.1), idet lys passerer fra et medium med brydningsindeks n_1 til et andet medium med brydningsindeks n_2 . Opgaven kan med fordel løses i følgende trin:



Figur 1.3: Propagation af lys gennem flere medier, med overgangspunkter $P_i = (x_i, y_i)$.

- a) Opskriv et udtryk for rejse-tiden T fra A til B , som funktion x og andre indgående parametre (se fig. 1.2).
- b) Vi har nu et minimerings problem i én reel variable x , og vi kan opstille ligningen:

$$\frac{dT}{dx} = 0 \quad (1.6)$$

Omskriv nu ovenstående ligning i termer af n_1 , n_2 , $\sin(\theta_1)$ og $\sin(\theta_2)$. Fig. 1.2 kan med fordel anvendes som reference for geometriske betragtninger.

- c) Generaliser til tilfældet hvor en lysstråle passere gennem flere medier med forskellige brydnings-indekser. Opstil rejse-tiden T som funktion af x_1, x_2, \dots, x_n . Opstil derefter ligningerne:

$$\frac{\partial T}{\partial x_i} = 0 \quad , \quad i = 1, 2, \dots, n \quad (1.7)$$

Omskriv ovenstående ligninger i termer af n_1, n_2, \dots, n_n og $\sin(\theta_1), \sin(\theta_2), \dots, \sin(\theta_n)$. Fig. 1.3 kan med fordel anvendes som reference for geometriske betragtninger.

1.3 Ekstremering af funktionaler

En funktion af typen $T : C^p(\mathbb{R}, \mathbb{R}^n) \longrightarrow \mathbb{R}$, der som input tager en anden funktion af typen $\gamma : \mathbb{R} \longrightarrow \mathbb{R}^n$ for $n \in \mathbb{N}$, kaldes ofte for en *funktional*. Der findes utallige situationer inden for både matematikken og fysikken, hvor man har brug for at finde ekstremal-funktioner for en given funktional, og det er derfor er stor praktisk interesse at kunne finde disse ekstremal-funktioner på en systematisk måde. For en generisk funktional er dette dog et meget svært problem, og en generel fremgangsmåde er, om ikke andet for denne forfatter, ikke kendt. Det viser sig imidlertid at der findes en mere medgørlig delmængde af alle funktionaler, for hvilke der findes en konkret strategi for bestemmelse af ekstremal-funktioner. Denne delmængde er funktionaler af følgende type.

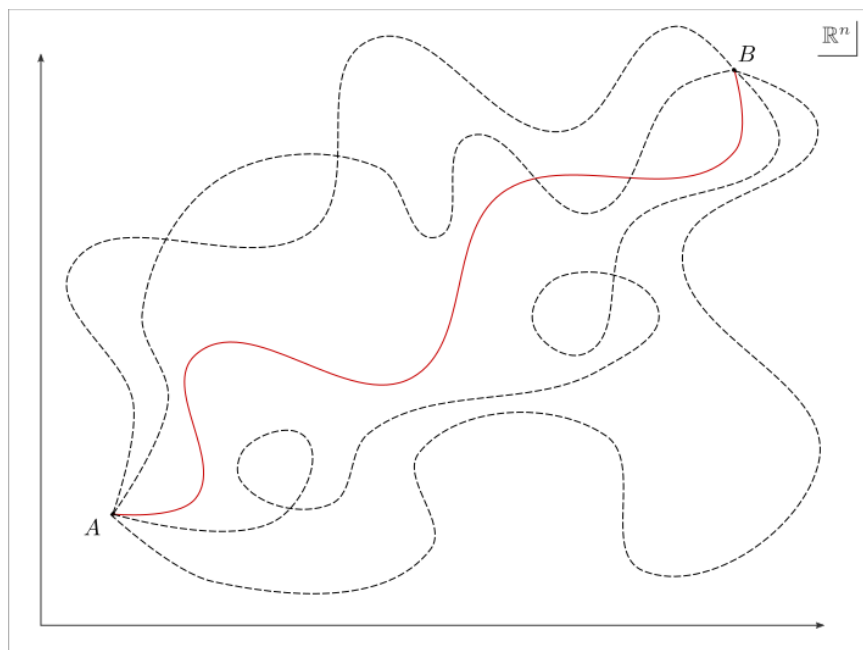
$$F[\gamma] = \int_{t_A}^{t_B} L(\gamma(t), \gamma'(t), t) dt \quad (1.8)$$

Hvor funktionen $L : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \longrightarrow \mathbb{R}$, kaldes funktionalens *Lagrange-funktion*. Funktionaler af ovenstående type kan for den uindviede læser virke som en meget begrænset delmængde af alle funktionaler. Det viser sig dog at et overvældende flertal af fysisk relevante funktionaler falder indenfor denne klasse. Vi har allerede set ét eksempel på en sådan fysisk relevant funktional; nævnligt rejse-tids funktionalen T , som præcis er på formen (1.8), med $n = 3$ og Lagrange-funktion:

$$L(x, v, t) := \frac{n(x)}{c} \sqrt{v \cdot v} \quad (1.9)$$

Vi vil senere i dette kapitel diskutere flere eksempler på fysiske problemstillinger, som involvere bestemmelse af ekstremal-funktioner for funktionaler af typen (1.8). I første omgang må vi dog diskutere præcis hvordan det lader sig gøre at bestemme ekstremal-funktioner for funktionaler af typen (1.8). I analog til problemet om bestemmelse af ekstremal-værdier for funktioner af typen $f : \mathbb{R} \longrightarrow \mathbb{R}$, vil vi som udgangspunkt fokusere på bestemmelse af *stationære* input for en given funktional F . For en funktion af typen $f : \mathbb{R} \longrightarrow \mathbb{R}$, kan stationærhed defineres på følgende måde:

$$x \text{ er et stationært punkt for } f \iff \left. \frac{df(x + \varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.10)$$

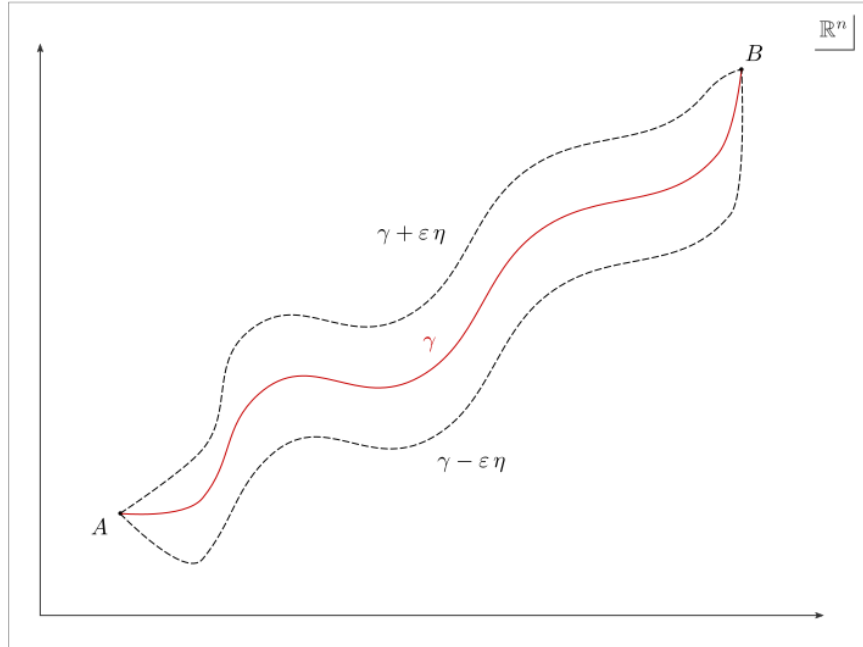


Figur 1.4: Eksempler på mulige bane-kurver gennem \mathbb{R}^n fra A til punkt B . Stiplede kurver repræsenterer ikke-stationære baner mht. funktionen F , mens den røde kurve repræsenterer en stationær bane.

Men hvad præcis udgøre en brugbar definition af stationærhed for funktioner? Svaret på dette spørgsmål vil vi nu forsøge at motivere så klart som muligt, og derefter udledningen hvordan stationære funktioner, for en type (1.8) funktional, kan bestemmes ved hjælp af Euler-Lagrange ligningerne for en sådan funktional.

1.4 Stationære funktioner: Euler-Lagrange ligninger

Lad os først forsøge at motivere definitionen af stationærhed for (1.8) funktioner. Antag først at der findes en funktion $\gamma \in C^p(\mathbb{R}, \mathbb{R}^n)$ for hvilken F er stationær, hvad end det vil sige. Vælg nu en *perturbations-funktion* $\eta \in C^p(\mathbb{R}, \mathbb{R}^n)$, for hvilken $\eta(t_A) = 0$ og $\eta(t_B) = 0$. Vi vil antage at $\eta(t_A) = \eta(t_B) = 0$, grundet at man som oftest ønsker at optimere F over delmængden $\{\gamma \in C^p(\mathbb{R}, \mathbb{R}^n) \mid \gamma(t_A) = A \text{ og } \gamma(t_B) = B\} \subset C^p(\mathbb{R}, \mathbb{R}^3)$. Dette var for eksempel tilfældet for Fermat's princip, hvor vi fokuserede på lysstråler startende fra punktet A og slutende i punktet B .



Figur 1.5: Eksempel på en variation af banen γ , med endepunkterne A og B , ved brug af funktionen $\eta \in C^p(\mathbb{R}, \mathbb{R}^n)$, hvor $\eta(t_A) = \eta(t_B) = 0$ og $\varepsilon \in \mathbb{R}$.

For en given stationær funktion γ og perturbations-funktion η , kan $F[\gamma + \varepsilon \eta]$ betragtes som en funktion fra $\mathbb{R} \rightarrow \mathbb{R}$, og $F[\gamma + \varepsilon \eta]$ kan dermed differentieres på sædvanlig vis. Et oplagt krav til en hver fornuftig definition af stationærhed for F vil dermed være:

$$\gamma \text{ er en stationær funktion for } F \quad \Rightarrow \quad \left. \frac{dF[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.11)$$

Taylor-udviklingen af funktionen $F[\gamma + \varepsilon \eta]$ omkring inputet γ , har altså intet 1. ordensled.

$$F[\gamma + \varepsilon \eta] = F[\gamma] + \mathcal{O}(\varepsilon^2) \quad (1.12)$$

Hvilket præcist er tilfældet for en funktion af typen $f : \mathbb{R} \rightarrow \mathbb{R}$, med stationært punkt x .

$$f(x + \varepsilon) = f(x) + \mathcal{O}(\varepsilon^2) \quad (1.13)$$

Den opmærksomme læser vil måske have bemærket, at implikationspil i (1.11) kun går den ene vej. Dette skyldes at der sagtens kunne findes funktioner η og ξ , opfyldende $\eta(t_A) = \eta(t_B) = 0$ og $\xi(t_A) = \xi(t_B) = 0$, således at.

$$\left. \frac{dF[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad , \quad \left. \frac{dF[\gamma + \varepsilon \xi]}{d\varepsilon} \right|_{\varepsilon=0} \neq 0 \quad (1.14)$$

Det vil i dette tilfælde ikke give mening at kalde γ for en stationær funktion for F , af den simple årsag at $F[\gamma + \varepsilon \xi]$ ikke er stationær for $\varepsilon = 0$. Løsningen på dette problem er simpelt. Vi definerer en funktion γ som værende en stationær funktion for funktionalen F , hvis og kun hvis følgende betingelse er opfyldt.

$$\gamma \text{ er en stationær funktion for } F \quad \Leftrightarrow \quad \forall \eta \in C^p(\mathbb{R}, \mathbb{R}^n) : \quad \left. \frac{dF[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.15)$$

Hvor det i ovenstående definition stadig er påkrævet at $\eta(t_A) = \eta(t_B) = 0$. Med denne definition er vi sikret at (1.12) er opfyldt, ligegyldigt hvilken perturbations-funktion η vi benytter. Lad os nu forsøge at anvende definitionen (1.15) til at bestemme de stationære funktioner for F . Ved anvendelse af (1.8) finder vi at.

$$\left. \frac{dF[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} = \int_{t_A}^{t_B} \left. \frac{dL(\gamma(t) + \varepsilon \eta(t), \gamma'(t) + \varepsilon \eta'(t), t)}{d\varepsilon} \right|_{\varepsilon=0} dt \quad (1.16)$$

For fast t kan vi betragte $\gamma(t) + \varepsilon \eta(t)$ og $\gamma'(t) + \varepsilon \eta'(t)$ som funktioner fra $\mathbb{R} \rightarrow \mathbb{R}^n$, hvor ε spiller rollen som uafhængig variabel. Da Lagrange-funktionen L tager $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$, sådan at $(x, v, t) \mapsto L(x, v, t)$, har vi i ovenstående udtryk brug for at anvende *kæde-reglen* for funktioner af mere end én reel variable. Lad os til dette formål indføre følgende notation:

$$x = (x_1, x_2, \dots, x_i, \dots, x_n) \in \mathbb{R}^n \quad , \quad v = (v_1, v_2, \dots, v_i, \dots, v_n) \in \mathbb{R}^n \quad (1.17)$$

$$\gamma(t) := (\gamma_1(t), \gamma_2(t), \dots, \gamma_i(t), \dots, \gamma_n(t)) \quad , \quad \gamma_i : \mathbb{R} \rightarrow \mathbb{R} \quad (1.18)$$

$$\eta(t) := (\eta_1(t), \eta_2(t), \dots, \eta_i(t), \dots, \eta_n(t)) \quad , \quad \eta_i : \mathbb{R} \rightarrow \mathbb{R} \quad (1.19)$$

Lad os for at øjeblik indføre de to hjælpe-størrelser $z = \gamma(t) + \varepsilon \eta(t)$ og $w = \gamma'(t) + \varepsilon \eta'(t)$, for at lette notationen. Ved brug af kædereolen for funktioner af flere variable på integranden i udtryk (1.16), finder vi nu følgende udtryk.

