

# Boundary value problems and the finite difference method MA8020 Lecture notes

Pererik Andreasson  
Email pererik.andreasson@hh.se

December, 2021

## 1 Introduction

This document describes (although in a slightly unstructured manner) what will be required on a written exam in the course *Tekniska beräkningar MA8020*. I will, rather boldly, claim that it does not contain any direct errors, but it is far from mathematically stringent. Sometimes, short cuts are taken to enhance *understanding* of the concepts. Well founded criticism is gladly accepted.

## 2 Boundary value problems (BVP)

Boundary value problems are differential equations, where the solution is known on the boundaries (edges of the interesting area). Here two typical problems will be treated, the catenary curve, i.e., a string that is attached in both ends and the example of heat conduction in a rod where both ends are held at constant temperatures. A more intricate example (which might come in the future): a string. The function which describes the deviations from the equilibrium for the string will satisfy the wave equation. The two last examples will both yield partial differential equations (PDE), but with sufficient simplifications we can turn them into relatively simple (but important) boundary value problems (this material does not really treat PDEs).

### 2.1 A little introductory mechanics: The catenary curve

The following section is taken more or less directly from the introduction chapter in Powers' book *Boundary Value Problems: And Partial Differential Equations*[1].

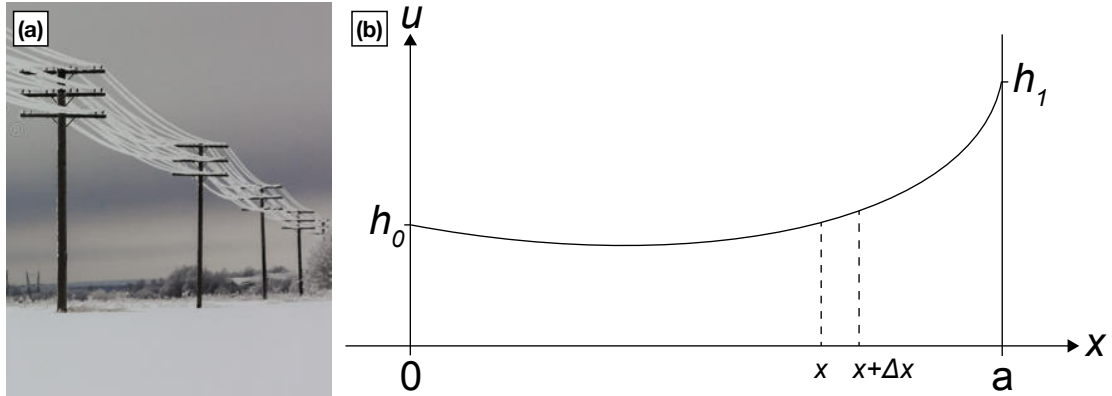


Figure 1: (a) Still hanging power cables hanging between posts (with snow). (b) Mathematical model for one portion of the cable hanging between two points (at  $x = 0$  and  $x = a$ ).

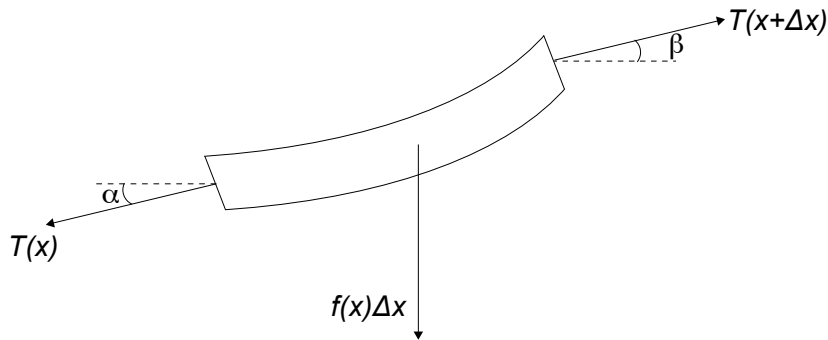


Figure 2: A small piece of the model of the cable which is presented in figure 1 (b). Here  $T$  are the tension force along the cable and  $f(x)\Delta x$  is the gravitational force.

