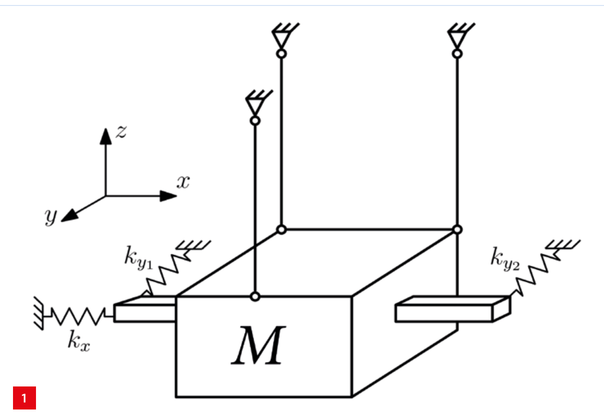
DIRECT OR INDIRECT? THAT'S THE QUESTION

What if it is possible to create an indirect-drive actuator with the benefits of a direct drive? Typical struggles in actuator selection are to balance upsides, such as high force, with the downsides of additional mass, volume and availability. Adding a transmission has its own downsides, such as friction, backlash and hysteresis. But what if a gear ratio can be created with the aid of flexure mechanics, eliminating all these downsides? How does one deal with the parasitic movements, the own guiding stiffness, dynamics and motor choice in a static use case? Here, a pragmatic implementation is presented of an (in)direct-drive actuator that is self-guided and compliant in multiple passive directions to be used in a steady-state multi-axis kinematically coupled system. Stiffness is dominant and accelerations are low, putting the linear motor outside its comfort zone of merely high-dynamic motion.

NIELS KOENRAAD, FRANK SPERLING AND PETER KLEIN MEULEMAN

Load case

Before diving into the principles of direct versus indirect drive actuation it is essential to fully define the load case at hand. This consists of a large mass (> 100 kg) that needs to be actuated in two degrees of freedom (DoFs), i.e. in the *x*- and *y*-direction. A third DoF, the rotation around *z*, needs to be constrained either passively or actively. Theoretically, this can be done by using an intermediate body or another construction element like a bellow, which in turn will introduce its own dynamics and increase the overall complexity. This direction is therefore actively constrained by a third actuator instead. This has as an extra benefit that two actuators (y_1 and y_2) can be used for the *y*-direction (and *Rz*), which for various reasons provides significantly more stiffness in the actuation direction. This in turn possibly allows for the use of the same actuator type for all three positions (one ~~size~~ 12NC fits all).



Load guided and passively constrained using flexure mechanics leaving three actively controlled DoFs.

A schematic depiction of the load case is shown in Figure 1. The mass is suspended from the top by three flexible elements that together constrain it in *z, Rx, Ry* and allow for movement in the other DoFs. The actuators are represented by a spring with servo stiffness k_x, k_{y1} and k_{y2}, respectively, with one *y*-actuator on each side to also control *Rz*, and a single *x*-actuator on the right side. From this layout it is apparent that the actuators need to be able to cope with movements perpendicular to the actuator direction: firstly, the functional stroke of the other actuators (millimeter range) and secondly, the parasitic move of the flexure mechanics (tens of microns).

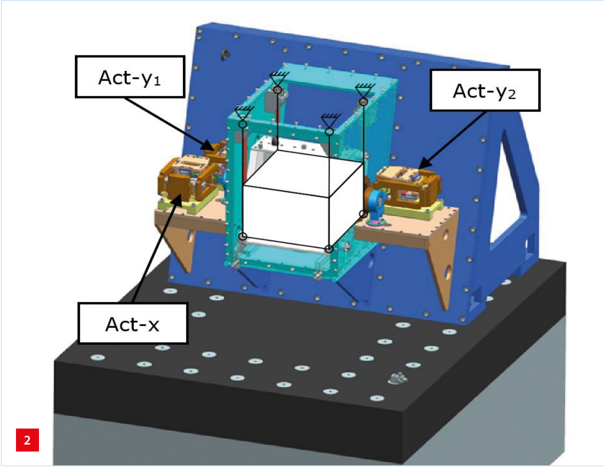
The load will be moved with low acceleration to setpoints where it will remain stationary for long periods of time. During this time, the main function of the actuators is to suppress disturbances from the environment, such as floor vibrations and any flow disturbances to keep a steady position (i.e. trajectories of < 100 µm, < 0.5 Hz). Flexure mechanics can maintain a stable position and enable submicron movements without hysteresis or (local) wear. This will result in a system with high actuation stiffness and low acceleration forces.