$$\frac{dL(z, w, t)}{d\varepsilon} = \sum_{i=1}^n \left[\frac{\partial L(z, w, t)}{\partial x_i} \frac{dz_i}{d\varepsilon} + \frac{\partial L(z, w, t)}{\partial v_i} \frac{dw_i}{d\varepsilon} \right] \quad (1.20)$$

Ud fra definitionerne $z = \gamma(t) + \varepsilon \eta(t)$ og $w = \gamma'(t) + \varepsilon \eta'(t)$, ses det nu umiddelbart at:

$$z|_{\varepsilon=0} = \gamma(t) \quad , \quad \left. \frac{dz_i}{d\varepsilon} \right|_{\varepsilon=0} = \eta_i(t) \quad (1.21)$$

$$w|_{\varepsilon=0} = \gamma'(t) \quad , \quad \left. \frac{dw_i}{d\varepsilon} \right|_{\varepsilon=0} = \eta'_i(t) \quad (1.22)$$

Ved brug af ovenstående udtryk, samt (1.20), kan vi nu opskrive følgende udtryk for den afledte af $F[\gamma + \varepsilon \eta]$ evalueret ved $\varepsilon = 0$.

$$\left. \frac{dF[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} = \int_{t_A}^{t_B} \sum_{i=1}^n \left[\frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial x_i} \eta_i(t) + \frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial v_i} \eta'_i(t) \right] dt \quad (1.23)$$

Ved delvis integration af andet led, og ved brug af $\eta(t_A) = \eta(t_B) = 0$, finder vi nu følgende.

$$\begin{aligned} \left. \frac{dF[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} &= \int_{t_A}^{t_B} \sum_{i=1}^n \left[\frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial x_i} - \frac{d}{dt} \frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial v_i} \right] \eta_i(t) dt \\ &\quad + \sum_{i=1}^n \left[\frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial v_i} \eta_i(t) \right]_{t_A}^{t_B} \end{aligned} \quad (1.24)$$

$$\Rightarrow \left. \frac{dF[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} = \int_{t_A}^{t_B} \sum_{i=1}^n \left[\frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial x_i} - \frac{d}{dt} \frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial v_i} \right] \eta_i(t) dt \quad (1.25)$$

Hvis ovenstående udtryk, per definition (1.15), skal være 0 for alle $\eta \in C^p(\mathbb{R}, \mathbb{R}^n)$ med $\eta(t_A) = \eta(t_B) = 0$, er vi nu tvunget til at konkludere, at kombinationen af partielt afledte i den firkantede parentes må være 0 for alle $t \in \mathbb{R}$.

$$\frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial x_i} - \frac{d}{dt} \frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial v_i} = 0 \quad , \quad i = 1, \dots, n \quad (1.26)$$

Dette sæt af koblede ordinære differential-ligninger er Euler-Lagrange ligningerne for funktionalen F givet ved (1.8). Vi ser på dette stadige i vores analyse, at $p \geq 2$, samt at Lagrange-funktionen L må kræves at være mindst én gang kontinuert differentiable.

1.5 Analyse af rejse-tids funktionalen

Lad os nu anvende vores nye værktøj, Euler-lagrange ligningerne, til at løse vores motiverende problem; nævnlige at finde de stationære funktioner for rejse-tids funktionalen T :

$$T[\gamma] = \int_{t_A}^{t_B} L(\gamma(t), \gamma'(t), t) dt \quad , \quad L(x, v, t) = \frac{n(x)}{c} \sqrt{v \cdot v} \quad (1.27)$$

Vi finder nu de partielt afledte af Lagrange-funktionen. Resultatet er følgende udtryk.

$$\frac{\partial L(x, v, t)}{\partial x_i} = \frac{1}{c} \frac{\partial n(x)}{\partial x_i} \sqrt{v \cdot v} \quad , \quad \frac{\partial L(x, v, t)}{\partial v_i} = \frac{n(x)}{c} \frac{v_i}{\sqrt{v \cdot v}} \quad (1.28)$$

Vi kan nu indsætte de ovenstående udtryk for de partielt afledte i Euler-Lagrange ligningerne (1.26). Resultatet er følgende differential-ligninger.

$$\frac{d}{dt} \left[n(\gamma(t)) \frac{\gamma'_i(t)}{\sqrt{\gamma'(t) \cdot \gamma'(t)}} \right] = \frac{\partial n(\gamma(t))}{\partial x_i} \sqrt{\gamma'(t) \cdot \gamma'(t)} \quad (1.29)$$

Alt vi nu, i principet, skal gøre for at de stationære funktionen for rejse-tids funktionalen T , er at løse det ovenstående sæt af differential-ligninger, med randbetingelserne $\gamma(t_A) = A$ og $\gamma(t_B) = B$. I praksis er det dog stadig et svært problem at løse disse ligninger, med mindre $n(x)$ er ekstraordinært simpel. Vi vil derfor nu diskutere et par koncepter, som signifikant letter vores analyse af rejse-tids funktionalen.

1.5.1 Hamilton-funktionen

En hver person, som har beskæftiget sig med problem analyse i den klassiske fysik, ved at bevarelses-love, for eksempel for impuls og energi, typisk simplificere en et givent problem dræstisk. På samme måde kan bevarelses-love simplificere vores funktional analyse. Specifikt kan det vises, at givet den såkladte *Hamilton-funktion* $H : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$:

$$(x, v, t) \mapsto H(x, v, t) := \sum_{i=1}^n \left[v_i \frac{\partial L(x, v, t)}{\partial v_i} \right] - L(x, v, t) \quad (1.30)$$

Da vil $H(\gamma(t), \gamma'(t), t)$ være en bevaret størrelse, hvis følgende 2 betingelser er opfyldte:

1. Funktionen γ er en løsning til Euler-Lagrange ligningerne (1.26).
2. Lagrange-funktionen L afhænger ikke explicit af t . Med andre ord:

$$\frac{\partial L(x, v, t)}{\partial t} = 0 \quad (1.31)$$

Lad os nu demonstrerer at $H(\gamma(t), \gamma'(t), t)$ er en bevaret størrelse, givet at de ovenstående betingelser er opfyldt. Ved brug af multi variablel kæde-reglen finder vi at:

$$\begin{aligned} \frac{dH(\gamma(t), \gamma'(t), t)}{dt} &= \sum_{i=1}^n \left[\gamma''(t) \frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial v_i} + \gamma''(t) \frac{d}{dx} \frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial v_i} \right] \\ &\quad - \sum_{i=1}^n \left[\frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial x_i} \gamma'(t) + \frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial v_i} \gamma''(t) \right] \\ &\quad - \frac{\partial L(\gamma(t), \gamma'(t), t)}{\partial t} \end{aligned} \quad (1.32)$$

Ved anvendelse af Euler-Lagrange ligningerne (1.26) samt vores antagelse om explicit t uafhængighed (1.31), kan vi nu konkludere at:

$$\frac{dH(\gamma(t), \gamma'(t), t)}{dt} = 0 \quad \Rightarrow \quad H(\gamma(t), \gamma'(t), t) = C \quad , \quad C \in \mathbb{R} \quad (1.33)$$

Vi kunne nu i princippet udregne Hamilton-funktionen for rejse-tids funktionalen T . På det stadige af vores analyse ville dette dog ikke hjælpe os meget. Dette skyldes at Hamilton-funktionen, udregnet på baggrund af Lagrange-funktionen (1.27), er identisk 0. Denne udregning er simpel, og vi opfordre læseren til udføre udregningen ved brug af definition (1.30). Årsagen til dette lidt mystiske resultat, skyldes i sidste ende såkaldt *reparameteriserings invariants* af rejse-tids funktionalen. Vi vil på dette tidspunkt ikke diskutere om denne form for invariants, da problemet med Hamilton-funktionen forsvinder efter indførelsen af vores næste simplificerende koncept.

1.5.2 Dimensionel reduktion

Når det kommer til evaluering af integraler, kan det ofte betale sig at foretage et passende koordinat-skift. Det samme kan siges ift. funktional analyse, hvor et passende koordinat-skift potentielt kan simplificere Lagrange-funktionen. Idéen bag såkaldt *dimensional reduktion*, er at foretage følgende skifte af integrations-variable.

$$x_1 = \gamma_1(t) \quad \Rightarrow \quad t = \gamma_1^{-1}(x_1) \quad (1.34)$$

Hor γ er input funktionen for den givne funktional, og γ_1 er 1. komponenten af γ . For at kunne foretage det ovenstående variable skift, må vi antage at γ_1 er bijektiv; altså at γ_1^{-1} eksisterer. Dette er ækvivalent med forudsætningen om at γ_1 er en monotont-voksende funktion. Dette er en yderst rimelig forudsætningen i mange sammenhænge, som vi snart skal se. Lad os først foretage variable skiftet fra t til x_1 i funktionalen:

$$T[\gamma] = \frac{1}{c} \int_{t_A}^{t_B} n(\gamma(t)) \sqrt{\gamma'(t) \cdot \gamma'(t)} dt \quad (1.35)$$

Vi kan nu med fordel definere nye bane-kurve funktioner $\tilde{\gamma} : \mathbb{R} \rightarrow \mathbb{R}^n$, sådan at $\tilde{\gamma} = \gamma \circ \gamma_1^{-1}$. Dermed ses det umiddelbart at:

$$\gamma(t) = \tilde{\gamma} \circ \gamma_1(t) = \tilde{\gamma}(\gamma_1(t)) = \tilde{\gamma}(x_1) \quad (1.36)$$

Ved brug af kæde-reglen finder vi desuden at γ' og $\tilde{\gamma}'$ er relateret på følgende simple måde.

$$\gamma'(t) = (\tilde{\gamma} \circ \gamma_1)'(t) = \tilde{\gamma}'(\gamma_1(t)) \gamma_1'(t) = \tilde{\gamma}'(x_1) \gamma_1'(t) \quad (1.37)$$

Vi kan nu på sædvanlig vis finde Jacobaint-faktoren der relaterer de to mål dt og dx_1 .

$$x_1 = \gamma_1(t) \quad \Rightarrow \quad dx_1 = \gamma_1'(t) dt \quad (1.38)$$

Dermed finder vi at output-værdien $F[\gamma]$ for funktionen F , kan omkrives til følgende.

$$\begin{aligned} T[\gamma] &= \frac{1}{c} \int_{x_{1;A}}^{x_{1;B}} \frac{n(\tilde{\gamma}(x_1))}{\gamma_1'(t)} \sqrt{\gamma_1'(t)^2 \tilde{\gamma}'(x_1) \cdot \tilde{\gamma}'(x_1)} dx_1 \\ &= \frac{1}{c} \int_{x_{1;A}}^{x_{1;B}} n(\tilde{\gamma}(x_1)) \sqrt{\tilde{\gamma}'(x_1) \cdot \tilde{\gamma}'(x_1)} dx_1 \end{aligned} \quad (1.39)$$

Hvor $x_{1;A} = \gamma_1(t_A)$ og $x_{1;B} = \gamma_1(t_B)$. Observer nu at 1. komponenterne for $\tilde{\gamma}$ og $\tilde{\gamma}'$ effektivt set er elimineret som frihedsgrader, da de i følge (1.36, 1.37) altid antager følgende værdier:

$$\tilde{\gamma}_1(x_1) = \gamma_1 \circ \gamma_1^{-1}(x_1) = x_1 \quad , \quad \tilde{\gamma}_1'(x_1) = \frac{\gamma_1'(t)}{\gamma_1'(t)} = 1 \quad (1.40)$$

Bane-kruve funktionen $\tilde{\gamma}$ og dens afledte $\tilde{\gamma}'$ er altså effektivt set funktioner fra $\mathbb{R} \rightarrow \mathbb{R}^{n-1}$, hvor vi i dette tilfælde har $n = 3$. Vi har altså reduceret dimensionen af output-rummet for bane-kruve funktionerne med 1. Rejse-tids funktionen T kan nu ækvivalent beskrives i termer af $C^p(\mathbb{R}, \mathbb{R}^{n-1})$ bane-kruve funktioner. Resultatet er følgende reducerede udtryk:

$$T[\gamma] = \frac{1}{c} \int_{x_{1;A}}^{x_{1;B}} n(x_1, \gamma(x_1)) \sqrt{1 + \gamma'(x_1) \cdot \gamma'(x_1)} dx_1 \quad , \quad \gamma \in C^2(\mathbb{R}, \mathbb{R}^{n-1}) \quad (1.41)$$

Hvor randbetingelserne for input funktionerne nu er $(x_{1;A}, \gamma(x_{1;A})) = A$, $(x_{1;B}, \gamma(x_{1;B})) = B$. Lad os nu forsøge at finde Hamilton-funktionen H hørende til ovenstående funktional. Vi finder først de nødvendige partielt afledte af Lagrange-funktion L .

$$L(x, v, x_1) = \frac{n(x_1, x)}{c} \sqrt{1 + v \cdot v} \quad , \quad \frac{\partial L(x, v, x_1)}{\partial v_i} = \frac{n(x_1, x)}{c} \frac{v_i}{\sqrt{1 + v \cdot v}} \quad (1.42)$$

Hvor vi nu har $x = (x_2, x_3, \dots, x_n)$ og $v = (v_2, v_3, \dots, v_n)$. Vi kan nu finde et udtryk for Hamilton-funktionen H ved brug af definitionen (1.30):

$$H(x, v, x_1) = \frac{n(x_1, x)}{c} \frac{v \cdot v}{\sqrt{1 + v \cdot v}} - \frac{n(x_1, x)}{c} \sqrt{1 + v \cdot v} = -\frac{n(x_1, x)}{c} \frac{1}{\sqrt{1 + v \cdot v}} \quad (1.43)$$

Såfremt brydnings-indeksfunktionen $n(x_1, x)$ ikke afhænger af parameteren x_1 (*vi skriver i dette tilfælde $n(x) \equiv n(x_1, x)$*), har vi nu formået at finde en bevaret størrelse for det dimensionelt reducerede rejse-tids problem:

$$H(\gamma(x_1), \gamma'(x_1), x_1) = -\frac{n(\gamma(x_1))}{c} \frac{1}{\sqrt{1 + \gamma'(x_1) \cdot \gamma'(x_1)}} = C \quad , \quad C \in \mathbb{R} \quad (1.44)$$

For nu et dette enden på vores analyse af Fermat's princip / rejse-tids problemet. Motiveret af dette problem har vi udledt *Euler-Lagrange ligningerne*, vist at *Hamilton-funktion* kan give os en nyttig bevaret størrelse, og kort diskuteret *dimensionel reduktion*. Alle disse værktøjer vil visse sig uundværlige når vi i senere kapitler vil begynde introduktionen til alternative formuleringer af den klassiske mekanik. I resten af dette kapitel vil vi beskæftige os med at løse en række interessante variations-problemer ved brug af vores nye værktøjer.