In figure 2 a very small piece of the hanging string is schematically (and by no means physically correct) presented. For the following discussions it is assumed that this piece is not accelerating sideways, which yields that the tension in the string has the same magnitude on both sides, although at a slightly different angle (possibly). In the figure, three forces are shown: the tension on the left side, denoted  $T(x)$ ; the tension on the right side, denoted  $T(x + \Delta x)$ ; and gravity, acting at the center of the piece where  $f(x)$  is the weight per unit length and  $\Delta x$  is the length of the piece.

From Newton's second law, we have  $\sum F = ma$ , if the piece is hanging "still" there is no acceleration and the sum of all the forces must equal zero:

$$\sum F_x = 0 \quad (1)$$

$$\text{and } \sum F_y = 0. \quad (2)$$

If the horizontal forces are summed up:

$$\sum F_x = T(x + \Delta x) \cdot \cos(\phi(x + \Delta x)) - T(x) \cdot \cos(\phi(x)) = 0 \quad (3)$$

and the vertical forces:

$$\sum F_y = T(x + \Delta x) \sin(\phi(x + \Delta x)) - T(x) \cdot \sin(\phi(x)) - f(x) \cdot \Delta x = 0. \quad (4)$$

Since the magnitude of the tension forces on each side are the same (no acceleration) we can simplify and use  $T(x + \Delta x) = \frac{T}{\cos(\phi(x + \Delta x))}$  and  $T(x) = \frac{T}{\cos(\phi(x))}$  to simplify the vertical expression (eqn. ) to:

$$T \cdot \frac{\sin(\phi(x + \Delta x))}{\cos(\phi(x + \Delta x))} - T \cdot \frac{\sin(\phi(x))}{\cos(\phi(x))} - f(x) \cdot \Delta x = 0. \quad (5)$$

Which almost automatically yields:

$$T \cdot (\tan(\phi(x + \Delta x))) - T \cdot (\tan(\phi(x))) - f(x) \cdot \Delta x = 0. \quad (6)$$

So far so good, but now we need to use a quite elaborate trick. **Pay attention:**  $\phi(x)$  measures the angle between the tangent to the string ("the center line" of the small piece in figure 2) and the horizontal. This gives us that the expression  $\tan(\phi(x))$  is just the "slope" of the string in the point  $x$ . If we call the curve that the string describes  $\mathbf{u}(\mathbf{x})$ , the slope, or the tangent is just the derivative of this curve with respect to  $x$ :

$$\tan(\phi(x)) = \frac{du(x)}{dx}. \quad (7)$$

If the tan is replaced in equation 6 by the derivative of  $u$  we get something that seems to be ridiculously close to the definition of the derivative:

$$T \cdot \left( \frac{du(x + \Delta x)}{dx} - \frac{du(x)}{dx} \right) = f(x) \Delta x. \quad (8)$$

In general, depending on  $f(x)$  this can be quite tricky to sort out, but if we choose  $f(x)$  to something simple, like  $f(x) = \frac{mg}{L}$  which is independent of  $x$  we should be able to simplify further and also quickly compute it analytically. Here  $m$  is the

total mass of the string,  $g$  is the acceleration due to gravity and  $L$  is the total length of the string. With all this taken into account, we end up with

$$\frac{u'(x + \Delta x) - u'(x)}{\Delta x} = \frac{mg}{LT}. \quad (9)$$

If then  $\Delta x$  is chosen sufficiently small, we get the definition of the derivative and the final differential equation that describes how the string is hanging to be:

$$\frac{d^2u}{dx^2} = \frac{mg}{LT}. \quad (10)$$

## 2.2 A little bit of thermodynamics: Heat conduction *to be completed*

*Sorry about the language switch, the opinion is that English will be more useful in the long run.* We will use a very simple example here to introduce an intuitive situation where the physics will lead to a boundary value problem. It will be introduced in one dimension to keep the math simple so we can focus on understanding the physics and the math of a typical boundary value problem. A uniform rod of a certain material will be kept with each end of the rod at different temperatures,  $T_1$  and  $T_2$ . A schematic illustration for the purpose of explaining the different physical parameters is shown in figure 3.