In trying to identify and verify all dynamics and mechanics, a functional model (fumo) was made of the load as well as for the different actuator concepts that will be discussed in this article. This fumo is shown in Figure 2, where a mass is suspended from four rods rather than three, yielding an over-determined system. This results from the optimisation process between eigenmodes, actuation stiffness, stresses

AUTHORS' NOTE

Niels Koenraad (system architect), Frank Sperling (director of technology) and Peter Klein Meuleman (lead mechanical engineer) all work at Nobleo Technology, an engineering company based in Eindhoven (NL).

niels.koenraad@nobleo.nl
www.nobleo.nl



2

Functional model set-up with the load suspended from four rods and actuated by three actuators, one for the x-direction and two controlling the y- and Rz-direction together.

and more that yielded major benefits at the cost of being overdetermined, which to a great extent can be solved in the manufacturing and assembly process. In addition, three actuators that will actuate the mass can be seen in Figure 3, showing their layout with respect to the load itself.

Direct vs Indirect drive

With the load case defined, an actuator can be selected. To sum up: a static use case with little to no peak forces, with parasitic movements in the same order of magnitude as the actuation stroke itself, micron-level movements, no wear, no friction and the ability to suppress high disturbance forces. To add one more requirement not related to the load case per se but to the overall architecture, there is a limited voltage supply (V_{bus} < 30 V) and current (I_{rms} < 3 A). Given this description a logical choice is a direct-drive actuator, which certainly eliminates wear and friction but does not necessarily make it suitable for static use and parasitic movements. Furthermore, it is not so trivial how this choice will affect the capability to suppress disturbances. An alternative might be the indirect drive, but this obviously has drawbacks.

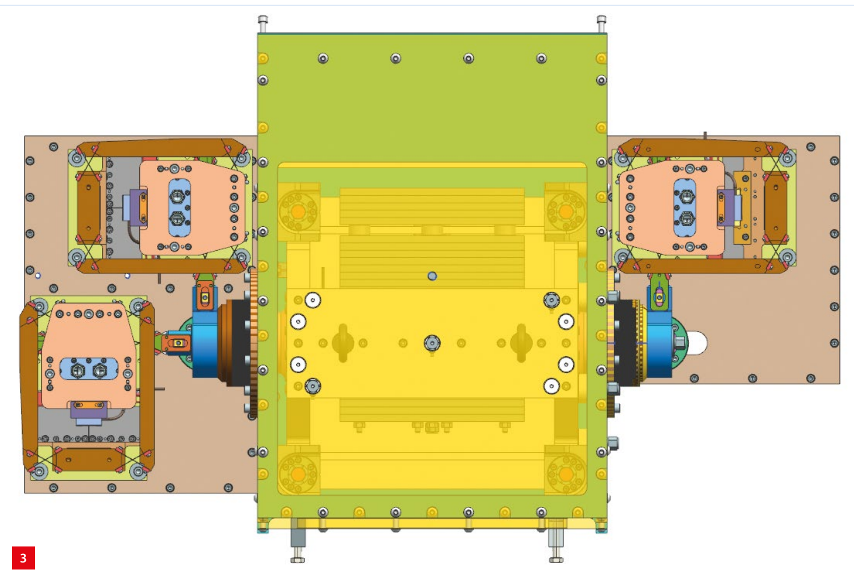
Direct drive

A direct-drive actuator is simply a motor directly connected to the load. For linear motion this is typically in the form of a voice-coil or linear 3-phase motor. The latter is the most suitable in this case, since a voice-coil is constrained in two DoFs by its airgaps, whereas the 3-phase motor is only limited in one DoF. The result is a simple solution with a motor directly fixed to the mass and the stator fixed to the force frame, with only a connecting element, i.e. the servo stiffness k_p as shown in Figure 4. The parasitic movements of one actuator will cause the coils of the perpendicular actuator to move out of the magnet yoke by a few millimeters, resulting in motor constant loss and an additional increase of the complexity of the sensor system. Here, an optical encoder is used to determine the actuation position, which requires a large lateral tolerance for the perpendicular motion

So far nothing is insurmountable, but another issue is the required force (> 100 N) that will have to be created with the limited amplifier supply. For this concept the only possibility to increase force is by increasing the motor size. This is a somewhat inefficient process, because larger motors produce more N^2/W but will see less and less of their potential used. For example, the design might end up with a 500-W motor for a 100-W application simply because there are no other knobs to turn.

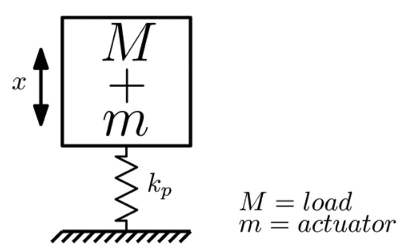
Indirect drive

In a lot of applications, fitting a transmission such as a planetary gear, toothed belt or spindle to the actuator is a solution that will make it possible to create high forces (or torques) with a relatively small motor. This typically has drawbacks, the worst being backlash (i.e. play) and friction, which make it especially ill-fitted for an application where submicron positioning is required. An alternative approach is to create a lever using elastic elements, where the distance