1.6 Flere eksempler på variations-problemer

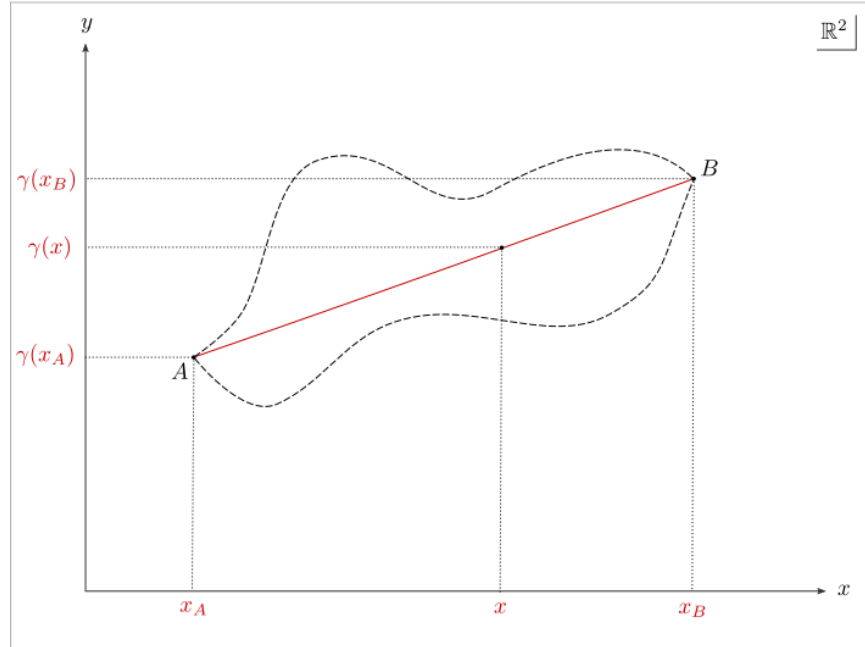
Vi vil nu analysere tre kendte variations-problemer. Alle tre problemer kan formuleres i 2D, hvilket vil sige at bane-kurverne γ er funktioner fra $\mathbb{R} \rightarrow \mathbb{R}^2$. Vi vil desuden se at alle tre problemer kan dimensionelt reduceres, sådan at vi kan betragte bane-kurve funktioner af typen $\gamma : \mathbb{R} \rightarrow \mathbb{R}$. Lad os nu for et kort øjeblik diskutere valg af notation. Når der arbejdes med problemer i planen \mathbb{R}^2 , benævner vi normalt et punkt som (x, y) i stedet for (x_1, x_2) . Desuden benævnes vektor-komponenter (v_x, v_y) i stedet for (v_1, v_2) . Med det sagt vil vi nu starte ud med at analysere det mest elementære problem indenfor variations-regningen.

Den korteste vej mellem to punkter: Vi vil demonstrere at den korteste vej mellem to punkter $A = (x_A, y_A)$ og $B = (x_B, y_B)$, er den rette linie som forbinder de to punkter. For en hver bane-kurve $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$, for hvilken $\gamma(t_A) = A$ og $\gamma(t_B) = B$, ved vi længden fra A til B målt langs γ , er givet ved følgende udtryk:

$$S[\gamma] = \int_{t_A}^{t_B} \sqrt{\gamma'(t) \cdot \gamma'(t)} dt \quad , \quad \gamma \in C^2(\mathbb{R}, \mathbb{R}^2) \quad (1.45)$$

Vi observerer nu at funktionalen S blot er et special tilfælde af rejse-tids funktional T (1.27), for hvilken:

$$\frac{n(\gamma(t))}{c} = 1 \quad (1.46)$$



Figur 1.6: Den røde bane-kurve γ repræsenterer den korteste vej fra A til B ; den rette linie. De stiplede bane-kurver repræsenterer ikke-stationære baner.

Vi kan derfor dimensionelt reducere funktional S , sådan at $S[\gamma]$ er givet som et specialtilfælde af (1.41).

$$S[\gamma] = \int_{x_A}^{x_B} \sqrt{1 + \gamma'(x)^2} dx \quad , \quad \gamma \in C^2(\mathbb{R}, \mathbb{R}) \quad (1.47)$$

Det er nu simpelt at identificere Lagrange-funktionen L for den ovenstående funktional.

$$L(y, v_y, x) := \sqrt{1 + v_y^2} \quad (1.48)$$

Vi kan nu anvende den ovenstående Lagrange-funktionen for S , til at opskrive Euler-Lagrange ligningerne for vores problem. Vi finder først de relevante partielt afledte for L .

$$\frac{\partial L(y, v_y, x)}{\partial y} = 0 \quad , \quad \frac{\partial L(y, v_y, x)}{\partial v_y} = \frac{v_y}{\sqrt{1 + v_y^2}} \quad (1.49)$$

Vi kan nu opskrive Euler-Lagrange ligningerne (1.26) for funktionalen S . Da vores bane-kurver $\gamma : \mathbb{R} \rightarrow \mathbb{R}$ kun har 1 komponent, vil vi for dette problem kun have 1 Euler-Lagrange ligning:

$$\frac{d}{dx} \frac{\partial L(\gamma(x), \gamma'(x), x)}{\partial v_y} = \frac{\partial L(\gamma(x), \gamma'(x), x)}{\partial y} \quad \Rightarrow \quad \frac{d}{dx} \left[\frac{\gamma'(x)}{\sqrt{1 + \gamma'(x)^2}} \right] = 0 \quad (1.50)$$

Det er nu en simpel sag at integrere ovenstående udtryk, hvormed vi finder følgende ligning.

$$\frac{\gamma'(x)}{\sqrt{1 + \gamma'(x)^2}} = k \quad \Rightarrow \quad \gamma'(x) = \pm \sqrt{\frac{k^2}{1 - k^2}} \quad (1.51)$$

Hvor $k \in (-1, 1)$ er en integrations konstant. Det er igen en simpel sag at integrere ovenstående udtryk. Resultatet er følgende.

$$\gamma(x) = ax + b \quad , \quad a \in \mathbb{R} \quad , \quad b \in \mathbb{R} \quad (1.52)$$

Integrations-konstanterne a, b kan nu bestemmes ved brug af de forudsatte randbetingelser:

$$\gamma(x_A) = y_A \quad , \quad \gamma(x_B) = y_B \quad (1.53)$$

Formen af en udspændt sæbehinde: Vi forestiller os nu en sæbehinde, udspændt mellem to cirkulære ringe som indikeret på figure 1.7. Grundet rotations-symmetri er det klart at den minimale sæbehinde må være et omdrejningsflade omkring x -aksen. Arealet af en sådant omdrejningsflade er da givet ved følgende integral.

$$A[\gamma] = \int_{t_A}^{t_B} 2\pi \gamma_y(t) \sqrt{\gamma'(t) \cdot \gamma'(t)} dt \quad , \quad \gamma \in C^2(\mathbb{R}, \mathbb{R}^2) \quad (1.54)$$

Vi observerer, også for dette eksempel, at funktionen A blot er et special tilfælde af rejse-tids funktional T (1.27), nu med:

$$\frac{n(\gamma(t))}{c} = 2\pi \gamma_y(t) \quad (1.55)$$

Vi kan derfor også dimensionelt reducere funktional A , sådan at $A[\gamma]$ er givet som et special tilfælde af (1.41).

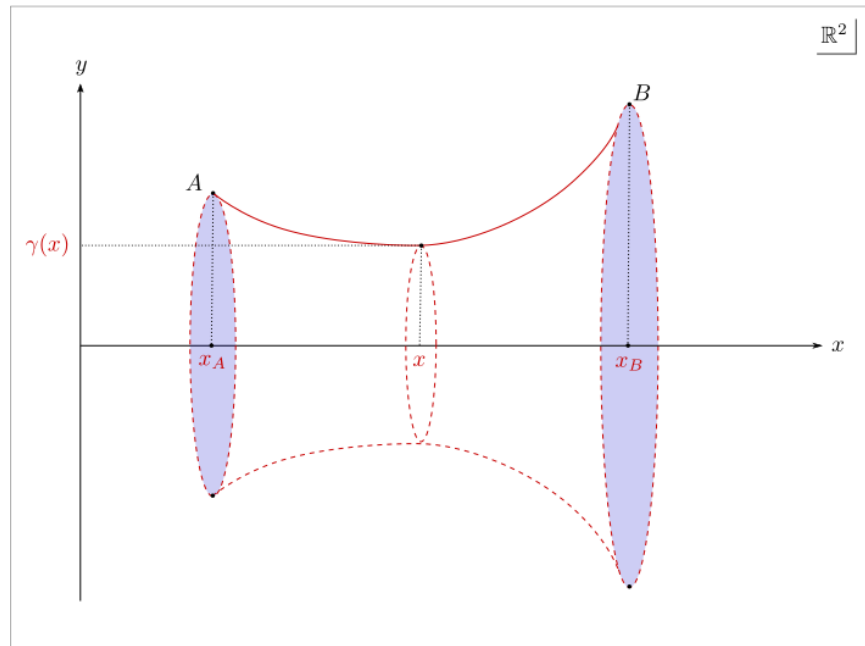
$$A[\gamma] = \int_{x_A}^{x_B} 2\pi \gamma(x) \sqrt{1 + \gamma'(x)^2} dx \quad , \quad \gamma \in C^2(\mathbb{R}, \mathbb{R}) \quad (1.56)$$

Det er nu simpelt at identificere Lagrange-funktionen L for den ovenstående funktional.

$$L(y, v_y, x) := \frac{n(y)}{c} \sqrt{1 + v_y^2} \quad , \quad \frac{n(y)}{c} = 2\pi y \quad (1.57)$$

Vi kan nu, i princippet, anvende den ovenstående Lagrange-funktionen for A , til at opskrive Euler-Lagrange ligningerne for vores problem. Lad os i stedet udnytte at $L(y, v_y, x)$ ikke afhænger af x , hvilket betyder at (1.44) med n givet ved (1.57), er en bevaret størrelse.

$$H(\gamma(x), \gamma'(x), x) = -\frac{2\pi \gamma(x)}{\sqrt{1 + \gamma'(x)^2}} = C \quad (1.58)$$



Figur 1.7: Den røde bane-kurve γ repræsenterer her meridian-kurven hørende til den minimale sæbehinde overflade. Punkterne A og B er valgt sådan, at deres separation i retningen vinkelret på (x, y) -planen er 0.

Hvor $C \in \mathbb{R}$ kan betragtes som en integrations-konstant. Ved brug af ovenstående udtryk, finder vi nu følgende 1. ordens differential-ligning for γ :

$$\gamma'(x) = \pm \sqrt{k^2 \gamma(x)^2 - 1} \quad , \quad k^2 \equiv \frac{4\pi^2}{C^2} \quad (1.59)$$

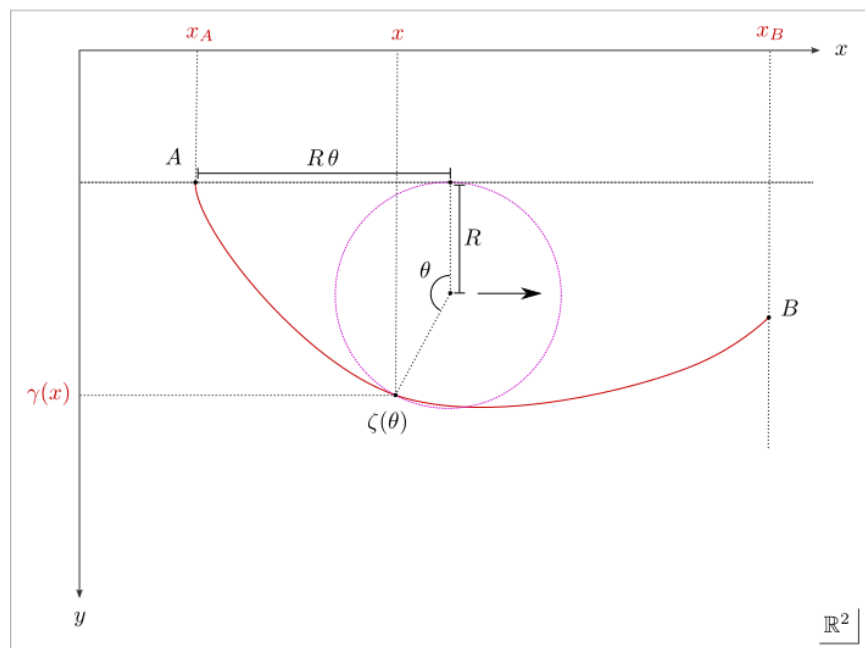
Da ovenstående ligning er separabel kan vi finde løsnings-kurven γ ved simpel integration:

$$\pm \int_{\gamma(x_A)}^{\gamma(x)} \frac{1}{\sqrt{k^2 z^2 - 1}} dz = \int_{x_A}^x dz \quad (1.60)$$

Hvor vi har indført integrations variabelen z . Integralet på højresiden er trivielt, mens integralet på venstresiden er givet i termer af den hyperbolske cosinus funktion \cosh . Vi finder dermed følgende resultat, hvor vi også har gjort brug af randbetingelsen $\gamma(x_A) = y_A$:

$$\frac{1}{k} \left[\cosh^{-1}(k\gamma(x)) - \cosh^{-1}(ky_A) \right] = x - x_A \quad (1.61)$$

Hvor $k \in \mathbb{R}$, og vi dermed ikke længere behøver \pm på venstresiden af det ovenstående udtryk.