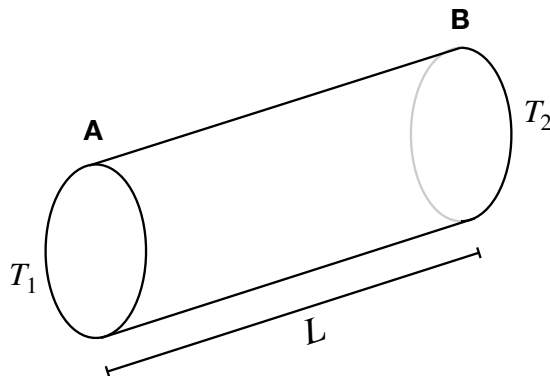


Figure 3: A uniform rod of length  $L$  with each end held at a constant temperature,  $T_1$  and  $T_2$  for the left (marked A) and the right (marked B) hand side respectively. In this case,  $T_1 > T_2$ . The side of the cylinder (along  $L$ ) is isolated from the surroundings.

The laws of thermodynamics tells us that if we have two connected regions with different temperatures, heat will flow *from* the region with higher temperature to

the region with lower temperature. If it is assumed that  $T_1 > T_2$  for the situation in figure 3, we then expect that heat will flow in the rod from the left side (marked A) to the right side (marked B). The physics here is governed by *Fourier's law* of heat transfer, rate of heat transfer is proportional to the negative temperature gradient. If we denote the unknown temperatures with  $u$  and the spatial coordinate along  $L$  with  $x$ , Fourier's law becomes:

$$\frac{\text{Rate of heat transfer}}{\text{area}} = -K_0 \frac{du}{dx}. \quad (11)$$

In equation 11 above,  $K_0$  pops up as the *thermal conductivity*, which is an intrinsic material property of all materials. For simplicity, we assume this to be constant along the rods length.

This will yield a partial differential equation (PDE). How do we force steady state?

We derive and set the time to infinity: yields a very simple solution but introduces boundary value problems nicely.

*Följande avsnitt är helt eller delvis inspirerat av Heath: Scientific Computing An Introductory Survey[2], där inte andra källor anges.*

### 3 Finita differensmetoden

Finita differensmetoden används till att reducera randvärdesproblem till ett algebraiskt problem. I denna metod ersätts derivator med *finita differenser*. (Jag håller med, den förklaringen hjälpte inte...) Enklaste exemplet är att vi ersätter en förstaderivata först, sen gräver vi ner oss något. Du *måste* komma ihåg derivatans definition, här enligt den ökända felkällan WIKIPEDIA<sup>1</sup>:

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}. \quad (12)$$

För att vara sträng och matematisk bör dessa härledas från Taylor-utvecklingar, men för att förstå poängen anser jag att det räcker med derivatans definition. Kom ihåg från Bertils/Mikaels föreläsning att det finns tre olika sätt för förstaderivatan, framåt-, bakåt- och centraldifferensmetod. Förstaderivatan använder vi centraldifferens:

$$y'(x) \approx \frac{y(x + h) - y(x - h)}{2h}. \quad (13)$$

För att åstadkomma en andraderivata används framåt- och bakåt differens *tillsammans* för att få en rimlig andraderivata. Jag kallar framåt differensderivatan för

---

<sup>1</sup>I denna frågan litar jag blint på WIKIPEDIA dock.

$D_+(x)$  och bakåtdifferensderivatan för  $D_-(x)$  så ges andraderivatan på ett snyggt sätt av:

$$f''(x) \approx \frac{D_+(x) - D_-(x)}{h} \quad (14)$$

$$= \frac{\frac{f(x+h)-f(x)}{h} - \frac{f(x)-f(x-h)}{h}}{h} \quad (15)$$

$$= \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}. \quad (16)$$

Här är egentligen hela iden till finita differensmetoden framställd i all sin effektiva enkelhet, det handlar alltså om att stega en viss bit i funktionen ( $h$ ) och genom att "gissa" lösningsvärden iterativt lösa problemet. När du tittar på ekv. 16 skall du alltså tänka  $x_i$ ,  $x_{i+1}$  och  $x_{i-1}$ . För att exemplifiera löser vi följande problem:

$$y'' = 6t, \quad 0 \leq t \leq 1, \quad (17)$$

med randvärdessvillkoren:

$$y(0) = 0, \quad y(1) = 1. \quad (18)$$

Vi nöjer oss nu i demo-syfte att beräkna funktionsvärdet i en endaste punkt i intervallet,  $t = 0.5$ . Om vi kallar punkterna  $y_0$  osv då  $t = 0$  osv kan vi använda oss direkt av ekv. 16 med  $h = 0.5$  och sätta in i ekv. 17:

$$\frac{y_2 - 2y_1 + y_0}{h^2} = 6t_1. \quad (19)$$

Här vet vi  $y_2 = 1$  och  $y_0 = 0$  och  $t_1 = 0.5$ , därför blir ekvationen vi skall lösa:

$$\frac{1 - 2y_1 + 0}{(0.5)^2} = 6(0.5) \quad (20)$$

$$4 - 8y_1 = 3 \quad (21)$$

$$y_1 = \frac{1}{8}. \quad (22)$$

Lösningen är alltså  $y(0.5) = \frac{1}{8}$  i den obekanta punkten. Hur nära sanningen är vi med detta då? Den inbyggda lösaren `dsolve()` i MATLAB ger mig lösningen  $y(t) = t^3$ , plottar jag den jämte vår finita differenslösning ser det ut som i figur 4:

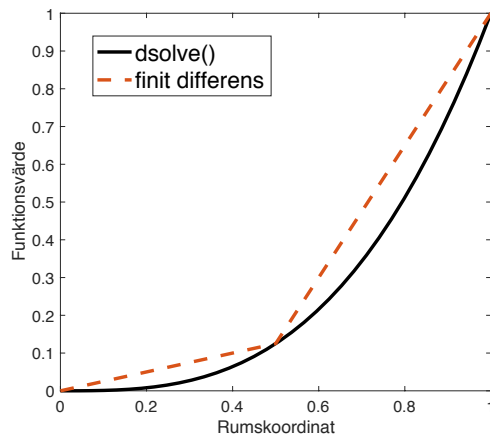


Figure 4: Lösningen till vårt randvärdesproblem, heldragen linje är från *dsolve()* och streckad är från vår finita differensekvationen enligt ekvation 22

### 3.1 Övning till finita differensmetoden

Lös följande randvärdesproblem med differensmetoden, gör för hand på papper först och implementera sedan i MATLAB eller MATHEMATICA.

1.  $y'' = y' + y + 2x$ , för  $0 \leq x \leq 0.4$  och med  $y(0) = 0.4$ ,  $y(0.4) = 1$ .
2.  $y'' - y = 0$ , för  $0 \leq t \leq 1$  och med  $y(0) = 0$ ,  $y(1) = \sinh(1)$ .
3.  $y'' - y = x^2 - 2$ , för  $0 \leq t \leq 1$   $y(0) = 1$ ,  $y(1) = \cosh(1) - 1$ .

### 3.2 Seminarieövning finita differensmetoden

Gör ett datorprogram i MatLab eller Mathematica som löser randvärdesproblemet

$$y'' + 2y' + y = 0, y(0) = 0, y(1) = 1 \quad (23)$$

Du skall leverera en muntlig presentation med hjälp av din dator som beskriver följande saker:

- En lösning av ekv. 23 med finita differensmetoden, presentera de olika delarna av ditt datorprogram på ett överskådligt sätt.
- Köra ditt program för väldigt grovt intervall (tre delar) något fler (kanske 20) och jättemånga (100). Se till så det inte tar för lång tid att köra för 100-delat intervall.

- Plotta din lösning som punkter över en linje som motsvarar en lösning av någon annan lösare (tex `bvp()` i MatLab, eller analytisk). Skall uppdateras för varje körning med intervallet i de olika delarna.