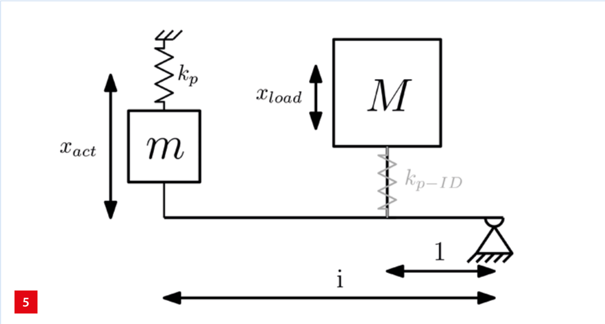
3

Top view of the functional model with the actuator lay-out clearly visible. The bottom left actuator controls the x-direction, the top left and top right actuators control the y-direction and Rz-rotation together.



4

Direct-drive actuator connected to the load, resulting in a servo stiffness to the fixed world and a load mass (M) that is increased by the actuator mass (m).



Indirect-drive actuator connected to the load via a gear ratio of 1:i, with stroke and force being scaled by this ratio. Stiffnesses drawn here are the servo stiffness, k_p, and that same stiffness scaled on the load side, $k_{p\text{-}ID}$. Stiffnesses omitted from this drawing are the gear stiffness and load stiffness, which are important to keep in mind for the force budget and will be discussed in the optimisation of the gear ratio itself.

of the motor to the (elastic-) pivot point and that of the load to the pivot point determines the gear ratio, as shown in Figure 5. The force output of the actuator can now be influenced by both the motor size and the gear ratio.

For this to work, a pivot point needs to be created that allows motion and guarantees sufficient stiffness. Additionally, this will have both upsides and downsides regarding the dynamics and control of the complete system. The creation of this pivot point and (optimal) lever will be discussed in the next section. First, a less trivial benefit will be discussed, that of the impact on the servo stiffness.

Servo stiffness amplifier

The main function of the actuator is to suppress disturbances from the environment. This requires a high servo bandwidth or rather high servo stiffness. The disturbances, the dynamics of the load and the stiffness of the reaction path (k_{fp}) are a given, so the last variable in the equation is the actuator itself. Figure 6 shows the load with both the direct-drive (grey) and the indirect-drive (black) actuator, for the first of these the achievable bandwidth will be dictated by the dynamics of the load (which will be negatively influenced by the additional motor mass, because the actuator mass is fully supported by the flexure mechanics of the load). For ease of argument, *m* is identical for the direct and the indirect drive. The servo stiffness of the direct drive is then:

$$k_p = k_{p\text{-DD}} = (M + m)(2\pi \cdot f_{bw\text{-DD}})^2$$

Note: the reaction path is the force path from the static part of the motor to the 'fixed' world, whereas the action path is that of the moving part of the motor to the load. Their superposition yields the overall dynamics typically shown in a Bode plot.

Assuming the load mass >> motor mass, there is little left to do on the actuator side in terms of dynamics. For the indirect drive, on the other hand, one additional variable, namely the gear ratio 1: *i*, opens up a field of possibilities:

$$k_p = (M/i^2 + m)(2\pi \cdot f_{bw\text{-ID}})^2$$

$$k_{p\text{-ID}} = k_p \cdot i^2 = (M + m \cdot i^2)(2\pi \cdot f_{bw\text{-ID}})^2$$

By increasing either the motor mass or the gear ratio, the servo stiffness as 'felt' by the load can be increased. Additionally, in contrast to the direct-drive case, this actuator carries its own weight (literally) by its own linear guiding, meaning that it will have little effect on the load dynamics. Of course, there is no such thing as a free lunch; additional mass will have to be moved and the own guiding and mass will introduce its own dynamics that might become dominant instead of the load dynamics. Under the hypothesis that this is not the case (i.e. drive train and guiding are infinitely stiff), the potential of this additional variable will be demonstrated, assuming a motor mass of 3 kg for both types of actuators, a 100-kg mass, a bandwidth of 25 Hz and a gear ratio of 1:5.

$$k_{p\text{-DD}} = 2.5 \cdot 10^6 \text{ N/m}$$

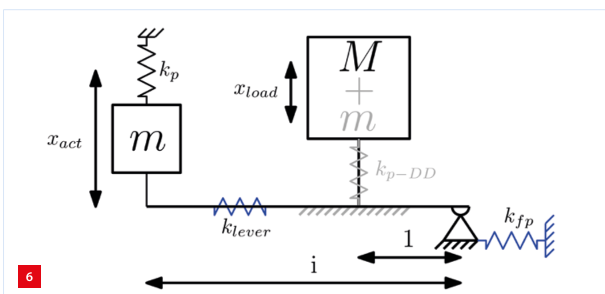
$$k_{p\text{-ID}} = 4.3 \cdot 10^6 \text{ N/m}$$

$$k_{p\text{-ID}} / k_{p\text{-DD}} = 1.7$$

Thus, introducing a gear ratio does not only amplify the force output but the servo stiffness as well. The above example is merely theoretical and in practice increasing the servo stiffness will prove to be more unruly, partly because stiffnesses like k_{lever} and k_{fp}, which have not yet been discussed, will play a role in the design of the indirect-drive actuator.