Figur 1.8: Den røde bane-kurve γ repræsenterer her cykloide-stykket som forbinder punkterne A og B . Cykliden genereres af den lille cirkel med radius R , ved at følge punktet $\zeta(\theta)$ på periferien når cirklen rulles langs den stiplede linie, liggende parallelt med x -aksen.

Det er nu en simpel sag at isolere $\gamma(x)$ i ovenstående udtryk. Resultatet ser ud som følger.

$$\gamma(x) = \frac{\cosh[k(x - x_A) + \cosh^{-1}(k y_A)]}{k} \quad (1.62)$$

Konstanten k kan nu bestemmes ved at gøre brug af den tilbageværende randbetingelse:

$$\gamma(x_B) = y_B \quad (1.63)$$

Det brakistokrone problem: Vi forestiller os nu en situation, hvor to punkter A og B er forbundet med en krave γ . En partikel med masse m placeres nu i punktet A og bevæger sig under påvirkning af tyngdekraften langs γ til punktet B . Vi ønsker nu at bestemme γ , sådan at partiklens fald-tid er mindst mulig.

Vi kan først udnytte at den mekaniske energi i 1-partikel systemet er bevaret, hvilket medfører at vi kan ækvivalere energien for tiden t_A og tiden t .

$$\frac{1}{2}m|\gamma'(t_A)|^2 - mg\gamma_y(t_A) = \frac{1}{2}m|\gamma'(t)|^2 - mg\gamma_y(t) \quad (1.64)$$

Hvor g er tyngdeaccelerationen og $|\gamma'(t)| = \sqrt{\gamma'(t) \cdot \gamma'(t)}$ benævner normen af hastighedsvektoren $\gamma'(t)$. Antager vi at partiklen startes i hvile, sådan at $\gamma'(t_A) = 0$, og anvender randbetingelsen $\gamma_y(t_A) = y_A$, finder vi nu følgende udtryk for fald-tiden.

$$t_B - t_A = \int_{t_A}^{t_B} dt = \int_{t_A}^{t_B} \frac{1}{\sqrt{2g(\gamma_y(t) - y_A)}} \sqrt{\gamma'(t) \cdot \gamma'(t)} dt \quad (1.65)$$

Hvor vi har anvendt energi-bevarelsen (1.64) til at indsætte et passende 1, i det ovenstående integral. Fald-tids funktionalen D kan dermed beskrives ved følgende udtryk:

$$D[\gamma] = \int_{t_A}^{t_B} \frac{1}{\sqrt{2g(\gamma_y(t) - y_A)}} \sqrt{\gamma'(t) \cdot \gamma'(t)} dt \quad , \quad \gamma \in C^2(\mathbb{R}, \mathbb{R}^2) \quad (1.66)$$

Vi ser nu, at også fald-tids funktionalen D er et special tilfælde af rejse-tids funktional T (1.27), i dette tilfælde med:

$$\frac{n(\gamma(t))}{c} = \frac{1}{\sqrt{2g(\gamma_y(t) - y_A)}} \quad (1.67)$$

Vi kan derfor også dimensionelt reducere funktional D , sådan at $D[\gamma]$ er givet som et special tilfælde af (1.41).

$$D[\gamma] = \int_{x_A}^{x_B} \frac{1}{\sqrt{2g(\gamma(t) - y_A)}} \sqrt{1 + \gamma'(t)^2} dx \quad , \quad \gamma \in C^2(\mathbb{R}, \mathbb{R}) \quad (1.68)$$

Det er igen simpelt at identificere Lagrange-funktionen L for den ovenstående funktional.

$$L(y, v_y, x) := \frac{n(y)}{c} \sqrt{1 + v_y^2} \quad , \quad \frac{n(y)}{c} = \frac{1}{\sqrt{2g(y - y_A)}} \quad (1.69)$$

Vi udnytter nu endnu en gang at $L(y, v_y, x)$ ikke afhænger af x , hvilket betyder at (1.44) med $n(y)$ givet ved (1.69), er en bevaret størrelse.

$$H(\gamma(x), \gamma'(x), x) = -\frac{1}{\sqrt{2g(\gamma(x) - y_A)}} \frac{1}{\sqrt{1 + \gamma'(x)^2}} = C \quad (1.70)$$

Hvor $C \in \mathbb{R}$ igen kan betragtes som en integrations-konstant. Ved brug af ovenstående udtryk, finder vi nu følgende 1. ordens differential-ligning for γ :

$$\gamma'(x) = \pm \sqrt{\frac{1 - k(\gamma(x) - y_A)}{k(\gamma(x) - y_A)}} \quad , \quad k \equiv 2g C^2 \quad (1.71)$$

Da ovenstående ligning er separabel kan vi finde løsnings-kurven γ ved simpel integration:

$$\pm \int_{\gamma(x_A)}^{\gamma(x)} \sqrt{\frac{k(z - y_A)}{1 - k(z - y_A)}} dz = \int_{x_A}^x dz \quad (1.72)$$

Hvor vi har indført integrations varibalen z . Integralet på højresiden er igen trivielt, mens dette umiddelbart ikke er tilfældet for integralet på venstresiden. Dette integral kan dog bringes til en mere medgørlig form, ved at gennemføre følgende skift af integrations-variabel:

$$k(z - y_A) = \sin^2\left(\frac{\theta}{2}\right) \Rightarrow dz = \frac{1}{k} \sin\left(\frac{\theta}{2}\right) \cos\left(\frac{\theta}{2}\right) d\theta \quad (1.73)$$

Vi kan nu udregne venstreside-integralet i termer af den nyeligt indførte integrations-variabel θ . Resultatet bliver da følgende udtryk:

$$\pm \int_{\gamma(x_A)}^{\gamma(x)} \sqrt{\frac{k(z - y_A)}{1 - k(z - y_A)}} dz = \int_0^{\theta(x)} \frac{1}{k} \sin^2\left(\frac{\theta}{2}\right) d\theta = \frac{1}{2k} [\theta(x) - \sin(\theta(x))] \quad (1.74)$$

Hvor $k \in \mathbb{R}$, og vi dermed ikke længere behøver \pm i ovenstående udtryk. Vi har i udregning ovenfor har gjort brug af følgende dobbelt-vinkel formel:

$$\cos(2\alpha) = 1 - 2\sin^2(\alpha) \quad (1.75)$$

Desuden er størrelsen $\theta(x)$, på baggrund af (1.73), implicit defineret på følgende måde:

$$k(\gamma(x) - y_A) = \sin^2\left(\frac{\theta(x)}{2}\right) \quad (1.76)$$

Ved at kombinere (1.72, 1.74) og (1.75, 1.76), finder vi nu følgende trigonometriske udtryk for x og $\gamma(x)$, i termer af $\theta(x)$:

$$x = \frac{1}{2k} [\theta(x) - \sin(\theta(x))] + x_A \quad (1.77a)$$

$$\gamma(x) = \frac{1}{2k} [1 - \cos(\theta(x))] + y_A \quad (1.77b)$$

Vi kan nu i princippet løse for $\theta(x)$ i ligning (1.77a), og derefter indsætte det fundende udtryk for $\theta(x)$ i ligning (1.77b). Ser man på formen af (1.77a, 1.77b) virker det dog umiddelbart mere naturligt at beskrive vores løsning i form af en ny bane-krue $\zeta : \mathbb{R} \rightarrow \mathbb{R}^2$, med komponent-funktioner ζ_x og ζ_y , sådan at:

$$\zeta_x(\theta) = \frac{1}{2k} [\theta - \sin(\theta)] + x_A \quad , \quad \zeta_y(\theta) = \frac{1}{2k} [1 - \cos(\theta)] + y_A \quad (1.78)$$

Vores løsnings-kurve ζ , beskriver dermed et stykke af en *cykloide*, startende i punktet A og sluttende i punktet B (se figur 1.8), med radius:

$$R = \frac{1}{2k} \quad (1.79)$$

Konstanten $k \in \mathbb{R}$ kan nu bestemmes ved brug af den tilbageværende randbetingelse, som nu tager følgende form når beskrevet i termer af ζ :

$$\zeta_x(\theta_B) = x_B \quad , \quad \zeta_y(\theta_B) = y_B \quad (1.80)$$

1.7 Opgave: geodæter på hyperboloide

Vi søger i denne opgave den korteste vej mellem to punkter A og B på overfladen af en *hyperboloide*. En hyperboloide med strob-radius $R = 1$, kan beskrives som mængden af punkter (x, y, z) i \mathbb{R}^3 , der opfylder ligning:

$$x^2 + y^2 - z^2 = 1 \quad (1.81)$$

Ovenstående ligning for hyperboloide-overfalden, kan løses ved følgende parameterisering:

$$x = \cos(\phi) \cosh(\omega) \quad (1.82a)$$

$$y = \sin(\phi) \cosh(\omega) \quad (1.82b)$$

$$z = \sinh(\omega) \quad (1.82c)$$

Hvor $\omega \in \mathbb{R}$ og $\phi \in [0, 2\pi)$. Opgaven kan nu med fordel løses i følgende trin:

1. Anvend ligningerne (1.82), og vis at linie-elementet ds^2 på hyperboloide overfalden er beskrevet ved:

$$ds^2 = d\omega^2 + \cosh^2(\omega) d\phi^2 \quad (1.83)$$

Givet at linie-elementet i Euklidisk \mathbb{R}^3 er beskrevet ved:

$$ds^2 = dx^2 + dy^2 + dz^2 \quad (1.84)$$

Følgende identiteter kan være nyttige at have i dette trin.

$$\cos^2(\phi) + \sin^2(\phi) = 1 \quad , \quad \cosh^2(\omega) - \sinh^2(\omega) = 1 \quad (1.85)$$

2. Fra linie-elementet (1.83) kan vi nu konkludere, at længden $S[\gamma]$ målt langs en bane-
krue $\gamma : \mathbb{R} \longrightarrow H^2$ på hyperboloiden, fra punktet A til punktet B , er givet ved:

$$S[\gamma] = \int_{t_A}^{t_B} \sqrt{\gamma'_\omega(t)^2 + \cosh^2(\gamma_\omega(t)) \gamma'_\phi(t)^2} dt \quad , \quad \gamma \in C^2(\mathbb{R}, H^2) \quad (1.86)$$

Hvor $H^2 = \mathbb{R} \times [0, 2\pi)$. Efter dimensionel reduktion, med parameter-skift $\omega = \gamma_\omega(t)$, finder vi at $S[\gamma]$ nu kan skrives som:

$$S[\gamma] = \int_{\omega_A}^{\omega_B} \sqrt{1 + \cosh^2(\omega) \gamma'(\omega)^2} d\omega \quad , \quad \gamma \in C^2(\mathbb{R}, S^1) \quad (1.87)$$

Hvor $S^1 = [0, 2\pi)$. Find nu Euler-Lagrange ligningen for ovenstående funktional, og vis at den tager følgende form:

$$\frac{d}{d\omega} \left[\frac{\cosh^2(\omega) \gamma'(\omega)}{\sqrt{1 + \cosh^2(\omega) \gamma'(\omega)^2}} \right] = 0 \quad (1.88)$$

3. Anvend nu ligning (1.88), og vis at $\gamma'(\omega)$ opfylder følgende 'differential-ligning'.

$$\gamma'(\omega) = \frac{k}{\cosh^2(\omega) \sqrt{1 - k^2 + k^2 \tanh^2(\omega)}} \quad (1.89)$$

Hvor $k \in \mathbb{R}$ er en integrations-konstant. Identiteterne (1.85) kan med fordel udnyttes til at vise ovenstående sammenhæng.