## 4 Finita elementmetoden (För kännedom, kommer inte på tentan.)

Finita elementmetoden används också till att reducera randvärdesproblem till ett algebraiskt problem. I denna metoden approximeras lösningen till randvärdesproblemet av en lineärkombination av ett begränsat antal basfunktioner, *element*. Dessa basfunktioner, eller element, brukar vara polynom och i Heaths bok kallas dessa  $\phi_i$ . Lösningen,  $y(t)$ , blir således på formen

$$y(t) \approx u(t) = \sum_{i=1}^n x_i \phi_i \quad (24)$$

Koefficienterna  $x_i$  bestäms genom att man sätter olika villkor på *residualen*, definierad som skillnaden mellan vänster- och högerledet i differentialekvationen. De huvudsakliga metoder som används är:

- Kollokation: Residualen är noll, differentialekvationen är exakt löst i  $n$  diskreta punkter.
- Galerkin: Residualen är ortogonal mot rummet som basfunktionerna spänner upp.
- Rayleigh-Ritz: Residualen är minimerad på ett viktat minsta-kvadrat-sätt.

De två senare metoderna är likartade, vi vill alltså hitta en lineärkombination av basfunktionerna som approximerar lösningen till problemet för att residualen blir så liten som möjligt. Alla tre metoder leder till ett system av ekvationer för att hitta koefficienterna  $x_i$ . Detta system kan vara lineärt eller icke-lineärt, beroende på om  $f$  är lineär eller inte. För att systemet skall vara glest, är det en fördel om basfunktionerna är lokala, dvs noll överallt utom på ett litet ställe. Typiska funktioner i en dimension är B-splines, bitvis lineära "häxhatt"-funktioner, se figur 5.

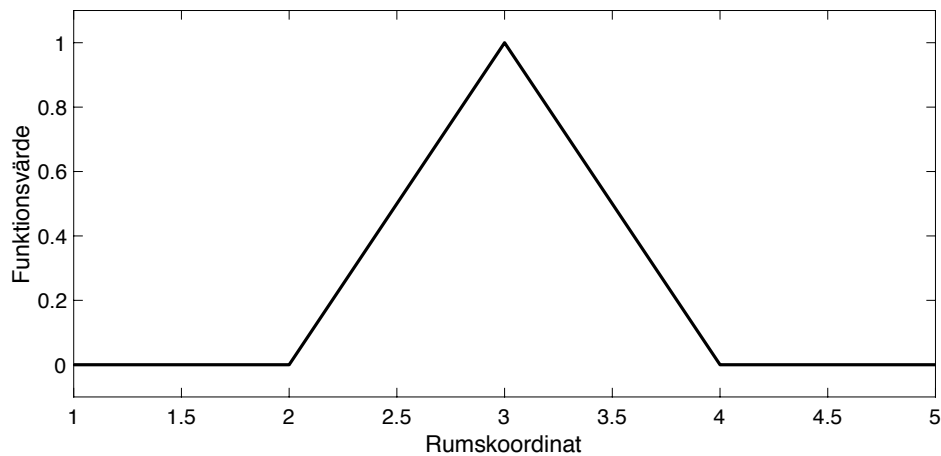


Figure 5: B-spline  $B_3^1$ , häxhattsfunktionen, typisk basfunktion (element) för finita elementmetoden i  $1D$ .

Den glesa matrisen som blir resultatet av detta kallas av traditionella anledningar för “styvhetsmatrisen<sup>2</sup>”. Om andra basfunktioner används än dessa enkla “häxhatt”-funktioner, kallas metoden att lösa spektral eller pseudospektral. (Det är alltså *inte* det vi skall göra i denna kursen.)

## Galerkinmetoden

I kursen kommer vi inte djupdyka i de olika metoderna och skillnaden mellan dem, det anses tillräckligt att du kan lösa enkla randvärdesproblem med Galerkinmetoden som beskrivs i detta avsnitt.

Vi vill här lösa randvärdesproblemet som beskrivs av differentialekvationen i ekv. 17 och använder Galerkinmetoden och bitvis lineära basfunktioner (häxhattar). Vi delar upp intervallet i tre punkter och har en häxhattstopp i varje punkt på intervallet, enligt figur 6.

---

<sup>2</sup>Vilket kan ha med mekanik tillämpningar att göra...