Action and reaction path dynamics

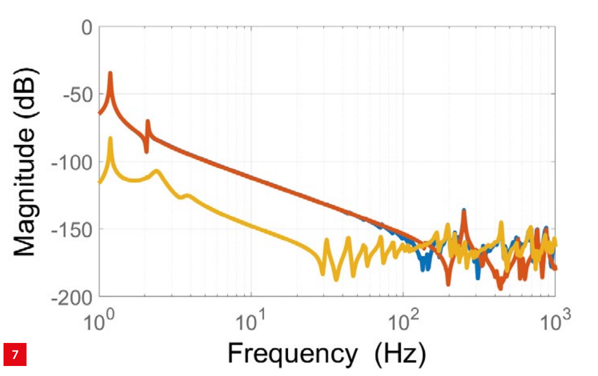
An additional differentiator between the two types of actuators are the relative amplitudes of the action path with respect to the reaction path (i.e. the amplitude difference).



The combined result of a direct-drive (grey) and indirect-drive (black) system can be seen here, with the relevant mechanism stiffnesses (blue) that were omitted from the earlier figures. These stiffnesses become relevant when looking at the action/reaction path dynamics, because in reality the hinge (on the right) is not a 'fixed' world, but a relatively large mass with its own dynamics that may influence the load.

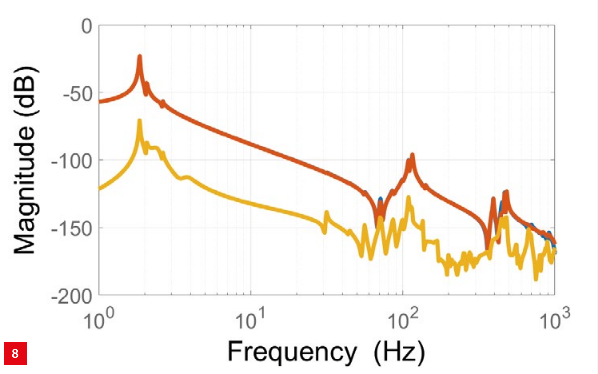
In Figure 7, the action, reaction and total dynamics (force-to-displacement) of the direct-drive actuator in the total system (load and environment) is plotted. This shows that at higher frequencies (> 100 Hz) the reaction path becomes dominant and will limit the maximum achievable bandwidth. However, these reaction path dynamics are often beyond the scope of the design, so limited influence can be exerted on them.

Figure 8 shows the same transfer functions for the indirect-drive actuator in the total system. Here, the reaction path is not dominant due to the higher relative amplitudes of the action path with respect to the reaction path. From a motor perspective, this can be intuitively explained by the load mass being reduced by the square of the gear ratio, resulting in a 'lifted' action path (see absolute numbers of magnitude in Figures 7 and 8). However, the reaction path is lifted too, because there has to be a force balance (e.g. 1 N on motor side is 5 N on load side) that results from the pivot point taking a large share of this balance. Since this effect is linear, there is ultimately a gain in the relative amplitudes. Here, the achievable bandwidth is limited by the dynamics of the actuator itself, more specifically the introduced compliancy of the drivetrain stiffness (i.e. the stiffness of the lever is not infinite as was hypothesised in the previous subsection).



7

Direct-drive and load dynamics, action (red), reaction (yellow) and total path (blue) transfer functions (force-to-displacement) plotted in a Bode magnitude diagram.



8

Indirect-drive and load dynamics, action (red) reaction (yellow) and total path (blue) transfer functions (force-to-displacement) plotted in a Bode magnitude diagram.

Whether this is an advantage or not depends on the design scope. When all dynamics in the action and reaction path are in control it might not be necessary to decouple the two. However, as in most cases, only part of the design is within scope here. In this case, this concerns the actuator and the load design; other dynamics are out-of-control and uncertain. The decoupling of the reaction path from the action path is therefore considered as a benefit here, since this makes it more insensitive to frame dynamics.

Pros and cons

To sum up, Table 1 lists the pros and cons of both drive concepts for this use case. Some characteristics will turn out better or worse in a different load case. To further analyse the differences between the direct- and indirect-drive concepts for this load case, both concepts have been designed, built and tested.