4. Efter integration af (1.89), finder vi at:

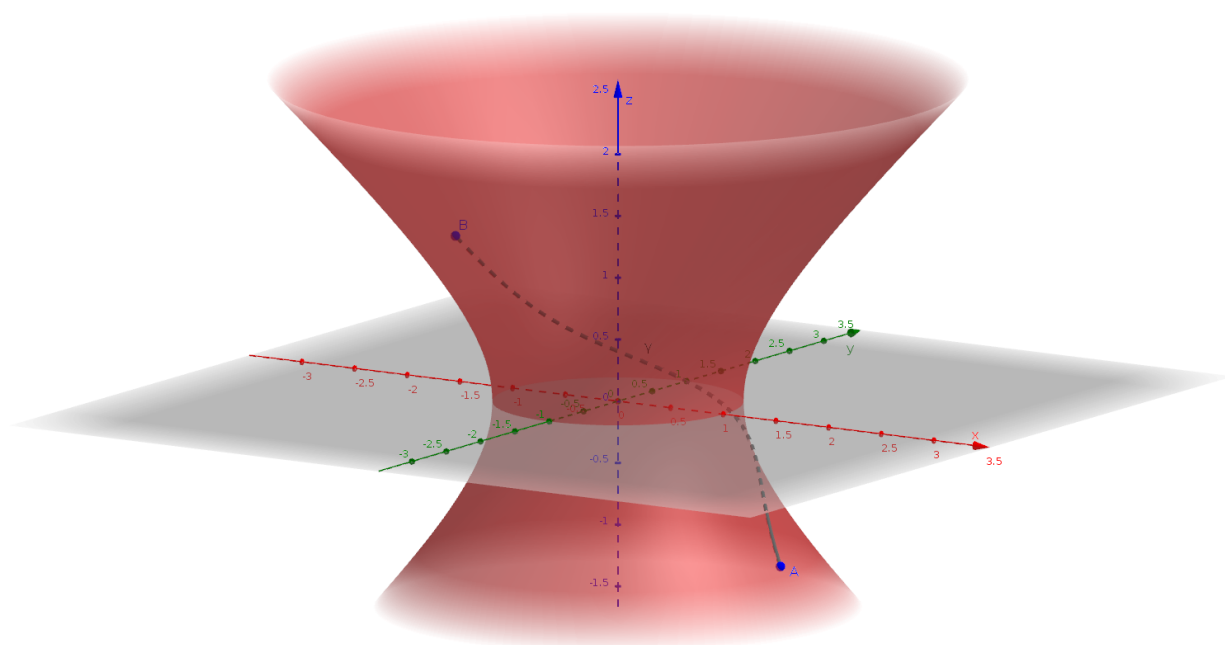
$$\gamma(\omega) - \gamma(\omega_A) = \int_{\omega_A}^{\omega} \frac{k}{\cosh^2(z) \sqrt{1 - k^2 + k^2 \tanh^2(z)}} dz \quad (1.90)$$

Hvor vi har indført integrations-variablen z . Vis nu, at hvis vi udføre følgende skift af integrations-variabel:

$$\alpha = \tanh(z) \quad (1.91)$$

Kan integralet i (1.90) bestemmes, og giver os følgende resultat:

$$\begin{aligned} & \int_{\omega_A}^{\omega} \frac{k}{\cosh^2(z) \sqrt{1 - k^2 + k^2 \tanh^2(z)}} dz \\ &= \sinh^{-1} \left(\frac{k \tanh(\omega)}{\sqrt{1 - k^2}} \right) - \sinh^{-1} \left(\frac{k \tanh(\omega_A)}{\sqrt{1 - k^2}} \right) \end{aligned} \quad (1.92)$$



Figur 1.9: Den røde overfalde repræsenterer her hyperboloiden givet ved $x^2 + y^2 - z^2 = 1$. Den sorte stiplede linie γ repræsenterer den korteste vej fra punktet A : $\omega = -1$, $\phi = 0$, til punktet B : $\omega = 1$, $\phi = \pi$, på hyperboloide-overfladen.

Følgende differential-kvotienter for hyperbolske funktioner kan med fordel anvendes til at vise det ovenstående resultat:

$$\frac{d \sinh(z)}{dz} = \cosh(z) \quad , \quad \frac{d \cosh(z)}{dz} = \sinh(z) \quad , \quad \frac{d \sinh^{-1}(\alpha)}{d\alpha} = \frac{1}{\sqrt{1 + \alpha^2}} \quad (1.93)$$

5. Anvend nu (1.92) samt (1.90), og vis at:

$$\cosh(\omega) \sinh(\gamma(\omega) - \bar{\phi}) = \frac{k}{\sqrt{1 - k^2}} \sinh(\omega) \quad (1.94)$$

Hvor konstanten $\bar{\phi}$ er givet ved:

$$\bar{\phi} = \gamma(\omega_A) - \sinh^{-1} \left(\frac{k \tanh(\omega_A)}{\sqrt{1 - k^2}} \right) \quad (1.95)$$

6. Find den koreste vej mellem punkterne A : $\omega = -1$, $\phi = 0$, og B : $\omega = 1$, $\phi = \pi$, på hyperboloide-overfladen, og skriv denne på formen (1.94).

1.8 Betinget optimering: Lagrange multiplikatorer

I mange praktisk relevante problemstillinger, ønsker man ikke at optimere en given funktional F over hele dens værdimængde $C^2(\mathbb{R}^n, \mathbb{R})$, men derimod over en given delmængde $\mathcal{U} \subseteq C^2(\mathbb{R}^n, \mathbb{R})$, defineret via en eller flere *betingelser*. For optimerings-problemer som involverer funktioner af typen $f : \mathbb{R}^n \rightarrow \mathbb{R}$, definerer vi oftest en analog delmængde $\mathcal{V} \subseteq \mathbb{R}^n$, via et sæt af m *betingelses-funktioner* $g_j : \mathbb{R}^n \rightarrow \mathbb{R}$, hvor $j \in \{1, \dots, m\}$:

$$\mathcal{V} = \{x \in \mathbb{R}^n \mid \forall j \in \{1, \dots, m\} : g_j(x) = c_j\} \quad , \quad c_j \in \mathbb{R} \quad (1.96)$$

For en funktional F , kan vi komplet analogt definere en delmængde $\mathcal{U} \subseteq C^2(\mathbb{R}^n, \mathbb{R})$ via et sæt af *betingelses-funktioner* $G_j : C^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathbb{R}$, hvor $j \in \{1, \dots, m\}$:

$$\mathcal{U} = \{\gamma \in C^2(\mathbb{R}^n, \mathbb{R}) \mid \forall j \in \{1, \dots, m\} : G_j[\gamma] = C_j\} \quad , \quad C_j \in \mathbb{R} \quad (1.97)$$

I variationsregningen kalder vi betingelser defineret via et sæt af betingelses-funktioner, for *globale betingelser*. I modsætning til optimerings-problemer som involverer funktioner af typen $f : \mathbb{R}^n \rightarrow \mathbb{R}$, kan vi for funktionaler også definere vores delmængde via såkaldt *lokale betingelser*. Denne type betingelser er defineret via betingelses-funktioner af typen $K_j : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$, hvor $j \in \{1, \dots, m\}$:

$$\mathcal{U} = \{\gamma \in C^2(\mathbb{R}^n, \mathbb{R}) \mid \forall j \in \{1, \dots, m\} \forall t \in \mathbb{R} : K_j(\gamma(t), \gamma'(t), t) = C_j\} \quad , \quad C_j \in \mathbb{R} \quad (1.98)$$

Man kan tænke på lokale betingelser som et ekstra sæt af differential-ligninger, for hvilket et hvert element $\gamma \in \mathcal{U}$ kræves at være løsning.

Det gode spørgsmål er nu hvordan vi i praksis bære os ad med at optimere en funktional over en delmængde $\mathcal{U} \subseteq C^2(\mathbb{R}^n, \mathbb{R})$. Lad os igen drage en parallel til optimerings-problemer for funktioner over \mathbb{R}^n . Det er for sådanne problemer et velkendt resultat, at en funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$, kan optimeres over en delmængde af typen (1.96) ved at løse følgende ligningssystem:

$$\nabla f(x) + \sum_{j=1}^m \lambda_j \nabla g_j(x) = 0 \quad , \quad g_j(x) = 0 \quad , \quad \forall j \in \{1, \dots, m\} \quad (1.99)$$

Hvor $\lambda_j \in \mathbb{R}$, er de såkaldte *Lagrange multiplikatorer* for problemet. De ubekendte i ovenstående ligningssystem er $x = (x_1, \dots, x_n)$ samt $\lambda_1, \dots, \lambda_m$. Antallet af ubekendte er altså lig

antallet af ligninger, og vi har dermed en chance for at finde en løsning til ligningssystemet. Det viser sig at en given funktional F kan optimeres over en delmængde $\mathcal{U} \subseteq C^2(\mathbb{R}^n, \mathbb{R})$ ved at løse et ligningssystem analogt til (1.99). Detaljerne for dette ligningssystem vil afhænge af om \mathcal{U} er defineret via globale eller lokale betingelser.

1.8.1 Globale betingelser

Vi vil i det følgende begrænse vores analyse til problemer med 1 global betingelse-funktional G , og til slut skitsere generalisering til problemer med mere en 1 global betingelse. Vi vil, lige som vi har gjort indtil nu, begrænse vores analyse til funktionaler F af typen:

$$F[\gamma] = \int_{t_A}^{t_B} L(\gamma(t), \gamma'(t), t) dt \quad (1.100)$$

Vi kræver ligeledes, at betingelses-funktional G er af samme form som vores funktional F .

$$G[\gamma] = \int_{t_A}^{t_B} K(\gamma(t), \gamma'(t), t) dt \quad (1.101)$$

Antag nu at funktionen γ er en stationær funktion for F , over en delmængde af $C^2(\mathbb{R}^n, \mathbb{R})$, defineret ved:

$$\mathcal{U} = \{\gamma \in C^2(\mathbb{R}^n, \mathbb{R}) \mid G[\gamma] = 0\} \quad (1.102)$$

Vælg nu 2 linært uafhængige pertubations-funktioner $\eta_1, \eta_2 \in C^2(\mathbb{R}^n, \mathbb{R})$, hvor det som tidligere kræves at $\eta_1(t_A) = \eta_1(t_B) = 0$ og $\eta_2(t_A) = \eta_2(t_B) = 0$. Da vi ønsker at optimere F over \mathcal{U} , skal der for enhver tilladt perturbation ξ gælde $G[\gamma + \varepsilon \xi] = 0$. Dermed følger.

$$\left. \frac{dG[\gamma + \varepsilon \xi]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.103)$$

For vores tilfældigt valgte perturbationer η_1 og η_2 , vil ovenstående betingelse generelt ikke være opfyldt. Vi vil derimod finde at.

$$\left. \frac{dG[\gamma + \varepsilon \eta_1]}{d\varepsilon} \right|_{\varepsilon=0} = \alpha_1, \quad \left. \frac{dG[\gamma + \varepsilon \eta_2]}{d\varepsilon} \right|_{\varepsilon=0} = \alpha_2 \quad (1.104)$$

Hvor $\alpha_1, \alpha_2 \in \mathbb{R}$. Vores strategi er nu at konstruere en tilladt perturbation, som linear kombination af η_1 og η_2 . Bemærk først at der for vilkårlige η_1, η_2 og $a_1, a_2 \in \mathbb{R}$ gælder.

$$\left. \frac{dG[\gamma + \varepsilon (a_1 \eta_1 + a_2 \eta_2)]}{d\varepsilon} \right|_{\varepsilon=0} = a_1 \left. \frac{dG[\gamma + \varepsilon \eta_1]}{d\varepsilon} \right|_{\varepsilon=0} + a_2 \left. \frac{dG[\gamma + \varepsilon \eta_2]}{d\varepsilon} \right|_{\varepsilon=0} \quad (1.105)$$

Dette kan vises ved brug af (1.101) samt kæde-reglen for funktioner af flere variable. For yderligere detaljer, se afsnit 1.4.

Ved at anvende (1.105), kan vi nu ved hjælp af η_1 og η_2 konstruere en tilladt perturbationsfunktion ξ . Vi postulerer nu, at følgende perturbation er tilladt.

$$\xi = \alpha_1 \eta_2 - \alpha_2 \eta_1 \quad (1.106)$$

Hvor α_1 og α_2 er defineret ud fra (1.104). Dette postulat bevises simpelt ved brug af (1.105).

$$\left. \frac{dG[\gamma + \varepsilon \xi]}{d\varepsilon} \right|_{\varepsilon=0} = \alpha_1 \left. \frac{dG[\gamma + \varepsilon \eta_2]}{d\varepsilon} \right|_{\varepsilon=0} - \alpha_2 \left. \frac{dG[\gamma + \varepsilon \eta_1]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.107)$$

Da vi ønsker at optimere funktionalen F over delmængden \mathcal{U} , må det ligeledes gælde at.

$$\left. \frac{dF[\gamma + \varepsilon \xi]}{d\varepsilon} \right|_{\varepsilon=0} = \alpha_1 \left. \frac{dF[\gamma + \varepsilon \eta_2]}{d\varepsilon} \right|_{\varepsilon=0} - \alpha_2 \left. \frac{dF[\gamma + \varepsilon \eta_1]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.108)$$

Ligningerne (1.107, 1.108) udgør tilsammen et system af lineære ligninger, som bekendt kan repræsenteres ved en enkelt matrix-ligning.

$$\begin{bmatrix} \left. \frac{dF[\gamma + \varepsilon \eta_2]}{d\varepsilon} \right|_{\varepsilon=0} & - \left. \frac{dF[\gamma + \varepsilon \eta_1]}{d\varepsilon} \right|_{\varepsilon=0} \\ \left. \frac{dG[\gamma + \varepsilon \eta_2]}{d\varepsilon} \right|_{\varepsilon=0} & - \left. \frac{dG[\gamma + \varepsilon \eta_1]}{d\varepsilon} \right|_{\varepsilon=0} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (1.109)$$

Antag nu at vi kan vælge η_1 og η_2 , sådan at enten α_1 eller α_2 er forskellige fra 0. Hvis ikke dette var tilfældet ville $\mathcal{U} = C^2(\mathbb{R}, \mathbb{R}^n)$, og vores betingede optimerings-problem er trivielt.

Vi kan nu avende følgende resultat fra den lineære algebra:

For enhver *invertibel* matrix \underline{A} , gælder det at $\underline{A}\underline{v} \neq \underline{0}$ hvis \underline{v} er forskellig fra nul-vektoren.