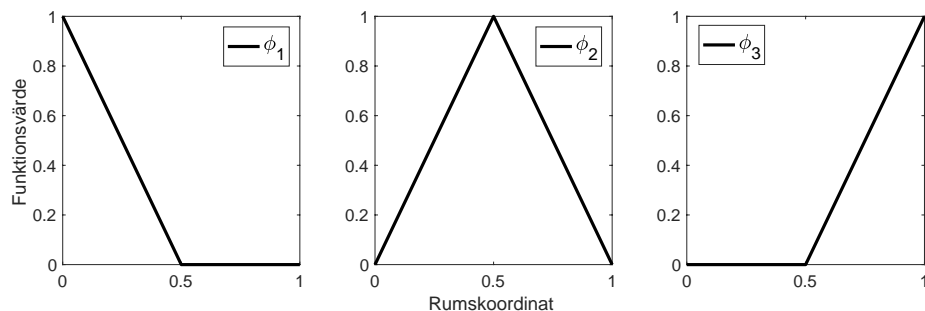


Figure 6: Hächhattsfunktionerna som vi använder för att lösa randvärdesproblemet i ekvation 17

Vår finita elementlösning till problemet ser alltså ut på följande sätt:

$$y(t) \approx u(t) = x_1\phi_1(t) + x_2\phi_2(t) + x_3\phi_3(t). \quad (25)$$

Från randvillkoren kan vi direkt bestämma  $x_1$  och  $x_3$ , där vet vi vad lösningen är,  $x_1 = 0$  och  $x_3 = 1$ . För att bestämma  $x_2$  använder vi oss av Galerkins metod. Ovan nämnde jag i förbifarten att Galerkinvillkoret är att residualen är ortogonal mot det rum som basfunktionerna spänner upp. Ett sådant uttalande gör ganska ont i ögonen men här kommer förklaringen på hur det ser ut i praktiken. Ortogonal betyder att de är vinkelräta, vilket för funktioner inte riktigt har någon betydelse (för vektorer är det väldigt logiskt dock). För två ortogonala vektorer är skalärprodukten noll, vi behöver alltså en *annan* definition av skalärprodukten, en som funkar för våra små basfunktioner (hächhattar). Detta kommer bli sköj! Två funktioner,  $f$  och  $g$ , definieras som ortogonala på ett intervall  $a \leq x \leq b$  om

$$\int_a^b f(x)g(x)dx = 0. \quad (26)$$

Residualen,  $\mathcal{R}$ , som vi skall använda är, från ekv. 17:

$$\mathcal{R} = y''(t) - 6t \approx u''(t) - 6t. \quad (27)$$

Med ortogonalitetsvillkoret från ekv. 26 mellan  $\mathcal{R}$  och basfunktionen  $\phi_2$  får vi integralen:

$$\int_0^1 (u''(t) - 6t)\phi_2(t)dt = \quad (28)$$

$$\int_0^1 u''(t)\phi_2(t)dt - 6 \int_0^1 t\phi_2(t)dt = 0. \quad (29)$$

Vi börjar med att lösa första intergralen i högerledet från ekv. 29, den löses lämpligast i två steg, först med partiell integration enligt:

$$\int_0^1 u''(t)\phi_2(t)dt = u'(t)\phi_2(t)|_0^1 - \int_0^1 u'(t)\phi_2'(t)dt. \quad (30)$$

Vi vet att  $\phi_2$  är noll i  $t = 0$  och  $t = 1$ , därför kan första delen av partialintegreringen sättas lika med noll. I andra delen av partialintegreringen måste vi sätta in vår approximation för  $u(t)$ , dvs det vi har i ekv. 25, här sätter jag in med summatecken för att komprimera raderna något:

$$\int_0^1 u'(t)\phi_2'(t)dt = \int_0^1 \left( \sum_{i=1}^3 x_i\phi_i' \right) \phi_2'(t)dt \quad (31)$$

$$= x_1 \int_0^1 \phi_1'(t)\phi_2'(t)dt + x_2 \int_0^1 \phi_2'(t)\phi_2'(t)dt + x_3 \int_0^1 \phi_3'(t)\phi_2'(t)dt. \quad (32)$$