Design

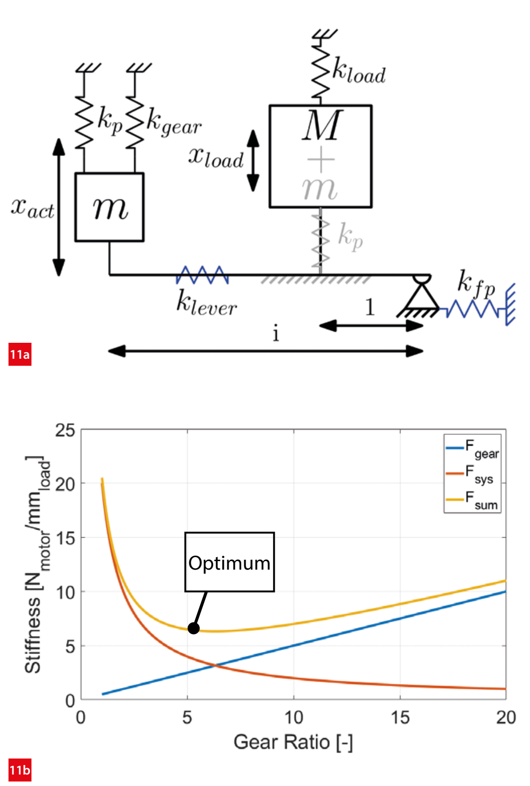
The following designs for the direct- and indirect-drive actuators are functional models and lack the details and maturity of a prototype. However, given that both concepts have been developed to this level, they provide a fair comparison. The design of the direct-drive actuator (Figure 9) consists of few elements, namely: motor(s), encoder(s), a body (i.e. moving platform) holding both motor and encoder, and the fixed world. In both actuators, the 'fixed world' holds the coils and is actively cooled to limit the motor temperature. The 3-phase motor shown here consists of two motors that are – parallelly – electrically connected, making it a single motor for the amplifier. For this to work, the alignment of the coils with respect to the yokes and with respect to each other is important, since misalignments will result in motor constant losses. However, this is a cosine effect, so small misalignments (a small percentage of the magnet pitch) are acceptable. A double encoder system is

Table 1
Pros (+) and cons (–) of the direct-drive and indirect-drive concept.

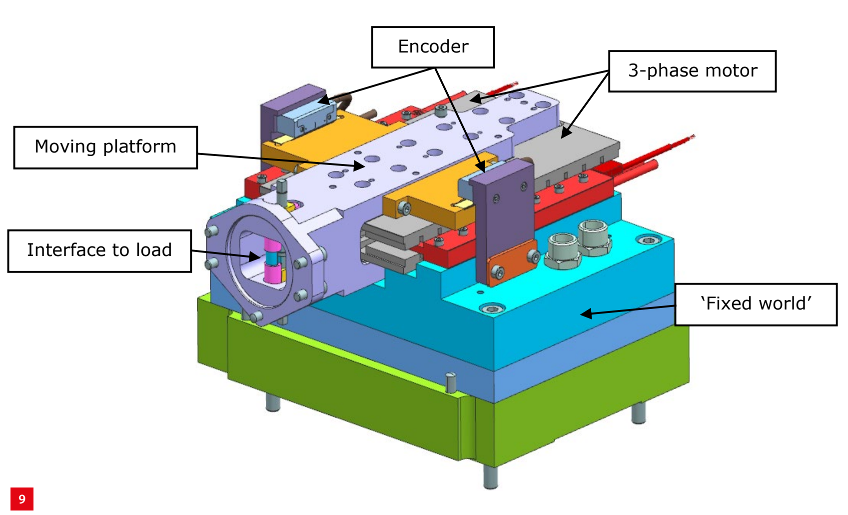
Direct drive	Indirect drive
+ Low part count	– Complexity / internal guiding
– Control bandwidth limited by reaction path dynamics	+ Insensitive to reaction path dynamics
+ High drivetrain stiffness	– Additional mechanics in drivetrain
	+ Parasitic movements can be decoupled from motor
+ Inherently linear and backlash- and hysteresis-free drivetrain	
– Force requirements determine motor size	+ Motor size and gear ratio can be tailored to force requirements (more knobs to turn)
– Dynamic stiffness function of servo stiffness only	+ Dynamic stiffness function of servo stiffness and gear ratio

Gear ratio

When working with a gear ratio that is accomplished with the use of an elastic lever, not every ratio is beneficial for the force budget. Increasing the ratio will reduce the required motor force to move the load. However, this yield diminishes as the ratio increases and will flatten out after a certain ratio has been reached. In contrast, the motor force required to move the actuator itself will increase proportionally, eventually leading to a system where more force is required to move the actuator than the load itself. In Figure 11a two additional stiffnesses are shown, k_{gear} and k_{load}, representing the stiffness of the load and actuator in the actuation direction, respectively. Figure 11b shows the trade-off between the linear and nonlinear effects. For this particular case, a ratio of 1:5 was chosen as an optimum, also taking into account volume, stresses, dynamics, etc.