Vi kan dermed konkludere at matricen:

$$\begin{bmatrix} \left. \frac{dF[\gamma + \varepsilon \eta_2]}{d\varepsilon} \right|_{\varepsilon=0} & - \left. \frac{dF[\gamma + \varepsilon \eta_1]}{d\varepsilon} \right|_{\varepsilon=0} \\ \left. \frac{dG[\gamma + \varepsilon \eta_2]}{d\varepsilon} \right|_{\varepsilon=0} & - \left. \frac{dG[\gamma + \varepsilon \eta_1]}{d\varepsilon} \right|_{\varepsilon=0} \end{bmatrix} \quad (1.110)$$

Ikke er invertibel. Hvis en matrix ikke er invertibel, følger det at rækkerne i matricen må være *linært afhængige*:

$$\begin{bmatrix} \left. \frac{dF[\gamma + \varepsilon \eta_2]}{d\varepsilon} \right|_{\varepsilon=0} & - \left. \frac{dF[\gamma + \varepsilon \eta_1]}{d\varepsilon} \right|_{\varepsilon=0} \end{bmatrix} + \lambda \begin{bmatrix} \left. \frac{dG[\gamma + \varepsilon \eta_2]}{d\varepsilon} \right|_{\varepsilon=0} & - \left. \frac{dG[\gamma + \varepsilon \eta_1]}{d\varepsilon} \right|_{\varepsilon=0} \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix} \quad (1.111)$$

Hvor $\lambda \in \mathbb{R}$. Da perturbation-funktionerne η_1 og η_2 er arbitrært valgte, kan vi på baggrund af den lineære afhængighed konkludere at.

$$\forall \eta \in C^2(\mathbb{R}, \mathbb{R}^n) : \quad \left. \frac{dF[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} + \lambda \left. \frac{dG[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.112)$$

Definerer vi nu funktionalen I , sådan at.

$$I = F + \lambda G \quad (1.113)$$

Kan (1.112) omskrives til følgende form.

$$\forall \eta \in C^2(\mathbb{R}, \mathbb{R}^n) : \quad \left. \frac{dI[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.114)$$

Det er altså lykkedes os at omskrive vores betingede optimerings-problem for funktionalen F , til et ubetinget optimerings-problem for funktionalen I ; dog med det krav at $G[\gamma] = C$. Vi kan til slut anvende Lagrange-funktionen J for funktionalen I , til at opskrive Euler-Lagrange ligningerne for optimerings-problemet (1.114).

$$\frac{\partial J(\gamma(t), \gamma'(t), t)}{\partial x_i} - \frac{d}{dt} \frac{\partial J(\gamma(t), \gamma'(t), t)}{\partial v_i} = 0 \quad , \quad i = 1, \dots, n \quad (1.115)$$

hvor

$$J(u, v, t) := L(u, v, t) + \lambda K(u, v, t) \quad (1.116)$$

Mere end 1 global betingelse: Lad os nu kort diskutere hvad man kan gøre, i det tilfælde man ønsker en funktional F optimeret under flere globale betingelser, givet ved $G_j[\gamma] = C_j$ for $j \in \{1, \dots, m\}$, sådan at:

$$G_j[\gamma] = \int_{t_A}^{t_B} K_j(\gamma(t), \gamma'(t), t) dt \quad (1.117)$$

$$\mathcal{U} = \{\gamma \in C^2(\mathbb{R}^n, \mathbb{R}) \mid \forall j \in \{1, \dots, m\} : G_j[\gamma] = 0\} \quad (1.118)$$

Antag, ligesom i tilfældet med 1 global betingelse, at γ er stationær funktion for funktionalen F . Vælg nu $m + 1$ linært uafhængige perturbationer: $\eta_{11}, \eta_{12}, \dots, \eta_{1m+1}$, hvor der kræves at $\eta_{1k}(t_A) = \eta_{1k}(t_B) = 0$ for $k \in \{1, \dots, m + 1\}$. Vi kan anvende disse $m + 1$ perturbationer til at konstruere m nye perturbationer $\eta_{21}, \eta_{22}, \dots, \eta_{2m}$, som alle opfylder:

$$\left. \frac{dG_1[\gamma + \varepsilon \eta_{2k}]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad , \quad k \in \{1, \dots, m\} \quad (1.119)$$

Konstruktionen af perturbationerne $\eta_{21}, \eta_{22}, \dots, \eta_{2m}$ er i komplet analog med konstruktionen af den tilladte perturbation i tilfældet med 1 global betingelse. Lad os som eksempel se på konstruktionen af η_{21} . Vi ved at der eksistere værdier $\alpha_{11}, \alpha_{21} \in \mathbb{R}$, sådan at:

$$\left. \frac{dG_1[\gamma + \varepsilon \eta_{11}]}{d\varepsilon} \right|_{\varepsilon=0} = \alpha_{11} \quad , \quad \left. \frac{dG_1[\gamma + \varepsilon \eta_{12}]}{d\varepsilon} \right|_{\varepsilon=0} = \alpha_{12} \quad (1.120)$$

Ved brug af samme logik som for tilfældet med 1 global betingelse, finder vi nu at:

$$\eta_{21} := \alpha_{11}\eta_{12} - \alpha_{12}\eta_{11} \quad \Rightarrow \quad \left. \frac{dG_1[\gamma + \varepsilon \eta_{21}]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.121)$$

Vi konstruere på samme måde η_{22} ved brug af η_{12} og η_{13} , η_{23} ved brug af η_{13} og η_{14} etc. På denne måde sørger vi for at $\eta_{21}, \eta_{22}, \dots, \eta_{2m}$ også er lineært uafhængige.

Vi har nu m perturbationer som alle opfylder den 1. betingelse. Vi kan nu anvende disse til at konstruere $m - 1$ nye perturbationer $\eta_{31}, \eta_{32}, \dots, \eta_{3m-1}$, som opfylder:

$$\left. \frac{dG_1[\gamma + \varepsilon \eta_{3k}]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad , \quad \left. \frac{dG_2[\gamma + \varepsilon \eta_{3k}]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad , \quad k \in \{1, \dots, m-1\} \quad (1.122)$$

Perturbationer $\eta_{31}, \eta_{32}, \dots, \eta_{3m-1}$ konstrueres ud fra perturbationer $\eta_{21}, \eta_{22}, \dots, \eta_{2m}$, i komplet analog til vores konstruktion af $\eta_{21}, \eta_{22}, \dots, \eta_{2m}$ ud fra $\eta_{11}, \eta_{12}, \dots, \eta_{1m+1}$, hvor G_2 nu er den relevante betingelse. På grund af (1.105), er vi desuden sikret at $\eta_{31}, \eta_{32}, \dots, \eta_{3m-1}$ stadig opfylder G_1 betingelsen.

Vi fortsætter nu denne procedure indtil vi står tilbage med 1 perturbation ξ , som opfylder alle de givne globale betingelser:

$$\left. \frac{dG_j[\gamma + \varepsilon \xi]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad , \quad j \in \{1, \dots, m\} \quad (1.123)$$

Da γ er antaget at være stationær funktion for F i delmængden \mathcal{U} , må vi desuden kræve:

$$\left. \frac{dF[\gamma + \varepsilon \xi]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.124)$$

Det er klart at perturbationen ξ vil være en lineær kombination af de først valgte perturbationer $\eta_{11}, \eta_{12}, \dots, \eta_{1m+1}$, som vi startede med. Lad os kalde koefficienterne for denne lineære kombination β_k med $k \in \{1, \dots, m+1\}$:

$$\xi = \beta_1\eta_{11} + \beta_2\eta_{12} + \dots + \beta_{m+1}\eta_{1m+1} \quad (1.125)$$

Hvor koefficienterne β_k kan bestemmes udfra α_{jk} , med $j \in \{1, \dots, m\}$ og $k \in \{1, \dots, m+1\}$. Ved anvendelse af (1.105), ser vi nu at (1.123, 1.124) tilsammen udgør et system af $m+1$ lineære ligninger:

$$\begin{bmatrix} \left. \frac{dF[\gamma+\varepsilon\eta_{11}]}{d\varepsilon} \right|_{\varepsilon=0} & \dots & \left. \frac{dF[\gamma+\varepsilon\eta_{1m+1}]}{d\varepsilon} \right|_{\varepsilon=0} \\ \left. \frac{dG_1[\gamma+\varepsilon\eta_{11}]}{d\varepsilon} \right|_{\varepsilon=0} & \dots & \left. \frac{dG_1[\gamma+\varepsilon\eta_{1m+1}]}{d\varepsilon} \right|_{\varepsilon=0} \\ \vdots & & \vdots \\ \left. \frac{dG_m[\gamma+\varepsilon\eta_{11}]}{d\varepsilon} \right|_{\varepsilon=0} & \dots & \left. \frac{dG_m[\gamma+\varepsilon\eta_{1m+1}]}{d\varepsilon} \right|_{\varepsilon=0} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{m+1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (1.126)$$

Af samme årsag som i tilfældet med 1 global betingelse, konkluderer vi at den ovenstående matrice ikke kan være invertibel, og at dens rækker dermed er lineært afhængige. Da perturbationerne $\eta_{11}, \eta_{12}, \dots, \eta_{1m+1}$ var arbitrært valgte, kan vi nu endeligt konkludere at:

$$\forall \eta \in C^2(\mathbb{R}, \mathbb{R}^n) : \left. \frac{dF[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} + \sum_{j=1}^m \lambda_j \left. \frac{dG_j[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.127)$$

Definerer vi nu funktionalen I , sådan at.

$$I = F + \sum_{j=1}^m \lambda_j G_j \quad (1.128)$$

Kan (1.127) omskrives til følgende form.

$$\forall \eta \in C^2(\mathbb{R}, \mathbb{R}^n) : \left. \frac{dI[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (1.129)$$

med de tilhørende Euler-Lagrange ligninger:

$$\frac{\partial J(\gamma(t), \gamma'(t), t)}{\partial x_i} - \frac{d}{dt} \frac{\partial J(\gamma(t), \gamma'(t), t)}{\partial v_i} = 0 \quad , \quad i = 1, \dots, n \quad (1.130)$$

hvor

$$J(u, v, t) := L(u, v, t) + \sum_{i=1}^m \lambda_i K_i(u, v, t) \quad (1.131)$$

1.8.2 Eksempel: formen af en hængende kæde

Vi forestiller os nu en kæde, med uniform masse-tæthed ρ , ophængt fra 2 givne punkter $A = (x_A, y_A)$ og $B = (x_B, y_B)$. Den potentielle energi af kæden i tyngdefeltet kan betragtes som en funktional V der som input tager bane-kurver $\gamma \in C^2(\mathbb{R}, \mathbb{R}^2)$ fra A til B .

$$V[\gamma] = \int_{t_A}^{t_B} \rho g \gamma_y(t) \sqrt{\gamma'(t) \cdot \gamma'(t)} dt \quad , \quad \gamma \in C^2(\mathbb{R}, \mathbb{R}^2) \quad (1.132)$$

Hvor $g \approx 9.8 \text{ m/s}^2$, er tyngde-accelerationen ved Jordens overflade. Vi observerer, også i dette tilfælde, at funktionalen V blot er et special tilfælde af rejse-tids funktional T (1.27), nu med:

$$\frac{n(\gamma(t))}{c} = \rho g \gamma_y(t) \quad (1.133)$$

Vi kan derfor også dimensionelt reducere funktional V , sådan at $V[\gamma]$ er givet som et special tilfælde af (1.41).

$$V[\gamma] = \int_{x_A}^{x_B} \rho g \gamma(x) \sqrt{1 + \gamma'(x)^2} dx \quad , \quad \gamma \in C^2(\mathbb{R}, \mathbb{R}) \quad (1.134)$$

I følge *princippet om minimal energi*, vil den hængende kæde antage den form som minimerer dens potentielle energi, og vi har dermed at gøre med et variations-problem. Vi kan dog ikke variere kædens form arbitrært, da vi antager kæden regid med længde ℓ . Dermed må enhver tilladt kurve-form γ opfylde:

$$S[\gamma] = \int_{x_A}^{x_B} \sqrt{1 + \gamma'(x)^2} dx \stackrel{!}{=} \ell \quad , \quad \gamma \in C^2(\mathbb{R}, \mathbb{R}) \quad (1.135)$$

Vi har dermed at gøre med et variations-problem med 1 global betingelse. Vi starter derfor med at definere funktionen I som:

$$I = V + \lambda S \quad (1.136)$$

Hvor $\lambda \in \mathbb{R}$, er den endnu ubestemte Lagrange-multiplikator for problemet. Det er nu en simpel opgave at identificere Lagrange-funktionen J for funktionen I :

$$J(y, v_y, x) := \frac{n(y)}{c} \sqrt{1 + v_y^2} \quad , \quad \frac{n(y)}{c} = \rho g y + \lambda \quad (1.137)$$

Vi kan nu, i princippet, anvende den ovenstående Lagrange-funktionen for V , til at opskrive Euler-Lagrange ligningerne for vores problem. Lad os i stedet udnytte at $J(y, v_y, x)$ ikke afhænger af x , hvilket betyder at (1.44) med n givet ved (1.137), er en bevaret størrelse.

$$H(\gamma(x), \gamma'(x), x) = -\frac{\rho g \gamma(x) + \lambda}{\sqrt{1 + \gamma'(x)^2}} = C \quad (1.138)$$

Hvor $C \in \mathbb{R}$ kan betragtes som en integrations-konstant. Ved brug af ovenstående udtryk, finder vi nu følgende 1. ordens differential-ligning for γ :

$$\gamma'(x) = \pm \sqrt{(a \gamma(x) + b)^2 - 1} \quad , \quad a \equiv \frac{\rho g}{C} \quad , \quad b \equiv \frac{\lambda}{C} \quad (1.139)$$

Da ovenstående ligning er separabel kan vi finde løsnings-kurven γ ved simpel integration:

$$\pm \int_{\gamma(x_A)}^{\gamma(x)} \frac{1}{\sqrt{(az+b)^2 - 1}} dz = \int_{x_A}^x dz \quad (1.140)$$

Hvor vi har indført integrations variabelen z . Integralet på højresiden er trivielt, mens integralet på venstresiden er givet i termer af den hyperbolske cosinus funktion \cosh . Vi finder dermed følgende resultat, hvor vi også har gjort brug af randbetingelsen $\gamma(x_A) = y_A$:

$$\frac{1}{a} \left[\cosh^{-1}(a\gamma(x) + b) - \cosh^{-1}(ay_A + b) \right] = x - x_A \quad (1.141)$$

Hvor $a \in \mathbb{R}$, og vi dermed ikke længere behøver \pm på venstresiden af det ovenstående udtryk.

$$\gamma(x) = \frac{\cosh [a(x - x_A) + \cosh^{-1}(ay_A + b)] - b}{a} \quad (1.142)$$

$$\gamma'(x) = \sinh [a(x - x_A) + \cosh^{-1}(ay_A + b)] \quad (1.143)$$

$$\gamma'(x_{min}) = 0 \quad \Rightarrow \quad x_{min} = x_A - \frac{1}{a} \cosh^{-1}(ay_A + b) \quad (1.144)$$

$$w := x - x_{min} \quad \Rightarrow \quad \gamma'(\tilde{x}) = \sinh(aw) \quad (1.145)$$