Integralen i högerledet av ekv. 32 löser vi med *kvadratur*, till exempel trapetsregeln:

$$I(f) = \frac{b-a}{2} (f(a) + f(b)) \quad (33)$$

Där  $f$  i vårt fall motsvaras av  $\phi_i'\phi_2'$ . Kom ihåg att våra häxhattar är väldigt lätta att derivera! (Även om det är klurig gränfallsräkning.) Dvs, vi har tex några givna,

$$\phi_1'(0) = -2 \quad \phi_1'(1) = 0 \quad (34)$$

$$\phi_2'(0) = 2 \quad \phi_2'(1) = -2 \quad (35)$$

$$\phi_3'(0) = 0 \quad \phi_3'(1) = 2. \quad (36)$$

Det hade sannolikt varit "snyggare" ur ett strikt numeriskt perspektiv att även approximera derivatorna med någon form, men det funkar även så här (vilket är matematiskt något slappt). Detta ger oss det ganska långa (men nu väldigt enkla):

$$x_1 \frac{1}{2} (-4 + 0) + x_2 \frac{1}{2} (4 + 4) + x_3 \frac{1}{2} (0 - 4) = \quad (37)$$

$$-2x_1 + 4x_2 - 2x_3. \quad (38)$$

Nu behöver vi bara lösa andra delen av vänsterledet i ekv. 29, vilket inte borde vara så svårt. Givetvis gör vi detta på samma sätt med kvadratur. Eftersom det är  $\phi_2$  får vi vara lite försiktiga med att använda hela intervallet  $0 \leq t \leq 1$ , eftersom den är noll i ändpunkterna. Vi delar upp i två intervall och använder trapetsregeln igen, där  $f(t) = t\phi_2(t)$ :

$$-6 \int_0^1 t\phi_2(t)dt \quad (39)$$

$$\approx -6 \left( \frac{0.5-0}{2} [f(0) + f(0.5)] + \frac{1-0.5}{2} [f(0.5) + f(1)] \right) \quad (40)$$

$$= -6 \left( \frac{0.5}{2} 0.5 + \frac{0.5}{2} 0.5 \right) = -\frac{3}{2}. \quad (41)$$

Denna något sifferrika framställning leder oss till att när vi sätter in våra kända  $x_1$  och  $x_3$  måste vi lösa den högst triviala ekvationen:

$$4x_2 - 2 = -\frac{3}{2} \Rightarrow \quad (42)$$

$$x_2 = -\frac{3}{8} + \frac{4}{8} \Rightarrow \quad (43)$$

$$x_2 = \frac{1}{8}. \quad (44)$$

Vår approximativa lösning till randvärdesproblemet i ekv. 17 med Galerkinmetoden fås slutligen till följande:

$$y(t) \approx u(t) = \frac{1}{8}\phi_2(t) + \phi_3(t) \quad (45)$$

Hur nära sanningen är vi med detta då? Den inbyggda lösaren *dsolve()* i MATLAB ger mig lösningen  $y(t) = t^3$ , plottar jag den jämte vår Galerkinlösning ser det ut som följer:

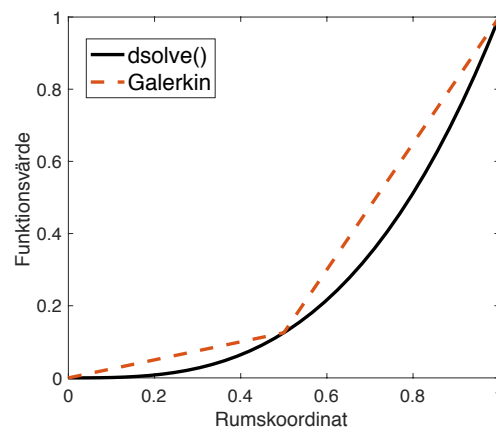


Figure 7: Lösningen till vårt randvärdesproblem, heldragen linje är från *dsolve()* och streckad är från vår Galerkinapproximation enligt ekvation 45. Observera att den streckade linjen endast innehåller information i punkterna  $x = 0$ ,  $x = 0.5$  och  $x = 1$ , däremellan är det en rät linje som inte har något med den verkliga lösningen (eller den approximativa) att göra.

I detta fall blir lösningen alltså exakt i punkten  $x_2$ .

## References

- [1] D. L. Powers. *Boundary Value Problems And Partial Differential Equations*. Elsevier Science & Technology, 2005.
- [2] M. T. Heath. *Scientific computing: an introductory survey*. Series in Computer Science. McGraw-Hill, 1997.