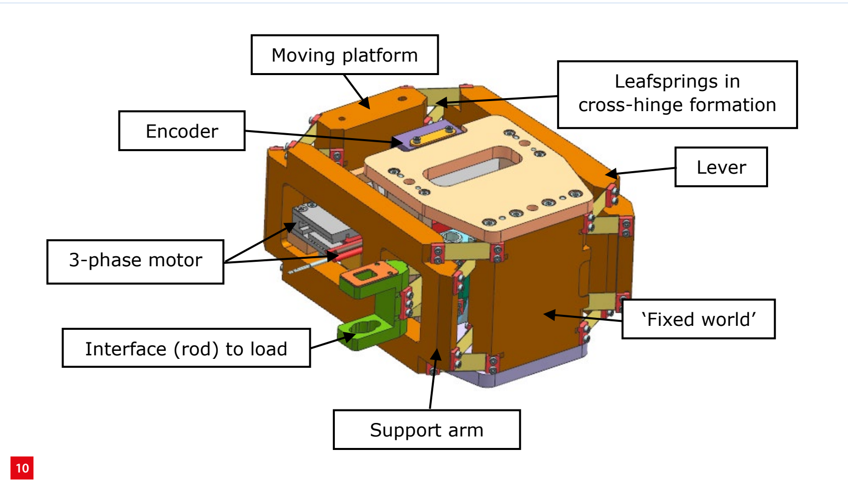
Finding the optimal gear ratio for the force budget.
(a) The system with stiffnesses on the load side (k_{load}) and on motor side (k_{gear}).
A ratio is chosen that results in the lowest motor force.
(b) The optimisation in the graph with the exponential curve of increasing the ratio for load stiffnesses, versus the force required to compensate the linear increase in motor stiffness, due to the increasing stroke on the motor side.



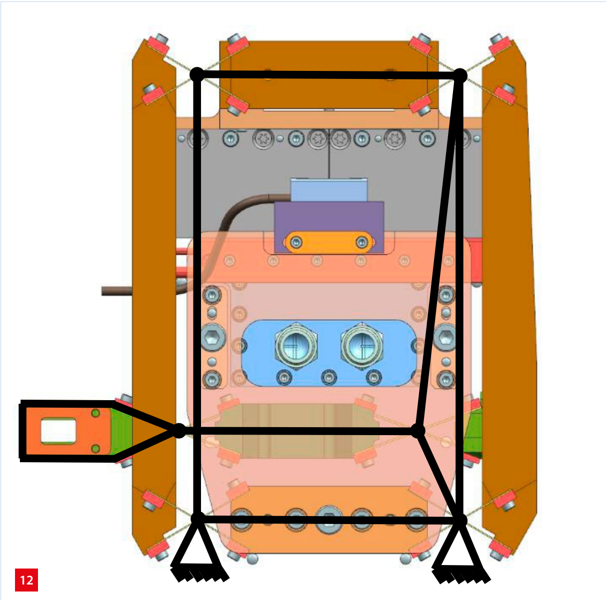
Direct-drive functional model design, with the two encoders enabling 'virtual encoder placement' along the central axis. Coil-to-magnet positioning depends on the frame/system environment.

present in this motor, so that the encoder can be on either side of the motor or virtually in the middle. This is not of significance for the actuator discussion and will not be elaborated in more detail here.

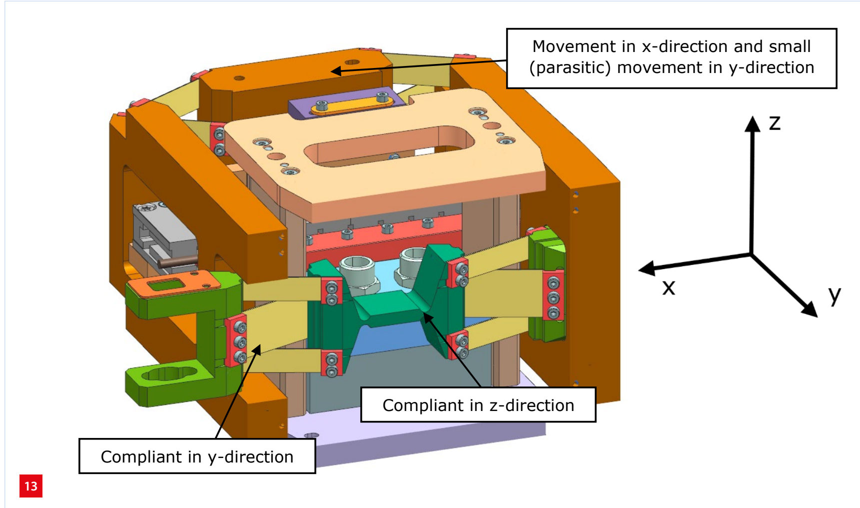
The design of the indirect-drive actuator requires a more elaborate explanation. It is obvious from Figure 10 that this design is more complex than its counterpart. As in the previous concept, this one also contains a motor, encoder, moving platform and fixed base. Additionally, there is the lever, the elastic pivot point(s) in the form of a cross-spring hinge and an interface rod. In this design a ratio of 1:5 is used (the optimisation is presented in the text box). Here, the rear body (or lever) acts as the force transfer and transmission. The front body (or support arm) does neither and only functions as an additional support pillar to complete the parallelogram shown in the next section; because of this it can be much leaner than the functional lever, if permitted by the dynamics.