Vi definerer $\Delta y \equiv y_B - y_A$. Ved brug af randbetingelserne $\gamma(x_A) = y_A$ og $\gamma(x_B) = y_B$, finder vi nu at:

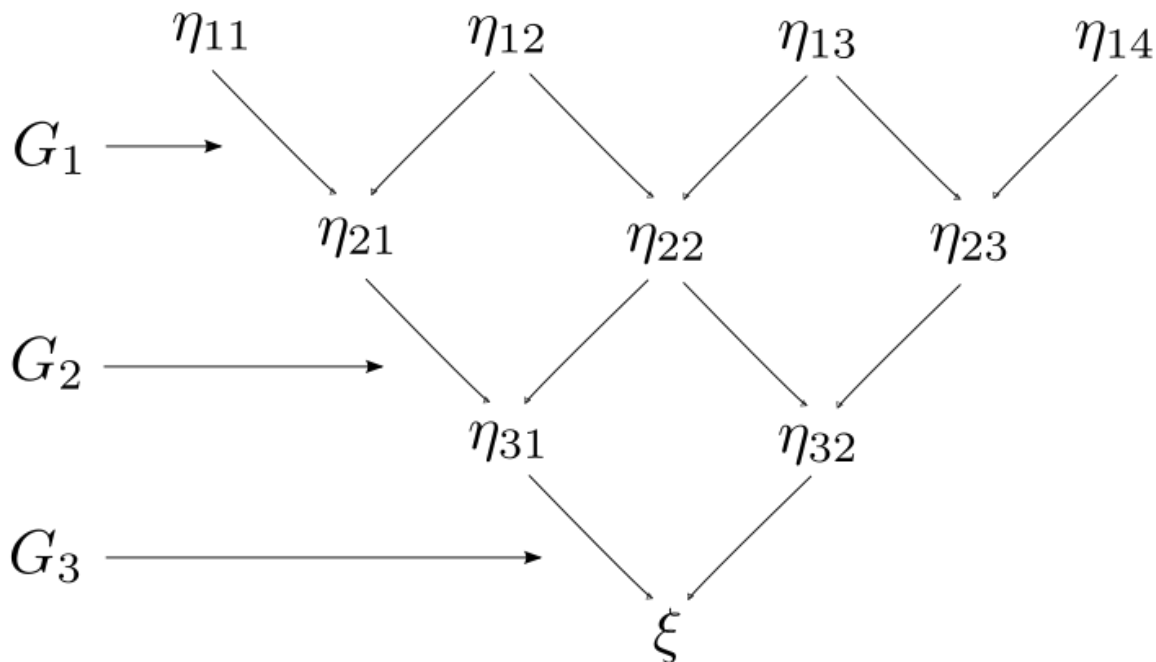
$$\Delta y = \gamma(x_B) - \gamma(x_A) = \frac{1}{a} (\cosh [a(x_B - x_{min})] - \cosh [a(x_A - x_{min})]) \quad (1.146)$$

$$\begin{aligned} S[\gamma] &= \int_{x_A}^{x_B} \sqrt{1 + \gamma'(x)^2} dx = \int_{x_A - x_{min}}^{x_B - x_{min}} \sqrt{1 + \sinh(aw)^2} dw = \int_{x_A - x_{min}}^{x_B - x_{min}} \cosh(aw) dw \\ &= \frac{1}{a} (\sinh [a(x_B - x_{min})] - \sinh [a(x_A - x_{min})]) \stackrel{!}{=} \ell \end{aligned} \quad (1.147)$$

$$\cosh(\alpha - \beta) = \cosh(\alpha) \cosh(\beta) - \sinh(\alpha) \sinh(\beta) \quad (1.148)$$

$$2 \sinh^2(\alpha) = \cosh(2\alpha) - 1 \quad (1.149)$$

$$\Rightarrow \quad \ell^2 - \Delta y = \frac{4}{a^2} \sinh^2 \left(\frac{a}{2} \Delta x \right) \quad (1.150)$$



Figur 1.10: En grafisk representation af processen til konstruktion af en perturbation som opfylder: $\gamma + \varepsilon \xi \in \mathcal{U} \subset C^2(\mathbb{R}, \mathbb{R}^n)$, for tilfældet med 3 globale betingelser.

1.8.3 Lokale betingelser

$$F[\gamma] = \int_{t_A}^{t_B} L(\gamma(t), \gamma'(t), t) dt \quad (1.151)$$

$$K(\gamma(t), \gamma'(t), t) = 0 \quad , \quad \forall t \in \mathbb{R} \quad (1.152)$$

$$\forall t \in \mathbb{R} : \quad K(\gamma(t), \gamma'(t), t) = 0 \quad \Leftrightarrow \quad (1.153)$$

$$\forall \lambda \in C^2(\mathbb{R}, \mathbb{R}) : \quad \int_{t_A}^{t_B} \lambda(t) K(\gamma(t), \gamma'(t), t) dt = 0$$

1.9 Ekstremums-test for funktionaler

Vi har indtil nu ignoreret spørgsmålet om hvorvidt en given stationær bane-kurve γ for en funktional F også er en ekstremums-funktion for funktionalen. I dette afsnit vil vi derfor

diskutere præcis hvordan man kan bære sig ad med besvare spørgsmålet om bestemmelse af ekstremums-funktioner.

Vores udgangspunkt er, ligesom for bestemmelse af stationære funktioner, at betragte output-værdien $F[\gamma + \varepsilon \eta]$ for F . Her antages $\gamma \in C^p(\mathbb{R}, \mathbb{R}^n)$ at være stationær funktion, da det ellers er klart at der findes en tilstrækkelig lille radius $r \in \mathbb{R}^+$, sådan at:

$$\exists \varepsilon_1 \in (-r, 0) \cup (0, r) : \quad F[\gamma + \varepsilon_1 \eta] - F[\gamma] = \left. \frac{dF[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} \varepsilon_1 + \mathcal{O}(\varepsilon_1^2) > 0 \quad (1.154)$$

og

$$\exists \varepsilon_2 \in (-r, 0) \cup (0, r) : \quad F[\gamma + \varepsilon_2 \eta] - F[\gamma] = \left. \frac{dF[\gamma + \varepsilon \eta]}{d\varepsilon} \right|_{\varepsilon=0} \varepsilon_2 + \mathcal{O}(\varepsilon_2^2) < 0 \quad (1.155)$$

Vi antager desuden at $\eta \in C^p(\mathbb{R}, \mathbb{R}^n)$ opfylder $\eta(t_A) = \eta(t_B) = 0$. Vores strategi er nu, inspireret af ekstremums-analyse for $\mathbb{R} \rightarrow \mathbb{R}$ funktioner, at betragte den 2. afledte af vores funktional F evalueret for den givne stationære funktion γ . Mere præcist ønsker vi at finde en test der sikre os at:

$$\forall \eta \in C^p(\mathbb{R}, \mathbb{R}^n) : \quad \left. \frac{d^2 F[\gamma + \varepsilon \eta]}{d\varepsilon^2} \right|_{\varepsilon=0} > 0 \quad (1.156)$$

eller

$$\forall \eta \in C^p(\mathbb{R}, \mathbb{R}^n) : \quad \left. \frac{d^2 F[\gamma + \varepsilon \eta]}{d\varepsilon^2} \right|_{\varepsilon=0} < 0 \quad (1.157)$$

Således at der findes en tilstrækkelig lille radius $r \in \mathbb{R}^+$ for hvilken vi er sikkert følgende:

$$\forall \varepsilon_1 \in (-r, 0) \cup (0, r) : \quad F[\gamma + \varepsilon_1 \eta] - F[\gamma] = \frac{1}{2} \left. \frac{d^2 F[\gamma + \varepsilon \eta]}{d\varepsilon^2} \right|_{\varepsilon=0} \varepsilon_1^2 + \mathcal{O}(\varepsilon_1^3) > 0 \quad (1.158)$$

eller

$$\forall \varepsilon_2 \in (-r, 0) \cup (0, r) : \quad F[\gamma + \varepsilon_2 \eta] - F[\gamma] = \frac{1}{2} \left. \frac{d^2 F[\gamma + \varepsilon \eta]}{d\varepsilon^2} \right|_{\varepsilon=0} \varepsilon_2^2 + \mathcal{O}(\varepsilon_2^3) < 0 \quad (1.159)$$

Vi antager, som vi har gjort indtil nu, at funktionalen F kan skrives på følgende form:

$$F[\gamma] = \int_{t_A}^{t_B} L(\gamma(t), \gamma'(t), t) dt \quad (1.160)$$

For at analysere den 2. afledte af $F[\gamma + \varepsilon \eta]$ har vi altså brug for at kende den 2. afledte af Lagrange-funktionen L . Per vores tidligere analyse af stationære funktioner for funktionaler af ovenstående form, ved vi at:

$$\frac{d^2 L(z, w, t)}{d\varepsilon^2} = \frac{d}{d\varepsilon} \sum_{i=1}^n \left[\frac{\partial L(z, w, t)}{\partial x_i} \frac{dz_i}{d\varepsilon} + \frac{\partial L(z, w, t)}{\partial v_i} \frac{dw_i}{d\varepsilon} \right] \quad (1.161)$$

Hvor $z = \gamma(t) + \varepsilon \eta(t)$ og $w = \gamma'(t) + \varepsilon \eta'(t)$. Ved brug af kæde-reglen for funktioner af flere variable, finder vi nu følgende udtryk for den 2. afledte af $L(z, w, t)$:

$$\begin{aligned} \frac{d^2 L(z, w, t)}{d\varepsilon^2} = \sum_{i,j=1}^n & \left[\frac{\partial^2 L(z, w, t)}{\partial x_i \partial x_j} \frac{dz_i}{d\varepsilon} \frac{dz_j}{d\varepsilon} + \frac{\partial^2 L(z, w, t)}{\partial v_i \partial v_j} \frac{dw_i}{d\varepsilon} \frac{dw_j}{d\varepsilon} \right. \\ & + \frac{\partial^2 L(z, w, t)}{\partial x_i \partial v_j} \frac{dz_i}{d\varepsilon} \frac{dw_j}{d\varepsilon} + \frac{\partial^2 L(z, w, t)}{\partial v_i \partial x_j} \frac{dw_i}{d\varepsilon} \frac{dz_j}{d\varepsilon} \\ & \left. + \frac{\partial L(z, w, t)}{\partial x_i} \frac{d^2 z_i}{d\varepsilon^2} + \frac{\partial L(z, w, t)}{\partial v_i} \frac{d^2 w_i}{d\varepsilon^2} \right] \end{aligned} \quad (1.162)$$

Ovenstående udtryk fremstår umiddelbart meget kompliceret, men kan dog simplificeres signifikant. Lad os først bemærke, at der per definitionen af z og w gælder:

$$\frac{d^2 z_i}{d\varepsilon^2} = 0 \quad , \quad \frac{d^2 w_i}{d\varepsilon^2} = 0 \quad (1.163)$$

Hvorned vi kan skrive den 2. afledte af $L(z, w, t)$ på følgende form:

$$\begin{aligned} \frac{d^2 L(z, w, t)}{d\varepsilon^2} = \sum_{i,j=1}^n & \left[\frac{\partial^2 L(z, w, t)}{\partial x_i \partial x_j} \frac{dz_i}{d\varepsilon} \frac{dz_j}{d\varepsilon} + \frac{\partial^2 L(z, w, t)}{\partial v_i \partial v_j} \frac{dw_i}{d\varepsilon} \frac{dw_j}{d\varepsilon} \right. \\ & \left. + \frac{\partial^2 L(z, w, t)}{\partial x_i \partial v_j} \frac{dz_i}{d\varepsilon} \frac{dw_j}{d\varepsilon} + \frac{\partial^2 L(z, w, t)}{\partial v_i \partial x_j} \frac{dw_i}{d\varepsilon} \frac{dz_j}{d\varepsilon} \right] \end{aligned} \quad (1.164)$$

Vi kan nu gøre brug af vores ny-fundende resultat (1.164) til at skrive den 2. afledte af funktional-værdien $F[\gamma + \varepsilon \eta]$ på følgende form:

$$\left. \frac{d^2 F[\gamma + \varepsilon \eta]}{d\varepsilon^2} \right|_{\varepsilon=0} = \int_{t_A}^{t_B} \left[\eta(t)^T Q(t) \eta(t) + \eta'(t)^T P(t) \eta'(t) + 2 \eta'(t)^T R(t) \eta(t) \right] dt \quad (1.165)$$

Hvor vi har defineret matricerne $P(t)$, $Q(t)$ og $R(t)$.

$$P_{ij}(t) = \frac{\partial^2 L(\gamma(t), \gamma'(t), t)}{\partial v_i \partial v_j} \quad (1.166)$$

$$Q_{ij}(t) = \frac{\partial^2 L(\gamma(t), \gamma'(t), t)}{\partial x_i \partial x_j} \quad (1.167)$$

$$R_{ij}(t) = \frac{\partial^2 L(\gamma(t), \gamma'(t), t)}{\partial v_i \partial x_j} \quad (1.168)$$

$$0 = \int_{t_A}^{t_B} \frac{d}{dt} \left[\eta(t)^T W(t) \eta(t) \right] dt = \int_{t_A}^{t_B} \left[\eta(t)^T W'(t) \eta(t) + 2 \eta'(t)^T W(t) \eta(t) \right] dt \quad (1.169)$$

$$\begin{aligned} \frac{d^2 F[\gamma + \varepsilon \eta]}{d\varepsilon^2} \Big|_{\varepsilon=0} &= \int_{t_A}^{t_B} \left[\eta(t)^T \left(Q(t) + W'(t) \right) \eta(t) + \eta'(t)^T P(t) \eta'(t) \right. \\ &\quad \left. + 2 \eta'(t)^T \left(R(t) + W(t) \right) \eta(t) \right] dt \end{aligned} \quad (1.170)$$

$$Q(t) + W'(t) = \left[R(t) + W(t) \right]^T P(t)^{-1} \left[R(t) + W(t) \right] \quad (1.171)$$

$$\frac{d^2 F[\gamma + \varepsilon \eta]}{d\varepsilon^2} \Big|_{\varepsilon=0} = \int_{t_A}^{t_B} \left[\xi(t)^T P \xi(t) \right] dt \quad (1.172)$$

$$\xi(t) = \eta'(t) + P(t)^{-1} \left[R(t) + W(t) \right] \eta(t) \quad (1.173)$$

$$R(t) + W(t) = -P(t) U'(t) U(t)^{-1} \quad (1.174)$$

$$\left(Q(t) - R'(t) \right) U(t) - P'(t) U'(t) - P(t) U''(t) = 0 \quad (1.175)$$

Kapitel 2

Lagrange mekanik og princippet om mindste virkning

Inspireret af Fermat's princip om lyssets bevægelse, påbegyndte den italiensk-franske matematiker og astronom *Joseph-Louis Lagrange*, i årene omkring 1780, udformningen af et lignende princip for tids-udvikling af mekaniske systemer. I 1788 kulminerede Lagrange's arbejde med formuleringen af den såkladte *Lagrange mekanik* og princippet om *mindste Virkning*. Før vi kan begynde at udforske den dynamiske side af Lagrange's idéer, må vi dog først diskutere et par essentielle kinematiske koncepter; nævnlige *uafhængige frihedsgrader*, *konfigurations-rum* og *tangent-rum* for mekaniske systemer.