Indirect-drive actuator functional model design.



The indirect drive as a parallelogram fulfils the same function as the simpler seesaw but with some practical benefits. In this top view of the actuator, the virtual hinges of the 'skeleton' (black) coincide with the pivot points of the cross-leafspring hinges. The connecting mechanics are shown as rigid bodies connecting the pivot points.



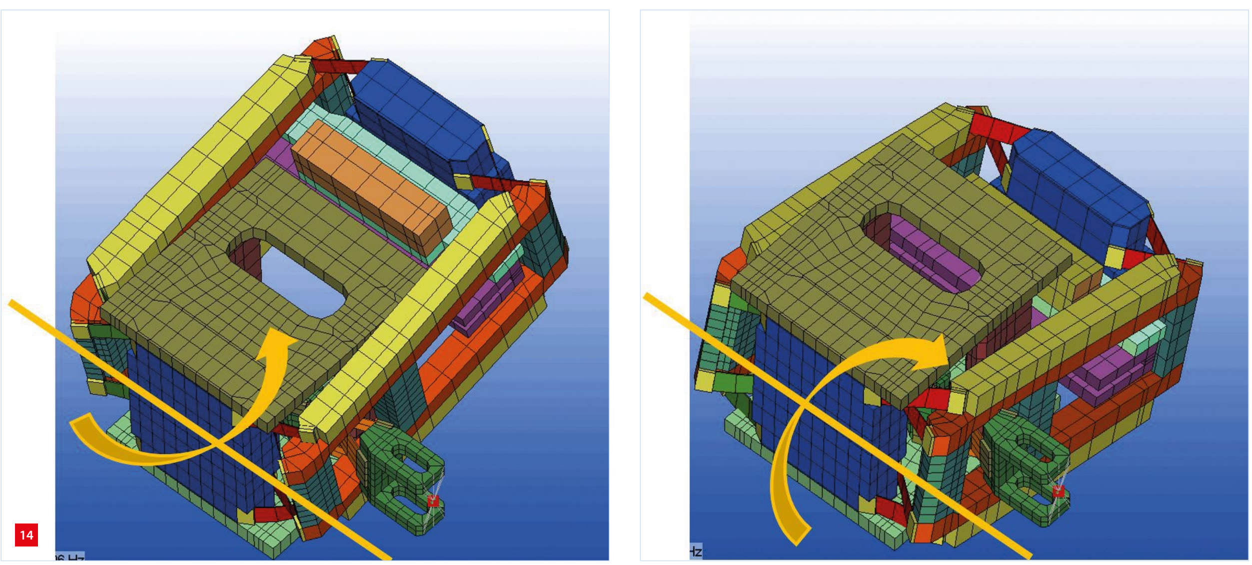
Section of the indirect-drive actuator, showing that the interface rod is compliant in two DoFs (y and z), decoupling the parasitic load movements from the motor.

Parallelogram

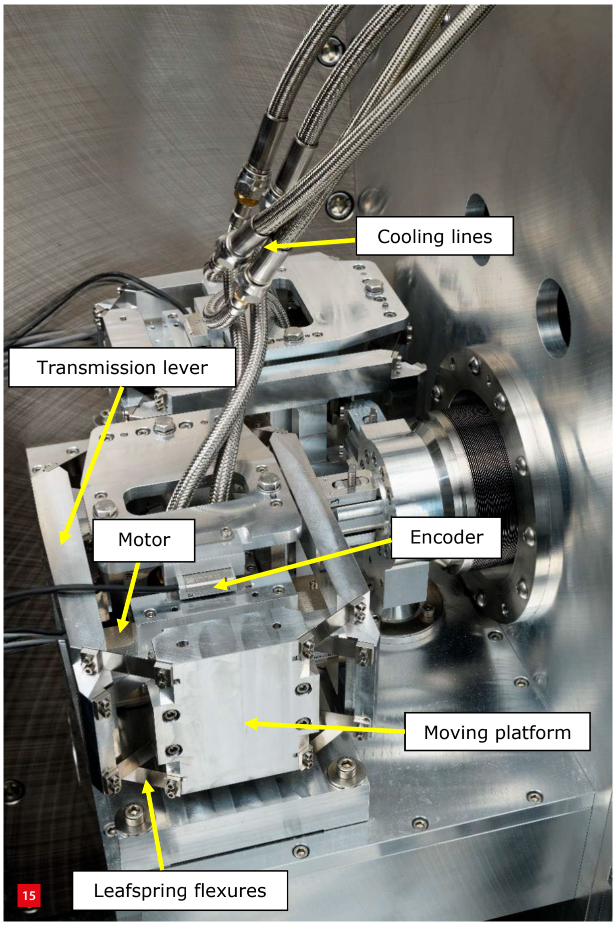
To create a functioning indirect drive the simple seesaw from Figure 5 was elaborated into a parallelogram as shown in Figure 12. The pivot points are in the centre of the cross-spring flexure and the moving platform is situated further from the fixed base hinges than the connecting interface rod. In fact, this difference in distance is 1:5, creating the intended ratio. From this image it can also be deduced why the support arm (left) has no function as a lever or force transmission: as the interface rod protrudes through this arm it is not in the force path. The movement of the parallelogram is not perfectly linear, but instead describes an arc with a radius equal to the lever length. Because of this, the magnet will move laterally over the coil as a function of the actuation position. This effect is smaller than the parasitic movements induced by the load.