2.1 Konfigurations-rum

Lad os først og fremmest diskutere begrebet *uafhængige frihedsgrader*. Dette er relativt simpelt; de uafhængige frihedsgrader for et givent mekanisk system er blot de uafhængige variable der behøves for at beskrive systemets tilstand til et givent tidspunkt.

Definition 4. *Konfigurations-rummet for et mekanisk system er givet ved \mathbb{R}^n , hvor n er antallet af uafhængige frihedsgrader for det givne system.*

Eksempel 1. *Et system bestående af én partikel med masse m har, i vores univers, 3 uafhængige frihedsgrader, beskrivende partiklens position i rummet. Disse er ofte benævnt*

x , y og z . Konfigurations-rummet for dette system er altså \mathbb{R}^3 .

Eksempel 2. For et system bestående af N partikler med masser m_i ($i = 1, 2, \dots, N$) vil der i alt være $3N$ frihedsgrader, og systemets konfigurations-rum er i dette tilfælde \mathbb{R}^{3N} .

Eksempel 3. Et system bestående af et idealt pendul med masse m og længde ℓ har 1 begrænset frihedsgrad; vinklen θ målt fra lodret. I dette tilfælde er konfigurations-rummet altså intervallet $[0, 2\pi]$.

2.2 Tangent-rum

Eksempel 4. Et system bestående af én partikel med masse m har, i vores univers, 3 uafhængige frihedsgrader, beskrivende partiklens position i rummet. Disse er ofte benævnt x , y og z . Konfigurations-rummet for dette system er altså \mathbb{R}^3 .

Endvidere har all tangent-vektorer (hastigheder) til et hvert punkt i konfigurations-rummet 3 ubegrænsede komponenter, ofte benævnt v_x , v_y og v_z . Tangent-rummet for systemet er altså ligeledes \mathbb{R}^3 .

Har vi i stedet et system bestående af N partikler med masser m_i ($i = 1, 2, \dots, N$) vil der i alt være $3N$ frihedsgrader, og systemets konfigurations-rum er da \mathbb{R}^{3N} . Det tilhørende tangent-rum bliver ligeledes \mathbb{R}^{3N} .

Eksempel 5. Et system bestående af et idealt pendul med masse m og længde ℓ har 1 begrænset frihedsgrad; vinklen θ målt fra lodret. I dette tilfælde er konfigurations-rummet altså intervallet $[0, 2\pi]$.

Tangent-vektorerne (vinkelhastighederne) for systemet er derimod ubegrænsede, og tangent-rummet er derfor givet ved hele \mathbb{R} .

Eksempel 6. Et system bestående af én relativistisk partikel med masse m har, præcis som i den klassiske fysik, 3 ubegrænsede frihedsgrader, x , y og z . Konfigurations-rummet for dette system er derfor igen \mathbb{R}^3 .

I modsætning til den klassiske fysik er tangent-rummet for dette system begrænset, siden lysets hastighed c ikke kan overskrides. Siden $\sqrt{v_x^2 + v_y^2 + v_z^2} < c$, danner mængden af mulige tangent-vektorer en 3D kugle (uden skal); ofte benævnt B^3 . Tangent-rummet er altså B^3 .

2.3 Mindste virknings princip

Definition 5. Virkingen S for et mekanisk system er en funktion der som input tager en bane-kurve q over systemets konfigurations-rum;

$$q : \mathbb{R} \longrightarrow \mathbb{R}^n \quad , \quad t \mapsto q(t) \quad (2.1)$$

og giver som output et element $S[q]$ i \mathbb{R} . Vi antager disse bane-kurver vilkårligt ofte differentiable. Mængden af alle vilkårligt ofte differentiable bane-kurver over systemets konfigurations-rum giver vi navnet $C^\infty(\mathbb{R}, \mathbb{R}^n)$. Virkingen S er da defineret som følger.

$$S : C^\infty(\mathbb{R}, \mathbb{R}^n) \longrightarrow \mathbb{R} \quad , \quad q \mapsto S[q] := \int_{t_A}^{t_B} dt L(q(t), \dot{q}(t), t) \quad (2.2)$$

Definition 6. Lagrange funktionen L for et mekanisk system er en funktion der som input tager: et punkt v i systemets konfigurations-rum, et punkt w i tangent-rummet over systemets konfigurations-rum, et element u i \mathbb{R} . Lagrange funktionen giver som output et reelt tal.

$$L : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \quad , \quad (v, w, u) \mapsto L(v, w, u) := T(v, w, u) - V(v, u) \quad (2.3)$$

Funktionen T er systemets kinetiske energi, og funktionen V systemets potentielle energi.

2.4 Symmetri og bevarede størrelser

2.4.1 Koordinat deformationer

En koordinat deformation er en funktion $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. En kontinuert familie af deformationer er en funktion $\phi : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$. Vi andender følgende notation: $\phi(\varepsilon, \cdot) \equiv \phi_\varepsilon(\cdot)$. Vi definerer nu deformationen af Virkingen S genereret af den kontinuert koordinat deformationen Φ , på følgende måde.

$$S_\varepsilon[q] := \int_{t_A}^{t_B} dt L(\phi_\varepsilon(q), \dot{\phi}_\varepsilon(q), t) \quad (2.4)$$

$$S_\varepsilon[q] = S[q] + \varepsilon \left[\Lambda(q, \dot{q}, t) \right]_{t_A}^{t_B} + \mathcal{O}(\varepsilon^2) \quad (2.5)$$

$$\Rightarrow \left[\Lambda(q, \dot{q}, t) \right]_{t_A}^{t_B} - \frac{dS_\varepsilon[q]}{d\varepsilon} \Big|_{\varepsilon=0} = 0 \quad (2.6)$$

$$\begin{aligned} \frac{dS[\phi_\varepsilon[q]]}{d\varepsilon} \Big|_{\varepsilon=0} &= \int_{t_A}^{t_B} dt \sum_{i=1}^n \left[\frac{\partial L(q, \dot{q}, t)}{\partial v^i} \frac{d\phi_\varepsilon^i(q)}{d\varepsilon} \Big|_{\varepsilon=0} + \frac{\partial L(q, \dot{q}, t)}{\partial w^i} \frac{d\dot{\phi}_\varepsilon^i(q)}{d\varepsilon} \Big|_{\varepsilon=0} \right] \\ &= \int_{t_A}^{t_B} dt \sum_{i=1}^n \left[\frac{d}{dt} \frac{\partial L(q, \dot{q}, t)}{\partial v^i} \frac{d\phi_\varepsilon^i(q)}{d\varepsilon} \Big|_{\varepsilon=0} + \frac{\partial L(q, \dot{q}, t)}{\partial w^i} \frac{d\dot{\phi}_\varepsilon^i(q)}{d\varepsilon} \Big|_{\varepsilon=0} \right] \\ &= \int_{t_A}^{t_B} dt \frac{d}{dt} \left[\sum_{i=1}^n \frac{\partial L(q, \dot{q}, t)}{\partial v^i} \frac{d\phi_\varepsilon^i(q)}{d\varepsilon} \Big|_{\varepsilon=0} \right] \end{aligned} \quad (2.7)$$

$$\left[\Lambda(q, \dot{q}, t) \right]_{t_A}^{t_B} - \frac{dS_\varepsilon[q]}{d\varepsilon} \Big|_{\varepsilon=0} = 0 \quad (2.8)$$

$$\Rightarrow \left[\Lambda(q, \dot{q}, t) - \sum_{i=1}^n \frac{\partial L(q, \dot{q}, t)}{\partial v^i} \frac{d\phi_\varepsilon^i(q)}{d\varepsilon} \Big|_{\varepsilon=0} \right]_{t_A}^{t_B} = 0 \quad (2.9)$$

2.4.2 Tidstranslation og Energibevarelse

Lad os nu betragte en specifik form for kontinuert transformation, kaldet tidstranslation, defineret på følgende måde.

$$T : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \quad , \quad T_\varepsilon(t) := t - \varepsilon \quad (2.10)$$

$$S_\varepsilon[q] = \int_{t_A}^{t_B} dt L(q, \dot{q}, T_\varepsilon(t)) = \int_{t_A}^{t_B} dt L(q, \dot{q}, t - \varepsilon) \quad (2.11)$$

$$\frac{dS_\varepsilon[q]}{d\varepsilon} \Big|_{\varepsilon=0} = - \int_{t_A}^{t_B} dt \frac{\partial L(q, \dot{q}, t)}{\partial u} \quad (2.12)$$

$$H(v, w, u) := \sum_{i=1}^n v_i \frac{\partial L(v, w, u)}{\partial w^i} - L(v, w, u) \quad (2.13)$$

$$\Rightarrow \frac{dH(q, \dot{q}, t)}{dt} = - \frac{\partial L(q, \dot{q}, t)}{\partial u} \quad (2.14)$$

$$\frac{dS_\varepsilon[q]}{d\varepsilon} \Big|_{\varepsilon=0} = 0 \quad \Rightarrow \quad \left[H(q, \dot{q}, t) \right]_{t_A}^{t_B} = 0 \quad (2.15)$$

Kapitel 3

Fra konfigurations-rum til fase-rum

3.1 Opgaver

3.1.1 Partikel på kugleskal

Find bevægelsesligningerne for en partikel med masse m begrænset til kugleskallen S^2 , givet ved følgende punktmængde:

$$S^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\} \quad (3.1)$$

Lad desuden partiklen være påvirket af en konstant tyngdekraft givet ved:

$$F_g = mg \quad (3.2)$$

Løsning Vi søger et sæt generaliserede koordinater som kan parameterisere partiklens bevægelse på kugleoverfalden. Et oplagt valg er polære kugle-koordinater.

$$x = \sin(\theta) \cos(\phi) \quad , \quad y = \sin(\theta) \sin(\phi) \quad , \quad z = \cos(\theta) \quad (3.3)$$

$$\phi \in [0, 2\pi) \quad , \quad \theta \in (0, \pi) \quad (3.4)$$

$$\dot{x} = \dot{\theta} \cos(\theta) \cos(\phi) - \dot{\phi} \sin(\theta) \sin(\phi) \quad (3.5)$$

$$\dot{y} = \dot{\theta} \cos(\theta) \sin(\phi) + \dot{\phi} \sin(\theta) \cos(\phi) \quad (3.6)$$

$$\dot{z} = -\dot{\theta} \sin(\theta) \quad (3.7)$$

$$\dot{x}^2 + \dot{y}^2 + \dot{z}^2 = \dot{\theta}^2 + \sin^2(\theta) \dot{\phi}^2 \quad (3.8)$$

$$L = T - V = \frac{m}{2} \left(\dot{\theta}^2 + \sin^2(\theta) \dot{\phi}^2 \right) - mg \cos(\theta) \quad (3.9)$$

$$\frac{\partial L}{\partial \phi} = 0 \quad , \quad \frac{\partial L}{\partial \dot{\phi}} = m \sin^2(\theta) \dot{\phi} \quad (3.10)$$

$$\Rightarrow \quad \frac{d}{dt} \left(\sin^2(\theta) \dot{\phi} \right) = 0 \quad (3.11)$$

$$\frac{\partial L}{\partial \theta} = m \sin(\theta) \cos(\theta) \dot{\phi}^2 + mg \sin(\theta) \quad , \quad \frac{\partial L}{\partial \dot{\theta}} = m \dot{\theta} \quad (3.12)$$

$$\Rightarrow \quad \ddot{\theta} = \sin(\theta) \cos(\theta) \dot{\phi}^2 + g \sin(\theta) \quad (3.13)$$

$$H(\phi, \theta, p_\phi, p_\theta) = \frac{1}{2m} \left(p_\theta^2 + \frac{p_\phi^2}{\sin^2(\theta)} \right) + mg \cos(\theta) \quad , \quad \dot{H} = 0 \quad (3.14)$$

$$Q(\phi, \theta, p_\phi, p_\theta) = \frac{p_\phi}{m} \quad , \quad \dot{Q} = 0 \quad (3.15)$$

$$H = \frac{m}{2} \left(\dot{\theta}^2 + \frac{Q^2}{\sin^2(\theta)} \right) + mg \cos(\theta) \quad \Rightarrow \quad \dot{\theta}^2 = \frac{2H}{m} - \frac{Q^2}{\sin^2(\theta)} - \frac{g \cos(\theta)}{2} \quad (3.16)$$