Interface rod

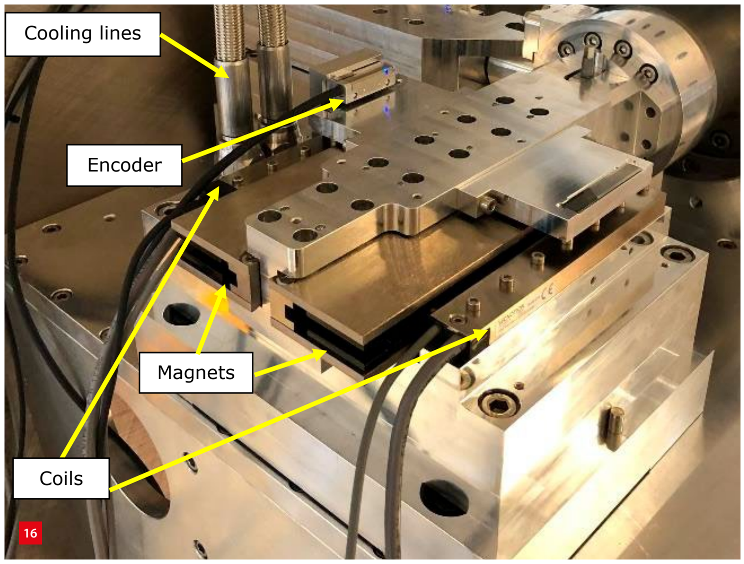
The lateral movement of the magnets with respect to the coil will be smaller than 1 mm, whereas in the original load case – and this is what the direct drive has to cope with – there is a parasitic movement up to 4 mm. The interface rod is compliant in two DoFs as shown in Figure 13, the leafsprings being compliant in the *y*-direction and the middle body in the *z*-direction. This allows the load to move in two directions and also allows for the parasitic movement of the flexures of the load in the *z*-direction. This completely decouples these parasitic movements from the motor and will only result in a small force.



The first eigenmode is the motor rotating around the x-axis (orange line).



Realisation of the indirect drive actuators, here showing two identical actuators for x and y_1.



Realisation of the drive actuators, here showing the actuator for the x-direction.

Dynamic considerations

To understand the geometry of the indirect drive it is necessary to consider the dynamics of the actuator. In the section on the servo stiffness of the actuators, the dynamics of the indirect drive were assumed to be infinitely stiff. In reality this is obviously not the case, nor do they need to be; they just have to be non-critical for the overall system. For the parallelogram implementation there are two parasitic eigenmodes that deserve some attention, the first being the rotational mode around the base as shown in Figure 14. The motor mass is a given, as is the arm on which it sits, so what remains is to increase the stiffness of the arms; this explains the height of these bodies as well as the wide placement of the leaf springs. In the second eigenmode the load and the motor move in anti-phase; this is due to the limited drive train stiffness and is a direct consequence of an indirect-drive concept. This can be improved, i.e. raising the frequency of the second eigenmode, by tuning the stiffness of the lever and the interface rod, and has been here to the point where the second eigenmode is no longer a limiting factor.

Lastly, there is an eigenmode that is not necessarily inherent to the parallelogram configuration but to the type of flexures. In the case of a normal reinforced leafspring, there is little to no lateral stiffness, but there is significant mass in the middle of the flexure. This typically starts resonating at much lower frequencies than when used with cross-spring flexures, because these do provide the lateral stiffness. This explains why the design described here contains cross-spring flexures.

Realisation and conclusion

Figures 15 and 16 show the realised functional models of the actuators, on which much of the data and insights discussed in this article are based. Returning to the initial question: direct or indirect drive? The answer to this is more nuanced and less trivial than common practice may suggest. In this specific case it was shown that even though direct drive is definitely not infeasible, indirect drive does have some clear and non-intuitive advantages, making it the clear winner for this case.